TRANSPORT AND TELECOMMUNICATION

Volume 5. N1

Proceedings

of the International Conference

RELIABILITY and STATISTICS in TRANSPORTATION and COMMUNICATION (RelStat'03)

Volume 1

2004

Transporta un sakaru institūts
(Transport and Telecommunication Institute)

TRANSPORT and TELECOMMUNICATION

Volume 5. N 1 - 2004

ISSN 1407-6160

Proceedings

Volume 1

International Conference

RELIABILITY and STATISTICS in TRANSPORTATION and COMMUNICATION (RelStat'03)

16–17 October 2003. Riga, Latvia

Organised by

Transport and Telecommunication Institute (Latvia)
The Kh. Kordonsky Charitable Foundation (USA)

in co-operation with

Latvian Transport Development and Education Association (Latvia)
Latvian Operations Research Society (Latvia)
Latvian Academy of Science (Latvia)
Air Navigation Service Company ALTA (Latvia)

ISBN 9984-668-65-7

Riga – 2004

Proceedings of the International Conference RELIABILITY and STATISTICS in TRANSPORTATION and COMMUNICATION (RelStat'03), 16–17 October 2003. Riga, Latvia. – Riga: Transport and Telecommunication Institute, 2004, Vol. 1. – 232 pages.
ISBN 9984-668-65-7

TRANSPORT and TELECOMMUNICATION, 2004, Vol. 5, No.1
ISSN 1407-6160
The journal of Transport and Telecommunication Institute (Riga, Latvia).
The journal is published since 2000.

The aim of the journal is to acquaint the readers with the national experience and research activities of various countries in the fields of transport and telematics as well as with the working materials of the European thematic networks and the research projects in this area.

This special edition is devoted to the International Conference "Reliability and Statistics in Transportation and Communication" (October 17–18, 2002, Riga, Latvia). The edition includes texts of conference reports written by the authors in a special journal version.

The authors have submitted the material in the English or Russian in accordance with the official languages of the Conference. Therefore, for the convenience of readers the editorial board has divided the selected materials into two editions – English (Vol.5, N1, 2004) and Russian (Vol.5, N2, 2004).

The authors themselves are responsible for context and grammar of their abstracts and reports.

Edited by: Igor V. Kabashkin

ISBN 9984-668-65-7

CONTENTS

Session 3. Computing Mathematics

Session 4. Management and Economics of Transport

Session 5. Reliability of Mechanical and Aviation Systems

Session 6. Efficiency of Electronics Systems and Devices

Poster Session

PROGRAMME

of the International Conference

"RELIABILITY and STATISTICS in TRANSPORTATION and COMMUNICATION (RelStat'03)"

16–17 October 2003. Riga, Latvia

Thursday, October 16

Tutorial day

Time	Transport and Telecommunication Institute Aud. 703
10:30–12:00	Chaos, fractals, the "neurofinancial theory" and quantum financial mathematics in the new risk management paradigm. *Mikhail Rogov* (Russia, International State University "Dubna")
12:15–13:45	The use of formal methods for integrated analysis of telecommunication and logistic systems. *Henrikas Pranevicius* (Lithuania, Kaunas Technical University)

Friday, October 17

8:30–9:30	**Registration** Air Navigation Services, 3 Floor, Foyer
9:30–11:00	**Opening Session (**Air Navigation Services, 3 Floor, Main Hall) Session Chair – *Igor Kabashkin* (Latvia) • *Igor Kabashkin,* Chairman of the Organising Committee, Vice-Rector of the Transport and Telecommunication Institute (Latvia) • *EugeneKopytov*, Rector of the Transport and Telecommunication Institute (Latvia) • *Inna Kordonsky-Frankel*, President of the Kh. Kordonsky Charitable Foundation (USA) • *Peter Maximilian de Heidendorf*, President of European Board of Trade (Austria) • *Edmundas Zavadskas,* Vice-Rector of Vilnius Gediminas Technical University (Lithuania) **Plenary Session** (Air Navigation Services, 3 Floor, Main Hall) Session Chair – *Igor Kabashkin* (Latvia) • Reliability Theory: with Applications to Preventive Maintenance. Presentation of new monograph by I. Gertsbakh. *Vladimir Rikov* (Russia) • Transportation Security Response, Reliability, Safety and Performance Issues. *Ernst G. Frankel* (USA)
11:00–11:30	**Coffee Break**
11:30–13:00	**Thematic Sessions**
13:00–14:00	**Lunch**
14:00–15:45	**Thematic Sessions**
15:45–16:15	**Coffee Break**
16:15–18:30	**Thematic Sessions**
18:45	**Closing Session**

Time	Hall #1	Hall #2	Hall #3
11:30-13:00	**Session 1. Transport** Session Chair – *Alexander Pankov,*	**Session 2. Statistical Applications** Session Chair – *Irina Yatskiv*	**Session 3. Computing Mathematics** Session Chair – *Eugene Kopytov*
	Model of Intelligent Transport System Development for City Agglomeration. *Igor Kabashkin* (Latvia) Mathematical Models with Application of Statistical Information in Transportation Management Process. *Adolfas Baublys* (Lithuania) Monte-Carlo Simulation of Resource Planning in Supply and Transportation. *Leonidas Sakalauskas* (Lithuania) Analysis of Resource Cycles in Marine Container Terminals. *Andrey Solomennikov, Vladimir Bardatchenko, Yuri Merkuryev, Fred Kamperman* (Latvia) Investigation of the Lithuanian Road Traffic Structure. *Alvydas Pikūnas, Vidmantas Pumputis, Vigilijus Sadauskas* (Lithuania) New Priority Ways in International Freight Transport and Estimation of Business Strategy of the Lithuanian Customs. *Aldona Jarasuniene* (Lithuania)	Parameter Estimation and Hypotheses Testing for Nonhomogeneous Poisson Process: Part 2. Numerical Examples. *I. B. Frenkel, I. B. Gertsbakh and L. V. Khvatskin* (Israel) Applications of Resampling Approach to Statistical Problems of Reliability Systems. *Alexander Andronov, Maxim Fioshin* (Latvia) Renewable Networks: Sequential Bounds on the Interval Reliability. *Y. Shpungin* (Israel) Preventive Maintenance Policy For Lamp Replacement Problem. *I. B. Frenkel, L. V. Khvatskin and A. Shendelis* (Israel) Application of the Packet Means for the Tasks of Regression Models Parameters Estimation. *Viktor Lumkis* (Latvia) Validation Methods of Cluster Decisions. *Lada Gusarova, Irina Yatskiv* (Latvia) Combinatorial and Probabilistic Properties of M. V. Lomonosov's "Turnip". *I. Gertsbakh, Y. Shpungin* (Israel)	To Over-Reliability Through Adequacy for Stochasticity Nature of Future. *A. V. Rutkauskas, V. Rutkauskas, L. Kaleininkaite* (Lithuania) Solution Making Support System under the Conditions of Uncertainty. *Edmundas Zavadskas, Leon Ustinovich* (Lithuania) Investigation of a Two-Dimensional Retrospective Reverse Thermal Conductivity Task. *Sharif Guseinov* (Latvia) Exact Analytical Solution and Some of its Approximations for Two-Dimensional Systems with Rectangular Fin. *A. Buikis, M. Buike, Sh. Guseinov* (Latvia) Invariant Problems in the Control Discrete Systems under the Regular Chaos Conditions. *Valery Nikolsky, Dmitry Pliachko* (Latvia) Control over Tasks Staying Time in One-Channel Mass Service System. *Valery Nikolsky, Alexander Zinoviev* (Latvia) Optimization of Aircraft Route for Assurance of Arrival Time. *Dalia Ziliene, Jonas Stankunas* (Lithuania)
13:00-14:00	**Lunch**		

Time	Hall #1	Hall #2	Hall #3
14:00-15:45	**Session 1. Transport** Session Chair – *Igor Kabashkin*	**Session 4. Management and Economics of Transport** Session Chair – *Rostislav Kopitov*	**Session 5. Reliability of Mechanical and Aviation Systems** Session Chair – *Vitaly Pavelko*
	Inventory Control Model for the Typical Railways Company. *Eugene Kopytov, Fedor Tissen, Leonid Greenglaz* (Latvia) Application of Functions with Flexible Structure for Forecasting of Freight Turnover in Logistical System. *Mindaugas Mazura, Olga Fadina* (Lithuania) Methodology for Multimodal Network Analysis by Using Concept of Virtual Links and Nodes. *Aidas Vasilis Vasiliuskas* (Lithuania) Application of Virtual Data Models in the Analytical System of the Latvian Railway. *Vasilijs Demidovs* (Latvia) Principles of Development of Complex Security Systems in Relational Databases on the Example of Latvian Railway. *Nataly Petoukhova* (Latvia) Investigation of the Possibility to Establish a Logistics Centre in Vilnius Region and its Potential Benefits. *Leonas Povilas Lingaitis, Olga Fadina* (Lithuania) Products Distribution Optimisation in the Logistics Network. *Ramunas Palsaitis, Kestutis Vislavicius* (Lithuania)	Cost of Business as an Advantageous Indicator of Assessing the Judgement of Shareholders about Keeping the Managers' Rights and Responsibilities. *Rostislav Kopytov* (Latvia) Development of a Complex System Providing Efficient Running of an Enterprise from the Point of View of Cost Increasing under the Conditions of Risk and Uncertainty. *Eugene Kolesnichenko* (Latvia) Impact of Electronic Commerce on Forming Competitive Advantages. *Daiva Novikoviene* (Lithuania) Diagnostics of Real Estate Object Cost Components. *Rostislav Kopytov, Alexander Krasov* (Latvia) Information Technologies and Public Procurement. *Vita Narnicka* (Latvia) Application of Self-Organizing Concepts in Organization Structures Modelling. *Roman Zakharov* (Latvia) Distribution of Market Shares Among Competitors in Lithuanian Transport Sector. *Jonas Lazauskas* (Lithuania)	Influence of Turbo-Prop Engine PW125B Parameters Measuring Errors on the Diagnostics Reliability of its Setting Condition. *Eugene Kopytov, Vladimir Labendik, Arvids Osis* (Latvia) Inspection Policies for Detection of Initial Cracks in Aircraft Structural Components. *Konstantin N. Nechval* (Latvia) The Design of Low Speed Wing Section for Safety and Performance. *Eduardas Lasauskas* (Lithuania) Estimation of Parameters of Fatigue Curve of Composite Material. *A. Yu. Paramonova, M. A. Kleinhof, Yu .M. Paramonov* (Latvia) The Investigation of Aerobatic Pilot's Fatigue. *Darius Ereminas* (Lithuania) The Statistical Criterion of Fatigue Crack Detection by Intensity of an Acoustic Emission. *Vitaly Pavelko, Erik Ozolinsh* (Latvia) Influence of Corrosion on the Required Number of an Airframe Inspection. *M. Shujauddin Wahab, Yu. M. Paramonov* (Latvia)
15:45-16:15	**Coffee Break**		

Time	Hall #1	Hall #2	Hall #3
16:15-18:30	**Session 1. Transport** Session Chair – *Leonidas Sakalauskas*	**Session 4. Management and Economics of Transport** Session Chair – *Ilia Frenkel*	**Session 5. Reliability of Mechanical and Aviation Systems** Session Chair – *Yuri. M. Paramonov*
	Evaluation of the DME Sitting Factors with Application to Palanga Airfield. *Algimantas Jakucionis, Romualdas Malinauskas, Jonas Stankunas* (Lithuania) Statistical Evaluations of Automobile Separate Models Reliability. *Valentinas Mickūnaitis* (Lithuania) Urban Traffic Flows Measurement in the Automated Centralized Transport Control Systems. *Alexandr Berezhnoj* (Latvia) Bus Research of the Reliability in the Joint Stock “Vilniaus Autobusų Parkas”. *Aurelijus Vestartas, Saulius Nagurnas* (Lithuania) Investigation of Dynamic Processes in the Hydraulic Braking System of Transport Vehicle with Antilock Brake System. *Marijonas Bogdevicius, Oleg Vladimirov* (Lithuania) Definition of a Trajectory of the Vehicle Movement, Taking into Account Interaction of its Wheel and Road Surface. *Olegas Prentkovskis, Marijonas Bogdevicius* (Lithuania) Operational Research of Front Mounting Reliability of „KAROSA“ Bus. *Kęstutis Lukoševičius, Egidijus Diržys, Nijolė Lukoševičienė* (Lithuania) Fleet of Lithuanian Road-transport means and actions on its renovation. *Alvydas Pikūnas, Valdas Valiūnas* (Lithuania)	Special Financing Form and Possibilities for Companies. *Peter Maximilian de Heidendorf* (Austria) Simulation Approach to the Weekend Effect's Problem. *B. I. Frenkel* and *D. Alberg* (Israel) Dynamics of Notions Mastering the Course "Educational Management". *Agris Upenieks* (Latvia) Social and Economic Principles of Developing an Integrated Public Transport Network in Vilnius. *Daiva Griskeviciene*, *Algirdas Griskevicius* (Lithuania) Diagnostics Concepts of Company Economical Activity on the Basic of its Cost Measurement Taking into Account Dynamics of Inner Processes. *Victor Siperkovsky* (Latvia) Research of Passenger Transportation Market in Lithuania. *Ionas Butkiavichus* (Lithuania) Enterprise Management with Franchising Advantages Application *Yury Senuto* (Latvia) Business Incubator as a Mean of Company Stable Development. *Cyril Petsevich* (Latvia) Review of Poster Session ▪ *Martin Richard G. Kristek* (Germany)	The Investigation of a Possible Trajectory of the Fatigue Crack Growth in Structure of Helicopter Mi-8. *Vitaly Pavelko, Jelena Goncharova* (Latvia) Aircraft Gas Turbine Engine’s Technical Condition Identification Technique. *A. M. Pashaev, R. A. Sadiqov, P. S. Abdullayev* (Azerbaijan) The Influencing of Sizes on Fatigue Durability of Sheet Details. *Vitaly Pavelko, Julia Timoschenko* (Latvia) Cross-Ply Bending And Stability of the Compressed Rod with Crack. *Igor Pavelko, Vitaly Pavelko* (Latvia) Application of Statistical Methods for Quality Evaluation of Asphalt Concrete Mixture Production. *Henrikas Siviliavichus* (Lithuania) Influence of Variable Stiffness of Cable for traction to vibrating parameters and Reliability of System "Trailed Vibrating Roller and Cable for Traction" Operating on the Inclined Surface. *Vilius Bartulis* (Lithuania) Review of Poster Session ▪ *Nickolas A. Nechval* (Latvia)
18:45	**Closing Session.**		

Time	Hall #4
14:00-15:45	**Session 6. Efficiency of Electronics Systems and Devices** Session Chair – *Yuri Shunin*
	Analysis of Difference Schemes in Modelling of Gyrotron equations. *Olgerts Dumbrajs* (Finland), *Harijs Kalis, Andrejs Reinfelds* (Latvia) Chaotic Phenomena in Nanostructures of Semiconductor Electronics. *Yu. Shunin, V. I. Gopeyenko, A. V. Gopeyenko* (Latvia) Exploratory Signal Analysis Using Fractal Dimension. *Alexander V. Grakovski, Alexander I. Alexandrov, Roman Kivlyonok* (Latvia) Homomorphy Suppression of Noise in Pulse Signals on the Basis of Decomposition on Elementary Waves. *Yuri Krasnitsky* (Latvia) Phase Filters with High Linearity of PFC. *Takhir Mamirov* (Latvia) Techkor Wireless Condition Monitoring System. *Victor Boytsov, Andrey Volkov* (Latvia) Recursive Polyphase Filters in Multi-Channel Signal Filtration. *Eduard V. Matosov* (Latvia) Studies of the Theory of Probability in Electronics and Telecommunications Specialists Training at TTI. *Oleg Schiptzov* (Latvia)

Main Hall

Poster Session

Finding Efficient Estimators on Small Data Samples*. Edgars K. Vasermanis, Nicholas A. Nechval, Uldis Rozevskis & Kristine Rozite* (Latvia)

Optimal Ordering Policy for Multi-Period Version of Inventory Control Problem. *Max Moldovan, Konstantin N. Nechval, Nicholas A. Nechval, Edgars K. Vasermanis, Uldis Rozevskis & Kristine Rozite* (Latvia)

Technique for Deriving Shortest Statistical Intervals. *Nicholas A. Nechval, Konstantin N. Nechval, Edgars K. Vasermanis* (Latvia)

The Distribution of Conditional on Realised Volatility Financial Returns. *Max V. Moldovan, Andrew C. Worthington, Nicholas A. Nechval and Hellen Higgs* (Latvia)

The Simulation-Free Simple Test for Normality: Some Comparisons of Size and Power. *Max V. Moldovan, Konstantin N. Nechval and Nicholas A. Nechval* (Latvia)

Health, Security and Environment (HSE) Engineer – a New Education Form for Latvian Business. *Martin Richard G. Kristek* (Germany)

HSE and the Flight and Transport Security. *Martin Richard G. Kristek* (Germany)

HSE Management System and the Conform Way the Right and Liability in the European Union. *Martin Richard G. Kristek* (Germany)

TRANSPORTATION SECURITY RESPONSE, RELIABILITY, SAFETY AND PERFORMANCE ISSUES

Ernst G. Frankel

Massachusetts Institute of Technology, Department of Ocean Engineering
77 Massachusetts Avenue, Building 5-222, Cambridge, MA 02139-4307, USA
Phone: 617 253 6763, Fax: 617 253 8125, E-mail: efrankel@mit.edu

Abstract

Since the introduction of technology, its performance, reliability and safety have been our major concerns. Performance is expressed in terms of usefulness effectiveness, and costs; reliability in teams of instant and uptime availability to perform; and safety in terms of its ability to assure user, operator, and the environment protection from any ill effects. These three major characteristics of technology were improved by using design and operational data to identify opportunities for advancement, and today we can generally assume that technology, and particularly transport technology, will perform to its design standards, achieve a high level of reliability, and be safe in use for users and the environment in which it works.

In recent years we were forced to confront anew issue not of serious concern before - security of and in the use of technology. While the basic concerns mentioned above were all dependent on the design, manufacture, use, operation, and the environment in which the technology worked and were therefore imbedded reactive characteristics, security is an externally imposed danger. Protection against security impacts cannot be provided by making technology more efficient, reliable or even safe, bur by introducing detection, barrier, fail proofing, and various proactive measures designed to protect the technology from purposely disruptive and destructive actions.

In this presentation we will discuss approaches to proactive and preventative measures under development designed to improve the security of transportation. We will note that these measures cannot be based on historic performance but must consider an array of potential interventions.

1. INTRODUCTION

Since the terrorist attacks of September 11th, security in transportation has become an ever serious concern and both national governments and international organizations have introduced major policies and requirements to assure safer transportation of goods and people. Although basic safety measures were in place before, most of these were designed to respond to problems introduced by physical faults, acts of God or human errors and not terrorism. The new phenomena of purposeful attacks that often include suicide of the perpetrator need a different approach to which organizations and agencies are only now trying to respond in a meaningful, organized way in the U.S. Intelligence, law enforcement, and security activities by government have traditionally been widely dispersed among numerous federal and local agencies with little coordination or information sharing among them. In fact, many of these agencies worked at cross-purposes.

International enforcement of safety and security measures is mainly by agencies of the United Nations such as the International Maritime Organization, UN Conference on Trade and Development, World Health Organization, and more.

Another issue is the development of safety and security requirements by countries that often reorganize their national security apparatus now. I will attempt to present the status of policies, initiatives, and measures current at this time, fully recognizing that the whole process is in a flux and subject to change as governments, organizations, and agencies at all levels and under various jurisdictions gain experience and converge on enforceable, defined rules, requirements, and measures.

1.1. Security Threats

General

Transportation security is a global confidence, tracking, inspection, and control problem that requires technological hardware and software as well as sophisticated confidence, personal evaluation, and effective background and current performance checks. Enormous relations and object oriented data bases are required, supported by reliable real time physical checks and tracking as well as effective analytical software for such problems including order entry checking, final destination, end use or user confirmation, driver, vehicle, and cargo consistency verification, cargo code standard verification, and more. Although physical threats dominate there are links between psychological and confidence measures as well as physical tracking and inspection. A global system for transportation security must include supply chain wide security measures from origin to destination and integrate cooperating standardized networks of confidence checks, tracking, inspection, and control procedures and technology use. Current weaknesses and needs for development of confidence building and tracking or inspection technology are identified and potential approaches to the solution of current gaps or problems are proposed, including increasing concerns with

- Piracy incidents are becoming more frequent. The connection between piracy and terrorism is alarming. Many actions can be taken to present acts of piracy and deter attacks on vessels.
- Harbor robbery incidents in many underdeveloped countries are becoming more frequent. Many actions can be taken to prevent acts of robbery and deter attacks on vessels and their cargo, including cargo transiting ports.
- The problem of stowaways is not new but the scope of the current problem is. The possibility of terrorist stowaways using ships for infiltration is even more alarming, particularly on containerships.
- Airports, docks, and harbor piers are the easiest way to access restricted areas. For some reason, one assumes that the water creates a well-protected barrier. The opposite is true, since water creates opportunities for stealthy penetration.
- Container cargo security has become a focal issue. The potential of import of nuclear, biological, and toxic materials and agents is on the increase and effective methods of detection and neutralization must be found and introduced.
- Port and port terminal security has become an important issue and effective means for boundary protection and area security must be developed to prevent sabotage and other criminal acts.
- Port and air terminal information and communication systems must be secured and protected from unauthorized access and manipulation.
- Port, air terminal, aircraft, and ship personnel and others with legitimate access to ports, aircraft, and ships must be more effectively scrutinized and access effectively controlled, and restricted to only those thoroughly checked, identified, and admitted to areas and facilities for legitimate purposes. Access and regress must be controlled, recorded, and reviewed.
- Port and airport equipment and facilities are complex, large, and expensive. Access, particularly to major equipment, must be strictly controlled and managed. Only specifically identified persons must have access.
- Aircraft, vehicles, and ships receive and discharge gaseous, dry, and liquid substances such as fuel, waste liquids, ballast water, etc. Effective monitors and other controls are required to assure that such transactions to not result in dangerous emissions, pollution, poisoning or environmentally harmful discharges and/or transfer of potential weapons of mass destruction.

- Shipboard and aircraft staff live and work in isolated quarters and could readily be incapacitated by gas or other toxic material intrusions designed to take over vessels or otherwise attain control of a ship. Similarly, operators of port and airport equipment are subject to potential attack and incapacitation, which can cause huge damage and/or penetration by terrorist threats.
- Aircraft high-jacking, demolishers, as well as truck or rail deviation is an increasingly important threat.

Considering security threats, more specifically as they relate to ports and port operations, the following lists of security threats and policies/technologies designed to guard against them is presented.

Air and Seaport Security Threats

1. Unsafe Ships and Aircraft: Aircraft and ships that constitute fire, explosion or toxic pollution hazards that can be triggered by terrorists, crew, as well as by external attacks.
2. Cargo: Hidden hazardous cargo in containers, boxes, holds or tanks such as nuclear, explosive, biological or toxic substances that when released or triggered cause large damage, loss of life, and an unsafe environment.
3. Air and Seaport Facilities: Penetrable, unsafe port fences, and boundaries including access from the waterside that is seldom controlled. Overflight and easy road access, high voltage lines crossing port areas, inflammable storage facilities, tanks, and other hazardous storage facilities.
4. Equipment: Easy access to critical port equipment controls. Access to rails, cables, wires, controls, and guides of major critical port equipment.
5. Air and Seaport Personnel and Persons in Ports: Ports require strict access and exit controls and effective port identification. Ports must have enforceable rules of admission and restrict all access to vicinity of critical equipment, operations, and storage except by essential people.
6. Air and Seaport Boundaries and Potential Penetrations: In addition to port fences and physical boundaries, ports are vulnerable to penetration by thrown or missive explosives, flammables, toxics, poisons, lubricants, etc. Similarly, port supplies such as power, communications, water, air, etc. can be interrupted, corrupted, etc.
7. Environmental (Air, Water, Distributed Systems) Attacks: Ports distributed systems including water, sewage, telephone, power, air, etc. are vulnerable to attack, corruption, interruption, and pollution.
8. Air and Seaport Information and Communication Security: Hand communications are today critical in port operations and management, as ports increasingly serve as info hubs for major supply chains. Much of modern port operations and equipment are digitally (electronically) controlled often using wireless LAN networks that can be penetrated by terrorists and others if not well designed and managed.

1.2. Security Measures and Policies

1. Direct threat recognition and terrorist risk analysis.
2. National laws and regulations and their impact on the possible solutions to be proposed.
3. International standards and conventions regulating marine trade and transit and their impact on the possible solutions proposed.

4. International perimeter environment - assessment of possible risks and "soft belly" areas.
5. External environment to the defined port perimeter - assessment of possible risks
6. Cargo security (incoming and outgoing)
   a. cargo profiling
   b. inventory of existing technologies covering bulk screening and sub-pressure simulations
   c. bulk explosive trace detection
   d. non-conventional (chemical, biological, and nuclear) threat
7. Aircraft and ships security
   a. ground protection
   b. air and sea protection
   c. catering and supplies screening
   d. sensitive carriers
   e. crew screening (incoming and outgoing)
   f. documentation screening and forgery detection
8. Transport personnel, worker, crew passenger security
   a. crew, port worker, passenger identification
   b. checking and security measures
   c. passenger handling procedures
   d. existing protocols
   e. existing training programs
   f. documentation
9. Access control
   a. definition of restricted areas
   b. vehicles control
   c. system for worker, crew, passenger recognition
   d. permit policy
10. Communications infrastructure
    a. examination of command and control communication facilities
    b. examination of contingency existing plans
    c. examination of existing communications protocols with security official services (police, FBI, others); identification of possible loopholes in case of emergencies.
11. Human resources
    a. examination of the human resources employed for serving the security needs of the ports
    b. training programs
12. Parking lots
13. Access roads
14. Obligatory checkpoints
15. Technology
16. Direct and indirect observations
17. Auditing policy

Data Security

a. Protection and security of the system's data itself
b. Data availability in real time
c. Vital elements running the Command and Control Center of any serious security system
d. Identifying vulnerability of data systems. Cyber terrorism and/or computers and data vandalism are real concerns.

Countering Security Threats in Ports

Steps in Developing Transport Security Management Systems

- System and specific risk assessment
- Building of protection and defense programs
- Mapping of computer systems, signaling sensitivity and access rights
- Planning and building of computer secured topology covering prevention, discovery, warning, and recovery
- Contingency plans in case of system collapse
- Simulation of cyber attacks to discover weaknesses of the system
- Simulation of physical penetration to discover weaknesses of the system
- Implementation of the program
- Auditing
- Training

2. TRANSPORT SECURITY TECHNOLOGY INITIATIVES

Governments and industry have recognized the need for new technology for transportation security. Among the technologies of concern are:

- non-intrusive luggage, cargo, and container inspection machines
- effective, secure, and reliable cargo and container tracking devices including content tracking
- container sealing devices
- container identification devices
- air and sea port boundary protection (land, air, water)
- ship hull protection against unauthorized pirate/terrorist boarding at sea, anchor or in port
- crew and port workers identification and access control
- secure wireless local port area networks – Port Info/Communications Security
- secure electronic data transmission storage and use
- biological terrorist attack prevention
- nuclear terrorist attack security
- port equipment securing devices
- security of port power, water, bunker supply, etc. systems

One of the issues of concern is that shippers in, and shipping to, the U.S. will be required to affix a high security manual or electronic seals to containers immediately upon the conclusion of the container stuffing process. An internationally accepted standard for such seals has been approved, and C-TPAT ocean carriers are encouraging their customers to affix such seals.

Carriers support expanding such a practice from a voluntary practice to a government requirement (Table 1).

Table 1. Container Seal Technology

Group	Technology
Sftvi Technology, Sunnyvale, CA	Electronic seals and smart cargo containers
NaviTag Technologies, North Quincy, MA	Electronic seals and tracking devices that use satellites
Isotwg, Addison, TX	Chemical-based intrusion detection using seals and handheld sensors
Argonnc National Laboratory, Argonne, IL	New detector to find nuclear materials in containers
MIT, Cambridge, MA	Radiation detector with imaging capability

Unlike high security manual seals, there is no agreed standard for e-seals, and the substantial difficulties in establishing such a standard have made it clear that such a standard is highly unlikely in the near future. Further, e-seal's limitations are becoming clearer. As the Los Alamos National Laboratory has concluded, e-seals can be defeated as easily as manual seals, and contrary to some of the marketing arguments for them, they cannot and do not inform about what a container's actual contents any more effectively than the carrier's manifest. They do not deter entry into a container any more effectively than a high security manual seal. Furthermore, as technology is developing, e-seals may be overtaken by other technologies that can provide better information, more economically. For example, some technology being tested would indicate whether the container doors have been opened, which while not perfect security information, would provide better security information than whether the seal on the container is the same seal that was originally affixed by the shipper. Furthermore, there is growing interest in sensors that might be economically placed in a container that could detect various security risks and monitor container contents.

The principal container security issue is: what has been loaded into the container? - not what does someone say was loaded in the container, and not whether the seal on the container is intact. Technology, in the form of non-intrusive inspection equipment, is a substantial step forward. Regarding technologies that might be attached to a container, it is important to recognize that ocean shipping containers do not operate in closed or dedicated services for a particular company or geography, but are interchangeable and are globally mobile.

What we may need is an intelligent container seal which can receive an alert signal generated as a data mining application flashing a warning with this alert indicator then used to flag boxes that should be scanned, non intrusively, with x-rays and gamma rays. Containers flagged for inspection should be stowed together. Technology such as SAIC and Heimann could in future scan entire boxes in 1 minute, thereby promising throughput rates of 50-60 boxes per hour, as well as increased revenues from Customs collection. The portability of the huge scanners is improving, enabling them to be quickly moved around a large yard. E-seals can now be permanently installed which use ultra-low power radio "Bluetooth Lite" transmitters.

As opposed to traditional electronic seals, this system will use a single globally available ratio frequency band, which is license-free. Since ALL-Track uses low cost Readers the reading infrastructure will apparently be cost efficient. The small size and high flexibility of the AllTrack Reader permits connection to a standard mobile phone, thus making it a 'mobile' reader.

The new ALLTrack e-Seal can be permanently installed in the container, and by the door hinge a small sensor registers door-opening events. The e-seal also has capabilities to connect a wide variety of external sensors to the inside of the container, enabling future 'smart container' capability at the same time without duplication of cost. A ten-year battery assures that e-seal will last for the entire lifetime of the container.

Considering scanning of containers, there are a number of x-ray, gamma ray, and pulsed neutron scanners now available. These are expensive, bulky, and heavy equipment and can be installed in drive-through facilities or truck mounted. There are now studies to develop container spreader-mounted scanners to permit use of the time a container is suspended on a spreader (usually 30-50 seconds) for scanning.

Active-pulsed neutron scanners (for nuclear, etc. material detection) obviously requires significant shielding. American Science and Engineering developed a low price ($2m) container inspection system that combines on a truck-mounted platform x-ray scanners with a highly sensitive array of radiation monitors able to register both gamma rays and neutrons. Most existing scanners cannot see through a container full of dense materials. An exception is the VACIS system (Science Application International Corp.). Scanners today are becoming more sophisticated, effectively dealing with back scatter and include detection software that spots anomalies in shipping manifests and minimizes failing layered systems.

Recent developments are to permit scanning from some distance and capturing content, etc. and valuation against listed contents. Effective container identification is part of the requirements. Different marking systems are now being listed.

Radio Frequency Identification hardware and software for item level trucking of containers (EFID) developed by Savi Technology is one approach. Containers uniquely identified with bar code or RF tag that are scanned at every milestone is another approach. There are now alternative data capture devices available that are fully web enabled. There are also magnetic container record markets that can be scanned for container content, origin, destination, and next/previous transfer point that again can be wen hosted (Tren-Star or Trencor Container System).

In addition various handheld testing and scanning equipment are now available from CO_2 testers to check container air samples to prevent illegal immigrant or terrorist border crossing to portable nuclear, biological, toxin, poison, and explosive material testing devices, The main issue in security technology development is in addition to effectiveness that it does not slow the flow of goods and/or introduce unacceptable costs.

Another issue that is being addressed is the problem of information tracking, checking, and updating that must include exception reporting. This must include monitoring of transport forwarders and consolidators in addition to shippers (a responsibility which is moving from the Federal Maritime Commission to U.S. Customs) as well as NVCCs which are consolidators and should also submit manifests to U.S. Customs (Automated Manifest System (AMS)). Finally, as containers are increasingly leased and not owned and therefore move among operators, lessor responsibility must be included,

CSI aims at tightening reporting requirements for all cargo coming into the U.S. with the 24 hour advance cargo manifest declaration rule. These security measures now act as a catalyst that forces all companies involved in the supply chain to adopt information technology and standards that will ultimately lead to more efficient business practices. We may find it attractive to extent this to the adoption of already developed e-commerce technology with a home port portal such as APL's GTN (Global Transport Network) portal that is already used by 12 carriers.

2.1. Transportation Security Strategy Consideration

The various maritime transportation and supply chain security initiatives both at the policy and strategy as well as at the technology level are in a flux. The new Department of Homeland Security is still trying to organize itself. International collaboration and support is being organized but is far from being defined. In the area of aircraft, vessel, and port security, international support has been demonstrated and implementation of new international rules as the best way to implement statutory requirements for aircraft, vessels, and ports.

In the area of container security, CSI agreements will be expected to transition from a sound concept to an effective operation. A lot needs to be done and greater clarity is required in terms of requirements and responsibilities. Most importantly, major improvements in technology are needed if we are to develop a secure and operationally efficient system. We must now perform

- assessment and analysis of the existing programs
- focused management on strengthening those aspects of existing initiatives that need improvement or more resources
- public explanation and demonstration of progress and results
- identify and support new technological opportunities

The various air transport maritime and supply chain security strategies clearly require greater international support if they are to work effectively.

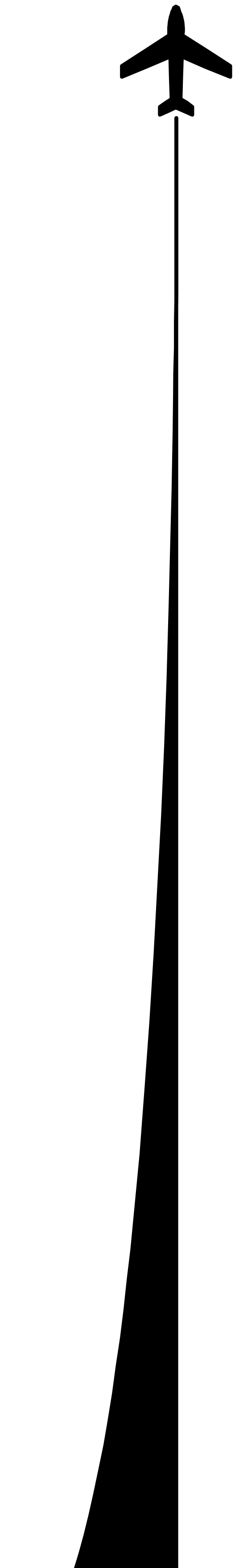
Ride assessment of security threats requires a unique analytical and statistical approach. Cause and effect and resulting failure mode analysis provide a basis but serve only as a structure. While the effects of security threats or causes can be identified and effects even quantified, the problem is determination of threat identification and probabilities. Threats, as noted before, are externally generated and independent of a transport systems technology, condition, application or performance. They are therefore identified as independent chance events with consequences. They are, furthermore, increasingly independent of geographic location or even ownership of transport. Threats are invoked or acted not just to cause damage and loss of life but as political means for undermining social, political, and economic systems.

Security threats in transportation are usually common cause events causing system failure that arc often catastrophic. Analytical network methods such as fault tree analysis are being investigated to develop rational approaches to relate primary threats and resulting failure events to common cause failures, using failure mode and effect analysis.

The main problem is the lack of predictive information on the type, likelihood, location, and timing of security threats that force us to continuously update or invent arrays of possible threats and their consequences. Yet there is no choice as security threats have become the single most important factor affecting transportation systems reliability, safety, and performance.

3. CONCLUSIONS

The performance and safety of all modes of transportation systems are today largely affected by security threats that must be considered in the planning, design, construction, and operation of transportation systems. This requires a different approach from that used in traditional safety and reliability analysis because it is based on completely random human actions with objectives to cause significant property, life, and political damage. There is no historic or performance data that allows meaningful prediction of such events.



Session 1

Transport

ANALYSIS OF RESOURCE CYCLES IN MARINE CONTAINER TERMINALS

Andrey Solomennikov[1], Vladimir Bardatchenko[2], Yuri Merkuryev[3], Fred Kamperman[4]

[1] *Transport and Telecommunication Institute*
Lomonosova 1, Riga, LV-1019, Latvia
Ph: +371-9455857, e-mail: andrey.solomennikov@btv.lv

[2] *Dept. of Modelling and Simulation, Riga Technical University*
Kalku 1, LV-1658 Riga, Latvia
Tel. +371 9118947, e-mail: vladimir.bard@btv.lv

[3] *Dept. of Modelling and Simulatio, Riga Technical University*
1, Kalku Street, LV-1658 Riga, Latvia
Tel. +371-7089514, e-mail: merkur@itl.rtu.lv

[4] *Baltic Container Terminal, Ltd.*
Kundzinsala 1, Riga LV-1822Latvia
Ph: +371 7076200, e-mail: fred.kamperman@bct.lv

Abstract

The aim of the article is to introduce the reader to the logistic model of the Baltic Container Terminal the possibilities of optimal use of existing resource pool of the terminal as well as analyzing marginal efficiency of available resources. One of the challenging tasks was to arrive at a reasonable level of detail abstraction of the logistic model. The article concentrates on aggregated operation blocks and adjustment of respective parameters to observed terminal productivity data. Simulation analysis yielded a set of optimal solutions. The solutions obtained proved stable with regard to the input data. The methodology discussed shows good potential in application for related diverse terminal processes.

Key words: marine, port, container, terminal, simulation, BCT

1. Project Overview

This work presents the general results obtained within the frames of the *Baltoports-IT IST-2001-33030* project *Simulation and IT Solutions: Applications in the Baltic Port Areas of he Newly Associated States.* The aim of this project was creating a simulation model of the *Baltic Container Terminal* using Rockwell Arena software. The key requirements brought forward by the management of the terminal was separate resource modeling and ability to monitor each single resource unit at any point of time.

2. Introduction

The container terminal logistics chain consists of four types of resources (numbers in parenthesis indicate the respective maximum units of the given resource type employed): quay crane (up to 2) yard crane (up to 2), forklifters (up to 2), and trucks (up to 6). This chain processes four types of inputs: 20ft containers, 40ft containers, vessel hatch covers, and restow containers (containers to be moved aside when accessing cargo on the vessel, irrespective of type). The logic of the model incorporates two logistic chains (both loading and discharging): 1) 40ft container chain and 2) 20ft container chain. Logical structure of each of these logistic chains above as well as the basic process blocks distinguished in the model are presented on FIGURE 1 below (only discharge illustrated in this case).

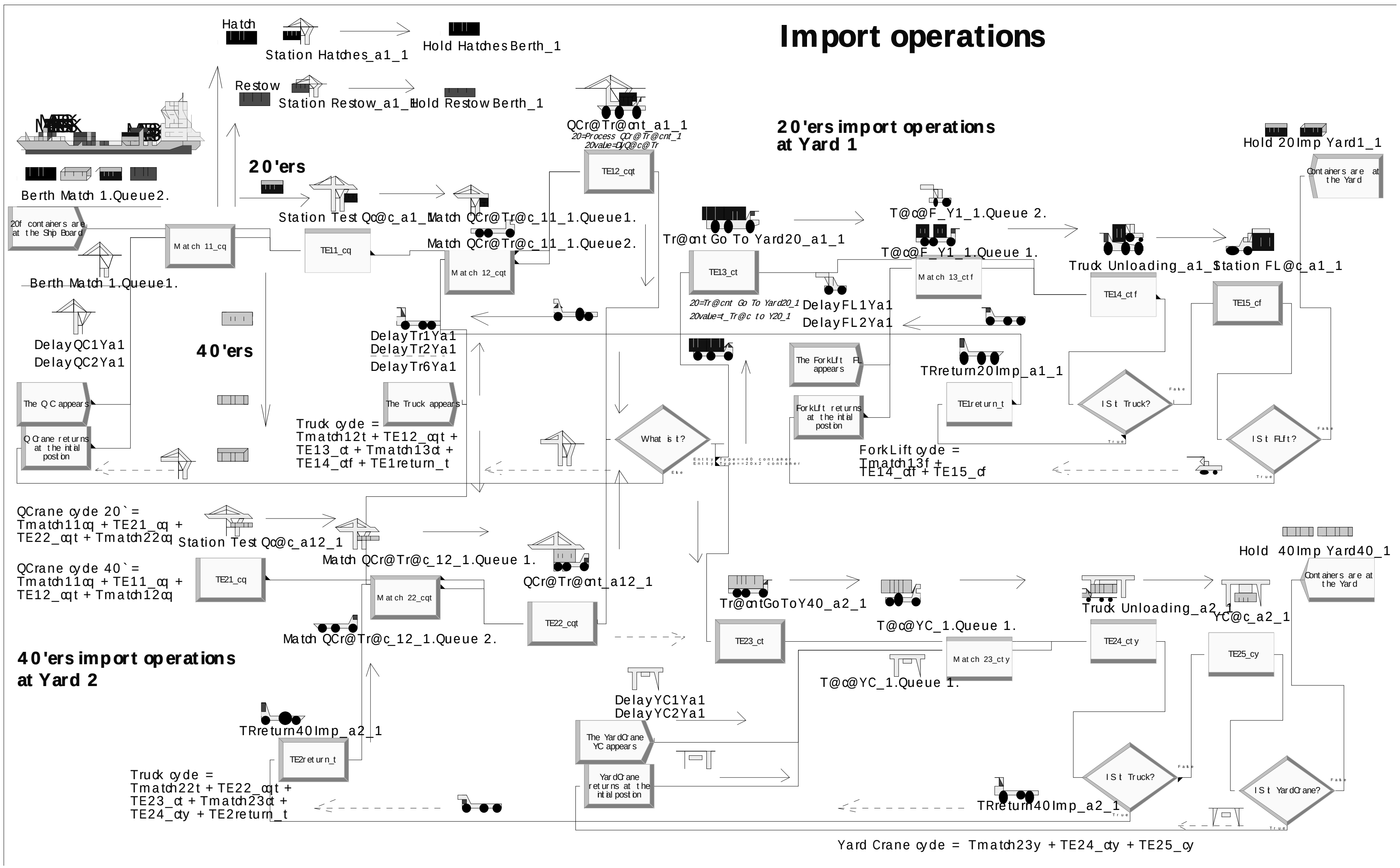


***FIGURE 1**. Model overview: discharge of 20ft and 40ft containers*

Due to the overview nature of the article, the authors delimit themselves to discussion of 40ft container chain solely to illustrate the basic underlying principles of modeling terminal logistics.

3. Resource Cycle Overview

As it follows from FIGURE 1 above, in the simulation among the available resources there can distinguished separate logical resource cycles. Thus, for the modelling purposes there emerged three logical resource cycles: the quay crane cycle, the truck cycle, and the yard crane cycle. Basing on these logical entities the performance of the model vis-a-vis real-life data is tested, as well as these allow easier adjustment of the model (see below for a more detailed discussion).

In order to introduce the outlined cycles, let us consider the terminal processes involved in a discharge of 40ft containers off vessel. Following the logistics chain of the port let us take a look at the first operation cycle, the quay crane.

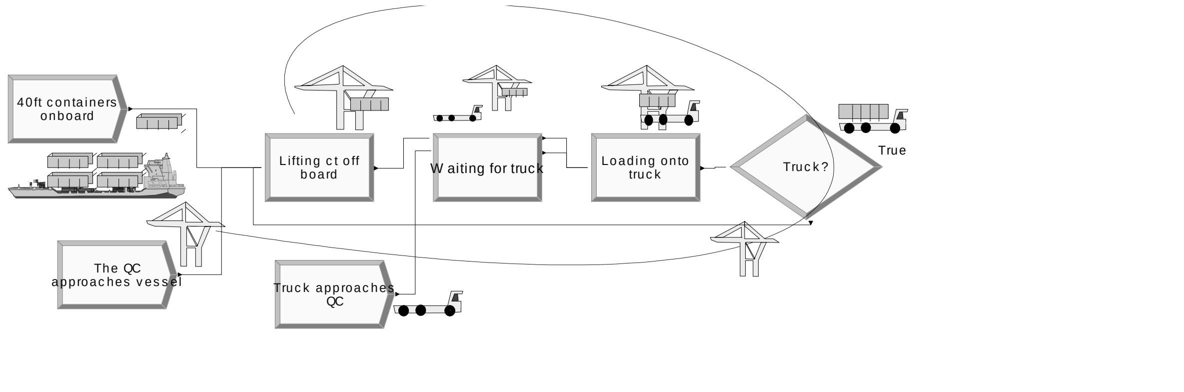


FIGURE 2. *Quay crane cycle*

The quay crane cycle breaks down into the following lower-level operations (see FIGURE 3 below): the crane head moves to container to be taken off the vessel, grabbing the container and moving it to truck position, waiting, positioning, and placing container onto truck, return to initial cycle position.

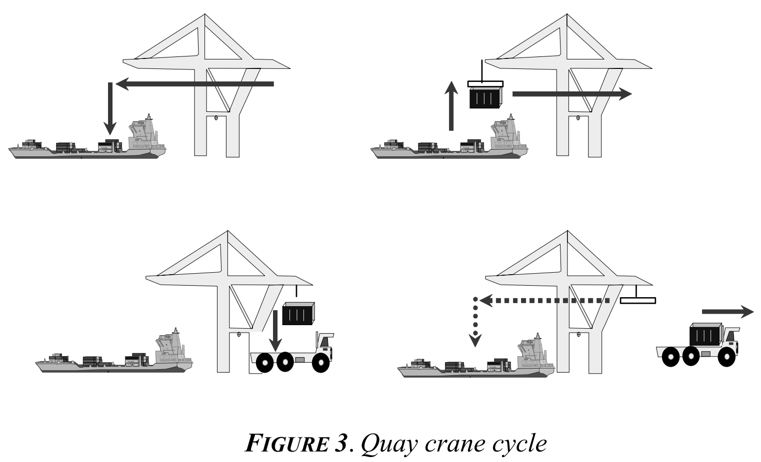


FIGURE 3*. Quay crane cycle*

In a similar manner, the operations as well as cycles for the other two types of resource (yard crane and trucks) are determined. Omitting some minor details, the truck cycle is portrayed on FIGURE 4.

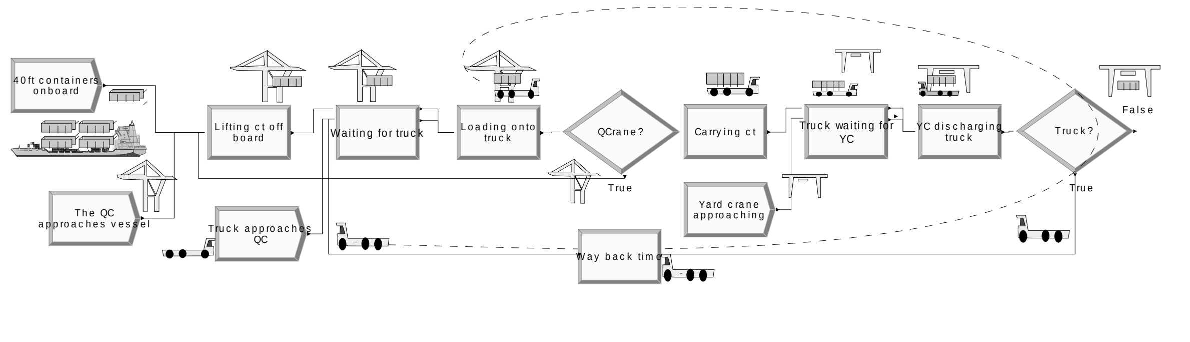


FIGURE 4. *Truck cycle*

Completing the resource cycle overview, the end-point of the 40ft container discharge chain is the yard crane cycle, presented on FIGURE 5 below.

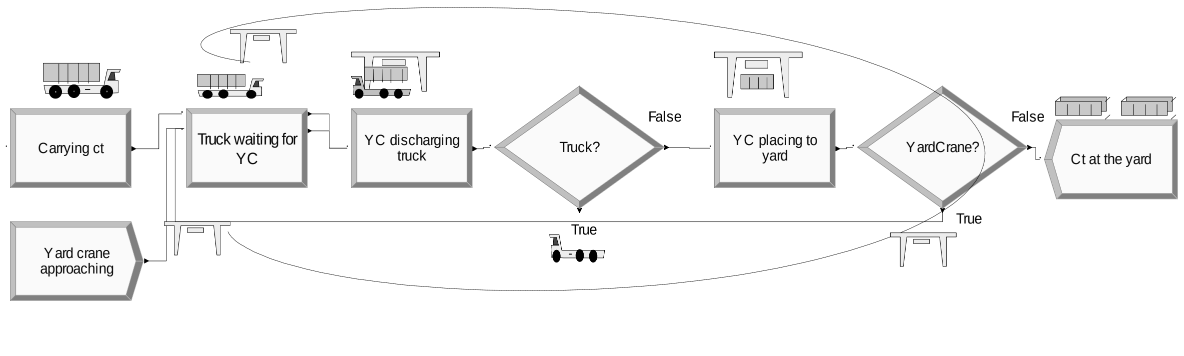


FIGURE 5. *Yard crane cycle*

4. Model Parameters

The model incorporates a two-tier resource parameter structure: 1) parameters of the lower technological level (durations and probability distributions of elementary operations for each type of resource), and 2) parameters of resource cycles outlined above (cycle durations and respective probability distributions). Parameters of the model also include number of simultaneously employed resources of each type, number of containers of different types to be discharged, and hatch covers necessary to be processed on the vessel.

The running parameters of the model are controlled with the help of hierarchically organized monitoring variables of each process modeled, which allows spot measurements. These monitoring variables are recorded in a dynamic database and can be traced for spot analysis of different "*what if…?"* scenarios.

5. Parameter Adjustment

Adjustment of model parameters involved indirect methods, whose attractiveness lies in much lower financial and time costs compared to direct measurements. Moreover, such an approach allows determining combinations of all model parameters, leading to optimal overall productivity and thus identifying existing inefficiencies in the terminal technological chain.

The simulation task formulated was to determine the values of the three resource cycles of the model Tq (quay crane cycle), Tt (truck cycle), and Ty (yard crane cycle) at a given level of terminal performance with different number of trucks available.

The calculations employed observed terminal productivity data NP(n)real. As indicated by the BCT management, the NP(n)real productivity value for the 40ft container load/discharge chain for the set of 3 trucks available constituted 24 +/-1.5 moves/hour. The productivity of bundle of resources with 4 trucks constituted 30 +/-1.5 moves/hour, and respectively 32 +/-1.5 moves/hour with 5 trucks available.

At the first stage the identification of the three cycles Tq(model), Tt(model), Ty(model) there was applied data of real BCT productivity with the given set of resources (1 quay crane, 1 yard crane, with 3, 4, and 5 trucks respectively). Thus, one of the reasonable options of analyzing model adequacy would be minimizing the least squares deviation of the model-yielded from the observed data with different number of trucks available (see FIGURE 6).

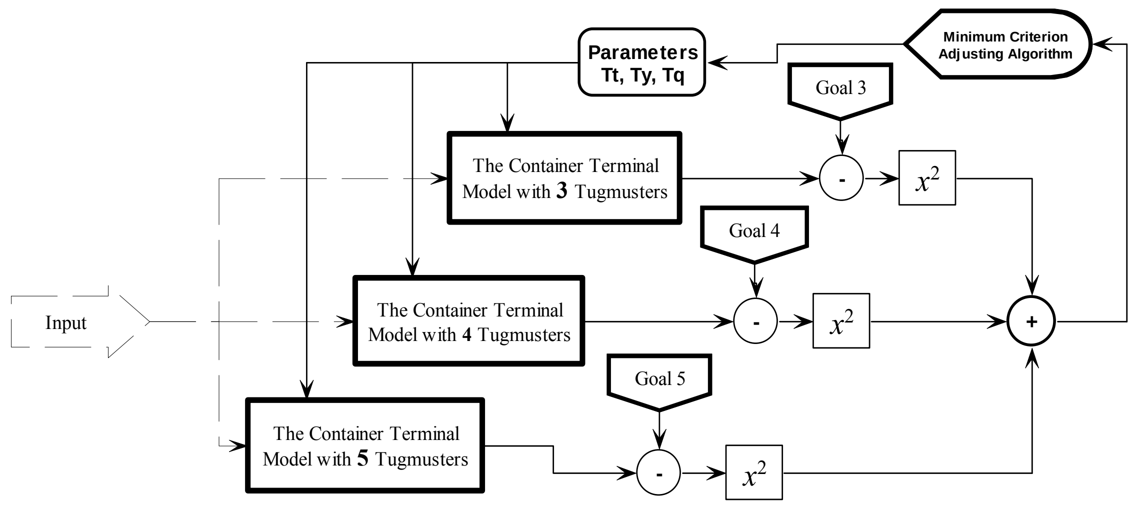


FIGURE 6. *Block-scheme of resource cycle parameter estimation (Tq, Tt, Ty)*

The elapsed truck cycle duration would then equal

Tttot = Tct&q + Tcqt + Tct + Tct&y + Tcty + Treturn

where:

Tct&q – truck waiting for the quay crane to pick the container off the truck.
Tcqt – container loading time
Tct – time necessary for the truck to carry the container to the yard where it is supposed to be stored
Tct&y – the waiting time of the yard crane for the truck carrying the container
Tcty – discharging container from the truck by the yard crane
Treturn – empty truck returning to pick another container

It should be pointed out that the period of the truck waiting for the quay crane to load the container onboard Tct&q as well as the Tct&y, the waiting time of the yard crane for the truck carrying the container, are non-stationary values as they are highly dependent on the number of the trucks. These values are not pre-determined, they are obtained during the modelling process since depend on involved resource co-ordination.

With an acceptable degree of precision it is possible to estimate and introduce the following values into the model: Tcqt, Tcty, Tct, and Treturn.

Whereas, it might be noted that Treturn =Tct *K, where K<1. Average values of Tcqt and Tcty as well as their respective statistical distributions are determined by the technological processes of loading and discharging by the quay crane and the yard crane. Therefore, the actual cycle duration at any point of time cannot be determined precisely, but can be modified manually by adjusting the Tct value.

By analogy to the to the truck cycle above, other resources' cycles are determined in a similar manner. It should be also noted that the resource cycle durations are correlated not only through the tightly distributed load/discharge micro-operations (e.g., Tcqt, Tcty) but are also affected by the highly dispersed time losses in queues. Since in practice the durations of elementary operations are not strictly determined, the time of every independent elementary operation was modeled by an individual uniformly distributed random generator whose boundaries are proposed by the BCT personnel.

Thus, in order to obtain an average productivity value as a model output, e.g. $NP(3)_{model}$, the model takes in the respective number of trucks (in this case: three) and for each set of resource cycle values Tq(model), Tt(model), Ty(model) it simulates discharge of 150 containers at a run (the average number of containers per vessel), basing on which the average productivity value is calculated. In the same manner the average $NP(n)_{model}$ productivity value is calculated for all the other input values. So, over the discrete subset of the values of four arguments **n** trucks, **Tq**(model), **Tt**(model), **Ty**(model) the value of productivity function NP(x, y, z, u) is determined. The task lies in finding the values y, z, u under which the value of the criterion will be minimal

$$\mathrm{Crit}(x, y, z, u) = \mathrm{sqrt}\,[(24 - NP(3, y, z, u))^2 + (30 - NP(4, y, z, u))^2 + (32 - NP(5, y, z, u))^2]$$

where y = **Tq**(model), z = **Tt**(model), **u** = **Ty**(model).

The methodology of the above criterion yielded rather different results, which indicated either existence of local minima or a trough (a continuous set of minima) on the criterion surface, indicating flaws in operation efficiency. Rather detailed analysis of criterion surface is represented on FIGURES 7, 8, and 9.

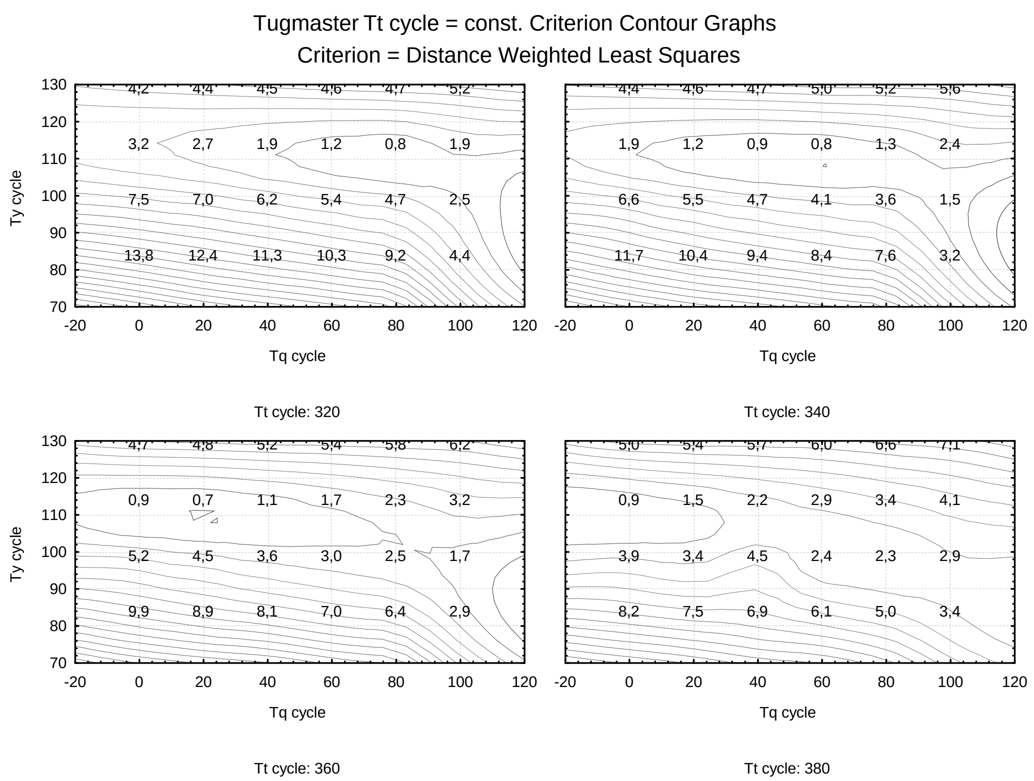


FIGURE 7. *Adequacy criterion contour graph, where Tt(model) truck cycles are equal to 320, 340, 360, and 380 secs. The criterion is represented as a function of two variables, X-scale: Tq(model), Y-scale Ty(model). As Tt(model) increases, the criterion minimum shifts outwards left from Tq(model)=80 to Tq(model)=20, with Ty(model) held constant at the level of 110 seconds*

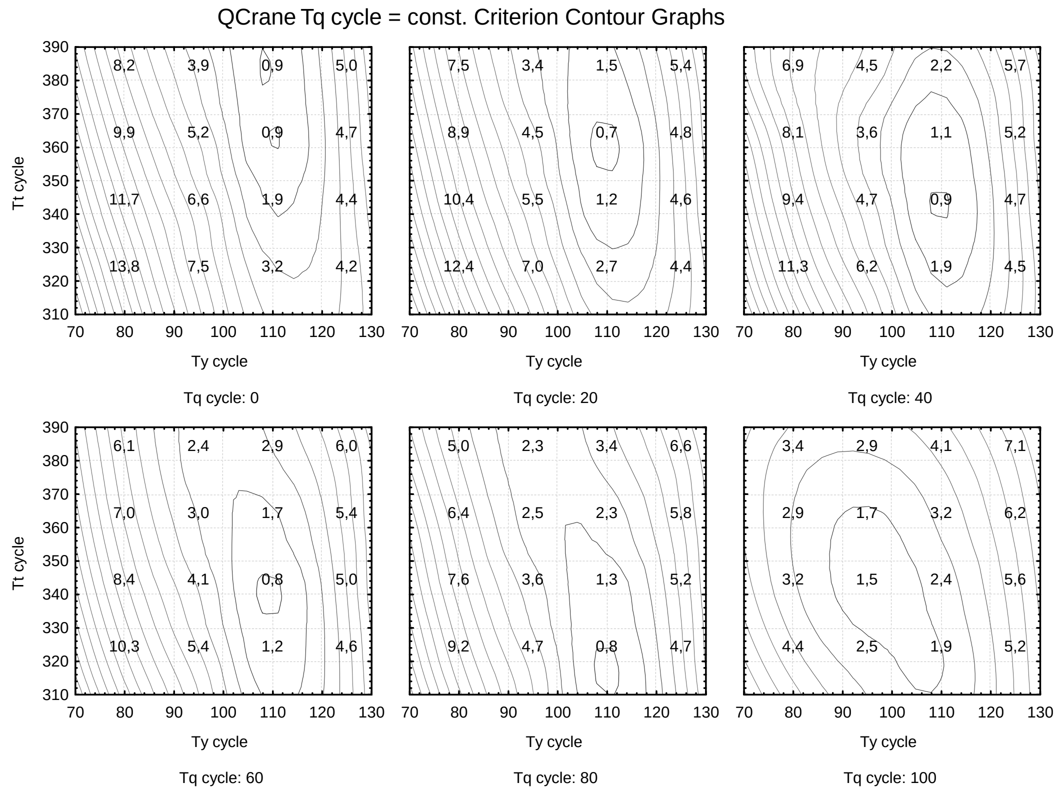


FIGURE 8. *Criterion contour graph, where Tq(model) quay crane's cycles are equal to 0, 20, 40, 60, 80, and 100 seconds. The criterion also represents a function of two variables, Ty(model) plotted on x-scale and Tt(model) on the y-scale. As Tq(model) increases, the criterion minimum shifts downards parralel to y-scale from Tt(model)=380 to Tt(model)=320, with Ty(model) held constant at the level of 110 seconds*

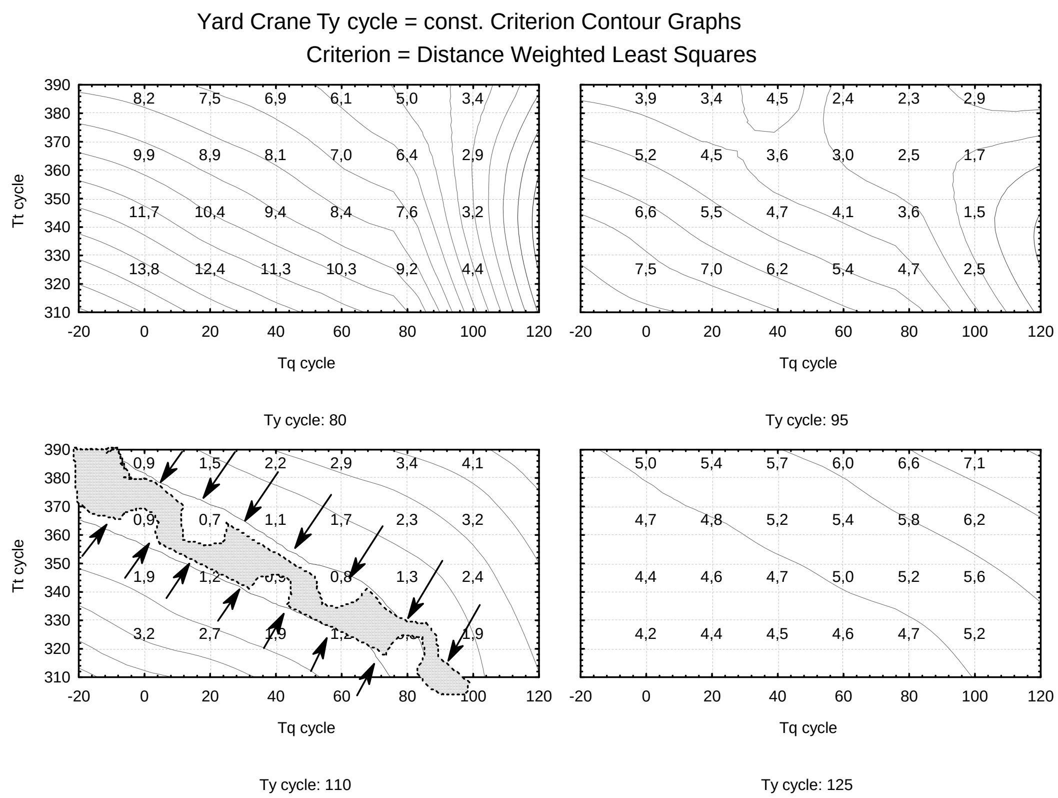


FIGURE 9. *Criterion contour graph, where Ty(model) yard crane cycles are equal to 80, 95, 110, and 125 seconds. Similarly to previous casees, the criterion is represented as a 2D function of two variables: Tq(model) plotted on x-scale and Tt(model) on y-scale. As Ty(model) increases, the criterion minimum appears at Y = Ty(model) = 110 seconds. The criterion is no local minimum but a trough and is approximated by Tt(model) = 380 – ¾ * Tq(model), Ty(model) = 110 seconds*

6. Testing Obtained Parameters for Sensitivity to Variations in Input Data

Now our task lies in analyzing the obtained values of y, z, and u bringing the criterion to its minimum and testing its sensitivity for variations in input productivity data

NP(3)real = 24 +/-1.5, NP(4)real = 30 +/-1.5, NP(5)real =32 +/-1.5.

Thus, there should be sought a dependence of the average criterion values on the randomly distributed parameters y, z, u:

AverCrit(x, y, z, u)] = Average{sqrt [(24 + delta(i) – NP(3, y, z, u))^2 + (30 + delta(i) – – NP(4, y, z, u))^2 + (32 + delta(i) - NP(5, y, z, u))^2]}

where y = **Tq**(model), **z = Tt**(model), **u = Ty**(model),
delta(1) = random value with uniform random distribution [-1, +1]
delta(2) = random value with uniform random distribution [-2, +2]
delta(3) = random value with uniform random distribution [-3, +3]

As follows from FIGURE 10, the graphs indicate that the criterion nature remains unchanged, whereas only the criterion level is being affected. Thus, it might be concluded that the obtained values delivering the criterion minima remain stable to variations in the productivity NP(i)real.

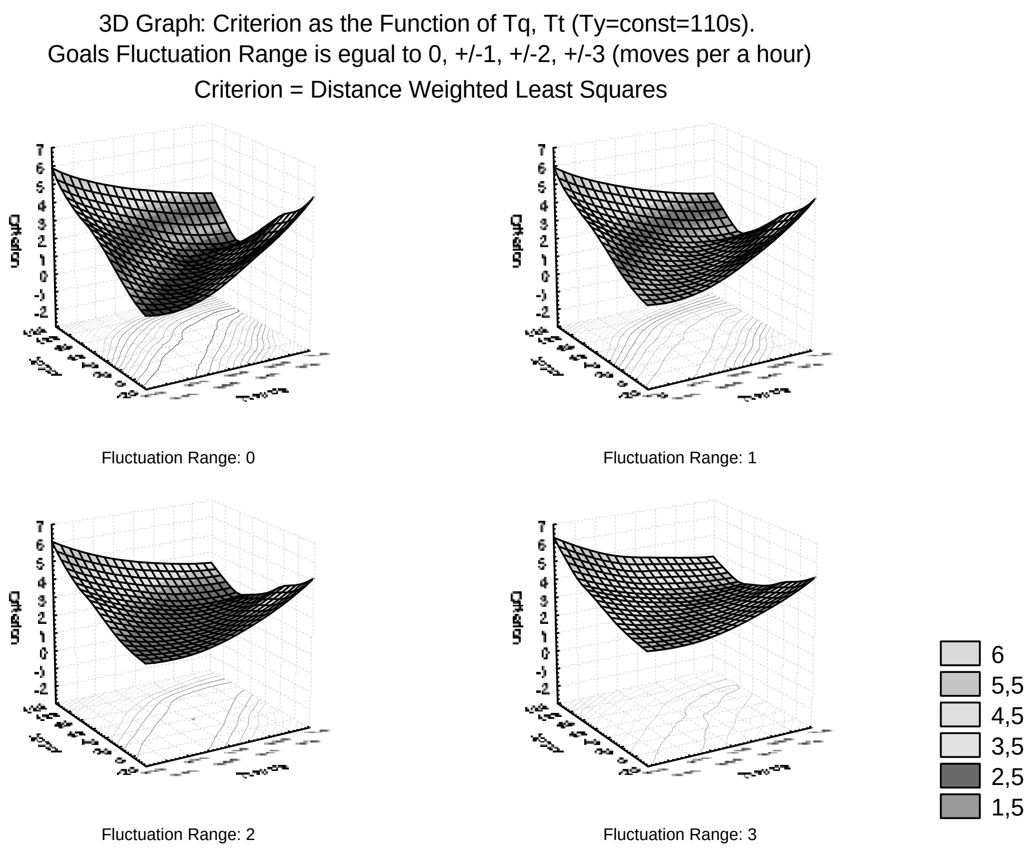


FIGURE 10. *(x-scale: Tt(model), y-scale: Tq(model)) illustrates the criterion surface transformation with different values of* delta *[NP(3)real = 24 +/- 1, +/- 2, +/- 3, NP(4)real = 30 +/- 1, +/- 2, +/- 3, NP(5)real =32 +/- 1, +/- 2, +/- 3] at Ty held constant at 110 seconds*

7. Conclusions

There are several essential conclusions that might be drawn from the project presented.

- The indirect method of parameter estimation of logistical models with load/discharge type of processes with variable number of carriers has been illustrated and proven practical. Moreover, this method proved cost- and time-efficient with a reasonable degree of precision.
- It has been determined that there exists a set of optimal solutions (in terms of adequacy criterion minimization). It was empirically found that all the optimal parameter sets are located along the line Tt(model) = 380 – ¾ * Tq(model), Ty(model) = 110 seconds. It should be noted that direct methods normally yield a single result.
- There has been performed a sensitivity analysis of the solutions obtained which revealed reasonable model stability with regard to variations in productivity values of NP(3)real = 24 +/- 1, +/- 2, +/- 3, NP(4)real = 30 +/- 1, +/- 2, +/- 3, NP(5)real = = 32 +/- 1, +/- 2, +/- 3.
- Due to the formal logical structure of the model, it was revealed that the underlying principles and methodology can easily be transferred to general terminal and warehousing process modeling, which represents practical value for future research in this direction.

References

1. Yuri Merkuryev, Vladimir Bardachenko, Andrey Solomennikov and Fred Kamperman. Multi-level simulation of logistics processes at the Baltic Container Teminal. International Workshops on Harbour, Maritime and Multimodal Logistics Modelling & Simulation, HMS2002, and Modelling and Applied Simulation, MAS2002. October 3-5, 2002, Bergeggi, Italy. Ed. by Agostino G. Bruzzone, Yuri Merkuryev, Roberto Mosca. Genoa University, 2002. P. 78-83.
2. Yuri Merkuryev, Vladimir Bardachenko, Fred Kamperman and Andrey Solomennikov. Multi-level simulation of logistics processes at the Baltic Container Terminal: Calibration study. Modelling and Simulation of Business Systems. Proceedings of the International Conference. May 13-14, 2003, Vilnius, Lithuania. Ed. By H.Pranevicius, E.Zavadskas and B.Rapp. Kaunas, Technologija, 2003. P. 274-278.
3. Yuri Merkuryev, Vladimir Bardachenko, Andrey Solomennikov and Fred Kamperman. Simulation of logistics processes at the Baltic Container Terminal: Model validation and application. Proceedings of the International Workshop on Harbour, Maritime and Multimodal Logistics Modelling and Simulation. HMS3003. September 18-20, 2003, Riga, Latvia. Ed. by Yuri Merkuryev, Agostino G. Bruzzone, Galina Merkuryeva, Leonid Novitsky and Edward Williams. Riga, RTU, 2003. 10 p. (yet to appear)

NEW PRIORITY WAYS FOR INTERNATIONAL FREIGHT TRANSPORT AND ESTIMATION OF BUSINESS STRATEGY OF LITHUANIAN CUSTOMS

Aldona Jarasuniene

Transport Research Institute, Vilnius Gediminas Technical University,
Plytinės 27, LT–2016, Vilnius, Lithuania,
E-mail: valca@ti.vtu.lt

Abstract

In this article there are presented new priority ways in international freight transport. In the intense of reconstruction of the international freight transport system in Lithuania, one of the main tasks is to guarantee effective transport activities oriented to satisfy the needs of Lithuanian economy, especially in international freight transport by road as a priority field.

In this article there is showed an implementation of new means border crossing for the improvement of whole freight transport system.

There is an analysis the of new computerised transit system that offers for traders and customs many advantages.

There is described business strategy vision, mission, strategic goals of Lithuanian customs.

There is showed a long term development strategy for Lithuanian transport system.

Key words: international freight transport; development; new computerized transit system; business strategy of the Lithuanian customs; long terms development strategy for Lithuanian transport system

1. Introduction

In the intense of reconstruction the transport system in the Republic of Lithuania one of the main tasks is to guarantee effective transport activities oriented forwards satisfying the needs of Lithuanian economy, especially in international freight transport by road as a priority field. In order to guarantee effective international freight transport it is very important to create favourable conditions for that. Those conditions are influenced a great deal by the customs which is an important public institution in charge of performing customs clearance operations and collecting taxes.

On this basis it is possible to identify measures for solving the problems of the busiest customs-houses, thus positively influencing Lithuanian economy and enabling a more precise view of inspection of customable freight carrying transport means' with the present number of customs officials. Very important to show general aspects of business strategy of the Lithuanian customs and to analyse situation of Lithuanian transport sector strengths, weakness, opportunities and threats.

The most important issue – to define precisely problems of international freight transport and to select properly as well as to adjust necessary means and new priority ways [1–3].

However, it must be born in mind that in preparation of means related to the improvement of international freight transport it is necessary to consider all the realisation consequences to the transport sector as well as to other sectors of national economy as well (including not only those of technological, economical aspect, but social, ecological and even political aspect as well) [5, 7].

Therefore, while speaking of new priority ways in international freight transport and other means ensuring freight transport and its development, attention should be focused on basic issues.

Very important guarantee effective transport activities oriented forwards satisfying the needs of Lithuanian economy, especially in international freight transport by road as a priority field.

In order to guarantee effective international freight transport it is very important to create favourable conditions for that.

The conditions should be created on the basis of analysis:

- of implementation of new means of border crossing for the improvement of whole freight transport system;
- of new computerised transit system (NCTS);
- of Lithuanian customs strategy;
- of Lithuanian transport development.

2. The New Computerised Transit System (NCTS)

The system offers traders many advantages, including:

– Improved quality of service:

- Less time spent waiting at customs, because the declaration will have been sent electronically beforehand;
- Greater flexibility in presenting declarations.

– Earlier discharge of the transit procedure because an electronic message is used instead of the return of the paper copy No 5 by mail, leading to a faster release of the guarantee.

– The high costs, incurred in relation with the paper-based system of declaring goods (lengthy procedures involving much time and effort), are reduced.

– A greater clarity of the transit operation, for the benefit of trade.

– Because customs will have decided well in advance of the arrival of the goods at the office of destination whether or not they want to check the consignment, the trader will not lose valuable time at the office of destination waiting for a decision.

Apart from these general advantages for trade, there is an additional advantage for authorised consignors linked to the NCTS system. They no longer have to carry out the cumbersome formalities that are necessary in a paper-based environment, because all the movements will be directly managed by the system.

The advantages of the NCTS for customs are:

– The communication and coordination between the customs administrations involved will improve.

– Repetitive activities will only have to be performed once; this saves time and eliminates the risks involved in the duplication of information.

– Creation of a more coherent system, which will speed up the processing of data and at the same time making the system more flexible.

– Harmonisation of operating criteria, which will do away with the plethora of subprocedures and divergent interpretations of how the rules have to be implemented.

– Availability of a system run directly by customs, which offers greater security and a higher tempo in managing transit, provides more reliable data and better monitoring of movements.

It is clear that the trader indirectly benefits from the advantages of the NCTS for customs, and vice versa [6].

Operation. Main items or messages used in a NCTS operation. Before going into the details it is useful to mention the main items and messages in a NCTS operation.

– The transit declaration, which is presented in a paper or electronic form.

– The movement reference number (MRN), which is a unique registration number, given by the system to the declaration to identify the movement.

– The transit accompanying document, which accompanies the goods from departure to destination.

– The “anticipated arrival record” message, which is sent by the office of departure to the declared office of destination mentioned in the declaration.

– The “anticipated transit record” message, which is sent by the office of departure to the declared office(s) of transit* to notify the anticipated border passage of a consignment.

– The “notification of crossing frontier” message, which is sent by the actual office of transit used after having checked the consignment.

– The “arrival advice” message, which is sent by the actual office of destination to the office of departure when the goods arrive.

– The “control results” message, which is sent by the actual office of destination to the office of departure after the goods have been checked.

Furthermore it is important to understand that the system covers all the possible combinations of normal and simplified procedures, at departure as well as at destination.

Office of departure. The transit declaration is presented at the office of departure, either in paper form (in which case the data is introduced in the system by the customs office – see Fig. 1) or in a computerised form (see Fig. 1, as well as Fig. 2 in case the simplified procedure is being used). Electronic declarations can be made from terminals made available to traders at the customs office of departure or from a trader’s own premises.

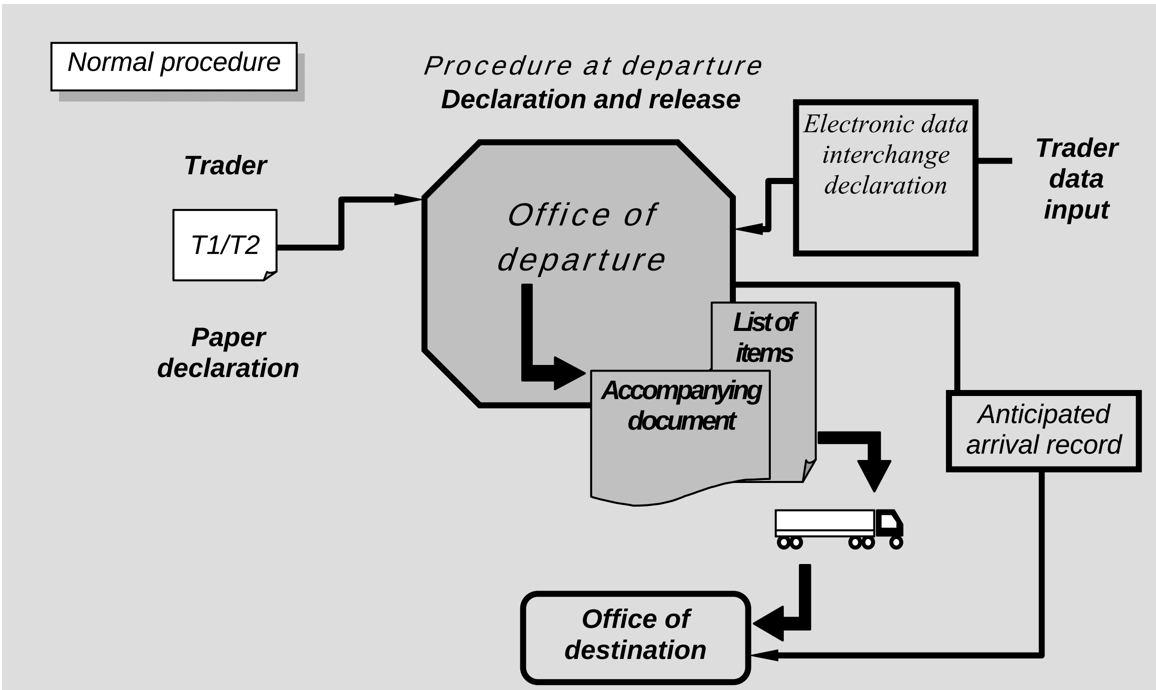


***Figure 1**. Office of departure – normal procedure*

Whatever the form of the presentation, the declaration must contain all the data required and comply with the system specifications, since the system codifies and validates the data automatically. If there is an inconsistency in the data the system will indicate this. The trader will be informed, so that he can make the necessary corrections before the declaration is finally accepted [6].

* An office of transit is a customs office situated a tone of the external land borders of the EU or one of the other participating countries of the Common Transit Convention.

Once the corrections have been entered and the declaration is accepted, the system will provide the declaration with a unique registration number, the movement reference number.

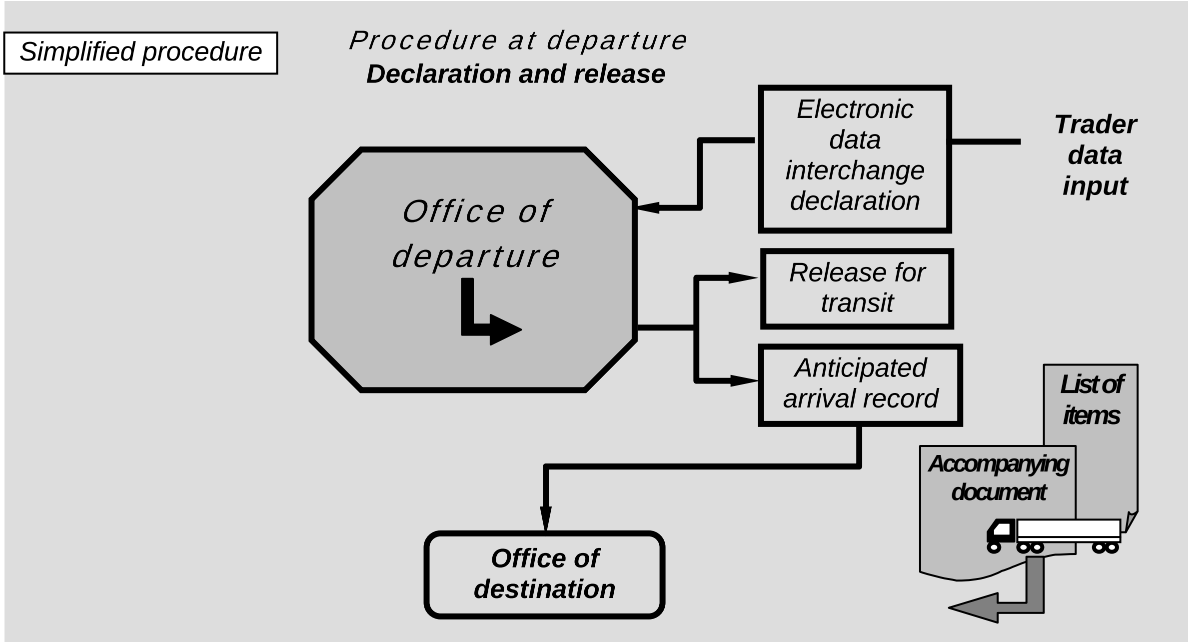


Figure 2. *Office of departure – simplified procedure (authorised consignor)*

Then, once any inspections have been carried out, either at the office of departure itself or at the authorised consignor's premises, and the guarantees are accepted, the goods will be released for transit. The system will print the transit accompanying document and, where appropriate, the list of items, either at the office of departure or at the authorised consignor's premises. The accompanying document and the list of items must travel with the goods and be presented at any office of transit and at the office of destination.

When printing the transit accompanying document and the list of items, the office of departure will simultaneously send an anticipated arrival record to the declared office of destination. This message will mainly contain the information taken from the declaration, enabling the office of destination to control the consignment when it arrives. The office of destination needs to have access to the best possible information about the transit operation to take a correct and reliable decision about what actions to take when the goods arrive.

Should the movement have to pass an office of transit, the office of departure will also send an anticipated transit record, so that any office of transit has prior notification of the consignment concerned and can check the passage of the movement.

Office of transit. When the goods pass by an office of transit, the goods, the transit accompanying document and, where appropriate, the list of items have to be presented to customs. The anticipated transit record, already available in the system, will automatically be located when the movement reference number is entered and subsequently the movement may be approved for passage. A notification of crossing the frontier is sent to the office of departure.

Change of office of transit or destination. If the goods go via an office of transit other than the declared one, the message that had initially been sent to the declared office of transit is of no use. In this case the actual office of transit will send a message to the office of departure, requesting the anticipated transit record, so that it can access the relevant information. Having checked the movement it will send the notification of crossing the frontier to the office of departure.

Likewise, the goods can be presented at an office of destination, other than the declared one. The actual office of destination will request the office of departure to send the anticipated arrival record so that the new office of destination may obtain the necessary information on the consignment.

If there is a change in office of transit or destination, the messages which have been sent to the declared offices are of no use and will remain open. To this end, the system will automatically send a message to the declared offices, notifying them where and when the goods have been presented, so that they can close the messages.

Simplified procedures: authorised consignor and authorised consignee. The use of both simplified procedures represents the optimal use of resources within the framework of the NCTS. The possibility of carrying out all the procedures at one’s own premises and exchanging information with customs electronically is clearly the most rapid, comfortable, secure and economic way of doing business.

Obviously in addition to satisfying the normal criteria to become an authorised consignor or authorised consignee, they will have to possess an adequate electronic data processing system for information interchange with their relevant customs offices. Of course this can only work if these offices are connected to the NCTS.

3. Implementation of Other Means in Border Crossing and Customs for the Improvement of International Freight Transport

For the freight transportation through the territory of the Republic of Lithuania the customs aims are following:

1. To develop a modern system for administering import and export duties and taxes.
2. To introduce advanced control technologies, based on sufficient prevention of law violations and authorisations to carry out investigations, application of risk analysis, use of information and intelligence, and modern technical equipment.
3. To expand the field of application of simplified customs procedures, creating necessary preconditions for the “electronic” declaration of goods.
4. To concentrate customs activities in the modern, well-equipped inland customs posts, provided with modern technology, while strengthening customs posts at the future external frontier of the European Union.
5. To introduce a modern customs information system based on up-to-date information and communication technologies.
6. To reorganise the transit system according to the principles of the Common transit system.
7. To expand co-operation with other Lithuanian institutions, the customs administrations of foreign states, institutions of the European Union, and other international organisations.

4. Business Strategy of the Lithuanian Customs

This strategy has been prepared taking into consideration the Communication from the Commission to the Council, the European Parliament and the Economic and Social Committee concerning a strategy for the Customs Union approved by the European Commission on 31 May 2001 and Council Resolution of 30 May 2001 (2001/C 171/01) concerning a strategy for the Customs Union.

The legal acts of the EU Acquis chapter “Customs Union”, regulating the procedure of trade for the EU and non-EU states, will not be practically transposed into the law of the Republic of Lithuania and will be applied directly upon Lithuania’s accession to the European Union. The Lithuanian Customs will be the main institution responsible for their implementation.

An important role in preventing the acts of terrorism, combating smuggling, money laundering, distribution of pornography, illegal trafficking of weapons, drugs, other prohibited and restricted goods, violations of application of international sanctions as well as in creating at the same time favourable conditions for the development of legal international trade falls on the Customs services of the EU member states.

After Lithuania becomes the EU member state, the role and responsibility of the Lithuanian Customs as of an administrator of duties and taxes will change, as it will become responsible for the collection not only of the Lithuanian state budget, but also of the part of the EU budget also. The Customs and tax administrations of the EU member states are responsible for the collection of taxes, i.e. EU traditional own resources, comprising the revenue of this budget.

After Lithuania joins the EU common market, the responsibility of the Lithuanian Customs will increase, since the goods having entered this market illegally can make harm to the whole market of the European Union. Therefore, the drawbacks of the Customs control will be very dangerous, in particular, where control is related with public health or environment protection.

The role of the Lithuanian Customs constantly increases in different fields of public protection (protection of intellectual property rights, environment, CITES goods, products safety, etc.).

The Lithuanian Customs has to participate in the activity of trade regulation and control. The complex or inflexible Customs procedures make negative influence on business competitiveness.

The Lithuanian Customs has to improve the service of businessmen, in particular, by reducing the costs occurring due to the different procedures applied. Therefore, the simplified Customs procedures and the speed of Customs clearance become increasingly important, and the application of risk analysis and selection methods becomes the main control criteria.

The introduction of the integrated Customs information system is one of the main priorities raised in the process of the Customs integration into the EU Customs system. The rapid development of new information technologies makes and will make in future a significant influence on the Customs activity as well.

In the process of the EU integration, new functions appear in the Customs requiring greater administrative capacities and additional human resources. For performing these functions, we have to reallocate the staff and to retrain employees by retaining, applying and improving their skills. In such conditions, a clear strategy for managing personnel changes and retraining is necessary.

The fight against corruption, i.e. cases of dishonesty and abuse of office by the Customs officials, should be strengthened. Special attention should be paid to the prevention of the manifestations of corruption and implementation of the principle of inescapable responsibility for illegal actions, education of intolerance with regard to such phenomena, development of international co-operation in the fight against corruption and other internal violations.

The experts of the Lithuanian Customs must be ready to work at the EU institutions, working groups and committees, since the success of the protection of Lithuania's interests in preparing the legal acts of the EU Acquis chapter "Customs Union" will depend on their preparation. Therefore, as the envisaged date of the EU membership approaches, it is necessary to learn to use effectively the possibilities given by this membership.

A modern infrastructure complying with the EU requirements is necessary for the efficient work of the Lithuanian Customs.

The development of the up-to-date Lithuanian Customs is impossible without close co-operation with the Customs services of foreign states and international organisations.

The updated strategy provides the milestones of the Customs activity covering the period of the preparation for and the firs years of the EU membership, i.e. a path to a modern Lithuanian Customs service positively estimated by the Lithuanian society and ready to act as a reliable partner of the Customs services of the EU countries.

The Lithuanian Customs understands clearly its objectives, sets itself serious tasks and makes every effort to fulfil them.

Vision. The vision of the Lithuanian Customs: a modern state institution properly performing all functions of the EU member state Customs service, ensuring the protection of the EU external border and EU Customs territory and serving as a model example for the Customs services of other states. This is flexible, efficient and reliable institution rendering high quality Customs services to travellers and business community, further improving its role as an institution protecting society and environment.

Mission. The Lithuanian Customs is a state institution performing the functions of tax administration and enforcement and exercising special authority to carry out investigation activities. Its main purposes are to:

- protect internal market and society by preventing illicit traffic in prohibited and restricted goods, combating the violations of laws and other legal acts regulating Customs activity, conducting criminal prosecution;
- ensure collection of import and export duties and taxes;
- collect, process and analyse foreign trade statistics while carrying out Customs operations.

In exercising its authority, the Lithuanian Customs observes the rule of law, principles of activity transparency, respects the rights and freedoms of individuals, applies legal acts regulating Customs activity correctly and consistently, implements international agreements and conventions, seeks to create favourable conditions for the legal international trade and speedy movement of goods and travellers flows as well as uniform conditions for competition, and to be transparent, honest, impartial and professional.

Strategic goals. In the field of management: to improve the public administration of the Lithuanian Customs based on the modern organisation management, ethics of service, a system for personnel motivation, career progression and promotion, for training, self-training and improving the qualification of Customs officials.

In the field of prevention of violations: to develop an efficient system for disclosing, pre-trial investigation and prevention of smuggling and violations of legal acts related to it.

In the field of preparation and implementation of legal acts and performing Customs formalities: to prepare by the date of the EU accession the legal acts regulating Customs activity aligned with the EU legal acts and to ensure their implementation; to apply more widely the simplified Customs procedures, to accomplish more rapidly Customs formalities, to ensure better service of clients.

In the field of administration of duties and taxes collected by the Customs: to develop and implement a system for administering duties and taxes complying with the EU requirements, to fight more effectively against tax evasion, to improve the accounting of income, their collection.

In the field of the Customs information systems: to introduce and develop a modern integrated Customs information system based on the newest information and communication technologies and its interfaces with the appropriate EU databases which would create conditions for fast production of data to users, increase the reliability of statistical information collected, reduce the duration of the accomplishment of Customs formalities and improve the service of clients.

In the field of infrastructure development: to develop a modern infrastructure, to supply the Lithuanian Customs offices with modern control measures and equipment, to grant special attention to the strengthening of the future external EU border.

In the field of international co-operation: on fulfilling the obligations assumed by international agreements, granting particularly great attention to the future EU membership, building upon relations with the EU institutions and Customs services of the EU member states, to develop co-operation with the World Customs Organisation and other international organisations.

In the field of public relations: to become a reliable partner of legal business entities, an institution open to the public and providing promptly information necessary to the public.

Tasks in the field of customs information systems. To develop, introduce and maintain the automated system ensuring the processing of all import and export declarations, presented for any Customs procedure electronically or in usual way.

To develop, introduce and maintain the automated transit control system complying with the requirements of the New Computerised Transit System (NCTS).

To develop, introduce and maintain the automated information systems:

- integrated tariff and related information management system;
- simplified Customs procedures processing system;
- system for accounting of taxes within the Customs competence and their control;
- office system;
- reference information management system;
- risk management system;
- Customs accounting and resource management system.

To adapt or to develop and introduce the independent distant training information systems.

To develop and constantly update the description of the integrated Customs information system's structure – the architecture of its information technologies (IT).

To develop, introduce and maintain the Customs data warehouse and to ensure the high quality data transfer to the interdepartmental data warehouse.

To develop, introduce and maintain an integrated Customs computers and communication network, covering the whole Customs territory of Lithuania, based on the newest information and communication technologies and creating possibilities to integrate the appropriate control systems.

To develop, introduce and maintain interfaces with the information systems of the European Union, Lithuania's institutions and Customs administrations of foreign states.

To relate the automated information systems into an integrated Customs information system and to ensure the integrity of the newly developed systems with it.

To create technical possibilities for legal and natural persons to accomplish Customs procedures via the public networks, to apply information technologies and information networks for simplifying the accomplishment of Customs procedures and formalities.

To improve the activity of the Customs Information Systems Centre ensuring uninterrupted, efficient and safe functioning of the integrated Customs information system.

To apply the internet and intranet data protection technologies in the Customs information systems, to seek to implement strategies of the state information society development and e-Customs.

Implementation of the strategy. For the implementation of this strategy, it is necessary:

- that the Government of the Republic of Lithuania and other state institutions would acknowledge the Customs as an important institution and would understand the significance of its strategy;
- that the officials of the Lithuanian Customs would understand the importance of strategic tasks and contribute to their implementation;
- that the sufficient number of highly qualified officials would work at the Lithuanian Customs;
- that the positive image of the Lithuanian Customs would consolidate in the society;
- to ensure sufficient funding for the Lithuanian Customs activity and its modernisation;
- to prepare or, if necessary, to update and implement the following documents:
  The Lithuanian Customs human resources management strategy;
  The Lithuanian Customs officials training strategy;
  Anti-corruption strategy of the Lithuanian Customs;
  Business strategy of Customs units performing control;
  Strategy for the use of information technologies at the Lithuanian Customs;
  Public relations development strategy.

5. Long Term Development Strategy for Lithuanian Transport System

There is introducing a strategy for Lithuanian transport sector development in the long term prospect. There was carried out an SWOT analysis – analysis of Lithuanian transport sector strengths, weakness, opportunities and threats.

The main strengths are:

- Geographical location favourable for transit movements
- Non-freezing port of Klaipeda, which has modern container terminal;
- Highly developed road transport network
- Good political and economical relations with the neighbour countries.

During the SWOT analysis were determined some ***weaknesses***, main of which are:

- Physically worn railway infrastructure and fleet of railcars;
- Insufficient depth of Klaipeda port channel;
- The "bottleneck" is the connections by road and railways with the EU countries through Poland;
- Not prepared juridical and organizational basement for development of multimodal transport;
- Not prepared juridical basement which regulates mechanism of transport infrastructure modernization and development;
- Damaging impact of transport to the environment;

The ***key opportunities*** for Lithuanian transport sector would be those:

- While preparing for the membership in EU to harmonize juridical standards of Lithuanian transport sector with those that rules in EU;
- To establish logistic centres in Klaipeda and Kaunas and integrate them in to the networks of transport logistics centres of Baltic see region;
- To reach trans-European network status for main transport arteries;
- To promote combined transport.

And finally, the ***main threats*** are considered to be those:

- Insufficient effective coordination of action with neighbour countries while developing TINA networks;
- Lack of funds for renewal of transport infrastructure;

The ***main driving forces***, that should determine development of transport sector of Lithuania in the period from 2002 till 2015, are considered to be those:

- Integration into EU;
- Growth of country economics;
- Growth of economical partnership between Baltic sea region countries;
- Growth of demand for transport and transit service;
- New juridical basement that is harmonized with ES juridical standards;
- Political stability.

That's why in the process of preparation of strategy for Lithuanian transport sector development, ***exceptional attention*** is paid to:

- Modernization of road, rail, Klaipeda sea port and main airports infrastructure in he frames of TINA program;
- Establishment of Kaunas and Klaipeda logistic centres and their integration into European logistic centres network;
- Integration of Lithuanian carriers into EU transport service market;
- Betterment of interaction between transport modes;
- International partnership in the Baltic Sea region.

As you can see the priority attention is given to the solution of transport infrastructure, as well as juridical problems. It is hard to imagine the growth of country economic without efficient transport system and its proper infrastructure. It must be noted, that after expansion of EU, Baltic countries gradually will take a double role: they will act like an integrate part of expanded EU and they will be like a connectors that join old EU member countries with CIS countries.

To our mind, whole that had been said, must be accomplished if we want to see Lithuanian transport system as part of integrated Europe, and what is of even great importance, to create a favourable environment for transit transport.

As you can see the ***long term goals for transit transport development*** are these:

- The establishment of strong relations with neighbour countries and the team wise perfection of freight delivery process;
- The reformation and simplification of border crossing procedures;
- The stimulation of terminal equipment modernization;
- The modernization of Klaipeda port structure, seeking to increase its competition ability with other Baltic sea ports;
- The harmonization of the main juridical acts and regulations with EU juridical system and standards applied in transit transport.

And last, but not least. We are sure, that future research needs in the transport field is of great importance. That's why in the nearest future research works of Lithuanian (and most likely it can be applied to all Baltic countries) transport sector should be directed in these particular ways:

Modelling of Lithuania transport system development, which ***objective is***:
Integration of the most important Lithuanian multimodal infrastructure objects into trans European networks.

Establishment of system for modelling and forecasting transit transport flows between EU and CIS countries, which ***objective is***:
to form the methodical basis for country transit transport policy, which would ensure attractive country transport system and business environment.

Modelling of intermodal freight transportation development which main ***objective is***:
to project the development of intermodal transportation on the national and regional level.

The establishment of new technologies in the rail transport. Main ***objective*** in this particular field is:
with the help of the establishment of modern technologies to speed up an integration into trans-European networks.

The establishment of new multi-functional freight loading technologies in transport terminals. Main ***objective***:
to establish the new multifunctional freight loading technologies in transport terminals, which would ensure the modernization of currently existing terminals, what, would help to rise the competitive ability of whole Lithuanian transport system.

6. Conclusions

1. For analysing new priority ways in international freight transport it is very important to create favourable conditions for that..

2. Those conditions are influenced a great deal by the customs which is an important public institution in charge of performing customs clearance operations.

3. Therefore while speaking of new priority ways in international freight transport, attention should be focused on the following basic issues – the new computerised transit system and their advantages, – the task in the field of customs information systems; – business strategy of the Lithuanian customs – long terms development strategy for Lithuanian transport system.

References

[1] Frank Southworth, Bruce E. Peterson, *Intermadal and intermodal freight network modelling.* Transportation research part c8. 2000, p. 147–166.

[2] Powell Warren B., Carvalko Tassio, A real-time optimisation of containers and flatcars for intermodal operations. *Transportations science*, Vol. 32, No 2 May 1988, p. 110–126.

[3] Coule J. Bordi E., Carinato J, *Transportations third edition* (USA, 1999).

[4] *Transport of fast changing Europe.* Vers un Reseau European des Systems de transport by group transport 2000 plus. (Brussels, 2000).

[5] A. Baublys, *Transport system: models and development and forecast.* (Vilnius: Technika, 2003)

[6] *New customs transit systems for Europe.* Luxembourg: Office for Official Publications of the European Communities, 2001. 55 p.

[7] A. Baublys, N. Batarlienė, A. Jarašūnienė, Implementations of advanced technologies and other means in freight transportation through the territory of Lithuania. *International conference-TST 2002.* 345 p.

INVENTORY CONTROL MODEL FOR THE TYPICAL RAILWAYS COMPANY

***Eugene Kopytov*[1], *Fedor Tissen*[1], *Leonid Greenglaz*[2]**

[1] Transport and Telecommunication Institute
1 Lomonosova str., Riga LV-1019 Latvia
phone: +371 9621337, fax: +371 7383066, Email: kopitov@tsi.lv, fedortissen@tsi.lv

[2] Riga International School of Economics and Business Administration
Meza iela 1, k. 2, Riga, LV-1048, Latvia, Email: gringlaz@riceba.lv

Abstract
The given paper is devoted to inventory control of spare parts for the railway companies. The stochastic model "Continuous Review Fixed-Order-Quantity System" is considered. Main stages of the problem solution are described. The task is solved using software package WinQSB.

Keywords: inventory control, railways, delivery planning, modelling

1. Introduction

One of the most important problems for any railways' enterprise is to maintain the rolling stock in working conditions. We have to organise some stock of rather wide range of spare parts and equipment for rolling stock maintenance. If we increase the stock, it makes more expensive the railway transport services; on other hand, it eliminates the losses from the idle rolling stock, expecting reparations and technical services.

It is quite complicate mathematical task to find the optimal solution for the necessary stock, if you are working in the railways industry [1]. The inventory control management model should take into account the random demand for the spare parts and consider the whole supply chain of the companies, which have influence on spare parts deliveries "manufacturer – supplier – intermediate company – transport company".

In the given paper the inventory control model for one type of product is considered. We describe the example of the inventory model of the most frequently consumed product – rubber bush. Altogether the proposed method can be used in the case of multinomenclate products order. And this case is popular in the supplying process of spare parts and equipment [2].

For the construction the effective inventory control system authors propose to use the modern mathematical tools: inventory management theory, time series and forecasting methods.

2. The stages of the Task Solving

The proposed method of inventory control problem solving consists from the following stages:

- *The development of the mathematical model.* As the basic in given report the stochastic model "Continuous Review Fixed-Order-Quantity System" is chosen [1].
- *The gathering and processing of statistics about the deliveries and spare parts utilization*. In practice the demand for spare parts is probabilistic. The key factor on this stage is the determination of the distribution of the demand for spare parts. In given work the probability distribution of demand on spare parts is described by normal distribution. With

such tear rolling stock on the railways the parameters of this normal distribution are modified yearly, in particular the average value and standard deviation of the demand are growing.

- *The forecast of the demand on spare parts and the technical and economic indices, which determine the effectiveness of work of any railways enterprise.* For this problem solving the exponential smoothing method is used. In further research we propose to use multiple regression model.
- *The determination of an optimal orders quantity.* The problem is solved using developed stochastic model.
- *The obtained results analysis.* Working out of practical recommendations for order planning is the main task of the stage.

3. Mathematical Model

Input parameters of the model are as follows: D – demand of goods per month; C_H – cost of storage of a good unit during a time unit (month); C_0 – ordering cost per order. The main feature of the considered model is the random demand. It is proposed, that the distribution of demand is known.In the next section it will be shown that the demand on the considered product has the normal distribution with parameters changing within the time. We assume that the ordered goods are delivered completely in time, and time period from time of placing an order till time of receiving the goods is constant. This time period is called *lead time*; let's denote it *l.t.* (abbreviation from *lead time*). Because the demand is random, there is the possible situation of deficit, when the demand exceeds the existence quantity of goods. This is most likely possible in period of time from the order of good till its delivery. Using the abbreviation C_{sh} we can denote the cost of shortage (deficit) of a good unit, it is assumed that C_{sh} does not depend on the time of the shortage existence. To decrease costs, associated with the possible shortage, we could create some safety stock of good S. On the other hand, the supplementary stock increases the costs for holding. Main task is the calculation of the repeat order Q and the supplementary stock S. We have to take into account that the sum of costs for good's ordering, holding and losses from deficit should be minimal. The quantity of goods in the stock, when it is necessary to form the new order, is called as *reorder point*; let's denote it as R (*reorder point*). The value R is determined by the demand within the time of delivery $D_{l.t.}$ and safety stock S, that is

$$R = \overline{D}_{l.t.} + S\,, \tag{1}$$

where $\overline{D}_{l.t.}$ – average value of demand within the time of delivery (i.e. the time of placing an order till time of receiving the goods).

Using formula (1) for R, we assume that the reorder point is not less than the average demand within the delivery time. Figure 1 shows the availability of good at each moment of time.

The curve line (Q+S)—A is defined by the random demand; the respective straight line shows the average value of stock, which should be available at the different moments of time t.

We have to keep in mind that small changes in a volume of order imply insignificant changes of summary costs near the optimal order point [1]. So, we could approximately calculate the quantity of the repeat order Q, using the formula:

$$Q = \sqrt{\frac{2C_o\overline{D}}{C_H}}\,, \tag{2}$$

where $\overline{D}$ – average demand per month.

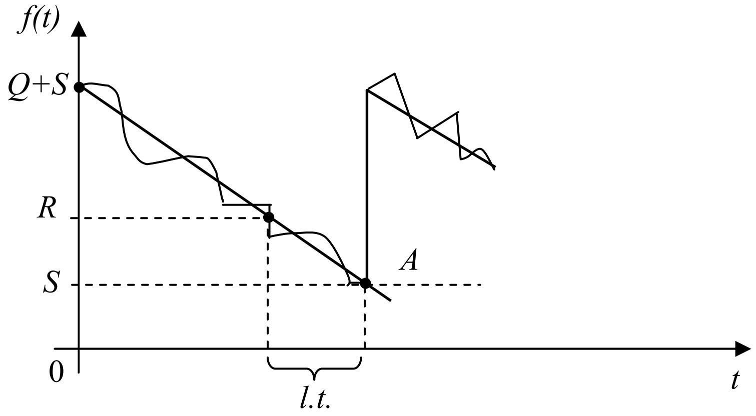


Figure 1. *Chart of availability of good in stock*

Hence, the considered task can be formulated as the problem of determination of the safety stock *S*. For solution of this problem the idea of the marginal approach is used (see, for example, [3]). This marginal approach is contained in the determination of the economical effectiveness from the safety stock increasing to a one unit.

Let us now suppose that at given moment the safety stock *S*, reorder point *R* and probability of the shortage occurrence $p_R = P(D_{l.t.} > R)$ are known. We assume that expected losses from the shortage of a good unit (if $D_{l.t.}$ will exceed *R*) within the month will be equal to

$$C_{sh} \cdot p_R \cdot n = C_{sh} \cdot p_R \frac{\overline{D}}{Q}, \tag{3}$$

where $n = \dfrac{\overline{D}}{Q}$ – quantity of orders during a time unit.

On the other hand, expenses, associated with the holding of the odd unit, are equal to C_H. It is reasonable to increase the safety stock until the expected losses, caused by shortage of a good unit, will be less than holding cost of a good unit, i. e. until

$$C_{sh} p_R \frac{\overline{D}}{Q} \geq C_H . \tag{4}$$

Using formula (4) we can find the optimal solution (optimal value of p_R):

$$p_R = \frac{C_H}{C_{sh}} \cdot \frac{Q}{\overline{D}} . \tag{5}$$

Let $F(R)$ is the known function of demand distribution, but $p_R = 1 - F(R)$; so we have

$$1 - F(R) = \frac{C_H}{C_{sh}} \cdot \frac{Q}{\overline{D}} . \tag{6}$$

If we know the distribution of demand, it is possible to determine the value of the reorder point *R* and the quantity of the safety stock *S*:

$$S = R - \overline{D}_{l.t.}$$

In the considered task the shortage of stock is divided into the two parts:

- part *p*, which is delivered to the customer (is reordered);
- part *q*, which is lost (customer could receive a good from another supplier or refuse the order).

It is clear that $p + q = 1$.

Let's denote by Π the losses from the reorder of a unit of good, by Θ the losses from the lack of a unit of order, and by *B* fixed losses in case of shortage appearance due, for example, the payment for extraordinary transportation of spare parts. Denote by *b* the average value of shortage within the delivery time. Generalisation of the formula (6) in the considered situation is the following equality [4]:

$$1 - F(R) = \frac{C_H Q - B \cdot D \cdot f(R)}{q C_H Q + (p \cdot \Pi + q \cdot \Theta) D}, \tag{7}$$

where $f(R)$ is density of the demand distribution, and Q is calculated from the formula

$$Q = \sqrt{2D(C_0 + (p \cdot \Pi + q \cdot \Theta)b + p_R \cdot B) / C_H} \,. \tag{8}$$

For the search of optimal values R and Q we will use the recursive procedure considered in [4].

4. Processing of the Statistical Data and the Demand Forecast

In given example we consider the weekly statistics of the demand for the rubber bushes. The statistics for the period October 2000 to September 2001 is presented in the Table 1.

Table 1. Weekly statistics of the demand for the good

Week	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*
Demand, units	552	559	450	562	552	727	719	704	347	650	764	690	919
Week	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	23	24	25	26
Demand, units	321	795	739	700	686	517	378	613	638	701	644	634	770
Week	27	28	29	30	31	32	33	34	35	36	37	38	39
Demand, units	790	855	770	739	843	575	678	570	553	594	749	682	578
Week	40	41	42	43	44	45	46	47	48	49	*50*	*51*	*52*
Demand, units	600	851	808	519	690	791	946	960	645	1055	758	872	639

We can see in Table 1, that the demand is random. Let us identify the distribution of the weekly demand within the first year. We'll solve this task using software package STATISTICA. On the first step of analysis a histogram of demand distribution was created (Figure 2). On the second step using Kolmogorov-Smirnov method hypothesis of normal distribution was tested (Figure 3). Using results of both steps we can assume that demand on spare parts is described by normal distribution.

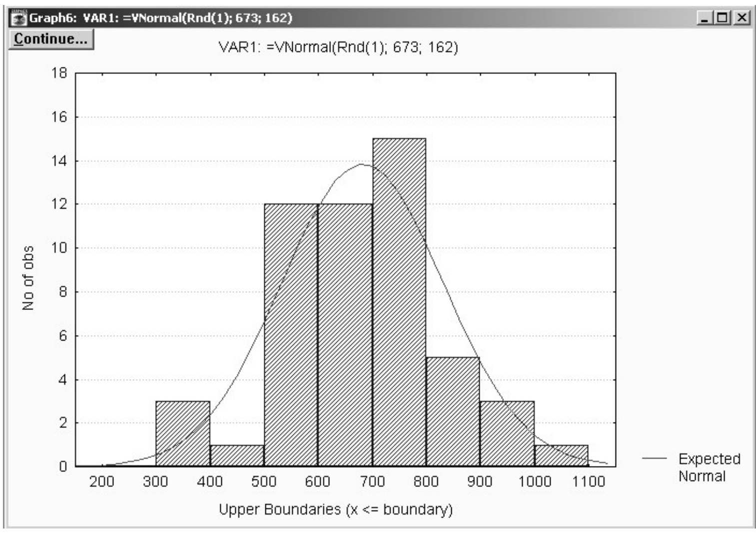


Figure 2. *Histogram of the demand distribution*

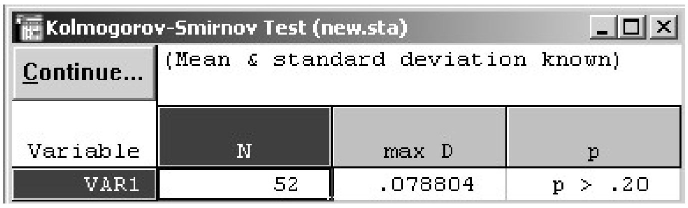


Figure 3. *The results of Kolmogorov-Smirnov test*

Similar conclusions have been received for other periods (April 2001 to March 2002, October 2001 to September 2002, etc). For each observed year numerical parameters of the normal distribution of demand (average value μ and standard deviation σ) were calculated (see Table 2). Obtaining the data in the table, wee see that the average value and standard deviation of the demand are increasing with time.

Table 2. Numerical parameters of the normal distribution of weekly demand of good in different periods

	October 2000 – September 2001	April 2001 – March 2002	October 2001 – September 2002	April 2002 – March 2003	October 2002 – September 2003
μ	673	817	948	934	990
σ	162	188	218	250	271

Our task is to make prediction of the numeral parameters of demand distribution into the future (the stock planning periods), and we need to identify the trends of parameters changing. There are several methods of time series forecasting, in our case we have used the exponential smoothing method. In further we try to use multiple regression methods, which gives more possibilities for forecasting in different economical situations. In our case the forecast of the weekly demand distribution for the good for the period October 2003 to September 2004 is the following: μ=1035; σ=302.

5. Calculation of the Optimal Orders for Spare Parts

We consider the inventory control system with time unit equal to 1 month. Using the forecast of the numerical parameters of the normal distribution of weekly demand of good (see Section 4), we can determine the forecast of the appropriate parameters ($\mu 1$ and $\sigma 1$) for a month period (approximately it is 4 weeks):

$$\mu_1 = 4 \cdot \mu = 4140, \text{ and } \sigma_1 = \sigma \cdot \sqrt{4} = 604 .$$

For problem solving we have used the software package WinQSB [4]. There are several methods of inventory control in the module "Inventory Theory and System" in the package. We have chosen the stochastic model of the repeated order (see formulas (7), (8) above), which is named in the module as "Continuous Review Fixed-Order-Quantity (s, Q) System". The form with input parameters of the model is presented on Figure 4.

We could see in the given table that the optimal order quantity is 12 140 units, value of the reorder point is 4723 units, at the same time the general expenses of the company per month, associated with the inventory control of this good, are equal to 139.66 USD (or 1675.92 USD per year). As the comparison, in the last year real the average demand for spare parts was smaller, but the general expenses of the company per year, associated with the inventory control of this good, were greater – 2051 USD.

DATA ITEM	ENTRY
Demand distribution (in month)	Normal
Mean (u)	4140
Standard deviation (s>0)	604
(Not used)	
Order or setup cost	178
Unit acquisition cost	0.18
Unit holding cost per month	0.01
Estimated % of shortage will be backordered	80
Unit backorder cost	0.29
Estimated % of shortage will be lost	20
Unit lost-sales cost	0.125
Fixed cost if shortage occurs	50
Lead time distribution (in month)	Constant
Constant value	0.7
(Not used)	
(Not used)	

Figure 4. *Input data*

The results of problem solving using "Continuous Review Fixed-Order-Quantity (s, Q) System" model are presented in Figure 5.

11-20-2003	Input Data	Value	Inventory & Cost Analysis (month)	Value
1	Demand distribution	Normal	Optimal reorder point (s)	4723.374
2	Average demand (month)	4140	Optimal order quantity (Q)	12140.18
3	Std. dev. of demand (month)	604	Average minimum on hand	1825.377
4	Unit acquisition cost	$0.1800	Average maximum on hand	13965.56
5	Order (setup) cost	$178.0000	Average on hand inventory	7895.468
6	Unit holding cost per month	$0.0100	Safety stock	1825.374
7	Estimated % of shortage	80%	Mean shortage during lead time	0.0166
8	Unit backordered cost	$0.2900	% of shortage during lead time	0.0153%
9	Estimated % of shortage lost	20%	Total order/setup cost	$60.7009
10	Unit lost-sales cost	$0.1250	Total holding cost	$78.9547
11	Fixed shortage cost	$50.0000	Total backorder cost	$0.0013
12	Lead time distribution	Constant	Total lost-sales cost	$0.0001
13	Average lead time (month)	0.7	Total fixed shortage cost	$0.0026
14	Std. dev. of lead time	0	Total shortage cost	$0.0041
15	Average lead time demand	2898	Total inventory relevant cost	$139.6596
16	Std. dev. of lead time demand	505.3427	Expected total acquisition cost	$745.1998

Figure 5. *Results of problem solving in the package WinQSB*

6. Conclusions

Using proposed method we can find the optimal solution in stock planning and decrease expenses in the inventory control system on the railways.

Given method of continuous review quantity in stock can find practical application in any transport company.

Further guidelines of the current research are the following: taking into account in the model the possible restrictions on the resources (stock capacity, capital and others), considering multinomenclate task, improving the demand forecast models.

References

1. Гринглаз Л., Копытов Е. Основы управления запасами с примерами решения задач на компьютере: Учеб. пособие. – Рига: ВШЭК, ИТС, 2002. – 91 с.

2. Тиссен Ф. Проблемы управления запасами на железой дороге. В кн.:. Программа и тезисы/ *Научно-практическая и учебно-методическая конференция «Наука и технология – шаг в будущее»*, Институт транспорта и связи, 7-11 апреля 2003, Рига, Латвия, с.119-120.

3. Bierman H., Bonini C., Hausmann W. Quantitative Analysis for Business Decisions. Irwin, 1991, 742 p.

4. Yih-Long Chang. WinQSB Version 2.0, John Wiley & Sons, Inc., 2003, 230 p.

METHODOLOGY FOR MULTIMODAL NETWORK ANALYSIS BY USING CONCEPT OF VIRTUAL LINKS AND NODES

Aidas Vasilis Vasiliuskas

Vilnius Gediminas technical university, Transport management department
Plytines 27, LT-2016, Vilnius, Lithuania
E-mail: aidasv@ti.vtu.lt

Abstract

This paper deals with the analysis of freight transportation over extensive multimodal networks. It is about the use of "virtual links and nodes" which enables to analyse network that consists of different types of roads, and where all transport operations (including all interface services in nodal platforms and terminals) are performed by different transport modes. Paper presents an overview of the applied methodology - the procedure for building the virtual network with a special notation of virtual links and nodes that permit to attach the cost functions specific to each virtual link or node.

Keywords: Transport Networks, Virtual Links and Nodes, Multimodal Transport

1. Introduction

Until a few years ago, transport models were essentially focused on passengers flows. More recently, some freight specific network models have been developed but they are essentially analysing networks from a "link" point of view rather than from a "node" point of view. Even if some of them deals with the different operations performed at the nodes, i.e. loading/unloading, transhipment or transit, their output still targets mainly transport flows on the networks.

As a result of a trend towards economic globalisation and road transports expansion, a great attention is paid to reorganisation of networks, especially to possibilities proposed by intermodal transports. Thus, there is a need for a better modelling of the functions assumed by nodes, and mainly for analysing minimum cost solutions to transportation flows over extensive international networks. Actually, with this presented model, they can be easily computed in a comprehensive way, taking into account all networks effects.

So this paper will present an overview of the methodology which is applied to model all transport operations: the systematic use of virtual links to model all transport operations on real links and in nodes, as well as the particular notation devised for connecting these virtual links to their relevant cost functions.

2. Building a Virtual Network

A geographical multimodal transport network is made of links like roads, railways or waterways, on which vehicles moves and nodes like terminals that connects mentioned above links and where goods are loaded, unloaded, transhipped or processed in different ways. To analyse transport operations over the network, costs or weights must be attached to these geographical links and to the connecting points. However, most of these infrastructures can be used in different ways and with different costs. For example, boats of different sizes and operating costs can use the same waterway; at a terminal a truck's load can be transhipped on train or simply unloaded. Normally, the costs of these alternative operations are different .

A simple geographic network does not provide an adequate basis for detailed analysis of transport operations where the same infrastructure is used in different ways. To solve this problem, the basic idea, is to create a virtual link with specific cost for a particular use of an infrastructure. So here is proposed a methodology and an algorithm which creates a complete "virtual network" with all the virtual links corresponding to the different operations which are feasible on every real link or node of a geographic network.

The solution can be presented first in way by using the example of a simple waterways network, as shown in figure 1.

This network consists of 4 nodes and 3 links. The links represent waterways that can support ships of 300, 1350 and 600 tons respectively. To get from node a to node d, the least expensive route may imply a transhipment of the goods at node b to a ship of 1000 or 1350 tons, and transferring the goods to a ship of 600 tons at node c. Another possibility is to start using ships of 600 tons at node b and keep using these to get to d. Finally, transport with ships of 300 tons only would also be possible. Hence, there is a clear problem of transport means choice given the different ways this waterway system can be used.

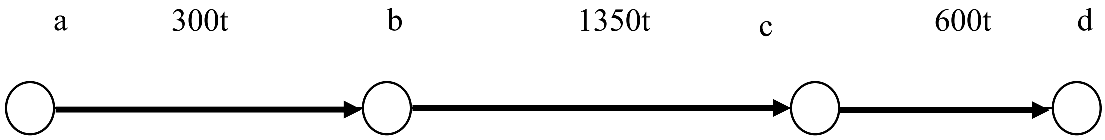


Figure1. *Simple network of waterways*

This problem can be successfully handled on the corresponding virtual network illustrated in figure 2, provided that the relevant costs are attached to each of the virtual links. As can be seen, the solution involves the creation of three set of virtual nodes b1-b5, c1-c5, d1-d2 and a set of virtual links connecting these nodes. Each real link has been split in as many virtual links as there are possible uses and their end-nodes are connected by new virtual links corresponding to simple transit or transhipment operations at the real nodes.

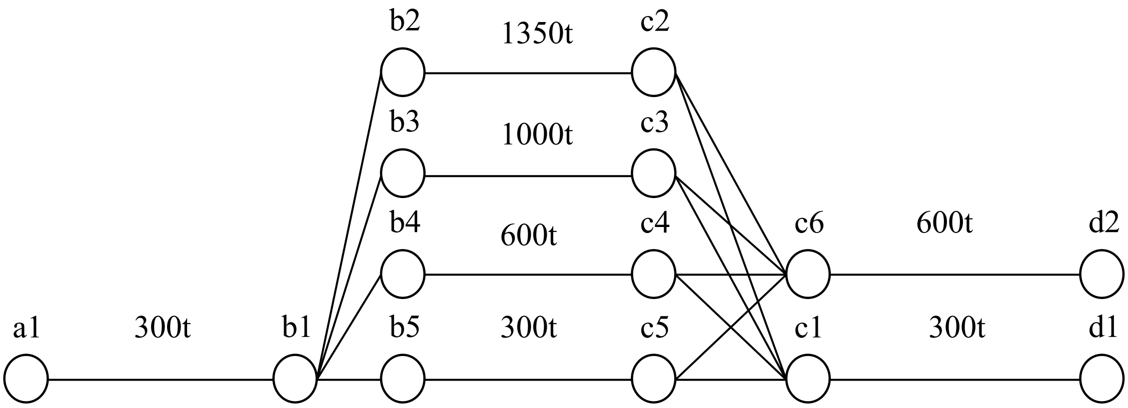


Figure 2. *Corresponding virtual network*

In this way, this network with multiple means use is represented by a unique but more complex network on which each link corresponds to a unique operation with a specific cost. Then, the cheapest path can be computed by means of an algorithm such as the Johnson's, an optimised version of the well-known algorithm of Dijkstra for low density graphs. The resulting solution is an exact solution, taking all the possible choices into account.

The same approach can be applied to multimodal networks where transhipment between modes must be analysed. The case is illustrated in figure 3, where there are 4 nodes (a, b, c and d) and 5 links (1 to 5) with two modes, inland waterway (W) and railway (R) which both offer two different transport means (W1, W2 and R1). We can use this example to demonstrate how the virtual network is build on the basis of simple, though somewhat complex notation, which provides a convenient way to link cost functions to virtual links.

Table numerates the elements of the real network.

Table 1. Real network

Link number	Node 1	Node 2	Type of link
1	a	b	W1
2	b	c	W2
3	b	d	R1
4	d	c	R1
5	a	d	R1

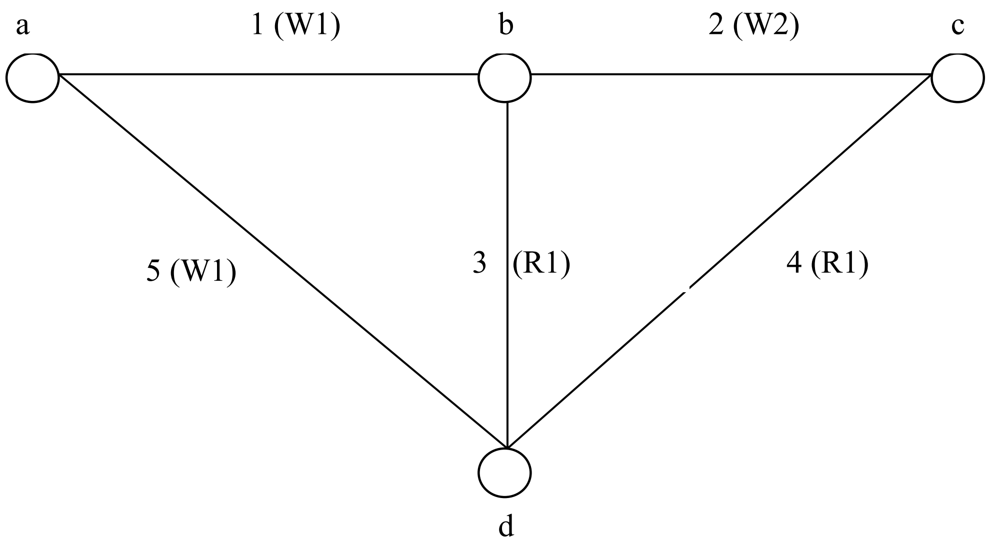


Figure 3. *Multimodal network*

In a first step, the virtual links corresponding to the real links must be generated. These are defined in table 2 by their end-nodes, the notation of which indicates successively the node, the real link, the mode and the means.

Table 2. Travelling (virtual) links

Real links	End-nodes of virtual links	
1	a1W1	b1W1
	a1W2	b1W2
2	b2W1	c2W1
3	b3R1	d3R1
4	d4R1	c4R1
5	a5W1	d5W1

In a second step, these virtual links must be connected by transit or transhipment virtual links. To keep things as simple a possible, we just enumerate in table 3 the connecting virtual links related to node b. They can be viewed in figure 4 this network, the boldfaced links represent the links of the real network, possibly split up. The dotted links represent the transit virtual links, while the transhipment links are indicated by a thin continuous line.

Table 3. Connecting virtual links to node b

Real node	End-nodes of virtual links	
b	b1W1	b1W2
	b1W1	b2W1
	b1W2	b2W1
	b1W1	b3R1
	b1W2	b3R1
	b2W1	b3R1

This network is not yet complete because it does not contain entry and exit nodes in/from the network. This can be done by the creation of additional virtual nodes associated to loading and unloading operations at nodes where they are possible. Those entry or exit points in the virtual network are referenced by adding "000" to the real node number. They must also be connected to other nodes by appropriate virtual links.

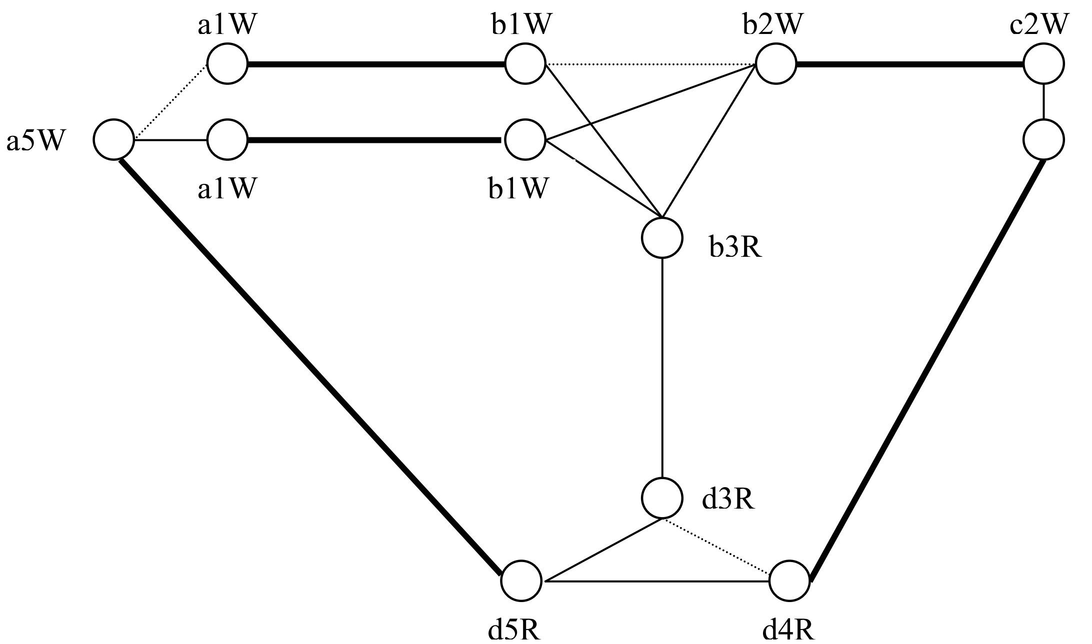


Figure 4. *Partial virtual network*

In general, the weight given on a link can vary with the direction it is used (loading and unloading operations for instance don't have necessarily the same cost). To solve that problem, all virtual nodes are doubled at generation time by adding a + or a – sign to their code. By the same, all links are split into two oriented arrows connecting these new nodes. This is illustrated for real node b in figure 5. Those "doubled" virtual nodes also make it possible to avoid "unwanted movements", like an unloading followed by a loading to circumvent a forbidden transhipment operation.

Obviously, the codification used in the preceding figures is not suitable for real applications; using a single letter as node label would limit the size of the real network to 26 nodes! That's why the virtual nodes are coded in the following way: a plus or minus sign, 5 digits for the node number, 5 digits for the link number, 2 digits for the mode and 2 digits for the means. Each label is thus represented by a 14 digits number preceded by a sign.

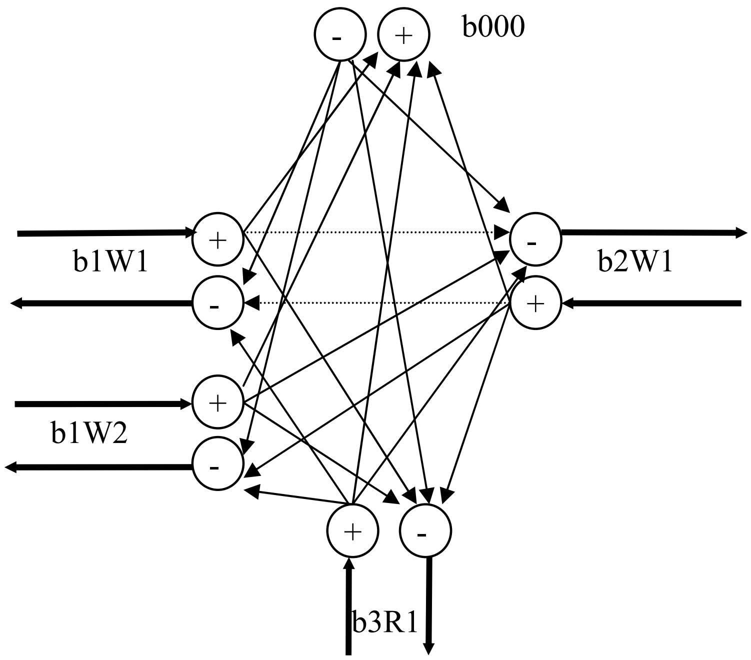


Figure 5. *Detailed virtual network at node b*

3. Cost Functions on Virtual Networks

After overview of the basic methodology, now it is necessary to explain how the cost functions are connected with the virtual network and which are their characteristics. As usual in transportation analysis again there is used the concept of "generalised cost" which allows to integrate all factors relevant for transport decision making in terms of monetary units. The specific cost functions which compose the generalised cost must be coherent across modes and means, but their functional forms can be freely chosen.

The virtual network requires the development of four types of cost functions, which are associated with specific virtual links through their notation:

- The use of a **travelling cost function** is indicated by a difference between the virtual link's indices denoting the real end-nodes;
- **Transit costs** are applied when the link's indices denoting the two connected real links vary, whereas the mode and means indices remain the same;
- For **transhipment costs**, the link's indices of the connected real links should vary, as well as the mode and/or means indices;
- For **loading/unloading** costs, one of the two indices of the real links is "00000".

These four types of function are made of the following elements:

1. All the costs related to moving a vehicle between a trip's origin and destination, like labour, fuel, insurance, maintenance costs, or tariffs;

2. The cost of inventory of the goods during transportation;

3. Handling and storage costs or tariffs, including packaging, loading and unloading and services directly linked to a transport;

4. All residual indirect costs like general administrative services which may be assigned to transports on an average basis.

Besides these costs, a full account of transport costs in a multimodal multimeans context should include some relative costs of transport quality differences, like differential reliability, safety, etc., if it is at all possible. These relative costs may vary from one category of goods to another. Unhappily, information about these factors is very scarce. Note, however, that these modal quality differences are taken into account to some extent by the adjustment made on cost functions to obtain the best fit of the model. To that effect, whenever possible, it is recommended to analyse separately different categories of goods with specific O-D matrixes and traffic assignments.

So it is no surprise, that real prices for transport services (especially in the case of intermodal transport) are very difficult to obtain, so that in most of applications rather is used transport production costs as an approximation. From a socio-economic point of view these are certainly the relevant parameters. However, in some cases, they can also be taken as acceptable approximation of prices or tariffs at least on the marginal contestable markets where modes compete for market share.

Since the network model imposes that the total cost function of a trip be an additive linear function with respect to distance, the vehicle's costs for transporting one weight-unit (ton) over successive links of distance s_j with a vehicle of type θ was defined as:

$$\sum_{i \in l_i} a^{\theta}.H_i^{\theta} + \sum_{j \in l_j} B^{\theta}.s_j \ , \qquad (1)$$

where

- $A^{\theta i}$ - the sum of loading and unloading costs plus the possible transhipping costs on a (not travelling) virtual link *I*;
- $B^{\theta}.s_j$ - the sum of all the costs of moving the goods on a virtual link *j* of the route.

We adopted the following formulas from Blauwens et al:

$$A^{\theta i} = a^{\theta}.H_i^{\theta} = \frac{F^{\theta}.H_i^{\theta}}{u.T^{\theta}} \text{ and } B^{\theta}.s_j = \frac{F^{\theta} + b^{\theta}.u}{u.v^{\theta}.T^{\theta}}.s_j, \qquad (2)$$

where :

- F^{θ} - annual fixed costs of the vehicle (capital annuity, insurances, maintenance and wages of the crew);
- H^{θ}_i - time needed for a particular handling operation on virtual link *i*;
- u - number of working hours per year;
- T^{θ} - average load of the vehicle;
- b^{θ} - energy consumption (per hour);
- v^{θ} - average speed.

Often, formula $H^{\theta}_i = a + b \, . \, Q^{c}$ was used for estimating the loading / unloading / transhipping times, *Q* being the handled quantity. The labour costs of handling the goods at different points on the route can be written as:

$$a^{\lambda} \sum_{i \in l_i} H_i^{\theta} \ , \qquad (3)$$

where a^{λ} is the labour cost per unit of time. The inventory cost of the goods during transport which is proportional to handling and shipping times can be expressed as

$$a^{\rho}.\left(\sum_{i\in l_i} H_i^{\theta} + \sum_{j\in l_j} \frac{s_j}{v^{\theta}}\right) = p.r.\left(\sum_{i\in l_i} H_i^{\theta} + \sum_{j\in l_j} \frac{s_j}{v^{\theta}}\right), \tag{4}$$

where a^{ρ} represents the opportunity cost of the goods per unit of time. Thus, altogether, the total transport cost of the transportation task defined by the O-D matrix can be written as

$$TC = \sum_{i\in l_i}\sum_{j\in l_j}\sum_{\theta}\left[\left(B^{\theta} + \frac{a^{\rho}}{v^{\theta}}\right).s_j + \left(a^{\theta} + a^{\lambda} + a^{\rho}\right)H_i^{\theta}\right].Q_{l\theta}\,. \tag{5}$$

This is the total cost which is then minimised to find the best choices of modes, means and routes on an all-or-nothing basis.

4. Conclusions

This paper has presented the methodology used to build transport models allowing complete and detailed analyses of multimodal freight transportation networks including intermodal solutions and bundling operations within terminals. It is based on the concept of "virtual link" which is applied in a systematic way to model all transport operations: travelling, loading, unloading, transhipping, waiting, handling, paying tolls, etc. with several modes and means of transportation. With this methodology and a suitable notation of virtual links and nodes, an appropriate algorithm automatically generates a complete "virtual network" representing all transport activities and interfaces on extensive multimodal networks. This "virtual network" can be seen as a monomodal representation of a multimodal network; it permits to solve problems of assignment to modes, means and routes over extensive networks in one single step.

References

1. P.T. Harker. *Predicting intercity freight flows*. VNU Science press, 1987.
2. T.G. Crainic, M. Florian, J. Guélat, and H. Spiess. Strategic planning of freight transportation: Stan, an interactive graphic system. *Transportation research record*, 1283, 1990.
3. B. Jourquin and M. Beuthe. Transportation policy analysis with a geographic information system: the virtual network of freight transportation in Europe. *Transportation Research C*, 4(6), 1996, 359-371.
4. D.B. Johnson. A note on Dijkstra's shortest path algorithm. *Journal A.C.M.*, 20, 1973, 385-388.
5. E.W. Dijkstra. A note on two problems in connection with graphs. *Numerische Mathematik*, 1, 1959, 269-271.
6. D.T. Kresge and P.O. Roberts. *Techniques of Transportation Planning: Systems Analysisand Simulation Models*. Brooking Institution, Washington DC, 1971.
7. W. Baumol and H. Vinod. An inventory theoretic model of freight transportation choice. *Numerische Mathematik*, 16, 1970, 413-421.
8. B. Jourquin, M. Beuthe, and L. Demilie. Freight bundling network models : Methodology and application. *Transportation Planning and Technology vol. 2*, 1999, 157-177.

INVESTIGATION OF THE POSSIBILITY TO ESTABLISH A LOGISTICS CENTRE IN VILNIUS REGION AND ITS POTENTIAL BENEFITS

***Leonas Povilas Lingaitis*[1], *Olga Fadina*[2]**

Department of Railway Transport, Vilnius Gediminas Technical University
J. Basanavičiaus g. 28, LT-2009 Vilnius, Lithuania
leonasl@ti.vtu.lt [1], olga@lineka.lt [2]

Abstract

The paper discusses the need for the establishment of a logistics centre in the region of the Lithuanian capital Vilnius and possible ways of implementing this idea. The European practice with the logistics centres and their benefits in providing freight transportation services and for the development of the regions is also studied. The role of the logistics centre for Lithuanian economy and its integration into the EU structures is outlined. Some financial aspects including sources of funding and the distribution of the investments in design and construction in per cent are considered.

It is emphasized that the concentration of various enterprises associated with transportation in the particular area produces a synergy effect thereby reducing the costs of services.

Keywords: logistics centre, cargoes, expansion, investments, funding, design, transport services

1. A Major Concept of Establishing a Logistics Centre in Vilnius Region

The geographical location of Lithuania as well as growing commercial and freight flows, the country's integration into the European Union and rapidly developing business centre Vilnius situated near the future Eastern border of the EU are the main factors determining the need for establishing a logistics (freight) centre in this region.

A logistics centre is a particular territory where such services as cargo transfer, storage, distribution over the territory of one or more countries, customs meditators, insurance, maintenance and repair of transport facilities, etc. are provided.

The main distinctive features of a logistics centre are as follows:

- Centres of logistics are usually established on the principle of business partnership of private and state enterprises. The local or central authorities usually initiate and support the development of logistics centres because such centres promote territorial and regional development.
- All kinds of services associated with transportation are provided at the logistics centres. They include the provision of operators, freight forwarders, brokers, loaders and customs meditators.
- Logistics centres usually occupy the area of 100 – 200 ha, in some cases, even reaching 1000 ha, where warehouses, terminals, special storage facilities for harmful materials, refrigerators, parking areas and other spaces used by the operators and other staff are found.
- According to the laws of free market competition, the centre should be accessible to all transport and logistics firms for developing their business.
- The concentration of firms at the particular area causes a synergetic effect allowing to reduce the costs of transport and logistics services.
- Large investments into information technology systems, storage of dangerous materials, provision of services and "know how" required considerably complicate the

performance of individual logistics enterprises. The concentration of enterprises in a special transport centre allows them to share the expenses as well as to provide services more effectively.

- All over Europe, centres of logistics are cooperating with the same national or international enterprises. This helps to optimize the flows and distribution of cargoes.
- Logistics centres make a part of a flexible and efficient transport system providing such services based on the advanced technology as centralized information systems, EDE communication systems, etc.

The competitiveness of a logistics centre is ensured by:

- The availability of means of communication with motor road network;
- The availability of means of communication with railway network;
- Optimization of usage of transport facilities;
- Cargo transfer facilities;
- Integrated logistics services;
- Accessibility of information technologies;
- Lower expenses.

2. The Demand for Logistics Centres in Europe and Lithuania

Logistics centres play an important role in solving problems of the efficiency and competitiveness of European transport.

In last decades the sector of freight transportation has been rapidly expanding. As a result, today the customers (passengers), operators and state institutions are faced with the problems of bottlenecks, pollution, accidents and low economic efficiency of investments.

In most of the EU member – states the logistics centres are working efficiently (there are 22 of them in Germany, 19 – in Spain, 7 – in Denmark) [1, 2]. The establishment of such centre in Lithuania is of primary importance for maintaining the competitiveness of its economy on EU market. In the Baltic sea region, Finland, Denmark and Germany are making efforts to create a network of logistics centres “NeLoC”. The EU supports the creation of logistics centres and their networks via “Marco Polo” and other programmes and foundations.

One of the most effective approaches to solving the above problems is the concept of intermodal transport, implying the integration of various transport systems, i.e. motor, rail, sea and air transport gaining the maximum advantage of each and maximizing its benefits to passengers.

In this context, a concept of the logistics centre is one of the main elements of the European transport network expansion scenario as well as a major tool of linking transport systems and logistics firms which need to have a developed infrastructure along the main motor- and railways. The most relevant factor of intermodal transport is the synergy of the integrated logistics centre.

The establishment of a logistics centre in Vilnius is in line with the national policy towards the expansion of the transport system.

In the National Transport Development Programme the priority is given to complex development and enhancement of transport infrastructure, technology and services in all areas, and, primarily, in the area of international – level transit relying on transport corridors [3, 4]. The Programme states that Lithuania will further adhere to the concept of integrating its transport facilities into the international multimodal transport system, because the effectiveness of the provided transport services largely depends on the advancement of various transport subsystems, including sea, railway, coordinated highway and air transport, as well as the integral market and logistics – based interrelationship.

For this purpose, the Programme provides for the creation of cargo distribution logistics centres at major transport junctions associated with a very important aspect of the European policy known as "intelligent" transport system development.

The geographical location of the logistics centre (at the future Eastern EU border) could attract the transitional flows of cargoes from the EU and ISU states.

3. The Role of the Logistics Centre in Economic Development of Lithuania

Centres of logistics not only help to increase the efficiency of transport but promote the development of the territory where they are established.

A logistics centre to be built in Vilnius region will contribute to the development of Lithuanian economy in the following way:

- *Promoting the integration into the EU structures*. The centre will be created taking into account the EU policy and requirements. It is aimed to maintain the competitiveness of Lithuanian economy on the EU market.
- *Introducing innovative technologies*. Since the purpose of the centre is to provide high – quality transport and logistics services, it will be based on the most advanced technologies. From the first stages the efforts will be made to cooperate with the International Standards Organization for obtaining ISO 9002 standard.
- *Reducing pollution*. A concept of intermodal transport and a logistics centre is an environmentally – friendly solution, allowing the concentration of the services provided now by individual enterprises spread over the country and combining the services of road, railway, air and sea transport.
- *Attracting the investments and creating new workplaces*. The experience of the existing logistics centres [5, 6, 7] has shown that after the completion of the first construction stages, about 1000 workplaces may be created, with up to 100 various enterprises located at the centre. This will attract new investments and promote the creation of new small- and medium- size enterprises.

4. Services Provided by a Logistics Centre and its Enterprises

The following services may be provided:

- Cargo transfer from one kind of transport (railway) into another;
- Customs warehouses;
- Warehouses – refrigerators;
- Packaging and marking of goods;
- Cargo distribution centre;
- Parking places for trucks;
- Petrol stations;
- Washing and repair of containers, trucks and semi – trailers;
- Services of customs meditators;
- Insurance, bank and other financial operations;
- Short – term office renting;
- Multimodal transportation;
- Carriage by road transport;
- Carriage by railways;
- Carriage by air transport.

5. The Establishment of a Logistics Centre

A public organization might establish and coordinate [8, 9] the performance of a logistics centre. The administration and business may also be carried out by an enterprise established by a public institution. The income obtained should be invested into the development of the centre and the infrastructure of roads.

The institutions which may establish a public enterprise are as follows:

- Ministry of Transport;
- Local government institution;
- Administration of the head of Vilnius region;
- Lithuanian Association of Automobile Operators “LINAVA”;
- EUROPLATFORM E.E.I.G. Association;
- “Lithuanian Highways” Association;
- National Association of Customs Warehouses and Import – Export Terminals;
- Lithuanian Association of Freight Forwarders “LINEKA”;
- VGTU Transport Research Institute;
- Other institutions.

The Ministry of Transport should be interested in developing transport services and in the increase of transit cargo flows in Lithuania. The establishment of the above centre may increase the competitiveness of Lithuanian transport firms on the EU market as well as being in line with the national policy in transportation.

The same goals are pursued by the “LINAVA” Association including over 1000 enterprises providing transport services.

The Association “Lithuanian Highways” and the Association of National Customs Warehouses and Import – Export Terminals and the National Association of Freight Forwarders “LINEKA” also support the idea being the members of the initiative group of the establishment of the logistics centre.

The local government should be interested in reducing the unemployment in the city (according to the rough estimate, about 1000 workplaces could be created) and in new investments. The administration of the head of Vilnius region should also be concerned with the future development of the region. They could be among the founders providing the right of using a lot.

The Association EUROPLATFORMS E.E.L.G. is a founder of a number of logistics centres in various European countries and in Lithuania. A mission of the Association is the creation of transport and logistics centres following the main principles of the development of the joint EU multimodal transport network.

6. Location of the Logistics Centre and the Investments Required for its Establishment

Since the centre will be related to multimodal transport, a plot chosen for it should be

- nearby major roads;
- not far from a railway junction.

The experience of other logistics centres as well as the preliminary calculations have shown that the area of about 200 hectare is needed for the establishment of a logistics centre in Vilnius region. A number of alternatives of establishing the centre in Trakai, Vilnius or Šalčininkai regions are being considered.

The administration of the head of Vilnius region is responsible for the provision of the plot because the establishment of the centre is in the interests of the wider public and state institutions will also take part in its establishment.

In the tables and figures presented below, the distribution of the investments into the logistics centre is shown.

Table 1. Investments into the logistics centre in 2004 – 2011
The plot is 100 hectare (with the area for expansion up to 200 hectare)

No	Item	Sum of money, %
	Design research work	
1.	VRLC engineering project	3,96
	Construction operations related to infrastructure development	
2.	Sites for container loading	5,20
3.	Parking places for cars	0,79
4.	Parking places for trailers	8,14
5.	Infrastructure	6,68
	Intermediate sum	**24,77**
	Construction operations	
7.	Terminals and warehouses	26,13
8.	Administrative spaces	3,85
9.	Commercial spaces	19,68
	Intermediate sum	**49,66**
	Railway	**25,56**
	Total	**100**

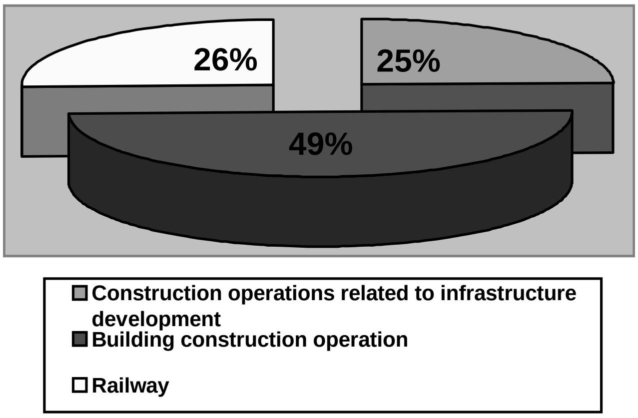


Figure 1. *Distribution of VLRC investments, per cent*

Table 2. Financing of the project of Vilnius region Logistics Centre

No	Issue	%
1.	PHARE	60
2.	Vilnius local government	20
3.	The logistics centre of Vilnius region	5
4.	Other partners	15
	Total	**100**

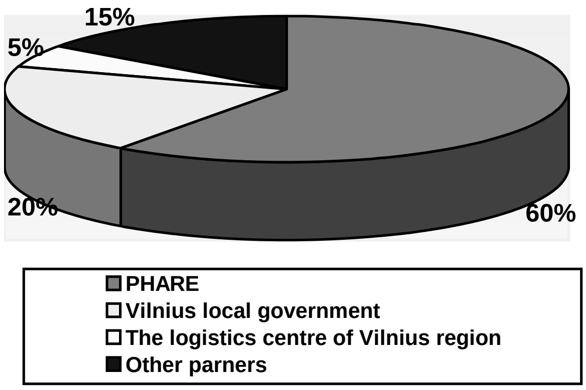


***Figure 2**. Distribution of VRLC funding sources*

Table 3. Distribution of investments into the project of the study of Vilnius region logistics centre
The plot is 100 hectare (with the area of up to 200 hectare for expansion)

No	Issue	%
1.	Project administration	*7*
2.	Study of possibilities	*15*
3.	Detailed plan	*23*
4.	Design solutions	*12*
5.	Developing the model of LC concept, services and management	*12*
6.	Search for foreign and business clients	*10*
7.	Establishing the links between EU logistics centres	*6*
8.	Conferences, presentations, image – making	*8*
9.	Project reserve	*7*
	Total	***100***

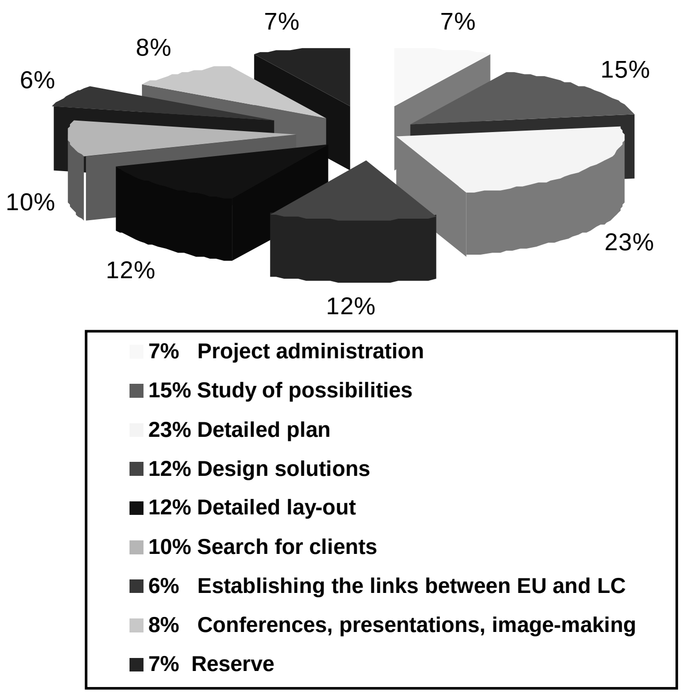


***Figure 3.** Distribution of VRLC finances, per cent*

As can be seen from the tables and diagrams, a major source of financing is the PHARE programme. About half of the money will be invested into the construction work.

7. Conclusions

- Based on the practice of the European countries, it may be stated that a logistics centre integrating various transport facilities should be established in Vilnius region.
- The above logistics centre should be integrated into a general system of the European centres of logistics.
- The logistics centre considered will be an active partner in the network of logistics centres in Lithuania, including the centres in Klaipėda, Kaunas and Vilnius regions.
- A so – called "from door to door" system should be created which will make it unnecessary for various business firms to have a warehouse in Vilnius.
- The logistics centre will help to create an environmentally – friendly transport system and concentrate the cargoes carried via international routes.
- The logistics centre created will help to avoid bottlenecks in Vilnius and on the highways.
- A logistics centre is a centre of business services which will create new workplaces.
- A logistics centre will ensure high – quality services, freight safety and high level of logistics operations.
- State institutions should support the establishment of a logistics centre, because it will help to maintain the competitiveness of Lithuanian economy on the unified EU market.

References

1. Nestler Steffen Deutsche GVZ-Gesellschaft mbH. New logistics solutions in Europe based on the German experience. Germany, 2003.
2. Burkovskis, R.; Palšaitis R. Interaction at the Klaipeda Sea Port and Railway transport. *Transport*, Vol XVII, No 2, 2002, 71–75 p.
3. Bentzen Kent. Network of Logistics Centres in the Baltic Sea Region Area: Development and Perspectives. *TRANSBALTICA-03*. Vilnius: Technika, 2003, 9-13 p.
4. Bentzen Kent, Hoffmann Tobias, Bentzen Lars, Nestler Steffen, Nobel Thomas. Best Practice Handbook for Logistics Centres in the Baltic Sea Region. Association Hoja Fondas, Lithuania, 2003.
5. The investment project of the logistics centre in Vilnius region for 2004 – 2014. Construction firm "Vilnius plan", 2003 (In Lithuanian).
6. Project of funds distribution for the study of a logistics centre in Vilnius region. Construction firm "Vilnius plan", 2003 (In Lithuanian).
7. Project of financing a logistics centre in Vilnius region. Construction firm "Vilnius plan", 2003 (in Lithuanian).
8. Bowersocks, D. D., Closs, D. J. Logistics. Integrated chain of shipments. "Olimp Business", Moscow, 2001, 640 p. (In Russian).
9. Anikin B. A. Logistics. The third edition. Moscow, INFRA-M, 2002, 347-356 p. (In Russian).
10. Armacost, A. P.; Barnhart, C., Ware, K. A. Composite Variable Formulations for Express Shipment Service Network Design. *Transportation Science,* Vol 36, No 1, 2002, 1–20 p..
11. Powell, W. B.; Carvalho, T. A. Dynamic Control of Logistics Queuing Networks for Large-Scale Fleet Management. *Transportation Science,* Vol 32, No 2, 1998, 90–109 p..
12. Nguyen, S.; Pallottino, S.; Gendreau, M. Implicit Enumeration of Hyperpaths in a Logic Model for Transit Networks. *Transportation Science,* Vol 32, No 1, 1998, 54–64 p.

PRODUCTS DISTRIBUTION OPTIMIZATION IN THE LOGISTICS NETWORK

Ramūnas Palšaitis, Kęstutis Vislavičius

Vilnius Gediminas technical university,
Plytines 27, Vilnius, LT-2016,
E-mail: trvadyba@ti.vtu.lt

Abstract

The main feature of all projects of products distribution in the logistics network is an abundance of parameters, conditions and limitations. So, the products distribution optimization is a multicriterion problem, and it is practically impossible to formulate it like a mathematical programming problem only.

An algorithm based on a dialogue between a planner and a personal computer using mathematical programming as well as engineering animation is recommended for optimization products distribution in the logistics network. The algorithm is illustrated by the project of sand and salt compound spreading during the winter time period in the road network On the basis of the proposed algorithm the programme for the personal computer is created.

Key words: network, multicriterion problem, mathematical programming, shortest way

1. Introduction

A rapid development of the calculation technology brought revolutionary changes in different fields of transport and logistics. This made it possible to formulate and solve problems which seemed to be unsolvable a few years ago. The algorithms based on mathematical programming became especially widely used. Unfortunately, many practical problems are mainly of a multicriterion character. As a rule it is impossible to derive the objective functions for such problems or they are purely academic. Multicriterion problems can successfully be solved when algorithms allowing to include the planner into the designing process are used. The dialogue between the planner and the computer often gives good results, especially, when the control of the designing process is based on a comprehensive information provided by the computer. Now using graphic possibilities of personal computer it is possible to create programs which enable the planner to see at sight the whole "battle-field" no matter whether it were a technological process, transport or logistics network. An algorithm based on a dialogue between a planner and a personal computer using mathematical programming as well as engineering animation is recommended for optimization products distribution in the logistics network.

2. Background

Products distribution in the logistics network projects have many common features no matter whether it were sand and salt compound spreading, snow clearing or any other process ([1], [2]). Therefore, the analysis of the algorithm proposed will be based on one of them, namely, on a project for sand and salt compound spreading, the algorithm of which has been best investigated (on the basis of this algorithm program for a personal computer has been worked out). The main feature of all road maintenance projects is an abundance of parameters, conditions and limitations. For example, in preparing a project for sand and salt compound spreading the following parameters and conditions must be estimated:

a) the variety of roads (according to their significance they may be local, state and international; according to pavement they may be asphalt and non asphalt; according to the conditions of strewing they may be strewed by one drive, strewed by two drives and non strewed);

b) the parameters of compound spreaders (their speed, when moving over non-strewed roads and that when moving over strewed roads and the speed during spreading; the load consisting of sand or a salt-sand compound; the strewing way of one loading);

c) the possibilities of sand and salt compound storehouses (the type of material to be loaded, the loading time);

d) the technological standards (asphalt roads can only be strewed with a sand and salt compound while only sand is suitable for non-asphalt roads).

e) time limitations must be satisfied: international roads must be ready for traffic before 6 o'clock, state roads should be prepared before 9 o'clock in the morning.

Thus, it is practically impossible to formulate a clear mathematical programming problem by estimating so many parameters, conditions and limitations.

3. Mathematical Analogue

Suppose that a road network consists of *m* nodes. The nodes represent sand and salt compound storehouses, the crossroads, towns and settlements as well as the points where asphalt link of the road begins (or end) or where sand and salt compound spreaders finish spreading their portions of mineral materials, i.e. the points, where the parameters of roads or mineral material spreaders are subject to intensive changes. By links we mean the distances joining two directly interconnected nodes. For links and their parameters two indexes will be used: the first index will mark a node which is the beginning of the link, the second index will be used for the node where a link ends. Thus, the indexes will denote not only the location of the link on a road but its direction too. For example, the marked symbol $t_{a,\,b}$ will mean the time required for going from the node *a* to the node *b*, and the marked symbol $t_{b,\,a}$ will be the time needed for going from the node *b* to the node *a*.

It is necessary to prepare the actions plan, which will satisfy all conditions and limitations and minimize the cost of road maintenance. As a rule, the cost depends from several criteria, which can not interconnected mathematically. For example, a project for sand and salt compound spreading includes three criteria: the amount of fuel used, the amount of spreaded mineral materials, the number and location of equipped sand and salt compound storehouses. In addition, it is necessary to take into consideration that some problem limitations are not strict or they are not easily included into a mathematically solvable problem. Thus, it should be noted again that the solution of the problem is practically impossible, when using only mathematical programming.

The first difficulty lies in the introduction of the planner's intention to move from one node to another into a mathematically solvable problem, i.e. to obtain the route (the sequence of links) joining these nodes. To solve this problem the following mathematical analogue is proposed:

$$\sum_i \sum_j \lambda_{i,j} \cdot r_{i,j} \to \min, \tag{1}$$

$$\sum_{k=1}^{p_i} r_{k,\,i} - \sum_{l=1}^{s_i} r_{i,\,l} = \begin{cases} -1 \text{ - } \textit{for an initial node} \\ \ \ 0 \text{ - } \textit{for an intermedial node} \\ \ \ 1 \text{ - } \textit{for a final node} \end{cases}, \tag{2}$$

$$i = 1, 2, 3, \ldots, m,$$

$$r_{i,j} = [0, 1]. \tag{3}$$

here: $\lambda_{i,j}$ – is the optimality criterion of the link *i-j*;

$r_{i,j}$ – the logical variable which can have only two meanings: if $r_{i,j} = 0$ it means that the link *i-j* is not included in the route, if $r_{i,j} = 1$, it means that the link *i-j* is included into the route;

p_i – the number of the nodes from which the links run to the node *i*;

s_i – the number of the nodes to which the links run from the node *i*.

The objective function (1) defines the sum of the optimality criteria of all links included in the route. It is possible to obtain different optimization problems by changing the optimality criteria. For example, if $\lambda_{i,j}$ is the length of the link *i-j*, problem (1)–(3) becomes a problem of searching for the short-cut ([3]).

Condition (2) determines three limitations: 1) only one link may run from the initial node; 2) only one link may run to an intermediate node and from it; 3) only one link may run to the final node. Thus, condition (2) can ensure the continuity of a route.

Condition (3) indicates what values can be obtained by the variables of the mathematical analogue. In addition, condition (2) shows, that they cannot be different. Thus, condition (3) can be replaced by condition $r_{i,j} \geq 0$, and in this way a problem of the Boolean programming may be transformed into a simple linear programming problem.

The solution of mathematical analogue (1)–(3) is a continuous route, i.e. a sequence of variables which are not equal to zero. Thus, mathematical analogue (1)–(3) is a fine tool for a dialogue between a planner and the computer. Besides, it can be successfully used for the optimization of road maintenance projects. Suppose a mineral materials spreader is in the storehouse and it must to go to the links of international roads (because they are to be prepared for the traffic first). In this case the final node of the route should be chosen approximately and the reduced values of link lengths should be related to the optimality criteria, for example, the lengths of the links can be reduced twenty times for international roads and ten times for state roads. Under these conditions a spreader moving from the storehouse to the final node will always pass through the links of international roads. Or suppose a spreader has been unloaded, and it must reach the nearest storehouse. Now the actual link lengths must be related to the optimality criteria. By solving problem (1)–(3) for each storehouse the shortest way to the storehouse will be found.

Mathematical programming is the core of the algorithm proposed. However, it gives only particular optimum routes, while the road maintenance project is the whole complex of such individual routes. For achieving good final result, i.e. for the optimization of the whole complex of individual routes engineering animation is recommended. It is a good means of showing the processes going on within an object. The planner obtaining both numerical and graphic information can make a quicker and better estimation of the situation and that leads to an optimum decision-making.

Thus for preparing an optimum road maintenance project an experienced planner, powerful personal computer and a high quality computer programme are needed. In this case experienced planner is not a planner possessing good skills for working with a personal computer but it is a planner with a good knowledge of the actual road network and local conditions and with an experience in creating analogous projects by conventional methods. A powerful computer is required because an actual object usually includes about 200 nodes and 600 links, and a matrix with large dimensions can be rapidly treated only by a computer with a high frequency and a great random-access memory. The third element, i.e. a programme for a personal computer, is the main part of the complex. As a rule, a programme of this kind should contain such indispensable elements as a menu, an editor, a system of dialogue's windows, a unit for the realization of the engineering animation programmes. It should be noted that since the algorithm is rather complicated it is not easy to combine all these elements into one programme. The authors hope that their created programme partly satisfies the requirements indispensable for this type of programmes.

4. Conclusions

1. Products distribution in the logistics network is a multicriterion problem.

2. The main difficulty of products distribution problem lies in the introduction of the planner's intention to move from one node to another into a mathematically solvable problem.

3. Mathematical programming is a fine tool for a dialogue between a planner and the computer.

4. Engineering animation is a good means for achieving good final result; i.e. for the optimization of the whole complex of individual routes.

5. For preparing an optimum products distribution in the logistics network project an experienced planner, powerful personal computer and a high quality computer programme are needed.

References

1. Baublys, *Introduction the theory of transport system* (Vilnius: Technika, 1997) (in Lithuanian).
2. T. L. Magnadi, R. T. Wong, Network design in transportation planning models and algorithms, *Transportation Science*, No 1, 1984, 3−55.
3. G. Dievulis, Method of finding shortest path in the transport network, *Transportas*, Vol XV, No 1, 2000, 3−11 (in Lithuanian).

EVALUATION OF THE DME SITING FACTORS WITH APPLICATION TO PALANGA AIRFIELD

Algimantas Jakucionis, Romualdas Malinauskas, Jonas Stankunas

Antanas Gustaitis Aviation Institute of Vilnius Gediminas Technical University
Rodunes kelias 30, 2038 LT-Vilnius, Lithuania
Ph: (+370)-2744809, Fax: (+370)-2745058, E-mail: avinst@ai.vtu.lt

ILS still remains in the use on the background of development of satellite navigation systems. Necessity of usage of the DME in lieu of ILS marker beacons arises especially for the ILS of coastal airfields. For this reason the main recommended DME siting criteria are discussed in this work. The model is presented for the calculation of distance measurement errors because of DME signal reflections. Some calculations are done uder the influence of worst case reflections from the objects of the Palanga airfield. The results of height measurements of Palanga airfield obstacles for DME signal are presented and analyzed with respect to some siting points of DME. Preferable DME siting point in airfield is evaluated.

Key words: DME siting, reflections model, Palanga airfield

Introduction

Instrument landing systems(ILS) still remains in the use on the background of development of satellite navigation systems. Necessity of usage of the distance measuring equipment(DME) in lieu of ILS marker beacons arises especially for the ILS of coastal airfields because of environmental requirements for recreation zones existing near the sea. As a rule coastal airfields have narrow flat zone and obstacles for the DME signal displaced in proximity to airfield boundaries. For this reason need of evaluation of DME signal reflections arises seriously.

Evaluation of Reflections

As approximation of possible DME error due to DME signal reflections from obstacles it was assumed that replay pulse is triangular type (fig. 1). Also it was supposed that a reflected signal is also a triangular type with the amplitude A_K and it is in antiphase to the signal from DME which amplitude is A_o. With reference to the fig.1 one may find that error in measured time due to the influence of reflections equal to the interval Δτ measured at the 0.5*A level.

Transformation of coordinates (U;t) into coordinates (U', t') with initial point ($U = \frac{A}{\tau} \cdot t_{vel}$, t = $t_{vel.}$) leads to the following expressions for the values of U of undestorsed and distorsed pulse corresponding the 0.5 A level of undestorsed pulse

$$U_0(t') = \frac{A - \frac{A}{\tau} \cdot t_{vel}}{\tau - t_{vel}} \cdot t', \tag{1}$$

$$U_{0K}(t') = \frac{A - \frac{A}{\tau} \cdot t_{vel} - A_K}{\tau - t_{vel}} \cdot t', \tag{2}$$

$\Delta\tau = t_2' - t_1'$

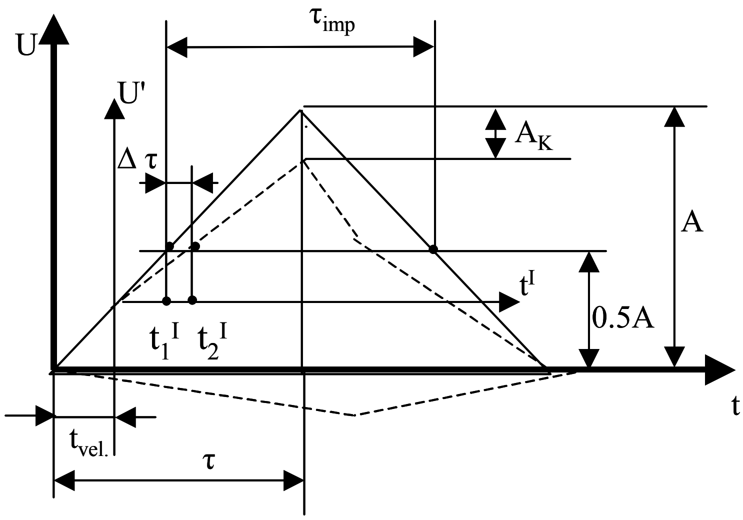


***Figure 1**. Approximation of the DME pulse by saw tooth type function*

$$t^I = t_1^/, \text{ when } U_0(t^/) = \frac{A}{2} - \frac{A}{\tau} \cdot t_{vel}, \tag{3}$$

$$t^I = t^I_2, \text{ when } U_{oK}(t^I) = \frac{A}{2} - \frac{A}{\tau} \cdot t_{vel}, \tag{4}$$

From equations (1) and (3):

$$\frac{A}{2} - \frac{A}{\tau} \cdot t_{vel} = \frac{A - \frac{A}{\tau} \cdot t_{vel}}{\tau - t_{vel}} \cdot t_1^/, \tag{5}$$

After multiplication of both sides of equation (5) by $\frac{\tau}{A}$ we have:

$$t_1^/ = \frac{\tau}{2} - t_{vel}, \tag{6}$$

From equations (2) and (4):

$$\frac{A}{2} - \frac{A}{\tau} \cdot t_{vel} = \frac{A - \frac{A}{\tau} \cdot t_{vel} - q \cdot A}{\tau - t_{vel}} \cdot t_2^I, \tag{7}$$

$$q \cdot A = A_K.$$

After the same procedure (multiplication by $\frac{\tau}{A}$) with equation (7) we get:

$$t_2^/ = (\frac{\tau}{2} - t_{vel}) \cdot \left[\frac{1}{1 - q \cdot \frac{1}{1 - (t_{vel}/\tau)}} \right]. \tag{8}$$

Finally:

$$\Delta\tau = t_2^I - t_I^I = (\frac{\tau}{2} - t_{vel})\left[\frac{1}{1 - q\cdot\frac{1}{1-(t_{vel}/\tau)}} - 1\right]. \tag{9}$$

Equation for $\Delta\tau$ is valid for $t_{vel} \le \frac{\tau}{2}$. When $t_{vel} \rangle \frac{\tau}{2}$, $\Delta\tau$=0.

According to expression (9) for Δτ the evaluations of possible distance errors for the range of values of delay time reflected signal amplitudes were carried out (tab.1).

Table 1

t_{vel}, μs	$\Delta\tau$,μs				Δr,m			
	q				q			
	0,1	0,05	0,01	0,005	0,1	0,05	0,01	0,005
0,9	0,133	0,061	0,0116	0,0058	40	18,4	3,5	1,7
1,1	0,111	0,051	0,0096	0,0048	33,3	15,3	2,9	1,4
1,5	0,053	0,024	0,0045	0,0022	16	7,2	1,3	0,7

For calculations it was assumed that duration of DME pulse is 3.5 microseconds at the 0.5*A level.

Requirements for DME Siting

For the purpose of evaluation of possible DME errors, we must discuss what requirements are for the site with tolerable obstructions for DME signal.

Although no absolute minimum practical requirements can be stated restrictions with slight variations are approximately the following [1; 2].

The terrain should be level within a radius of 60.0 meters. Between 60.0 meters and 300 meters, a downwards slope is acceptable if the rate of descent is not more than 1.2 meters in 30 meters and contour lines are generally circular around the site. Beyond 300 meters the terrain should be below the horizontal plane of the antenna.

There should be no obstructions in the horizontal radiation patterns of the antenna within 230 meters of antenna. There should be no metallic structures in the radiation patterns of the antenna inside an area from the ground to an angle 1.2^o above a horizontal plane at the base of the antenna. There should be no wooden structures with negligible metal content in the radiation pattern of the antenna inside an area from the ground at an angle 2.5^o above a horizontal plane at the base of the antenna. The antenna should be situated so that large structures like aircraft hangers or other buildings would present the narrowest aspect to a radial from the antenna. Single trees less than 10.5 meters high may be tolerated beyond 230 meters. There should be no groups of trees or groves within 300 meters and no overhead power or control lines within 230 meters of the antenna.

Beyond the 600 m large continuous metallic objects such as overhead power lines, masts, water towers or large metal – clad buildings which will penetrate beyond 0.5^o above the horizontal plane as measured from the array center or which subtend an angle of greater than 1.2^o are to be analyzed for potential interference prior to being approved.

In same localities it is not possible to meet all these site requirements.

In accordance with these requirements some Palanga airfield points with their surroundings were analyzed as possible points for DME antenna siting.

Analyzis of potential points to DME siting

Palanga airfield local coordinate systems with some important points in the airfield area are shown in Fig.2. These important points are: AKT – the aerodrome reference point; (KTC) – runway geometric center, (P), (P1), (P2) – points, which were analyzed as candidates for points positioning DME antenna.

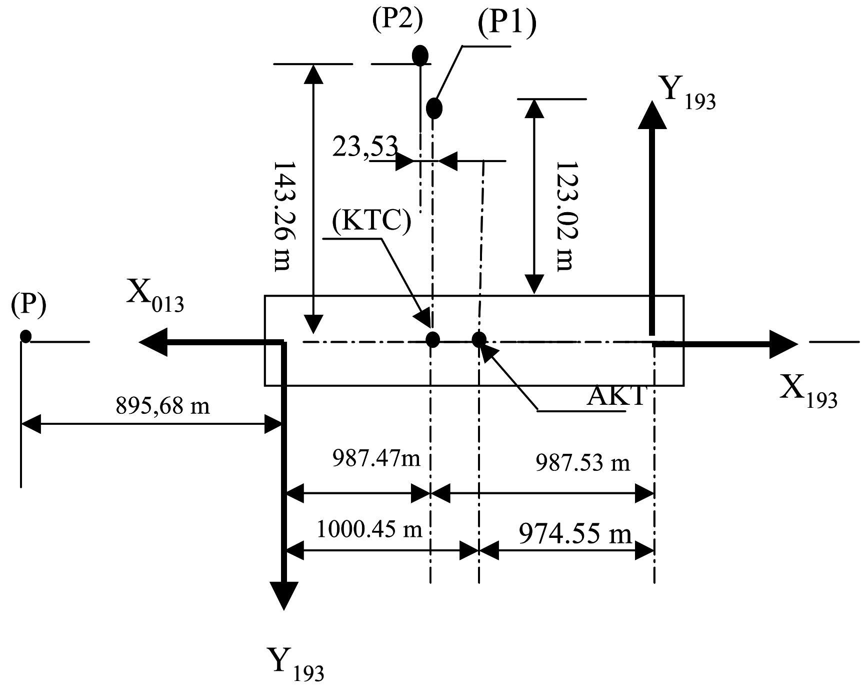


Figure 2. *Airfield local coordinate systems and points*

As the undertaken aerodrome obstruction survey show point (P) with its surrounding area would be quite appropriate to the terrain features and obstruction limitation requirements.

But two opposite directions of approach and take–off are used in Palanga airfield. For this reason points (P1) and (P2) were analyzed. The position of point (P1) was chosen in accordance with such things. Its position is of perpendicular to the center line of runway at the (KTC) point. This perpendicular is to opposite site from the apron and main airfield buildings. Y coordinate of point (P_1) was chosen to meet national requirements of the aerodrome design [10] in the part of minimum valid distances from the runway centerline. On the other hand the distance of (P1) from the nearest obstacles (groups of trees) must has to be as long as possible.

Fig.3 shows the results of the survey and calculations for angles witch create most critical obstacles above horizontal plane at the base of antenna with respect to antenna elevation.

These relationships show that for obstacles with relatively big heights but which are at some hundred meters or longer distances from the antenna, critical angles decreases very slowly with the elevation of the antenna. Unfortunately, these angles stay of undesirable values if the antenna base height is even 18 meters.

In fig. 3 is shown the second horizontal axis with the values of absolute altitudes. The altitude of point (P) on the aerodrome surface is 8.4 meters and the altitude of roof of the radio center building situated near the point (P) is 14 meters. This value is shown by a dashed line in fig.3.

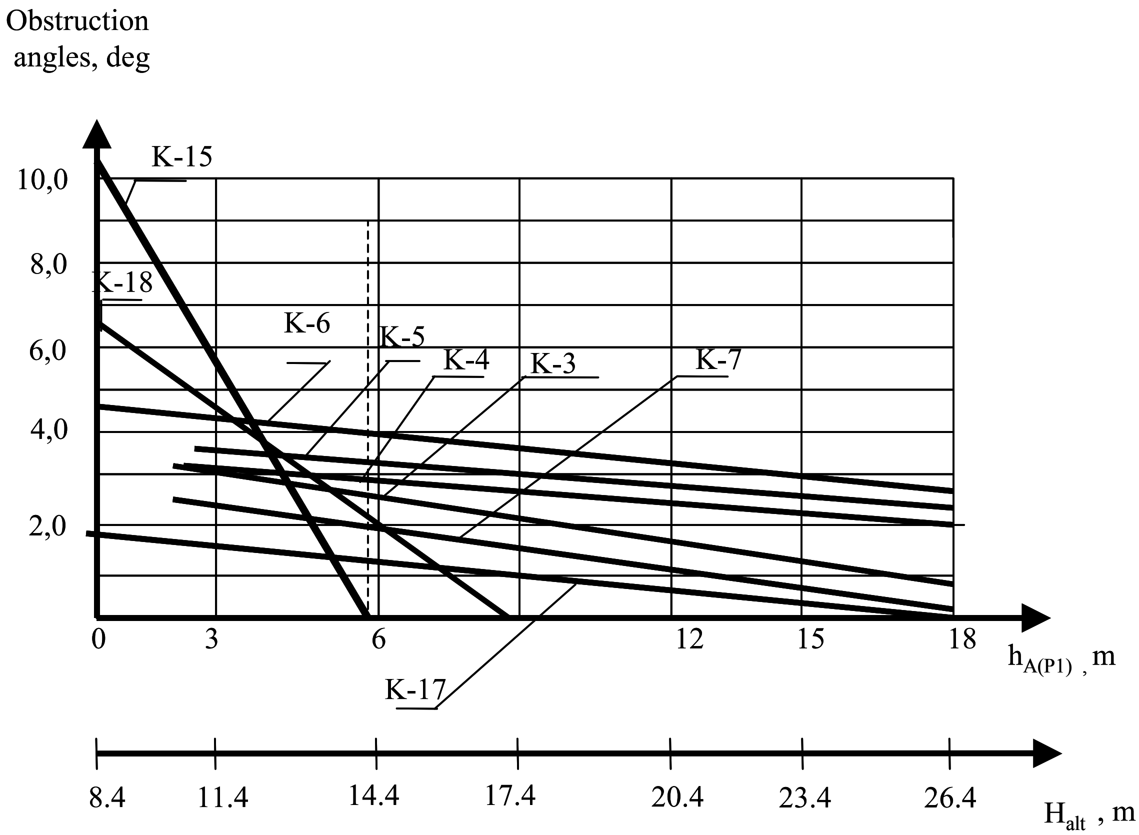


Figure 3. *Relationship of angles of most critical obstacles with respect to antenna elevation's $h_{A(P1)}$ height of the base of antenna at the point (P_1) attitude respectively*

From these relationships one can conclude:

a) for relatively high obstructions such as lights supporting constructions (K-3, K-4, K-5, K-7) or radio communication antenna (K-6) obstruction angles decrease very slowly with elevation of DME antenna. Unfortunately, these angles stay of undesirable values if antenna base height is even 18 meters. So the problem of reflections remains and will be discussed later.

b) In the closest to DME antenna zone in which obstructions are not tolerable e.g., such as building of the radio center (K-15) at the distance of 30.5 meters from point (P1) or contour of the nearest grove (K-18) at distance of 76 meters obstruction angles decrease at a big rate and at the heights more than 9 meters of DME antenna base the problem of these obstacles does not exist.

The situation of obstacles for point (P2) also was analyzed (see fig. 2). This point as a possible position for DME antenna was selected taking into account the following: 1) it is closed near a point (P1), concequently, it is quite close to the perpendicular from the center of runway; 2) the difference in altitudes of points (P2) and (P1) is 5,6 meters as benefit to point (P2); 3) this point is meets a requirement for the displacement of DME antenna i.e. requirement to the angle between the line of precision approach and directions to the point where distance information is needed for landing when DME is paired with ILS [3; 5] .

Evaluation of this angle maximum value which must not be higher than 20 degrees is illustrated in fig. 4.

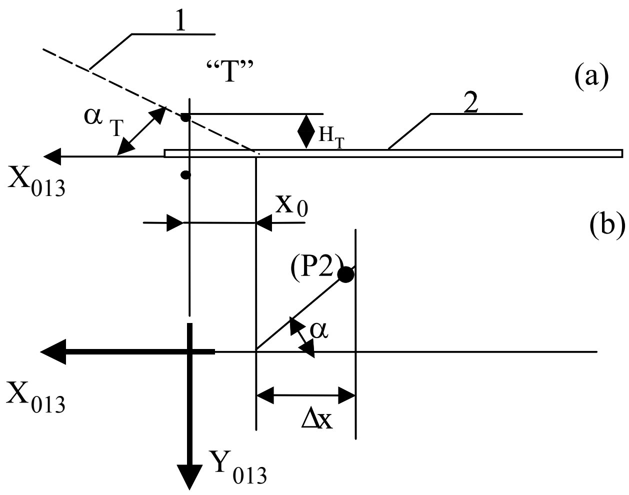


Figure 4. *Illustration for calculation of maximum value of angle α.*
(a) – vertical plane; (b) – horizontal plane; 1 – nominal glideslope; α_T – nominal value of glideslope angle; 2- runway; H_T – nominal value of height of point "T" according to ICAO

From fig. 4 we have:

$$X_o = H_T / tg\alpha_T;\ tg\alpha = \left|Y_{P_2}\right| \big/ (\Delta x).$$

After substitution of values for Palanga airfield: $H_T = 15$ m; $Y_{P_2} = 143{,}26$ m; $(\Delta X) = (\left| X_{(P_2)} \right| - \left| X_0 \right|)$, one can calculate that $\alpha = \alpha_{\max} \approx 12^o < 20^o$.

4) In point (P2) the requirements that higher point of DME antenna would not exceed the transitional surface are met as well (fig. 5) [3; 4].

Above point (P2) heights of points will be:

a) limited by inner transitional surface:

$$h_{A1} = [Y_{(P2)} - (D/2)]\ tg\alpha_{VP;}$$

b) limited by transitional surface:

$$h_{A2} = [Y_{(P2)} - (D/2)]\ tg\alpha_{PP.}$$

Evaluating for the worst case situation (D=155 m, $\alpha_{VP} = 33{,}3^0$, $\alpha_{PP} = 14{,}3^0$ one can find that the altitude of the upper point of DME antenna may be bigger than the altitude of aerodrome reference point by $h_{A1} \leq 43{,}2$m; $h_{A2} \leq 16{,}75$m.

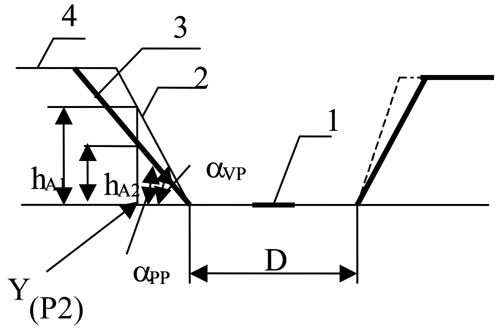


Figure 5. *Limitations of elevation of DME antenna.*
1 – runway; 2 – inner transitional surface; 3 – transitional surface; 4 – horizontal surface; D – width of inner approach surface

Additional advantage of point P2 would be absence of need for additional cabin construction for DME antenna.

Obstruction degrees as a function of DME antenna elevation for some more critical obstructions in point (P2) area are shown in fig.6.

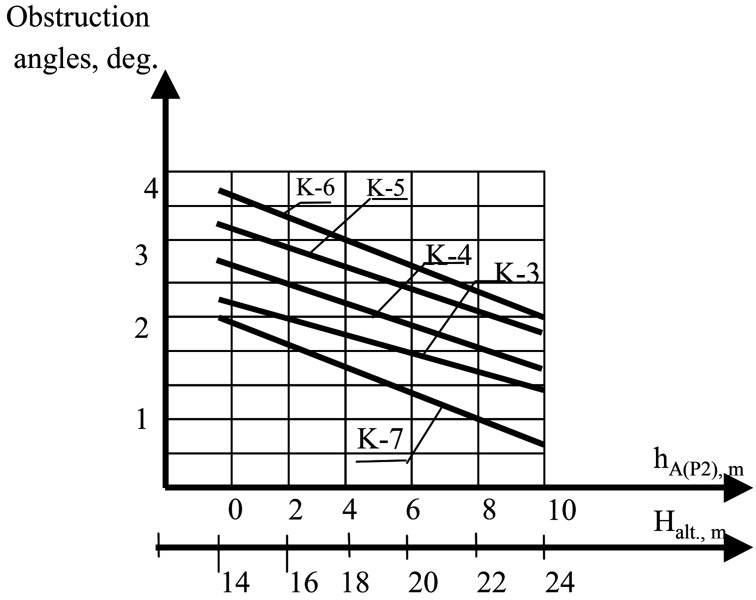


Figure 6. *Obstacles angles relationship to DME antenna base height in respect of point (P2) height and in respect of antenna base altitudes.*

One may see from fig. 6 that some obstructions remain problems analogous to the point (P1) problems. For obstacles K-3….., K-7 their angles remain more than required maximum 1.2^o despite big antenna elevations

Signal's Multipath Evaluation for Obstacles K-3,…., K-7

These obstacles there are constructions of lights supporting and communication facilities antenna. They are at distances 380…670 meters and by azimuth angles of 156^o 98^o with respect of the position at point (P2). Distance measurement errors due to the multipath arise at the interrogator. For the first approximation of the worst case evaluation of relative intensity of reflections we can do some assumptions. The width of radiating pattern of DME antenna in vertical plane is of φ_{AV} (Fig.7) and angle of radiation maximum is φ_A. As approximation we assume that antenna's radiated power, which shares in the cylinder with the center at point (P2). Radius S_K of this cylinder is equal for the distance to obstacles.

We may assume reflections from obstacles as secondary radiation sources radiating power

$$P_K = \frac{P_{Asp} \cdot Q_K}{2\pi S_K \cdot h_{SK}} .$$

Taking into account that obstacles K-3……K-6 are metallic sticks type structures we assume that reradiated power will be scattered in the cylinder with height h_{SK} and angle width ± 45 degrees. In this case at a receiver position powers ratio of signal reflection with the signal will be:

$$\frac{G_K}{G_A} = \frac{(P_K)/(0{,}25 \cdot 2\pi \cdot S_{SK})}{(P_{ASP})/(2\pi \cdot S_{XA} \cdot h_{XA})} = \frac{2}{\pi} \cdot \frac{d_K}{S_K} \cdot \frac{S_{XA}}{S_{XK}} ,$$

here d_K – effective width of an obstacle, S_{XA} and S_{XK} -distances of DME interrogator antenna from DME station antenna and from reflecting the obstacle respectively.

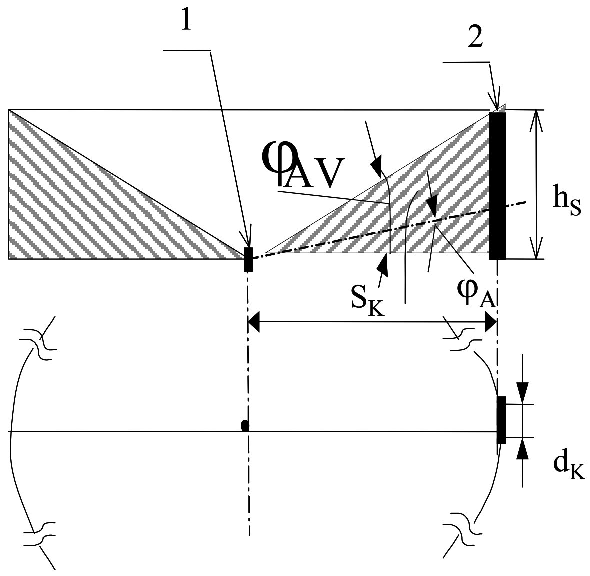


Figure 7. Approximated radiation pattern of DME antenna for evaluation of relative intensity of reflections; 1- DME antenna; 2-reflective obstacle; h_{SK} - width of radiation pattern of an obstacle at distance S_K from antenna; d_K -width of an obstacle; φ_{AV}-width in degrees of radiation pattern in vertical plane; φ_A – angle of maximum radiation

Approach zone and zone for missed approach for Palanga airfield are on the opposite side of the runway than obstacles K-3…K-7 are. For this reason we may assume $S_{XA} \approx S_{XK}$. Then at an interrogator antenna relation of reflection and signal powers $A_K = 10 \lg (G_K/G_A)$ for assumed reflection coefficient equal to one, $d_K \approx 3.5$ meters and $S_K \approx 500$ meters gives value of $A_K \approx -20$ dB. This corresponds to value of quantity $q \approx 0.01$ used in table 1.

Phase and Delay Time of Reflections

For the first stage evaluation of the influence of reflections to signal one may assume the worst case situation at an interrogator antenna point. It will be when the phase of radio oscillations of reflections will differ by 180^o with respect of the o signal radio oscillations phase. The worst case situation for reflections delay time would be at zones where reflections delay time would be close to the value of code interval T_K between DME pulses in pair. For delay time evaluations we may refere to fig. 8.

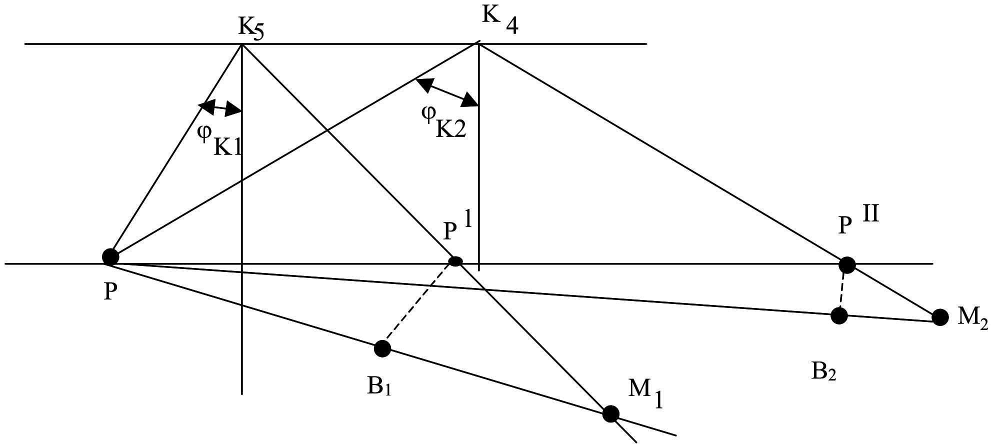


Figure 8. Multipath consideration for evaluation of delay of reflected signals; P is a possible position of DME antenna (point (P2)); K-4,K-5 – obstacles

As may be seen from fig.8 reflected signal path lengths to points M_1, M_2 in the receiving area will be ($PK_5 + K_5M_1$) and ($PK_4 + K_4M_2$) and for signal it will be lengths PM_1 and PM_2. As one can see $PP' \approx B_1P$, $P' M_1 \approx B_1M_1$, $P P'' \approx P B_2$ and $P'' M_2 \approx B_2 M_2$. Delay of reflected signals and the signal will define differences in paths $\Delta\tau_1 \approx (PK_5 + K_5 P') - P P'$ and $\Delta\tau_2 \approx (PK_4 + K_4 P'') - P P''$. As follow from fig. 8 and with allowance for the fact that incidence angle is equal to reflection angle of DME frequencies waves for these differences we may write

$$\Delta r_i \approx 2 S_{Kli} (1-\sin\varphi_{Ki}),$$

S_{Kli} – distance of i-th obstacle from point P; φ_{Ki} – angle defining directions of obstacles. Taking into account that displacement of investigated obstacles are approximately in parallel with the runways axis we may write for φ_{Ki}:

$$\varphi_{Ki} = \varphi_{Kazi} - (\varphi_{azx}+90),$$

where φ_{Kazi} – values of true azimuth of obstacles; φ_{azx} – true azimuth of runway axis. Values of S_{Kli}, φ_{Kazi}, φ_{azx}, were defined from the survey and initial calculations after the survey results. With this in mind probable values of signal reflections delays for obstacles K-3, K-4, K-6, K-7 with respect of the signal in approach direction are presented in table 2.

Table 2

Obstacle	K-7	K-6	K-5	K-4	K-3
S_{Kli}, m	383	412,5	470	556,5	673,5
$\varphi_{Ki,}$degrees.	5	0	24	41	53
Δr_i, m	699	825	557,7	348	271
$\Delta t_{vèli,}$ µs	2,33	2,75	1,86	1,16	0,9

One can see from this table that when φ_{Ki} increase delay time of reflected signals decrease and it never overcomes the value 2 S_{Kli} / v, where v – velocity of DME signal. The conclusion follows that $\Delta t_{veli} << T_K$ and reflections may give distortions for rise time of the first DME pulse in pairs not for the second one. For the analyzed airfield the worst conditions in respect of the distance errors may arise by reflections from obstacles K-3 and K-4 at distances of ILS middle marker area. Foreseen values of DME measured distance errors due to the influence of analyzed obstacles may be about 10 meters.

Conclusions

1. Simple models are proposed for evaluation of delay time and relative intensity of DME signal reflections from obstructions with a model incorporating these quantities for evaluation of DME measured distance error.

2. By survey, calculations and modeling of influence of obstructions to DME siting in the Palanga airfield analysis is carried out for three possible points of DME antenna positioning.

3. For intolerable obstructions possibilities to reduce their influence to DME signal by appropriate siting and elevation of DME antenna are evaluated.

4. Methods and models of this work for evaluation DME errors may be generalized and used for solving DME siting problems in other airfields.

References

1. Instructional Manual : Distance Measuring Equipment, Model 5690 – WILCOX, 1993, p: 6-1.
2. http://www.tc.gc.ca/CivilAviation/Aerodrome/Environment/TP1247E/Part2/Part2-4.htm
3. (section 2.4-Navigational Aids)
4. Lietuvos Respublikos civilinių aerodromų projektavimo, statybos ir naudojimo specialieji reikalavimai (Patvirtinti LR susisiekimo ministro ir aplinkos ministro 2000 02 23 įsakymu Nr. 42/69).
5. (Special requirements of Lithuania Civil aerodromes design, building and exploatation, signed by Lithuania Republic Minister of Transportation and Minister of Environmental affair by directive N42 69 in 02/23/2000).
6. Aerodromes. Annex 14 to the Convention of ICAO, vol. 1: Aerodrome design and Operations, 2nd Edition; July 1995,p.6 and p.28…30 (Chapter 2. Aerodrome data).
7. Aeronautical Communications. Annex 10 to the Convention of ICAO, Vol.1, 1996, 5 th Edition, section 2.1, 7.1.6.1…7.1.6.3.

MEASUREMENT AND MONITORING MEAN SELECTION TASK IN THE AUTOMATED CENTRALIZED TRANSPORT CONTROL SYSTEMS

Alexander Berezhnoy

Transport and Telecommunication Institute
1 Lomonosova st., Riga, LV-1019, Latvia
Ph.: (+371)-7100509, Fax: (+371)-7100510, E-mail: avb@junik.lv

Key words: traffic monitoring and control systems, transport detection and measurements

1. General Information

Any city, due to the permanent increase of transport means quantity, sooner or later faces the problem of road jams, lack of road network bandwidth, problem of redistribution of traffic flows control etc. In order to cope well with the abovementioned tasks it is needed to implement the entire set of actions, mainly aimed on the increase of moving mobile objects controllability level. Alongside, in order to determine the required correct managing influence it is necessary to possess enough of the authentic initial data on existing dynamics for movings of transport streams and knowledge, as well as clear understanding of trends for situation development. One of the most significant tasks for this context that needs to be realized at the city-scale is considered to be the unified urban traffic flow measurement and control procedure.

For this purpose, in the frames of the city transport infrastructure, stationary measuring networks are constructed, based on the systems of transport control and video observation, as an important structural element of the developed network. This kind of information obtaining complexes and monitoring systems are often integrated as a part of road traffic centralized control systems. The most common example of such sort of systems is UTMC (Urban Transport Management and Control) system.

2. Traffic Stream Characteristics

As the measuring information in the transport theory of traffic streams it is traditionally accepted to allocate the following parameters describing a stream of vehicles and allowing to estimate its change in time: rates of flow (vehicles per unit time), speeds (distance per unit time), travel time over a known length of road, occupancy (percent of time a point of the road is occupied by vehicles), density (vehicles per unit distance), time headway between vehicles (time per vehicle), spacing (distance per vehicle) and concentration, measured by density or occupancy [9].

Gathering the corresponding data needs use of essentially various methods and means for measurements as well as for its further processing, from the point of the listed characteristics practical measurements. Besides, end-point gauges and the devices maintained for measurements of the given values will strongly differ on a design, used functioning principles, a degree of complexity and final intellectual abilities.

On the other hand, practically measured characteristics sometimes differ from theoretically accepted indexes. The most common of them are: traffic count, average speed of traffic flow, density and intensity of the flow, distribution of flow among the lanes and parked car quantity.

Information completeness of gathering data subject is defined, mainly, by the reason of motives of data gathering at the investigated area, namely, - whether obtained data is intended for local needs, or centralized control systems of city traffic will be to be used, whether the control over infringements with the subsequent penalty policy is put as the aim, or the opportunity of vehicles detection is considered to be of a primary priority. That is the way things are sometimes realized in the systems of car parking either in the systems, used for signaling and the stolen vehicle tracing.

3. Measurement Zones

It is well known, that in various areas of the city, presence of pre-selected supervision and control points is strongly recommended or required. As a result, there is gathered the data of a various degree of informational saturation. So, essentially important area for getting the measuring data are considered to be the city boundaries, since reception of the information on mobility inside and outside of researched area enables to estimate approximately the current value of vehicles concentration inside the city. Besides, external perimeter, the point of heightened interest refers to the city centre, possessing the ultimate strategic value from the position of transport system resources usage.

It is convenient enough to take the benefit of using the existing accepted administrative zone division structure A-B-C of the city that shows relative remoteness and the cost of the corresponding ring-bound zone from city centre. The other way is to accept using the similar principle of zone division used for streets classification.

Table 1. Administrative street division

Classes of streets, roads and the areas	Arrangement of buildings and constructions
A	The backbone of city value, streets and squares: major lines, central streets, terminal areas, with a huge transport capacity, adjoining to bridges and multipurpose transport nodes.
B	The main streets of regional value, squares in front of large public buildings and constructions (stadiums, theatres, exhibitions, shopping centers, the collective-farm markets and other places of mass visiting)
C	Streets and roads of local value, smaller roads, countryside streets, the areas and squares before public buildings and constructions of countryside value.

On the same time, measurement and monitoring control areas won't completely fall over the same areas of importance comparing to the traditional zonal division for the city or street classification. It is explained by the fact of existence a number of points of interest in terms of transportation matter. These objects are as follows:

- Railway crossings;
- Linear sites of road with increased high-speed restrictions;
- Descents from bridges;
- Passing-tracks and multilevel junctions;
- Bottlenecks;
- Incident dangerous crossroads;
- Entrances to parking and parking zones;
- etc.

By analogies, it is useful to introduce the term of information zone division according to the criteria of the collected data significance level depending on the city area:

Table 2. Information zone classes

Information zone classes	Class priority and its description
A	Reference points located at the A-area according to the city and street classification; city boundaries, city key intersections and multilevel junctions, descents from bridges and dangerous areas.
B	Reference points located at the B-area according to the city and street classification; complex intersections, assumed and potential bottlenecks, sites with increased high-speed restrictions.
C	Reference points located at the C-area according to the city and street classification, railway crossings, selected non-prior measurement and vehicle-counting points, etc.
D	Control and monitoring points aimed on classification of vehicles in the stream and video observation of mobile objects. This zone is distributed, based on-demand principle all over the city.

The totally independent problem, which is not considered in given clause is definition of the minimal, sufficient and superfluous quantity of reference points of information collection in order to provide the permanent entrance dataflow of the city transportation dynamics to the centralized transport management and control center. The task of setting-up the control points, as a rule, is considered in common with a problem of topological distribution of measurements reference points regarding to cartography of the certain city.

4. Transport Measurements and Monitoring Systems

For the purpose of data gain on transport stream there have to be selected the appropriate means of transport measurements or measurement complexes. There exists a great variety of such systems nowadays. This is one of the obvious difficulties to understand the true information demand and make a decision on evaluation criteria while choosing the right measurement method and system. Main difference deals with type of sensor, which becomes the end-point receiver for measured characteristics.

The choice is to be made among the following most popular traffic detection systems: traffic loop detectors, video detection systems, microwave systems, infra-red sensor based systems, laser-based systems, radar systems, sonic detection systems, piezoelectric systems and finally, the magnetic detection systems. Even for the brief overview everyone could notice the existing difference regarding the type of data gain – at least with a rough separation of direct and indirect measurement classes. In order to make a first impression of this collection let's consider the short functional description for these systems with the advantage mark-up.

Table 3. End-point system sensor part possibility comparison

Name of the system	Performed tasks	Advantages	Disadvantages
1. Traffic loop	▪ Traffic volume calculation ▪ Vehicle classification ▪ Lane occupancy determination ▪ Vehicle speeds, average speeds calculation	▪ Weather immunity ▪ Low cost ▪ Out of reach of vandals	▪ Unreliable ▪ Require lane closures during installation
2. Video detection systems	▪ Traffic flow monitoring ▪ Vehicle identification ▪ Incident detection ▪ Vehicle tracking ▪ Traffic analysis ▪ Queue length measurement ▪ Speed enforcements ▪ Vehicle pass control	▪ Flexibility ▪ Image verification with original + Digital processing upgrade + Wireless image transmission + Low cost digital silhouette system	▪ Correct placement of camera ▪ Weather dependent ▪ Expensive

Name of the system	Performed tasks	Advantages	Disadvantages
3. Microwave systems	▪ Vehicle counting ▪ Motion detection at signal intersections ▪ Microwave-based systems for pedestrian acknowledgement	▪ Small ▪ Low power consumption ▪ Lightweight ▪ Easy to install ▪ Long range ▪ Cost effective	▪ Needed potential for interference (on frequencies)
4. Infra-red sensor based systems	▪ Vehicle classification ▪ Object determination with $T >$ abs.zero ▪ Motion detection ▪ VICS* vehicle navigation system	▪ Small ▪ Sensitive	▪ Unreliable within weather conditions (rain, fog, snowfall)
5. Laser-based systems	▪ Vehicle detection ▪ Vehicle classification ▪ Enforcement applications (speed, acceleration)	▪ Operation speed ▪ Precision ▪ Extremely high effectiveness ▪ Compatibility with Video detection systems	▪ Expensive
6. Radar systems	▪ Speed enforcement ▪ Vehicle classification ▪ Vehicle presence ▪ Traffic volume ▪ Lane occupancy	▪ Detection quality	
7. Sonic detection systems (ultrasound)	▪ Traffic detection ▪ Queue distance control	▪ Extremely reliable ▪ Require little maintenance ▪ Small ▪ Portable/Perm.	
8. Piezoelectric systems	▪ Vehicle counting ▪ Vehicle detection ▪ Vehicle classification ▪ Weight-inmotion	▪ High level of accuracy	
9. Magnetic detection systems	▪ Vehicle detection ▪ Vehicle determination ▪ Data on traffic flows		▪ Pioneer method

*VICS – Vehicle Information and Communication System

At the first look many of the described systems actually do the same tasks. Eventually it becomes the matter of comparison between the prefer ability of functional quality index group, in which terms the objective rating should be realized. The correct choice may be only made according to the existing aim function and limitations put in forward of the exact urban management and control system.

The general comparison is performed for the real existing systems, because it is much more complicated to find out any information about the precision and reliability characteristics of the measurement method by itself. That's the major reason for making the expert choice, based on the evaluations in the field of the possibilities for the top competitive automated traffic control systems, but not the clear theoretical background.

5. Selection of the Right Measurement or Monitoring System

During the stage of the data gathering system type selection it is useful to find out the exact position and the correlation between existing transport management and control center requirements and built-in technical functions of the given end-point system.

Practically, it is necessary to deal with such aspect, like comparison among characteristics of various systems by themselves, instead of the comparison of their separate components.

It would be theoretically possible to shatter a problem by means of applying a method of decomposition that allows dividing a problem into a number of more simple tasks, which can be solved separately. I.e. the question of comparison between gauges, the devices responsible of the primary data selection, devices of the data processing, send-receive blocks, modules of the analysis and decision-making modules, the specialized software is discussed, instead of proceeding systems as a whole. Modern intellectual systems, irrespective of used detectors type most commonly go through the construction process in the combined kind and often are of polytypic origin. Nevertheless they are still allowing to carry out similar measuring functions. Though, the more correct approach is to perform the final comparison of the entire systems.

As it has already been marked in section of information zone division, for various sites of a city transport network the set of characteristics priorities from the point of measuring system will strongly differ. For example, for a non-controllable railway crossing, there can be quite enough installation of inductive loop as a sensor of an end-point system for the task of vehicles quantity calculation on the crossing. At the same time, when we are speaking about a complex crossroads, or a city entrance road can be expediently equipped with more functional microwave radars that are not only capable of the simple counting up of automobiles quantity by lanes, but also can collect data on intensity and density of streams, average speed of movement, etc. It is obvious, that some control points, such as way-outs from bridges, complex crossroads or multilevel junctions it is necessary to equip with the systems of supervision allowing not only to take the necessary data on a stream, but actually, also to take the benefit of using the video information. Later it may be applied for carrying out of any complex decisions or may be helpful for the detailed analysis of some road events and certain cases.

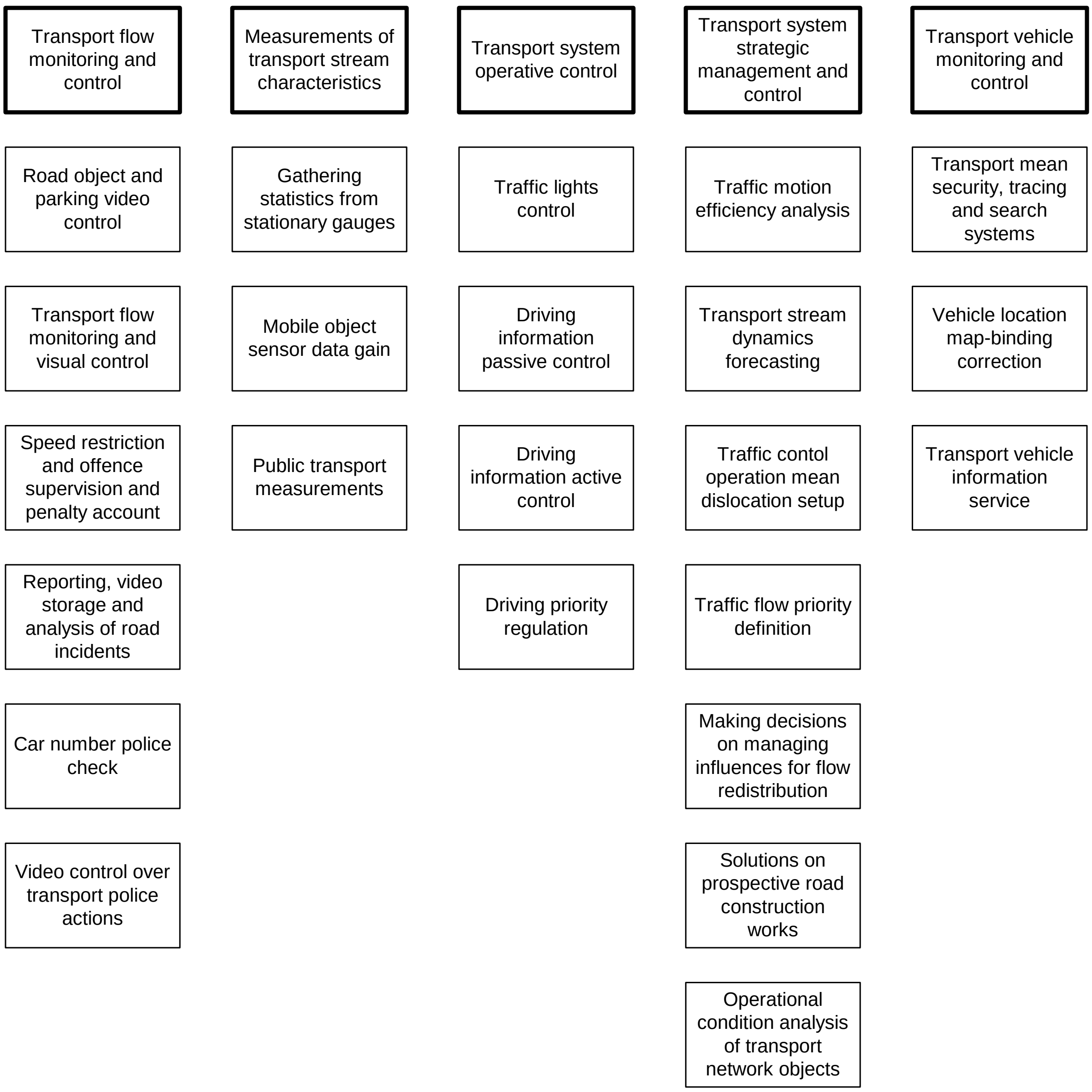


Figure 1. *Urban transport management and control center functions*

According to the scale of functional possibilities all listed systems could be conditionally divided into four respective groups: 1 – Traffic counting systems for the selected time period by lanes, 2 – Advanced traffic flow measurement systems, 3 – Combined measurement and monitoring complexes, performing classification tasks, 4 – Monitoring, video detection and identification systems.

Table 4. Measurement and monitoring systems functional groups

Measured characteristics	1	2	3	4
Transport count for the pre-defined time period, by lanes	X	X	X	
Traffic stream speeds		X	X	
Travel time		X	X	
Flow intensity		X	X	
Flow density		X	X	
Road surface occupancy, by lanes		X	X	
Time distance between vehicles	X	X	X	
Distance between vehicles			X	
Vehicle identification				X
Vehicle classification			X	X

So, the provided set of performed tasks is ranged from the pure counting-measurement task to the final stage of specific monitoring needs. The last category is potentially capable of performing any kind of pointed tasks, but actually it gets rarely used for these objectives, except the specific case of identification, classification and simple monitoring.

On the other hand, the choice of the required system inside the group is made, basing upon quality index groups, which could be common for the constructively different devices aimed on performing the similar or even the same tasks. Most general that could be mentioned here are: electric feed requirements (like tension stability, or autonomy), size and mass characteristics, resolution, the operative range, dependence on external influences, reliability of detection and data (non-failure operation characteristic, durability, maintainability), convenience of installation and adjustment, types of interfaces and a format of data presentation, additional opportunities (like the possibility of remote control).

The problem of a choice of the end-point measurements or supervision system is considerably complicated by a fact that the case of real functioning systems comparison is taken for the certain pre-selected manufacturers and doesn't mean the equivalency of possibilities for technologies. It means that absolute values of individual parameters of quality can differ substantially in the same systems for different manufacturers. It is also important to focus more attention on the delivered software specialization and distinctions that could influence on the functional completeness and reliability of the represented advanced data.

Conclusion

As a result of research of a question of the information gathering and supervision over transport system of city, the analysis of existing up-to-date types of measuring systems has been taken. Proceeding from existing priorities, recommendations for the choice of the most suitable measurement and control system are developed. The approach to the construction of the centralized system of the information gathering about the road network dynamics is offered on the basis of accommodation of various types of measurement means at the reference points.

Literature

1. "Elemental Approach" Intelligent Transport Systems (ITS), July/August 1999.
2. Specification for Urban Traffic Management and Control (UTMC) Systems, TS001: Part 1:1997, Department of Environment, Transport and the Regions, Crown Copyright, 1997.
3. Report: Critical Review of Air Quality Monitoring Technologies for Urban Traffic Management and Control (UMTC) Systems, K.Ropkins and R.N.Colvile, November 2000.
4. Л.И. Турчак Основы численных методов, – М.: Наука, главная редакция физико-математической литературы, 1987.
5. И.Ф. Шишкин, В.М. Станякин Квалиметрия и управление качеством, – М.: Издательство ВЗПИ, 1992.
6. Igors Kabaškins Intelektuālā transporta sistēmu izmantošanas ilgtermiņa programmas konceptuālā modeļa izstrāde Rīgas pilsētas aglomerācijas vajadzībām, Transport and Telecommunication, 2003, Vol. 4, No.3.
7. Inguss Vircavs The clash between mobility and automibilization in Riga, Riga City Development Department, 2000.
8. A.Berezhnoy Identification and Obtaining Statistical Information on the Traffic Flow with Radar Measuring System. Case Study: Latvia // Proceedings of the "Nordic-Baltic Transport Research Conference". April 13-14, 2000. Riga, Latvia. – Vol. 2, 4th sect.
9. Fred L. Hall Traffic stream characteristics
10. Ф. Хейт Математическая теория транспортных потоков, – М.: Мир, 1966.

RESEARCH OF RELIABILITY OF THE SUPPLY SYSTEM IN THE PSC "VILNIAUS AUTOBUSAI"

Kęstutis Lukoševičius, Aurelijus Vestartas, Saulius Nagurnas

Vilnius Gediminas technical university, Department of Automobile transport,
J. Basanaviciaus str. 28, Vilnius, LT-2006 Lithuania
Ph.: (+370)-5 274 4790, Fax.: (+370)-5 274 5068, E-mail: vaida@ti.vtu.lt

Abstract

This paper analyses reliability of the pneumatic system of "Karosa" buses in the PSC "Vilniaus Autobusai". A search of structural solutions is carried out for improvement of reliability of weakest units of this system.

The work includes a statistically processed fault distribution and laws of such distributions are determined. The main fault reasons are enumerated. At last, solutions of problems of improvement of the pneumatic system are presented.

It was established, that the unfailability indications of compressors, made by "Knorr" manufactory, can be increased by alignment of the safety valve so, that it would open at pressure 1.1 – 1.25 MPa. At the same time it will increase and reliability indications of valves located in the high pressure part of this system.

Key words: compressor, valve, reliability, pneumatic system

Introduction

Until 1995 in the fleet of buses of PSC "Vilniaus Autobusai" there were dominating "Ikarus – 280", "Ikarus – 260" and "Mercedes – Benz" buses. They are out – of – date morally and technically (were used 7 and more years). Therefore it was decided to purchase the "Karosa B – 741" bus for experiments.

Exploitation of these buses is rather cheap, parts are not expensive if compared with the analogs of Western producers. "Karosa" buses have many units of famous firms inside, their reliability as high. For example, engines of "Renault" trade – mark, which are used by the majority of vehicle producers, are being mounted into "Karosa" buses. Because of these reasons since 1995 there were purchased 60 more buses B – 741 and 6 B – 732.

Now run of many "Karosa" buses is about 300 000 km.

This paper examines the reliability of one important system (air supply), which drives brakes, suspension, door mechanism a. o. After purchase of new "Karosa" buses there were found that reliability of the air supply system is not high. It shows the necessity of such research.

Repair expenses of the air supply system are rather high. Particularly expensive is breakage of the compressor, which occurs comparatively often. Also there gets blocked up pipes, fails brake chamber a.o.

Research of the Air Supply System

In total 61 "Karosa" bus was examined. Initially there were examined repair orders, registered faults in the air supply system of "Karosa" buses and their repair costs. There were picked out the extremely unreliable units, namely: compressor, valve of pressure regulator, multilinear valve, reducing valve. There were formed variational series and statistical series of initial data.

For reliability evaluation and comparison of the air supply system there were calculated average values of resources and fault intensities. There were used the following formulas [1]:

- Average resource value:

$$T_{vid} = \sum_{i=1}^{n} T_{vi} \cdot p_i \,;$$

- Fault intensity:

$$\lambda(T) = \frac{f(T)}{1 - F(T)};$$

- Mean square deviation:

$$\sigma = \sqrt{\sum_{i=1}^{n} \left(T_{vi} - \overline{T}_{vid}\right)^2 * \frac{m_i}{N}} \,;$$

here: T_{vi} – resource value, corresponding to medium of the i – th interval; p_i – statistical probability of a random value appearance; $f(T)$ – differential function of a distribution law; $F(T)$ – integral function of a distribution law; m_i – frequency; N – total number of realizations of a random value.

Results of the calculations are presented in Table 1 and Figure 1.

The calculations showed that faults of multilinear and reducing valves are distributed according to the normal law; faults of compressor, valve of pressure regulator and brake chamber are distributed in accordance with the Weibull law.

Table 1. Average resource value, mean square deviation.

Meanings / Parts	Average resource value $\overline{T}_{vid}$, km	Mean square deviation σ, km
Compressor	274780	59340
Valve of pressure regulator	274080	50130
Multilinear valve	296350	51170
Reducing valve	296230	50650

Further there are carried out calculations of differential (probability density) and integral (distribution) functions [5]. These functions for the compressor as the main unit are shown in Fig. 1 and Fig. 2.

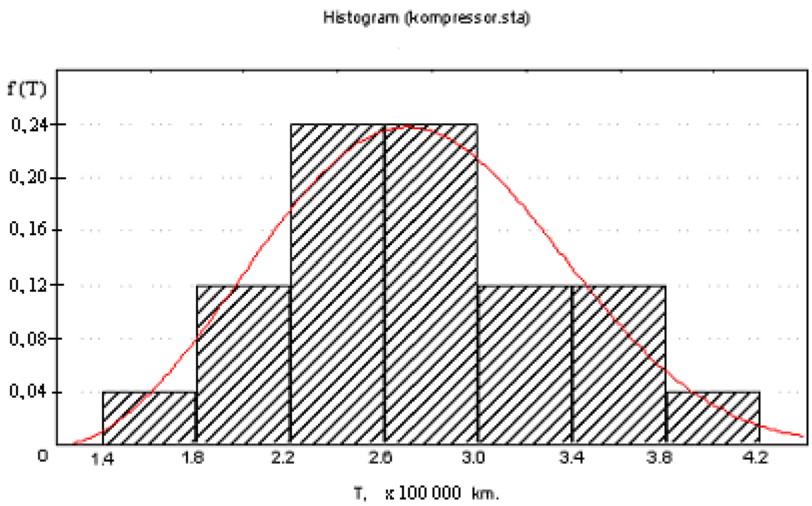


__Figure 1__. Histogram of the compressor fault distribution and graph of its differential function

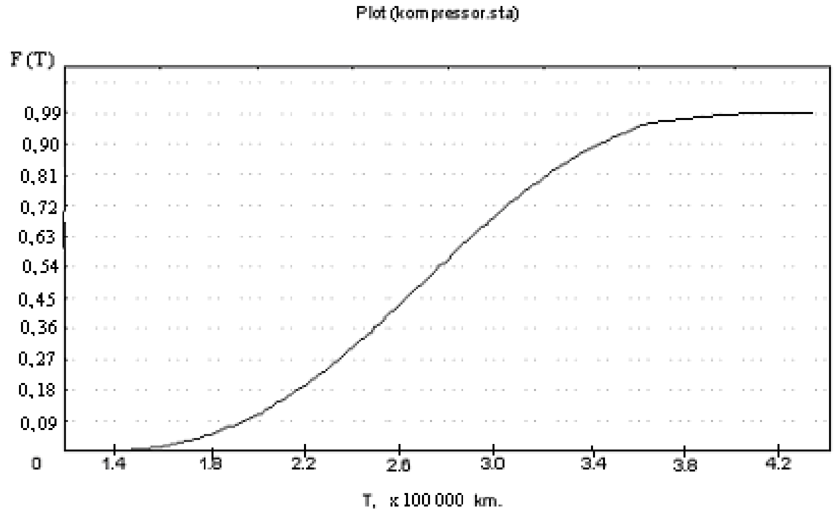


Figure 2. *Theoretical graph of the integral function of the compressor fault distribution*

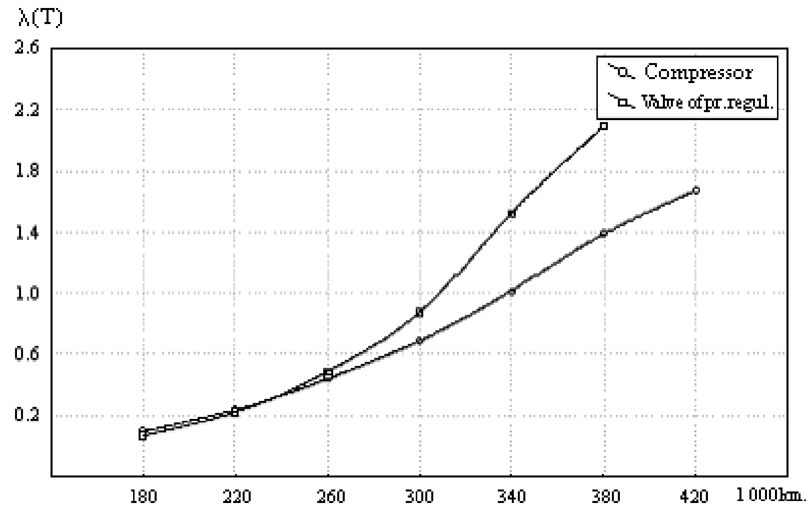


Figure 3. *Graphs of fault intensities of the compressor and of the valve of pressure regulator*

It is seen that fault intensity graphs of either compressor or valve of pressure regulator both coinside at the initial run intervals. But later the graph of the valve goes upwards faster. It can be explained because of aging of rubber parts caused by the agressive working conditions.

Examination of parts of the air supply system during their repair in service enterprises there were observed, that many of faults are connected with the compressor. For instance, the used up rings exhaust to the compressor output line more oil than permissible [4]. The exhausted oil covers valve seat and pipe network. Therefore the valves do not close hermetically, and they became unfit.

It is worth of underlining that compressors, mounted in „Karosa" buses, are not powerful enough, therefore they often fail. Their counter – pressure reaches 1.25 MPa. As for one – stage two – cylinder compressor such indication is rather high. But at such pressure the compressed air exhausts the great volume of heat which must be transferred to the environment intensively. Also the compressor works in the same kind of an enforced mode if there exists such high counter – pressure [2].

When such situation exists, there can be used the following ways of solution of this problem:

1. Cooling of the existing compressor must be enforced. If so, an oxidic using up would take place, which would allow the using up of cylinders and piston rings. To make cooling more intensive there is possible by enlargement of surface of cooling ribs of cylinders, also by

enlargement of surface of waterpockets of the cylinder head or enlargement of the cooling liquid volume across it. But then there would be necessary to reconstruct all the cooling system. As the cooling system of the compressor is connected with the cooling system of the engine, it would disturbe the work of the engine negatively.

2. The most appliable and effective way of improvement of the compressor reliability is use of a more weak safety valve. Now in „Karosa“ buses the safety valve has the regulatorial ability, i.e. for reduction of the opening pressure there is necessary to turn backwards the certain screw, which releases a certain spring inside. So a less pressure will open the valve. As there were mentioned before, the valve of a serial „Karosa“ bus was adjusted to the 1.5 MPa pressure. But the compressor is designed for pressure 1.25 MPa. Therefore the safety valve must be adjusted so that it would open at pressure egual to 1.1–1.25 MPa. Then the compressor will work not enforced, but in conditions, for what it was designed. After adjustment or change of the safety valve, the compressor heating at working conditions improves significantly and oxidic using up in the cylinder – piston pair takes place.

Such reconstruction reduces pressure in the high pressure line to 1.1 – 1.25 MPa, but in the middle line pressure of 0.8 MPa is sufficient. So pressure transferred to executive subsystems is not a bit less than before the reconstruction. It only increases the reliability of all system in common. This reconstruction is not complicated, not expensive and it is within reach of the Vilnius fleet of buses. It is the optimal solution of the problem of reliability increase for the compressor.

PSC „Vilniaus Autobusai“ modified its own „Karosa“ buses in such a way. Preliminary evaluation shows, that after modification the system reliability increased approximatety twice. As this modification was introduced recently, there are no more so far exact data.

When the compressor works rightly, the valves themselves fail seldom. Reduction of compressor failure frequency reduces oil exhaust into the output lines, therefore rubber valve parts remain cleaner, their lifetime becames longer. Besides their repair is not expensive or complicated. Often there is enough to change the worn out rubber parts.

Conclusions

1. The main reason of „Karosa“ bus air supply faults is inadeguacy of working conditions to character of certain units of the system. Most freguently it is related with the compressor and with the output line of high pressure. It was established that this compressor develops pressure up to 1.25 MPa, but the safety valve in the high pressure line is adjusted to pressure 1.5 MPa.

2. For increase of reliability of the compressor there is proposed change of its working conditions from the existing enforced to nominal ones: to adjust opening of the high pressure valve in the compressor output line at pressure 1.1 – 1.25 MPa. Such modification will result also in the reliability increase of other valves (multilinear and reducing) since they will work in cleaner conditions, without oil pollutants from the failed compressor.

3. In accordance with the preliminary data reliability of the air supply system increased approximately twice.

References

1. V Sakalauskas, *Statistika su Statistica*, Vilnius, 1998, 228 p.
2. Br Spruogis, *Hidraulinės mašinos ir pavaros*, Vilnius, 1991, 325 p.
3. F. Blanchini, P. Giannattasio, In: Adaptive control of compressor surge instability, *Automatica*, Vol 38, No 8, 2002, p 1373 – 1380.
4. Т. А. Сырицин, *Надежностъ гидро- и пневмо- привода*, М.,1984, 216 с.
5. В. П. Боровиков, И П. Боровиков, *STATISTICA*, М., 1998, 608 с.
6. AO KAROSA, *KAROSA пневматическая система*, Vysoke Myto, 1997, 82 с.

INVESTIGATION OF DYNAMIC PROCESSES IN THE HYDRAULIC BRAKING SYSTEM OF TRANSPORT VEHICLE WITH ANTI-LOCK BRAKE SYSTEM

Marijonas Bogdevicius [1], Oleg Vladimirov [2]

[1] *Vilnius Gediminas Technical University, Faculty of Transport Engineering,*
Department of Transport Technological Equipment, Plytines Street 27-308, LT-2016 Vilnius, Lithuania,
Telephone +370 5 2744782, Fax +370 5 2745060, E-mail: marius@ti.vtu.lt,
http://www.vgtu.lt/english/faculties/transport/technolog

[2] *Vilnius Gediminas Technical University, Faculty of Transport Engineering,*
Department of Transport Technological Equipment, Plytines Street 27-301, LT-2016 Vilnius, Lithuania,
Telephone +370 5 2744783, Fax +370 5 2745060, E-mail: olegvl@ti.vtu.lt

Abstract

The modern hydraulic brake system consists of various functional systems and elements such as ABS, EDB, cylinders, electromagnetic valves, filters, pipelines, microprocessors and etc. The processes, occurring in these systems, quickly varied, i.e. speed and pressure of fluid quickly vary in time. The speed of a sound in hydraulic system is much greater, in comparison with the speed of a sound in pneumatic system; therefore the dynamic processes in hydraulic system occur much faster. But if some air occurs in hydraulic system, the dynamic characteristic of it is rapidly changes. The account of physical processes in hydraulic systems, depending on properties of the system and fluid, has the importance evaluating reliability of break systems. It's also very important to know that processes are there in the system when any mode of failure occurs (such as the leakage of system and etc.). The simplest way to do this is to use simulation of dynamic processes in the hydraulic systems. In the article simulation of hydraulic brake system with ABS using the simulation of fluid movement in pipelines is considered. The following things are taken into account in the present model: distribution of sound waves in fluid, air influence to the speed of pressure waves. The received system of the equations of fluid movement is solved by a method of the characteristics. Nonlinear algebraic equations are solved by Newton method. The differential equations of piston movement are solved by trapezoid method. Air influence to the system work is evaluated. An example of the hydraulic brake system with simple ABS is considered. The relations of brake forces and pressures in cylinders are presented. Also results of the simulation of braking system on the different type of road surface are presented.

Keywords: anti-lock brake system, dynamic, hydraulic system, numerical methods

1. Introduction

Brake means the part in which the forces opposing the movement of the vehicle develop. It may be of 4 main types. A *friction* brake, when the friction between two parts of the vehicle moving relatively to one another generates the brake forces. An *electrical* brake, when the forces are generated by electromagnetic action between two parts of the vehicle moving relatively to but not in contact with one another. A *fluid* brake, when the forces are generated by the action of a fluid situated between two parts of the vehicle moving relatively to one another. An *engine* brake, when the forces are derived from a controlled increase in the braking action of the engine transmitted to the wheels [1].

To start simulating of automobile hydraulic brake system with ABS the first step is to create a simple properly working model of simple brake system. The additional functional elements (such as hydraulic amplifier, electromagnetic valves and etc.) are required to evaluate all the systems work as one unit and to establish the possibility to add more complicated components to the system.

2. Mathematical Model of Hydraulic Break System

In this article automobile hydraulic break system consisting of two contours is considered. The main components of the system are shown on fig.1. When the driver presses the brake pedal, it pushes down the piston in the master cylinder 1, so creating pressure in the fluid in pipeline 2. The ideal variant is when the fluid is incompressible, but in real model it is not so. The fluid comes to the hydraulic amplifier 3, where the pressure grows up propotionally to the crossections of pistons. Two contour separator 4 is intended to prevent failure of all system, when one of the cylinders or pipelines do not work properly. When one countour lost its functions, the other continues to brake. This requirement is essential for all brake systems [1, 2]. During braking the pressure of fluid is transmitted to the wheel cylinders 5-8, which forses the brake pads against the revolving disc. The amplifier 3 and the difference in crossection of the cylinders allow to use a relatively small force applied on the pedal to get a large force on the brake pads.

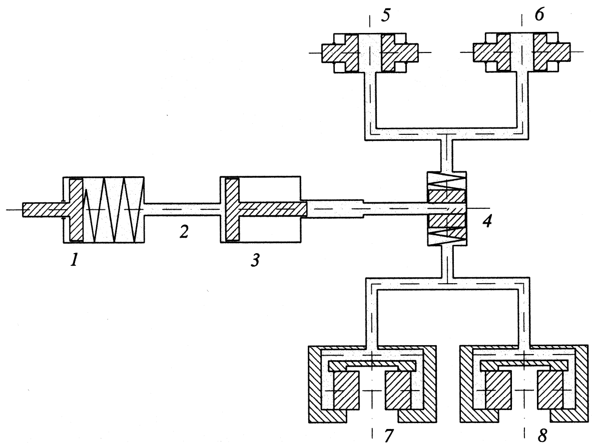


Figure 1. *Principle scheme of the system:*
1 – main cylinder, 2 – pipeline, 3 – hydraulic amplifier, 4 – two contours separator,
5, 6 – rear axle brake cylinders, 7, 8 – front axle brake cylinders

Let's consider the system from the main cylinder. The equation of piston number 1 movement can be written:

$$m_1\ddot{x}_1 = \sum F = F_{sp} - k\left(x_0 + x_1\right) - F_{tr} sign\left(\dot{x}_1\right) - S_1\left(p_1 - p_{atm}\right), \tag{1}$$

where m_1 is weight of piston, $\ddot{x}_1$, $\dot{x}_1$, x_1 are acceleration, velocity and position of piston 1, F_{sp} is pressing force, k, x_0 are stiffness and initial compression of the spring, F_{tr} is friction force, S_1 is area of the cylinder cross-section, p_1, p_{atm} are pressure in 1-st cavity and atmospheric pressure.

This equation (1) is solved by the Euler method, taking in account boundary conditions of piston *1* movement. Fluid pressure in *1*-st cavity is solved using equation of debit:

$$Q_1 = S_1 \cdot \dot{x}_1 - \frac{V_{10} - S_1 \cdot x_1}{K} \cdot \frac{dp}{dt} - S_2 \cdot v_{21} = 0\,, \tag{2}$$

where V_{10} is initial volume of 1-st cavity, v_{21} is fluid velocity in 1-st point of pipeline 2, K is modulus of elasticity of fluid.

From equation (2) we will receive the pressure in 1-st cavity at the moment of time $t + \Delta t$:

$$p_{1,t+\Delta t} = p_{1,t} + \Delta t \left[\left(S_1\dot{x}_1 - S_2 v_{21}\right) \frac{K}{V_{10} - S_1 \cdot x_1} \right]_t . \tag{3}$$

The equations of movement of other pistons and the equations of debit in the cylinders are described in the same form.

The movement of fluid in a hydraulic pipeline is accepted as one dimensional and unsteady, i.e. all local speeds are considered equal to the average speed and depend on time. The pressure, also, is considered identical at all points of the cross section and depends on longitudinal coordinate of the pipeline and time. Such movement of fluid is characterized by occurrence of a wave of increased and lowered pressure, which is distributed from the place of pressure change and deformation of the pipeline walls with the velocity of sound [3, 4, 5, 7, 8].

Equation of the fluid continuity can be written in a differential form as follow [5]:

$$\frac{\partial}{\partial t}\left[S(x)\rho\right]+\frac{\partial}{\partial x}\left[S(x)\rho v\right]=F_1(x). \tag{4}$$

where ρ, v are density and velocity of fluid, $S(x)$ is cross section area of the pipeline, $F_1(x)$ is discharge of fluid mass to the unit of the length, in the pipeline.

Equation of fluid flow impulse (momentum)

$$\frac{\partial}{\partial t}\left[S(x)\rho v\right]+\frac{\partial}{\partial x}\left[S(x)\ \left(p+\rho v^2\right)\right]+\Pi(x)\ \tau+S(x)\rho a_x=F_2(x)+p\frac{\partial S}{\partial x}. \tag{5}$$

where p is fluid pressure, $\Pi(x)$ is perimeter of cross section of the pipeline, τ is tangential fluid stress in the inner surface of the pipeline, a_x is acceleration along x axis, $F_2(x)$ is kinetic energy of the fluid flow in the pipeline to the unit of area.

The system of equations (4) and (5) can be written as the system second-order quasi-linear differential equations

$$[A]\left\{\frac{\partial u}{\partial t}\right\}+[B]\left\{\frac{\partial u}{\partial x}\right\}=\{f\}. \tag{6}$$

where a_{ij}, b_{ij}, f_i are matrices $[A]$, $[B]$ and elements of vector $\{f\}$ which depend on t, x and elements u_i of vector $\{u\}$.

The differential equations of fluid movement in pipelines are solved by the method of characteristics [3, 4, 5]. Other cylinders and pipelines are calculated in the same manner.

3. Mathematical Model of Anti-Lock Brake System

In this article automobile hydraulic ABS, consisting of hydraulic brake system (fig. 1), electromagnetic valve and wheel brake cylinder, is considered. The main components of the system are shown on fig. 2.

The equation of dynamic of electromagnetic valve can be written as follow:

$$L\frac{di}{dt}+Ri=U, \tag{7}$$

where L, R are induction and active resistance of coil; i is current; U is voltage.

When the control signal of coil (U) arrives, electromagnetic valve closes or opens depending on the stage of braking. The equation of electromagnetic valve movement:

$$m_S\frac{d^2X_S}{dt^2}=F_{em}-F\left(X_S\dot{X}_S\right), \tag{8}$$

where m_s is weight of piston; X_S is position of piston; $F\left(X_S\dot{X}_S\right)$ is valve's mechanical resistance force; F_{em} is force which is induced by electromagnets coil,

$$F_{em} = ci^2, \tag{9}$$

c is constant of electromagnet.

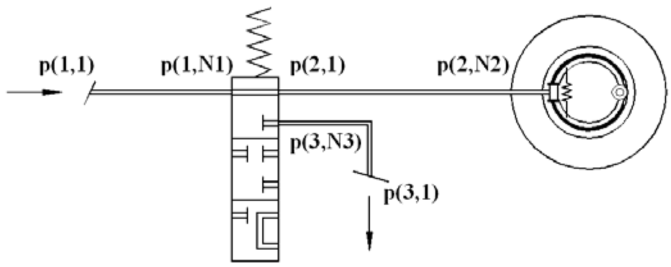


Figure 2. *Part of automobile anti-lock brake system*

Pressure in valve cavity:

$$\frac{dp_v}{dt} = \frac{K}{V}\left(Q_{input} - Q_{output}\right), \tag{10}$$

where K is modulus of elasticity of fluid; V is volume of valve cavity; Q_{input}, Q_{output} are input and output debits.

ABS operates by evaluating two main parameters. These are the changes in the wheel angular deceleration and wheel slip. The wheel deceleration is determined by differentiation of the wheel angular velocity captured by speed sensor, and the tyre slip ratio is estimated by comparing the current wheel angular velocity with the angular velocity calculated from the absolute speed of the vehicle which is called the reference speed. The equation of tyre slip λ can be written as follow:

$$\lambda = 1 - \frac{\omega(t) R(t)}{v(t)}, \tag{11}$$

where ω is wheel angular velocity; R is wheel radius; v is automobile speed.

The derivation of tyre slip of wheel i can be written as follow:

$$\frac{d\lambda_i}{dt} = -\frac{1}{v_i^2}\left[\left(\dot{\omega}_i R_i(t) + \omega_i \dot{R}_i(t)\right) \cdot v_i - \omega_i R_i(t) \cdot \dot{v}_i\right]. \tag{12}$$

In mathematical model automobile wheel angular velocity and radius, and linear velocity and deceleration are determined from the automobile movement equations' system [6]. So, combining automobile movement equations' system with the automobile hydraulic anti-lock brake system and the derivation of tyre slip of each wheel, we'll get full system of equations of automobile braking process.

4. Rezults of Simulation of Automobile Brake System with ABS

Using created mathematical model of automobile hydraulic brake system with ABS, the changes of most essential parameters of the system were determined. The relations of fluid pressures and velocities in typical systems points (shown on fig. 2) are presented on fig. 3 and 4. The relations of wheel brake cylinder pistons position and velocity are shown on fig. 5. The achieved braking force and the electromagnetic valves working periods are presented on fig. 6

and 7. Also results of the simulation of braking system on the different type of road surface are presented on fig. 8.

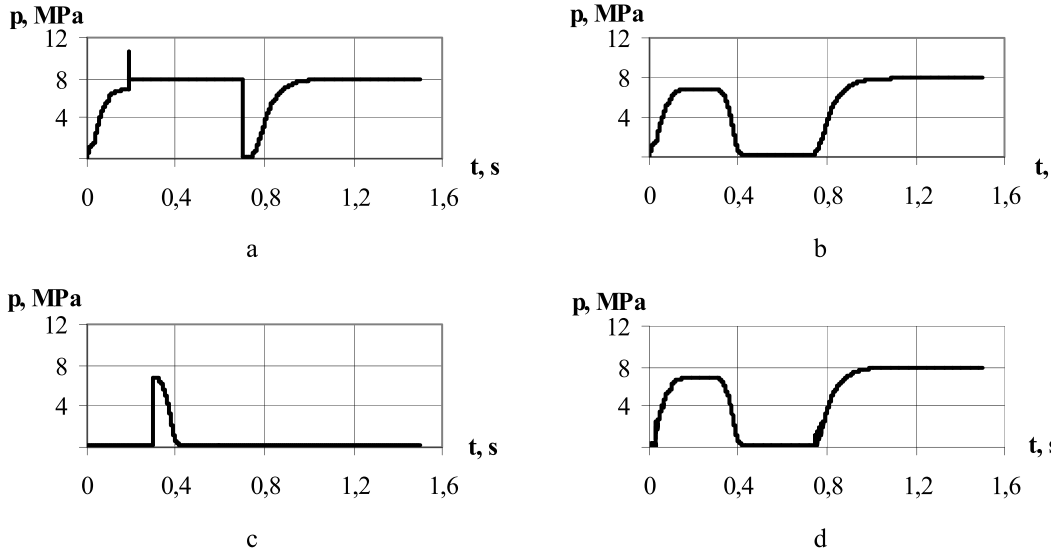


Figure 3. *Dependence of fluid pressure on time in typical ABS points:*
a $- p(1, N1)$; *b* $- p(2,1)$; *c* $- p(3, N3)$; *d* $- p(4,1)$

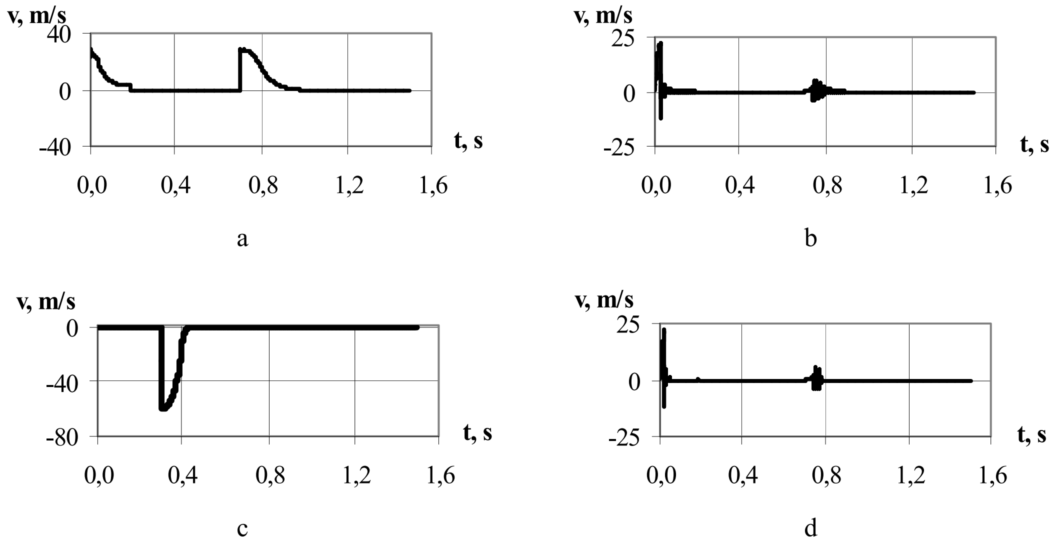


Figure 4. *Dependence of fluid velocity on time in ABS points:*
a $- v(1, N1)$; *b* $- v(2,1)$; *c* $- v(3, N3)$; *d* $- v(4,1)$

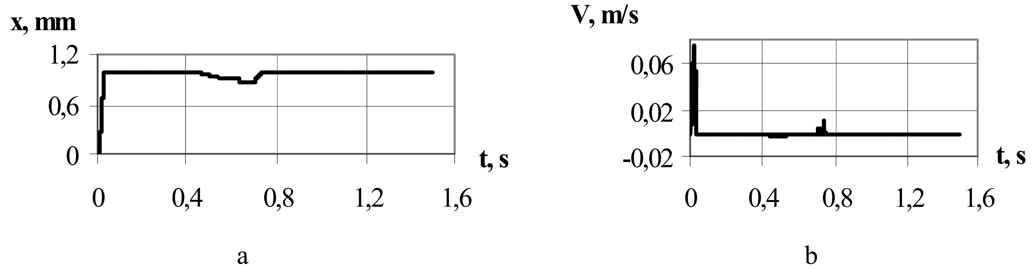


Figure 5. *Dependence of wheel brake cilinder pistons position and velocity on time:*
a –position; b –velocity

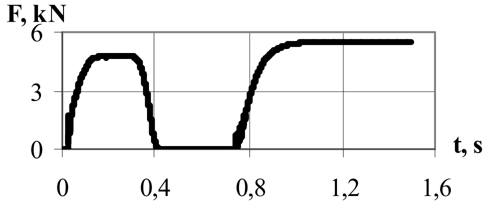


Figure 6. Automobile wheel braking force

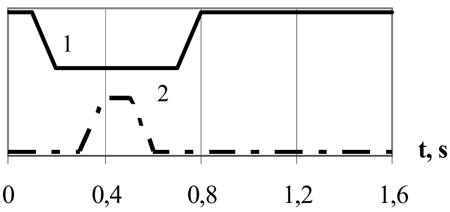
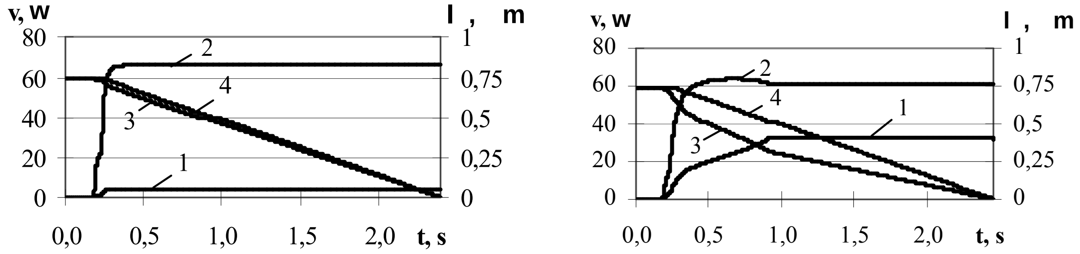


Figure 7. Periods of electromagnetic valve work: 1 – inlet; 2 – outlet



a b

Figure 8. *Results of the simulation of braking system on the different type of road surface: a – dry asphalt; b – wet asphalt; 1 – slip λ, 2 – friction μ, 3 – wheel angular velosity (ω, rad/s), 4 – automobile velosity (v, km/h)*

5. Conclusions

Mathematical model of simple hydraulic brake system with ABS is developed. The model shows properly work when the air quantity in fluid is taken into account. The proper work of hydraulic brake system with such components as hydraulic amplifier, two contour separator and electromagnetic valve shows the possibility to add more complicated components. Mathematical model of this system was attached to the model of automobile. The developed model allows to investigate any failure of hydraulic brake system by changing the boundary conditions (fluid pressure in any point of pipelines, motion of pistons, air quantity and etc.) and to simulate the behavior of automobile on the road.

References

1. Official journal of the European Communities 71/320.
2. Zeromskas R., Mitunevicius V. The Basic Requirements to Road Vehicles. Vilnius: Technika, 1998.
3. Aladjev V., Bogdevicius M. Maple 6: Solution of the Mathematical, Statistical and Engineering –Physical Problems. Moscow: Laboratory of Basic Knowlegdes, 2001, p.864 (in Russian).
4. Bogdevičius M., Simulation and Interaction of Mechanical and Hydraulic System, *Proceedings of Tenth World Congress on the Theory of Machines and Mechanisms* (Oulu, 1999, p. 2110-2115).
5. Bogdevičius M. Simulation of Dynamic Processes in Hydraulic, Pneumatic and Mechanical Drivers and Their elements. Vilnius: Technika, 2000, p.96.
6. Transport: Technologies, Economy, Environment, Health. Monograph. Vilnius: Technika, 2003, p.876 (in Lithunian).
7. Bogdevicius M. Non-stationary movement of fluid in the elastic, visco-elastic, and plastic pipe. Applied Mechanics. 1994, No 2, p 117 – 124 (in Lithuanian).
8. Bogdevicius M. Simulation of Hydraulic Systems of the Pump by the Method of Characteristics (Моделирование гидравлической системы насоса методом характеристик). Transport Engineering (Transportas). Vilnius, 1997, No. 2, p 30 – 37 (in Russian).

DEFINITION OF A TRAJECTORY OF THE VEHICLE MOVEMENT, TAKING INTO ACCOUNT INTERACTION OF ITS WHEELS AND ROAD SURFACE

Olegas Prentkovskis [1], Marijonas Bogdevičius [2]

[1] *Vilnius Gediminas Technical University, Faculty of Transport Engineering,*
Department of Transport Technological Equipment,
Plytines Street 27-301, LT-2016 Vilnius, Lithuania,
Tel. +370 5 2744784, Fax +370 5 2745060, E-mail: olegas@ti.vtu.lt,
http://www.vgtu.lt/english/faculties/transport/technolog

[2] *Vilnius Gediminas Technical University, Faculty of Transport Engineering,*
Department of Transport Technological Equipment,
Plytines Street 27-308, LT-2016 Vilnius, Lithuania,
Tel. +370 5 2744782, Fax +370 5 2745060, E-mail: marius@ti.vtu.lt,
http://www.vgtu.lt/english/faculties/transport/technolog

Abstract

In this paper the authors focus on the simulation of the vehicle on a certain road and propose their specific solution to this problem. Mathematical model of the system "vehicle – road" is presented. The vehicle is simulated by concentrated masses interconnected by elastic and dissipative links. The presented model of the vehicle evaluates the movement of the vehicle body in space; movement and turning of front and rear suspensions with respect to the body; interaction of the wheel with the road pavement surface; blocking of the wheel; changing cohesive forces which influence the vehicle. The investigated road pavement surface is simulated by triangular finite elements, a certain height of road pavement surface roughness and cohesion coefficients of road pavement surface and the vehicle wheel in the longitudinal and transverse directions of the wheel are selected in each finite element nodal point. Presented results, illustrating the vehicle movement trajectories when braking at various initial conditions and a certain pavement surface of the road section under investigation; the vehicle driving on the speed reduction bump ("sleeping policeman").

Keywords: vehicle, movement trajectory, dynamics, interaction, wheel, road pavement surface, mathematical simulation

1. Introduction

The vehicle is the most comfortable and popular means of transportation; however, it causes a lot of problems for all road users.

Accidents occur every 15 seconds, in which people suffer. One person is injured every 1.5 minute, i.e. one thousand people are injured in 24 hours. Thus, ca one million inhabitants of the Earth are lost in the period of three years. The number of old cars, exceeding its own resources several times, has increased in Lithuania. Following statistics, one person is killed every 8 hours, and one person is injured each hour on Lithuanian roads [1].

Each year, 4,000 – 6,000 registered road accidents occur in Lithuania (Fig. 1). The most typical are as follows: running over pedestrians, collision, overturning, hitting an obstacle, running over cyclists, hitting a parked vehicle [2].

The types of accidents vary. To study accidents of various types accurately, first of all, the universal dynamic model of the vehicle as well as the model of the road shall be constructed. Simulation and description of the vehicle movement have been studied by scientists for several years [3-9]. The dynamic models of vehicles presented by other scientists are very complicated, and they can be simplified for simulating the movement of the vehicle on the road.

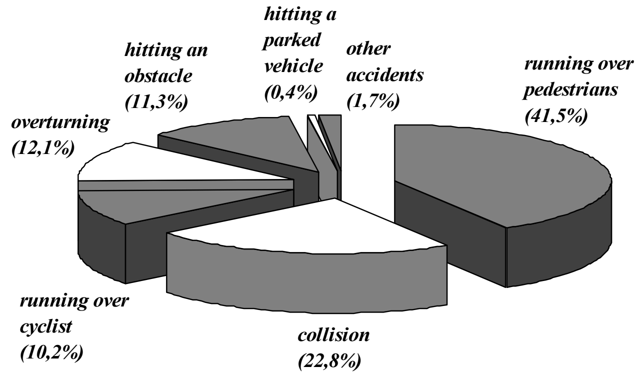


Figure 1. *Statistical data on registered accidents in Lithuania*

The authors of this article focus on the simulation of the vehicle on a certain road and propose their specific solution to this problem.

Movement of the vehicle on the road is a very complicated process which depends not only on the dynamic properties of the vehicle but also on the conditions of the road section under investigation.

The system "vehicle – road" was selected for research, it consists of:

- a certain vehicle with known geometric parameters and physical mechanical properties (masses of individual elements, moments of inertia, size, stiffness coefficients, mechanical energy damping coefficients, etc.);
- a road section with known geometric parameters and physical mechanical properties (length, width, pavement surface, etc.).

2. Mathematical Model of the Road Pavement Surface

To describe the road pavement surface roughness and cohesion coefficients of the road pavement and vehicle wheels on each nodal point of the road pavement surface, the method of finite elements is applied [2, 9, 10, 11].

The total pavement surface of the road section is divided into triangular finite elements (Fig. 2). A certain height of the road pavement surface roughness and cohesion coefficients of road pavement surface and vehicle wheels in the longitudinal and transverse directions of the vehicle wheels are selected in each nodal point of finite elements.

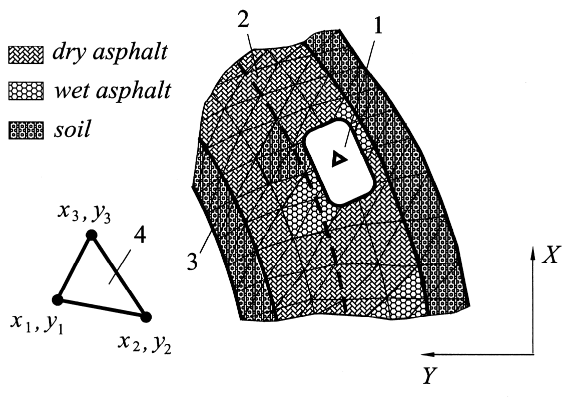


Figure 2. *Expansion of the pavement surface of the road section into finite elements:*
1 – vehicle; 2 – road carriageway; 3 – shoulder; 4 – triangular finite element

3. Mathematical Model of the Vehicle

The vehicle is simulated by concentrated masses interconnected by elastic and dissipative links (Kelvin-Foight elements). The vehicle model consists of seven concentrated masses: body, front and rear suspension and four wheels (Fig. 3) [2, 9].

In the geometrical centre O of the vehicle the system of coordinates $\zeta_v - \eta_v - \xi_v$ is introduced and in the masses centre C of the vehicle the system of coordinates $X_v - Y_v - Z_v$ is introduced. The masses centre C of the vehicle is remote from geometrical centre O at distances a_x, a_y, a_z in the directions of axes ζ_v, η_v, ξ_v. This discrepancy between vehicle masses and geometrical centre can be explained by the asymmetric distribution of passenger and load mass in the vehicle. The body can move in space in the directions of global axis X, Y and Z, and turn around local axes X_v, Y_v, Z_v of the vehicle. Front and rear suspensions can move in the direction of axis Z_v and turn on the vertical plane with respect to the body. Wheels can move in the direction of axis Z_v with respect to the body.

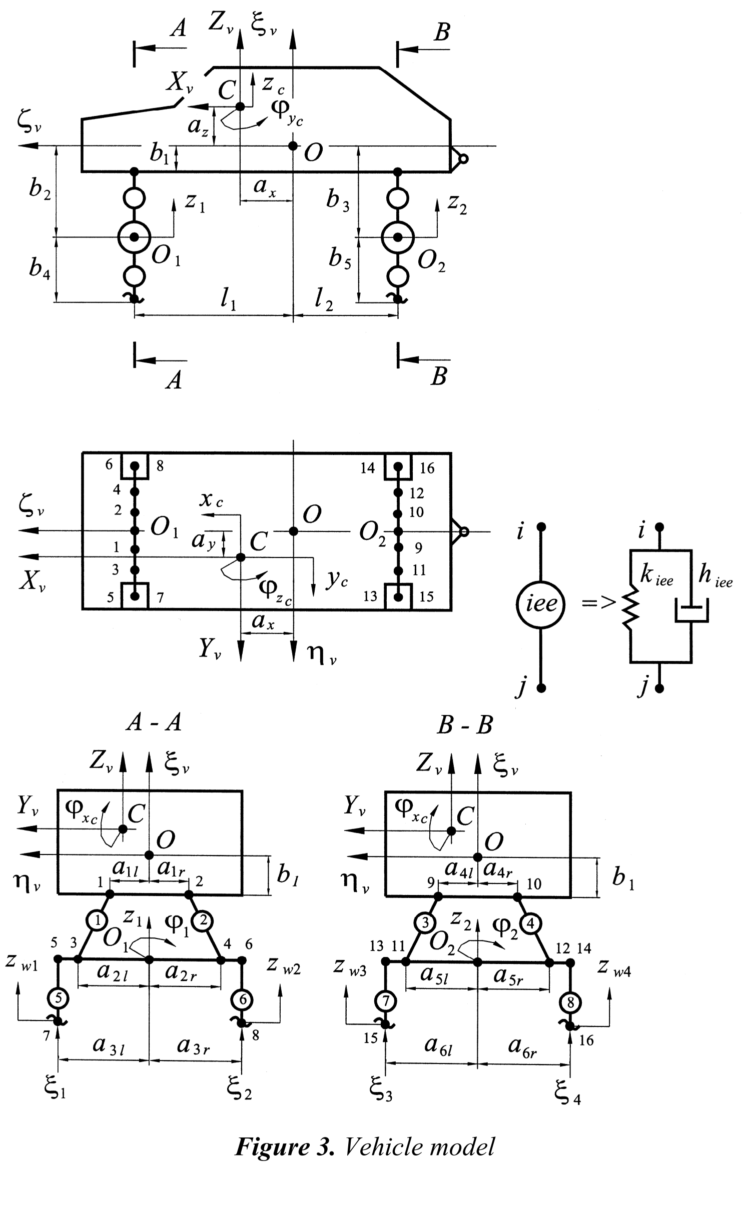


Figure 3. *Vehicle model*

To describe the movement of the vehicle, the following generalized coordinates are introduced:

$$\{q_v\}^T = \begin{bmatrix} x_C & y_C & z_C & \varphi_{x_C} & \varphi_{y_C} & \varphi_{z_C} & z_1 & \varphi_1 & z_2 & \varphi_2 & z_{w1} & z_{w2} & z_{w3} & z_{w4} \end{bmatrix}. \quad (1)$$

Movement of the vehicle, as movement of an intricate mechanical system, is described by second degree LaGrange equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\{\dot{q}_v\}}\right) - \frac{\partial L}{\partial\{q_v\}} + \frac{\partial H}{\partial\{\dot{q}_v\}} = \{F_v\}; \quad (2)$$

where: L – vehicle (mechanical system) LaGrange function,

$$L = \mathrm{T} - \Pi; \quad (3)$$

where: T – kinetic energy of the vehicle; Π – potential energy of vehicle elastic elements; H – dissipative function of vehicle dissipative elements; $\{q_v\}$ – vector of generalized coordinates; $\{\dot{q}_v\}$ – vector of generalized speeds; $\{F_v\}$ – vector of generalized forces, influencing the vehicle.

After having inserted the expressions of the vehicle kinetic and potential energies, dissipative functions and the vector of generalized forces, influencing the vehicle, into the second degree LaGrange equations (2), the system of the vehicle movement equation is obtained, which can be written in the matrix form:

$$[M_v]\{\ddot{q}_v\} = \{Q_v\}; \quad (4)$$

where: $[M_v]$ – matrix of vehicle masses; $\{\ddot{q}_v\}$ – vector of generalized accelerations; $\{Q_v\}$ – loading vector of the vehicle.

The system of equations (4) is solved by the method of Runge-Kutta [2, 12]. For this purpose, it has to be rearranged from differential equations of the second order to differential equations of the first order:

$$\begin{cases} \dfrac{d}{dt}\{q_v\} = \{\dot{q}_v\}; \\ \dfrac{d}{dt}\{\dot{q}_v\} = [M_v]^{-1}\{Q_v\}. \end{cases} \quad (5)$$

4. Conditions of the Contact of the Bottom Nodal Point of the Vehicle Wheel and road Pavement Surface

When solving system of equations (5), the contact between the vehicle wheel and road pavement surface has to be evaluated in every time step [2, 9].

The following conditions of contact between the vehicle wheel and road pavement surface are selected (Fig. 4):

$$z_{wj} = \begin{cases} z_{wj}, & when\ z_{wj} \geq \xi_j; \\ \xi_j, & when\ z_{wj} < \xi_j; \end{cases} \qquad \dot{z}_{wj} = \begin{cases} \dot{z}_{wj}, & when\ z_{wj} \geq \xi_j\ and\ \dot{z}_{wj} \geq 0; \\ 0, & when\ z_{wj} < \xi_j\ and\ \dot{z}_{wj} < 0; \end{cases}$$

$$\ddot{z}_{wj} = \begin{cases} F_{eq\ j+10}, & when\ z_{wj} \geq \xi_j\ and\ F_{eq\ j+10} \geq 0; \\ 0, & when\ z_{wj} < \xi_j\ and\ F_{eq\ j+10} < 0; \end{cases} \quad (6)$$

where: ξ_j – heights of road pavement surface roughness under the vehicle wheels, $j = 1 \div 4$; $F_{eq\ j+10}$ – vehicle movement equations right side obtained with respect to $\dot{z}_{wj}$.

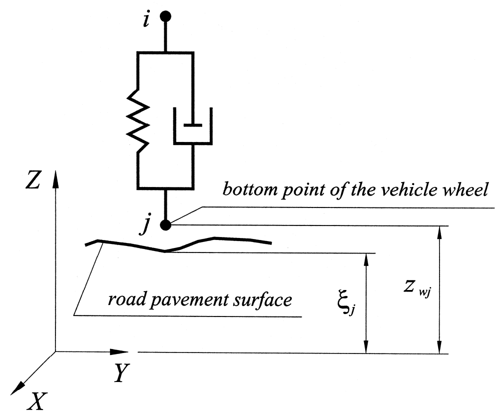


Figure 4. *Identification circuit of the contact between vehicle wheel and road pavement surface*

To find the heights of road pavement surface under the vehicle wheel, global coordinates of wheels in the system of coordinates $X - Y$ have to be known.

When global coordinates of the wheels are known, dependence of a certain wheel on the exact finite element of the road pavement surface and the height of roughness of the road pavement surface under it can be determined [2, 9].

5. Identification of Road Pavement Surface Roughness and Cohesion Coefficients with the Vehicle Wheels

The roughness of the surface on every nodal point x, y of road finite element (Fig. 2) are approximated as follows [2, 9]:

$$\xi(x,\ y)=[N_1 \quad N_2 \quad N_3]\begin{Bmatrix}\xi_1\\ \xi_2\\ \xi_3\end{Bmatrix}=[N(x,\ y)]\{\xi\}; \tag{7}$$

where: $\xi(x,\ y)$ – road pavement surface roughness on finite element nodal point x, y; N_1, N_2, N_3 – shape functions of triangle finite element; ξ_1, ξ_2, ξ_3 – surface roughness on finite element nodal points.

Let us consider that the cohesion coefficients of the vehicle wheel and road pavement surface are distributed according to the law of ellipsis (Fig. 5) [2, 9]:

$$\left(\frac{x_{wj}}{a}\right)^2+\left(\frac{y_{wj}}{b}\right)^2=1; \tag{8}$$

where: x_{wj}, y_{wj} – cohesion coefficients on every nodal point of ellipsis; $a=\varphi_{cohes}^{\max}$ – maximum cohesion coefficient (in the longitudinal direction of the wheel); $b=\varphi_{cohes}^{\min}$ – minimum cohesion coefficient (in the transverse direction of the wheel).

In Fig. 5 the distribution of the cohesion coefficient of wheel j are presented, maximum cohesion coefficient – in the direction of axis X_{wj}, minimum – in the direction of axis Y_{wj}. $X - Y$ – global system of coordinates, where movement of the vehicle is investigated; $X_v - Y_v$ – system of local coordinate of the vehicle, the beginning of this system of coordinates in the vehicle masses centre; $X_{wj} - Y_{wj}$ – system of local coordinates of the wheel j, the beginning of the system of coordinates is in the vehicle wheel masses centre; φ_{z_C} –

angle between the systems of coordinates $X-Y$ and $X_v - Y_v$ (turning angle of the vehicle around the vertical axis, which crosses the its masses centre); γ_{wj} – turning angle of wheel j with respect to the vehicle body; $\{v_{wj}\}$ – sliding speed vector of wheel j; χ_{wj} – angle between axis X_{wj} of wheel j and sliding speed vector $\{v_{wj}\}$; δ_{wj} – angle between the local axis X_v of the vehicle and sliding speed vector $\{v_{wj}\}$ of the wheel j.

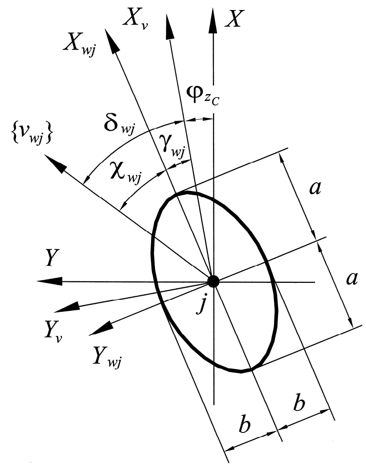


Figure 5. *Distribution of the cohesion coefficient of wheel j*

Equation (8) can be written parametrically:

$$x_{wj} = a\cos\chi_{wj}; \quad y_{wj} = b\sin\chi_{wj}. \tag{9}$$

Then dependence of cohesion coefficient on angle χ_{wj} equals to:

$$\varphi_{cohes}\left(\chi_{wj}\right) = \sqrt{\left(a\cos\chi_{wj}\right)^2 + \left(b\sin\chi_{wj}\right)^2}; \tag{10}$$

where: $\varphi_{cohes}\left(\chi_{wj}\right)$ – cohesion coefficient of wheel j and road pavement surface in the direction of the wheel sliding speed $\{v_{wj}\}$.

$$\chi_{wj} = \delta_{wj} - \gamma_{wj}; \tag{11}$$

where: γ_{wj} – turning angle of wheel j with respect to the body of the vehicle anti-clockwise; δ_{wj} – angle between the local axis X_v of the vehicle and sliding speed vector $\{v_{wj}\}$ of the wheel j,

$$\delta_{wj} = arctg\left(\frac{v_{wjy_v}}{v_{wjx_v}}\right); \tag{12}$$

where: v_{wjx_v}, v_{wjy_v} – projections of the wheel sliding speed vector $\{v_{wj}\}$ to axes X_v and Y_v respectively;

Cohesion coefficients of the wheel and the road pavement surface on any finite element nodal point x, y are approximated as follows:

$$\varphi_{cohes}^{\max}(x,\ y) = \left[N(x,\ y)\right]\left\{\varphi_{cohes}^{\max}\right\} = \left[N(x,\ y)\right]\{a\};$$
$$\varphi_{cohes}^{\min}(x,\ y) = \left[N(x,\ y)\right]\left\{\varphi_{cohes}^{\min}\right\} = \left[N(x,\ y)\right]\{b\}; \tag{13}$$

where: $\varphi_{cohes}^{\max}(x,\ y)$, $\varphi_{cohes}^{\min}(x,\ y)$ – maximum and minimum cohesion coefficients of the wheel and the road pavement surface on finite element nodal point x, y; $\left[N(x,\ y)\right]$ – shape functions of triangle finite element; $\{a\}$, $\{b\}$ – vectors of maximum and minimum cohesion coefficients on finite element nodal points.

Road pavement surface roughness and cohesion coefficients of the wheel and the road pavement surface are divided linearly on the triangle finite element, therefore, the net of finite element shall be denser in the places, where surface roughness or cohesion coefficients change rapidly (increase or decrease).

To perform fewer calculations, it shall be estimated to which of the finite element each wheel belongs and heights of road pavement surface roughness and cohesion coefficients of the each wheel and the road pavement surface are estimated only on four selected nodal points [2, 9].

6. Identification of Cohesive Forces

When studying the movement of the vehicle, vehicle wheel and road pavement surface cohesive forces cannot be neglected. Cohesive forces under different wheels can have different values depending on different cohesion coefficients of the vehicle wheels and road pavement surface [2].

If the direction of the vehicle speed and wheel speed coincide and the wheel is not blocked (is not braked), cohesive force between the wheel and road pavement surface is equal to zero. Cohesive force will have certain value provided the directions of speeds mentioned above do not coincide or the wheel is not blocked (braking occurs).

When braking the vehicle or when it is sliding, redistribution of the vehicle mass on the front and rear suspensions occurs, which influences the movement of the vehicle. The more intensively the vehicle is braked, the greater part of the mass influences the front suspension and the front wheels are pressed harder (when the vehicle moves forward).

In a general case, cohesive force, which influence wheel j, equals to:

$$F_{cohesj} = R_{zj}\varphi_{cohesj}; \tag{14}$$

where: R_{zj} – normal reaction, which influence road pavement surface; φ_{cohesj} – cohesion coefficient of the vehicle wheel and road pavement surface, which depends on the road pavement surface; $j = 1 \div 4$.

Normal reaction, which influence the road pavement surface, is estimated as follows:

$$R_{zj} = \begin{cases} k_j\left|\Delta l_j\right| + h_j\left|\Delta \dot{l}_j\right|, & when\ \Delta l_j \le 0; \\ 0, & when\ \Delta l_j > 0; \end{cases} \tag{15}$$

where: k_j – stiffness coefficient of suspension elastic element *j*; h_j – mechanical energy damping coefficient of suspension elastic element j; Δl_j – contraction of suspension elastic element *j*; $\Delta \dot{l}_j$ – contraction speed of suspension elastic element *j*.

If $\Delta l_j \le 0$, the vehicle wheel *j* contacts with the road pavement surface; provided $\Delta l_j > 0$, the vehicle wheel *j* does not contact with the road pavement surface.

Cohesive forces, which influence the vehicle, are shown on Fig. 6.

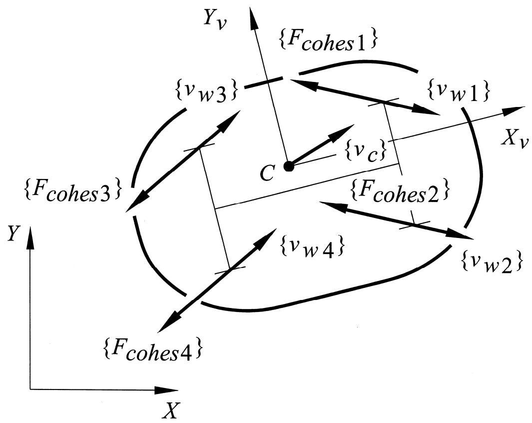


Figure 6. *Cohesive forces, which influence the vehicle wheels*

Cohesive force, which influences wheel *j*, is inserted in vector:

$$\{F_{coh esj}\} = \begin{Bmatrix} -sign\left(v_{wjx_v}\right)\ \left|F_{cohesj}\cos\left(\delta_{wj}+\varphi_{z_C}\right)\right| \\ -sign\left(v_{wjy_v}\right)\ \left|F_{cohesj}\sin\left(\delta_{wj}+\varphi_{z_C}\right)\right| \\ 0 \end{Bmatrix}; \tag{16}$$

where: v_{wjx_v}, v_{wjy_v} – projections of vehicle wheel *j* sliding speed vector $\{v_{wj}\}$ to local axes X_v and Y_v of the vehicle respectively (see Fig. 5); δ_{wj} – angle between the local axis X_v of the vehicle and sliding speed vector $\{v_{wj}\}$ of the wheel *j* (see Fig. 5 and expression (12)).

Cohesion moment of cohesive forces from wheel *j*, which influence the vehicle, is equal to:

$$\{M_{cohesj}\} = \{r_{cj}\}\times\{F_{cohesj}\} = \left[r_{cj}\right]\{F_{cohesj}\}; \tag{17}$$

where: $\{r_{cj}\}$ – local vector from the masses centre of the vehicle to wheel *j*; $\left[r_{cj}\right]$ – matrix of the vehicle wheel *j*.

General cohesive force vector, which influence the vehicle, and general cohesive moment, are equal to:

$$\{F^g_{cohes}\} = \sum_{j=1}^{4}\{F_{cohesj}\};\quad \{M^g_{cohes}\} = \sum_{j=1}^{4}\{M_{cohesj}\}. \tag{18}$$

General cohesive force vector $\{F^g_{cohes}\}$, which influence the vehicle, and general cohesive moment $\{M^g_{cohes}\}$, will be included to vector of generalized forces $\{F_v\}$ (see expression (2)), influencing the vehicle.

7. Identification of Dependence of the Bottom Nodal Point of the Vehicle Wheel Interacting with the Road Pavement Surface for Each Road Finite Element

As we have mentioned in Chapter 2, the road section under investigation is divided into triangular finite elements. On each finite element nodal point, a certain height of the road pavement surface roughness and cohesion coefficients of road pavement surface and vehicle wheels in the longitudinal and transverse directions of the vehicle wheels are selected. Then the mentioned parameters can be calculated in each nodal point x, y of finite element by applying the method of finite elements [2, 9, 10, 11].

To reduce the number of calculations, it shall be estimated to which finite element each of the vehicle wheels belongs.

It is only on these estimated four nodal points that the height of road pavement surface roughness as well as cohesion coefficients of the vehicle each wheel and road pavement surface are calculated (as we have mentioned in Chapter 5).

The position of a vehicle wheel on the road shall be expressed by the nodal point with coordinates x_{wj}, y_{wj}. Triangular finite element containing this nodal point is estimated.

Nodal point x_{wj}, y_{wj} belongs to triangle ife finite element with nodal points 1, 2, 3 provided condition:

$$\left|S^{ife}_{1,2,3}\right| = \left|S^{ife}_{1,2,wj}\right| + \left|S^{ife}_{2,3,wj}\right| + \left|S^{ife}_{3,1,wj}\right| \text{ is met;} \qquad (19)$$

where: $S^{ife}_{1,2,3}$ – area of ife triangle finite element; $S^{ife}_{1,2,wj}$, $S^{ife}_{2,3,wj}$, $S^{ife}_{3,1,wj}$ – areas of auxiliary triangles (Fig. 7) [2, 9].

If condition (41) is not met, nodal point x_{wj}, y_{wj} does not belong to ife triangle finite element.

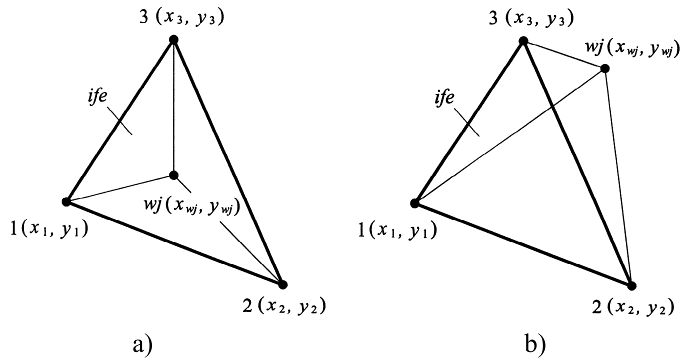
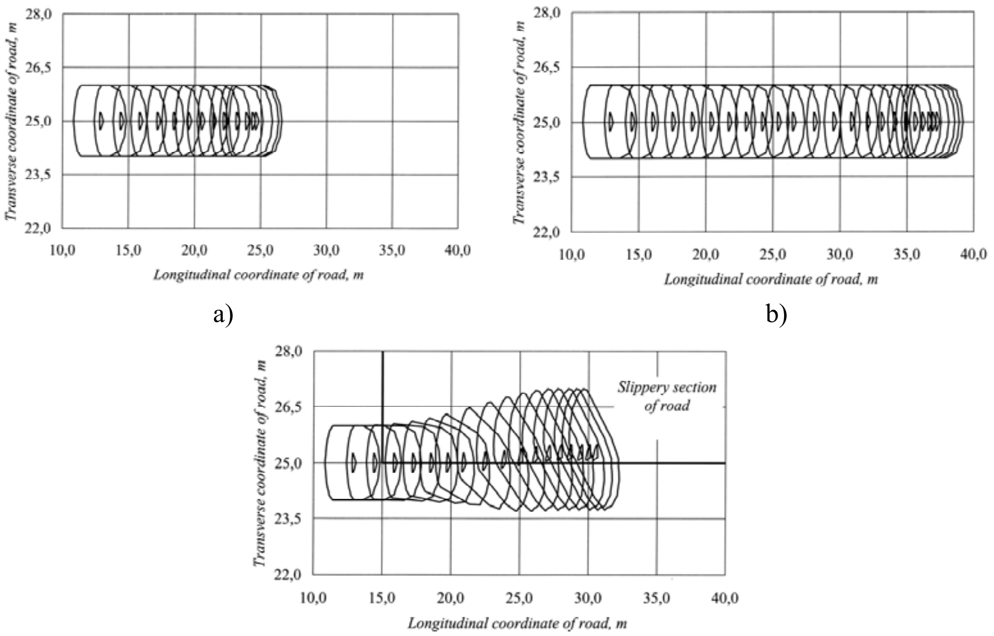


Figure 7. *The identification circuit of nodal point $wj(x_{wj},\ y_{wj})$ dependence on ife triangle finite element: a – nodal point belongs to a finite element; b – nodal point does not belong to a finite element*

8. Computer Aided Test Results

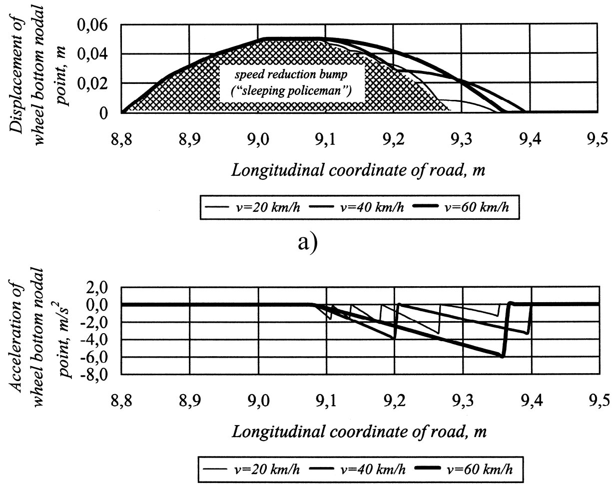
To solve the dynamic model of the system "vehicle – road", the following results are obtained during the computer experiment when using application packages *Compaq Visual Fortran Professional v 6.1* [13] and *Waterloo Maple 7.0* [14]:

- vehicle movement trajectories when braking at various initial conditions and a various certain pavement surface of the road section under investigation (Fig. 8);
- dependences of vertical displacement and acceleration of vehicle wheel bottom nodal point on the longitudinal coordinate of the road, when the vehicle drives on the speed reduction bump ("sleeping policeman") (Fig. 9);
- dependences of normal reactions, which influence road pavement surface on the longitudinal coordinate of the road, when a vehicle drives on the speed reduction bump ("sleeping policeman") (Fig. 10).



c)

Figure 8. *Vehicle movement trajectories (view from top), the initial driving speed is 60 km/h: a – road pavement is dry asphalt, braking all wheels; b – road pavement is dry asphalt, braking two front or rear wheels; c – when there is a slippery section on the road, braking all wheels*



b)

Figure 9. *Dependences of the vehicle wheel bottom nodal point vertical displacement and acceleration on the longitudinal coordinate of the road, when the vehicle drives on the speed reduction bump (“sleeping policeman”): a – displacement; b – acceleration*

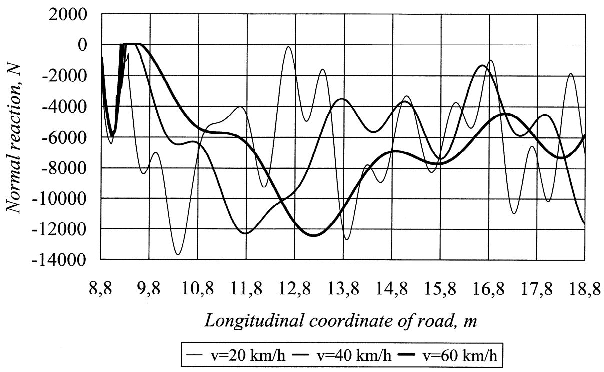


Figure 10. *Dependence of the vehicle wheel normal reaction, which influence road pavement surface, on the longitudinal coordinate of road, when the vehicle drives on the speed reduction bump (“sleeping policeman”)*

9. Conclusions

1. Mathematical model of the system "vehicle – road" is presented. It consists of the dynamic model of the vehicle and the model of the road section under investigation.

2. The vehicle is simulated by concentrated masses interconnected by elastic and dissipative links (Kelvin-Foight elements). The presented model evaluates the movement of the vehicle body in space; movement and turning of front and rear suspensions with respect to the body; interaction of the wheel with the road pavement surface; blocking of the wheel; changing cohesive forces which influence the vehicle.

3. The investigated road pavement surface is simulated by triangular finite elements, a certain height of road pavement surface roughness and cohesion coefficients of road pavement surface and the vehicle wheel in the longitudinal and transverse directions of the wheel are selected in each finite element nodal point.

4. Graphs in Fig 8 a-b, illustrating the braking of the vehicle on the road, show that the less number of wheels of the vehicle brake (e.g., two or four), the longer the vehicle braking way is. Graph in Fig 8 c, illustrating braking of the vehicle on the road, shows that if there is a slippery section of road and all wheels of the vehicle are braked, the vehicle starts turning around its vertical axis passing through the its masses centre, when two wheels of the vehicle drive on the slippery section of road.

5. Graphs in Fig 9-10, illustrating the movement of the vehicle on the speed reduction bump ("sleeping policeman"), show than is less initial driving speed of vehicle, the wheel will jump less on the speed reduction bump ("sleeping policeman"), and the weaker the wheel normal reaction as well as loading to the bearing part will be. In other words, the speed shall be reduced when the vehicle driving on the speed reduction bump ("sleeping policeman").

6. The presented mathematical model of the system "vehicle – road", as one part, can be used, when simulating certain traffic scenarios (for example, to study the existing or to design new dangerous road sections; to study the interaction of the vehicle with various obstacles, etc).

References

1. O. Lukoševičienė, The accidents analysis and simulation. Monograph (Autoįvykių analizė ir modeliavimas. Monografija) (Vilnius: Technika, 2001, 244 p.) – in Lithuanian.
2. O. Prentkovskis, Interaction between the Vehicle and Obstacles: Summary of Doctoral Dissertation, Technological Sciences, Transports Engineering (Vilnius: Technika, 2000, 56 p.).
3. A. A. Chachaturov et al, Dynamics of the System Road – Tyre – Automobile – Driver (А. А. Хачатуров и др. Динамика системы дорога – шина – автомобиль – водитель) (Moscow: Mashinostrojenije, 1975. 535 p.) – in Russian.
4. A. Rutka, J. Sapragonas, The Role of a Tire in Vehicle and Road Interaction, Transport, XVII, 2, 2002, 39-45. Vilnius: Technika.
5. S. J. DiMaggio, M. P. Bieniek, Vehicle dynamics using a limit surface treatment of the tyre – road interface, Proceeding Instn Mech Engrs, 212, D, 1998, 347-356.
6. J. Ambrósio, An Integrated Finite Element and Multibody Method for Frontal Oblique Vehicle Impact Analysis. European Conference on Computational Mechanics "ECCM – 2001", June 26 – 29, 2001, Cracow, Poland, (CD electronic form).
7. T. D. Williams, A. de Pennington, D. C. Barton, The frontal Impact response of a Spaceframe Chassis Sportscar, Proceeding Instn Mech Engrs, 214, D, 2000, 865-873.
8. J. Giergiel, M. Śmieszek, Modeling of Automated Guided Vehicle Movement Along Curvilinear Trajectory, Collection of Scientific Works of Department of Applied

Mechanics (Modelowanie ruchu automatycznie kierowanego pojazdu transportsowego na zakręcie. Zeszyty naukowe katedry mechaniki stosowanej), 15, 2001, 101-106. Gliwice: Wydawnictwo Katedry Mechaniki Stosovwanej – in Polish.

9. O. Prentkovskis, M. Bogdevicius, Dynamics of a Motor Vehicle Talking into Consideration the Interaction of Wheels and Road Pavement Surface, Transport, XVII, 6, 2002, 244-253. Vilnius: Technika.
10. R. Barauskas, Fundamentals of a Finite Element Method (Baigtinių elementų metodo pagrindai) (Kaunas: Technologija, 1998, 376 p.) – in Lithuanian.
11. M. A**.** Crisfield, Non-linear Finite Element Analysis of Solid and Structures, 1-2 (John Wiley & Sons, 1997).
12. G. Korn, T. Korn, Reference Book on Mathematics for Scientific Workers and Engineers. (Справочник по математике для научных работников и инженеров) (Moscow: Nauka, 1968. 720 p.) – in Russian.
13. W. Schick, G Silverman, Fortran 90 and Engineering Computation (John Wiley & Sons, 1995, 425 p.).
14. V. Aladjev, M. Bogdevicius, O. Prentkovskis, New Software for Mathematical Package Maple of Releases 6, 7 and 8. Monograph (Vilnius: Technika, 2002. 404 p.).

EXPLOITATIONAL RESEARCH OF RELIABILITY OF THE FRONT SUSPENSION OF „KAROSA" BUSES

Kęstutis Lukoševičius, Egidijus Diržys, Nijolė Lukoševičienė

Department of Automobile Transport, VGTU
Plytinės 27, LT–2040, Vilnius, Lithuania
Tel. +370 5 2744784, E-mail: vaida@ti.vtu.lt

Abstract

This paper presents the reliability analysis of front suspension of „Karosa" buses for working conditions existing at the PLC „Vilniaus autobusai". The weakest units of the front suspension are being looked for, laws of their fault frequencies and their reasons are being examined and some conclusions and propositions to the fleet of buses about the repair and exploitation conditions are made.

The paper presents a brief description of the front suspension of „Karosa" buses, statistically processed distribution according to probabilistic distributions of the most frequent unit faults and prognosis of their average working resource. The main reasons of such faults are also presented.

Also there was established that working conditions of front suspension of „Karosa" buses are hard and therefore its reliability is rather low. The average resource of some units of the suspension equals to 31- 34 thousand kilometres only. The structure of the independent suspension of this type is complicated and a continuous maintenance is necessary.

It is proposed to Vilnius fleet of buses to organize a technical service of „Karosa" buses every 30- 35 thousand kilometres.

Key words: Suspension, reliability, average resource, fault

Introduction

Efficiency of an enterprise of public transport is determined mainly by economical and ecological characteristics of its fleet of buses, by safety and comfortability of passenger transportation. Comfortability also includes the regularity of communication.

Such results can be achieved only by the enterprise having and expediently introducing a program for retaining reliability indications in the fleet of automobiles on the necessary level and so constantly improving its technical exploitation.

For this purpose there must be analysed data from test bench or exploitational test or statistical data gathered during exploitational in enterprises. Test bench researches often are being performed in industrial companies, producing such vehicles [1], [2]. Abroad for this purpose specialised centres and laboratories are built, which are properly supplied with modern equipment [3]. But data obtained from test bench research are perhaps of highest value for the automobile design. Data of exploitational tests are more useful for improvement of automobile exploitational characteristics [4], [5].

But for this purpose at first it is necessary to gather, systemize and analyze information about automobile faults also with data about technical service, repair and exploitation conditions in the given enterprise and after - to prepare recommendations.

This paper deals with one problem of the named ones - with reliability of the front suspension of „Karosa" buses, reasons of its faults and recommendations or their prevention.

Reliability of the front suspension of the bus, which is designed for passenger transportation in urban areas, is of highest importance in the sense of comfortability, safety and also for retaining the technical state of the automobile (lower overload probability of units and parts). This shows the urgency of the problem.

This work investigates the whole complex of structural peculiarities and mechanics of the front suspension of „Karosa" buses, statistical data of faults, suggestions on their reasons and models of statistical calculations of reliability.

Structure of the Front Suspension

The front axle with an independent suspension consists of the left and right parts. It is of trapezium type. This trapezium by help of linkage ball joints and by dismountable holders is fixed to the chassis frame.

In pneumatic suspensions pneumatic balloons are used as suspension springs (with compressed springy air inside). The proper pressure inside such balloons is regulated by the valve of body height regulation. This valve is tied with the movable parts of body by special rods and regulates pressure inside the balloons so, that body height is retained constant independently of the bus loading.

The balloons are made of a special oil- and aging- resistant rubber with a strengthening framework inside. This framework increases the balloon rigidity and makes the balloons resistant to high loadings and pressures.

In the front suspension of „Karosa" buses there are mounted two telescope-type hydraulic two- directional shock-absorbers for suppression of undesirable body oscillations.

Reliability Calculation for Suspension Units, Based on Statistical Models

As source of statistical data related to the front suspension faults of „Karosa" buses there were used orders for repair, registration of which in the enterprise is sufficiently reliable. Repair orders received during 8 months for a group of 104 buses were analyzed by help of statistical methods. The most frequent there were faults of the following units: silent blocks, shock-absorbers, hub bearings, body height regulator valve, sleeve of vertical swivel axle pin, hub seal rings, sleeves of horizontal cams, suspension air bags.

The results were used for creation of fault frequency variational series (Table 1).

Table 1. Variational series

Intervals, km / Units	0-8000	8000-16000	16000-24000	24000-32000	32000-40000	40000-48000	48000-56000	56000-64000
Silent-bloc	9	7	3	22	30	16	17	8
Shock-absorber	13	19	15	27	26	20	18	11
Hub bearing	11	12	45	11	8	25	18	6
Body height regulator valve	6	14	10	20	18	8	12	10
Sleeve of vertical swivel axle pin	22	31	25	39	33	36	20	14
Hub seal ring	5	12	8	5	10	9	1	4
Horizontal sleeve of cam	51	45	42	61	73	64	45	47
Air bag	0	1	2	5	12	6	3	0

The variational series shows the size of every interval and amounts m_i of values of random quantities which belong to a set of N variables and fall into i-th interval [6]. In accordance with data from fleet of buses, the average run of an express bus equals to 8,000 km per month, so 64,000 km per 8 months. The average bus 8 month run is divided into 8 intervals. The variational series is formed from the existing data. In the series there is inscribed the fault number m_i for every month.

From the data of the variational series the mean mathematical value of bus run $\bar{T}_{vid}$ can be calculated as follows:

$$\bar{T}_{vid} = \sum_{i=1}^{n} T_{i,vid} \cdot \frac{m_i}{N}; \tag{1}$$

where $T_{i,vid}$ – run value, corresponding to medium of the *i*-th interval (see Table 1);

n – number of intervals in the statistical series (see Table 1).

The run deviation as compared with $\bar{T}_{vid}$ is characterized by the mean square deviation σ, which can be calculated in accordance with the formula:

$$\sigma = \sqrt{\sum_{i=1}^{n} \left(T_{i,vid} - \bar{T}_{vid}\right)^2 \cdot \frac{m_i}{N}}. \tag{2}$$

Dispersion of a random quantity is characterized by the coefficient of variation ν. It can be calculated by evaluation of the initial shift of information:

$$\nu = \frac{\sigma}{\bar{T}_{vid} - T_p}; \tag{3}$$

where T_p – initial shift of information, equal to the beginning of the first interval.

All mean mathematical values, mean square deviations and coefficients of variation of resource of all units are presented in the Table 2.

Table 2. Mean mathematical value, mean square deviation and coefficients of variation of the units resource

Values / Units	Mean mathematical value $\bar{T}_{vid}$, thous. km	Mean square deviation σ, thous. km	Coefficient of variation ν
Silent-bloc	35.2143	15.0413	0.4271
Shock-absorber	31.6455	14.9353	0.4720
Hub bearing	30.0000	16.0807	0.5360
Body height regulator valve	32.4082	16.3282	0.5038
Sleeve of vertical swivel axle pin	30.4727	16.2434	0.5330
Hub seal ring	28.0000	16.1475	0.5767
Horizontal sleeve of cam	32.4673	17.4379	0.5371
Air bag	36.0000	9.3955	0.2610

Usually the initial information contains some data of a limited sample about the variation of the random quantity. Therefore the distribution series reflects some real phenomena with a certain error. Those errors can be eliminated in the process of evaluation of reliability indications by use of some distribution laws, which characterize the relation between values of a random quantity and their probabilities [7]. The theoretical law of a distribution is known if theoretical differential and integral distribution functions are computable.

In such case some theoretical distribution law can be chosen according to the value of the coefficient of variation:

for $v \le 0,33$ – normal law;

$v > 0,33$ – Weibull law.

The calculated values of coefficient of variation shows that values of the random quantity or fault frequencies of the first seven units, listed in the tables are distributed in accordance with the Weibull distribution law. Values of the random quantity of the air distributed in accordance with normal law.

The differential function of the Weibull distribution law or its probability density is calculated using the formula:

$$f(T) = \frac{b}{a}\left(\frac{T}{a}\right)^{b-1} \cdot e^{-\left(\frac{T}{a}\right)^{b}}. \qquad (4)$$

The integral function is expressed by the equation:

$$F(T) = 1 - e^{-\left(\frac{T}{a}\right)^{b}}; \qquad (5)$$

where *a* and *b*– are distribution parameters, determined by the coefficient of variation v.

The differential function of the normal distribution law or its probability density is expressed by the equation:

$$f(T) = \frac{A}{\sigma} \cdot f_0\left(\frac{T_{vi} - \overline{T}_{vid}}{\sigma}\right); \qquad (6)$$

where T_{vi} – the average value of the random quantity in the given interval *A*;

$f_0(T)$ – value, found from the tables for probability density function calculation for the normal distribution law [8].

The integral function or distribution function *F(T)* is found by integration of the differential function *f(T)* and by making some changes:

$$F\left(T_i^v\right) = F_0 \cdot \left(\frac{T_i^v - \overline{T}_{vid}}{\sigma}\right); \qquad (7)$$

where T_i^v – the value of the end (the highest value) of the statistical series interval.

Them the fault intensity is calculated according to formula:

$$\lambda(T) = \frac{f(T)}{1 - F(T)}; \qquad (8)$$

where *f(T)* – the differential function of the distribution law;

F(T) – the integral function of the distribution law.

Results of reliability calculations of the suspension units are presented in the Table 3.

Table 3. Results of suspension units reliability calculations

Intervals, km		0-8000	8000-16000	16000-24000	24000-32000	32000-40000	40000-48000	48000-56000	56000-64000
Silent-bloc	*f(T)*	0.0029	0.0105	0.0185	0.0242	0.0248	0.0199	0.0126	0.0065
	F(T)	0.0383	0.0999	0.2508	0.4469	0.6436	0.8026	0.9081	0.9642
	λ(T)	0.0031	0.0116	0.0247	0.0437	0.0695	0.1009	0.1373	0.1815
Shock-absorber	*f(T)*	0.0047	0.0145	0.0226	0.0261	0.0235	0.0165	0.0093	0.0046
	F(T)	0.0444	0.1500	0.3315	0.5400	0.7260	0.8587	0.9373	0.9755
	λ(T)	0.0049	0.0170	0.0338	0.0567	0.0856	0.1167	0.1492	0.1870
Hub bearing	*f(T)*	0.0069	0.0184	0.0246	0.0246	0.0202	0.0142	0.0086	0.0046
	F(T)	0.0562	0.2013	0.3948	0.5898	0.7515	0.8652	0.9341	0.9717
	λ(T)	0.0073	0.0230	0.0406	0.0599	0.0814	0.1052	0.1306	0.1612
Body height regulator valve	*f(T)*	0.0054	0.0153	0.0219	0.0239	0.0214	0.0160	0.01021	0.0058
	F(T)	0.0462	0.1646	0.3389	0.5294	0.7001	0.8277	0.9111	0.9580
	λ(T)	0.0057	0.0183	0.0331	0.0508	0.0713	0.0929	0.1153	0.1381
Sleeve of vertical swivel axle pin	*f(T)*	0.0067	0.0179	0.0240	0.0243	0.0203	0.0145	0.0090	0.0049
	F(T)	0.0545	0.1959	0.3853	0.5784	0.7402	0.8567	0.9286	0.9681
	λ(T)	0.0071	0.0222	0.0391	0.0577	0.0781	0.1011	0.1254	0.1529
Hub seal ring	*f(T)*	0.0107	0.0221	0.0255	0.0231	0.0178	0.0120	0.0073	0.0039
	F(T)	0.0840	0.2577	0.4595	0.6432	0.7850	0.8812	0.9401	0.9718
	λ(T)	0.0116	0.0298	0.0471	0.0647	0.0827	0.1009	0.1211	0.1398
Horizontal sleeve of cam	*f(T)*	0.0069	0.0186	0.0256	0.0270	0.0237	0.0181	0.0121	0.0072
	F(T)	0.0478	0.1750	0.3490	0.5332	0.6961	0.8193	0.9028	0.9518
	λ(T)	0.0072	0.0226	0.0393	0.0578	0.0782	0.0999	0.1240	0.1486
Air bag	*f(T)*	0.0034	0.0128	0.0800	0.2367	0.3397	0.2367	0.0800	0.0128
	F(T)	0.0010	0.0170	0.1010	0.3350	0.6650	0.8990	0.9830	0.9990
	λ(T)	0.0034	0.0130	0.0890	0.3560	1.0141	2.3437	4.7081	12.7721

Calculation of the Predicted Mean Run-to-fault for the Suspension Units

The mean run-to-fault or the mean resource is the total mean run of unit from the beginning of its operation or renewal after repair up to its transition to some limiting state. For predictive calculations of interval resources of suspension units the introduction of a correction coefficient k is necessary. This coefficient can be calculated based on the ratio of the total number of certain units of the same type of „Karosa" buses to the number of damaged ones during the considered time interval in accordance with the formula:

$$k = \frac{M}{N}; \tag{9}$$

where M – the total number of certain units of the same type in the fleet of buses;

N – the number of damaged units of this type during the time under consideration (8 months).

Then the predicted mean resource of the suspension units can be calculated accordingly to formula:

$$\overline{T}_{T,vid} = \overline{T}_{vid} \cdot k. \tag{10}$$

The results of the predicted mean run-to-fault for the suspension units are presented in the Table 4.

Table 4. Results of calculations of the predicted mean run-to-fault for the suspension units

Unit / Item	Silent-bloc	Shock-absor-ber	Hub bearing	Body height regulator valve	Sleeve of vertical swivel axle pin	Hub seal ring	Horizon-tal sleeve of cam	Air bag
k	3.7143	2.7919	3.0588	1.0612	1.8909	3.8519	0.9720	3.6207
$\overline{T}_{P.vid}$ $\cdot 10^3$ km	130.796	88.3524	91.7647	34.3923	57.6212	107.851	31.5570	130.344

Research of the Fault Nature and Reasons

After processing the data about the suspension units by statistical methods there were found that reliability of the front independent suspension is comparatively law. The weakest suspension units from the reliability point of view are the horizontal sleeves of cam. They are made of the polyamide plastic. The plastic sleeve is impressed into a metallic ring. Then the sleeve is put on the vertical swivel axle pin and from the outer side the metallic ring is pressed around by the iron ring of the reactive rod. When the bus runs, this unit is moving up and down continually, it joins the moving metallic parts of the suspension and so saves them from using up. But working in such hard conditions they serve about 31.6 thousand kilometres only. Since the bus runs about 8 thousand kilometres per month, they must be changed every 4 months.

Another unit which is among the most frequently damaged ones is the body height regulator valve. It retains some constant body height independently of the bus loading. When the height regulator of the front axle is damaged, the front height of the bus body remains unregulated. The most frequent result of this is that the bus front remains constantly law or high. This reduces comfortability; bus steering and control became worse. During exploitation of such bus there increase using up of tyres and air bags of the pneumatic suspension.

There are met two kinds of the height regulator valve damages. The most frequent damage is sticking of the regulator roller in the sleeve, because of rust or aggressive surroundings reasons. When it sticks, the operating lever can not turn it and operation of the valve fails. The bus frontal heights remain permanent and equal to the one before the valve sticking.

Another but not so frequent damage is loss of valve sealing. In time valve rubber parts such as seal rings, gaskets because of aging become hard, cracked and unfit for use. Therefore the air leak to other canals or to environment begins.

The predicted run resource of the valves of body height regulators is 34.4 thousand kilometres. It means that the valve of the every normally exploited „Karosa" bus has 3-4 faults per year. New valves are bought sufficiently seldom in the fleet of buses except perhaps they have mechanical defects. Most frequently they are cleaned off rust and other sewage and the defected rubber parts are changed or repaired. Repair of the valve is neither expensive nor time consuming, but its reliability is very poor.

Now in PLC „Vilniaus autobusai" there is being performed the seasonal technical service during of which the front suspension is being diagnosed. This system not guarantee the good working order of bus, therefore there exists certain risk with passenger safety. Thanks to so poor internal resources of the body height regulator valve and of the horizontal sleeves of cam it is proposed to organize a technical service of the front suspension for „Karosa" buses every 30 - 35 thousand kilometres. Such system of technical service would increase the bus safety and would prevent sudden and unexpected damages in the bus suspension. Such planning of repair strategy also enables better organization of bus exploitation, better prediction needs of spare parts, manpower and better economy of time [9].

Conclusions

1. By application of statistical methods to the data about damages in the suspension units it was found, that reliability of the bus front suspension is rather law. The weakest suspension units in the sense of reliability are horizontal sleeves of cam. Their average resource is only 31.6 thousand kilometres. Another weakest unit is the valve of body height regulator. Its average resource is 34.4 thousand kilometres.

2. Plastic horizontal sleeves of can are working in very hard conditions. They in average must be changed 3–4 times per year. The reason of such low their resource is their bad design, bad sleeve material. Metallic sleeves would work longer time in this unit. But production of plastic sleeves is cheaper. For this reason the fleet of buses refused idea of improvement by use of metallic sleeves.

3. The valve of body height regulator breaks 3–4 times per year. The most frequent fault is sticking of the regulator roller in the sleeve. It sticks because of corrosion, which appears because of bad sealing of the valve.

4. Because of law average resource of some suspension units the technical service of the front suspension of „Karosa" buses is proposed to be carried out every 30–35 thousand kilometres. Such system of technical service would increase bus safety and would prevent sudden and unexpected damages in the bus suspension. The proposed repair strategy would enable to improve the bus exploitation, predict needs of spare parts, manpower and save time.

References

1. J. M. Mortimer, Getriebe- Prüffeld bei Ricardo für Zwei- und Allradantriebe, *Autom. Eng. Partners*, 5, 2001, 24-26.
2. J. H. Marrill, J. Ritter, P. Onesti, Automated for the people, *Test. Technol. Int*., V, 2001, 60-61.
3. Moderner Klimawindkanal eingeweiht, *ATZ: Automobiltechn. Zeit*., 103, Sonderansg., 2001, 7.
4. E. Ezbl, Glut im Bauch, *Aut. Mot. Und Sport*., 4, 2002, 36-38.
5. А. Н. Москалюк, Ж. Ю. Моисеенко (А. О.„ Автоваз "), Управление надежностью автомобиля (прямая и обратная связъ этапов проектирования и эксплоатации). *Привод. техн*., 4, 2001, 27-33.
6. H. Medekšas, *Gaminių kokybė ir patikimumas* (Kaunas: Smaltija, 2003).
7. L. Jonaitis, R. Žeromskas, *Mašinų patikimumas* (Vilnius: Mokslas, 1988).
8. R. Žeromskas, *Mašinų patikimumas* (Vilnius: Technika,1997).
9. L. Jonaitis, *Mašinų servisas* (Kaunas: Smaltija, 1998).

THE FLEET OF MOTOR VEHICLES IN LITHUANIA AND THE ACTIONS ON ITS RENOVATION

A.Pikūnas, V.Valiūnas

Department of Automobile transport, Vilnius Gediminas technical university,
J. Basanaviciaus str. 28, Vilnius, LT-2009 Lithuania
Ph / Fax.: (+370)-5212 5551. E-mail: vaida@ti.vtu.lt

Abstract

On the basis of collected statistical data the following items were investigated the structure of Lithuanian motor fleet, the influence of motor vehicles on the environment.

The major part of the work is devoted to the analysis of the necessity of motor fleet renovation in Lithuania. The research has been done on the basis of Lithuanian and foreign literature on the subject and using the experience of foreign countries.

The results of the work are presented in tables and illustrations. The conclusions and proposals are made at the end of the work.

Key words: Programmes of renovation, vehicle fleet, system of utilization

1. The Introduction

Transport is of a vital importance in the life of the society, however, it is one of the main sources of pollution of nature. Its harmful impact upon the environment depends on the quantity and the type of the consumed fuel. Striving to reduce it, we should use not only ecologically pure fuel, but also modern and economic internal combustion engines.

This object may be realized by encouragement of selling new modern vehicles in Republic of Lithuania and regulation of import of used vehicles to our country.

2. The Structure of the Fleet of Motor Vehicles in Lithuania

The number of motor vehicles in Lithuania continually grows. Upon the present rates of growing of the fleet, it is forecasted that the total number of vehicles will exceed 1.5 million in 2005. In 2002, the number of vehicles registered in the country was 1’479’099 (Fig. 1).

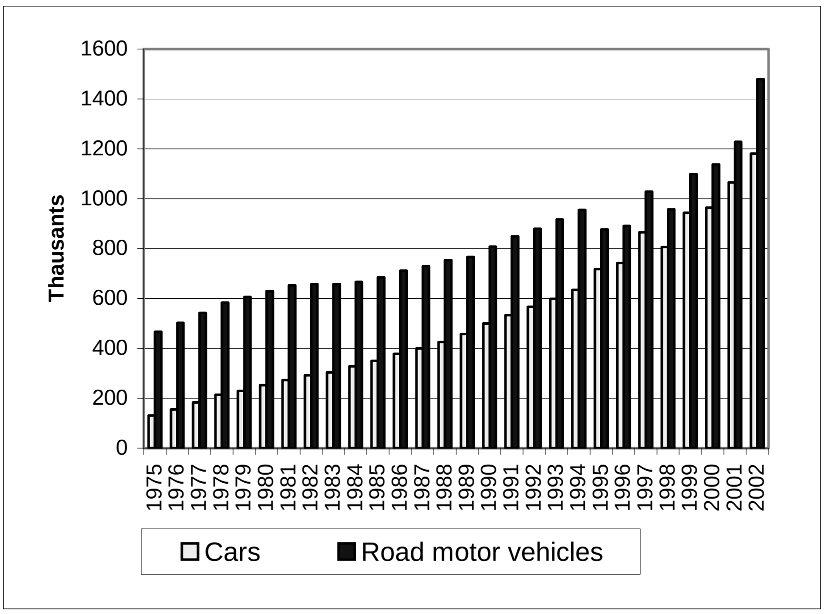


***Figure 1.** Number of motor vehicles and passenger cars 1975-2002*

The number of vehicles grew considerably, however, most vehicles of the fleet were used ones, i.e. 10-15 years aged vehicles. Last year, the fleet of vehicles became some more obsolete. The average age of vehicles presented for a technical checkup in 2002 was 14.9 years and in 2001 – 14.5 years [2]. In 2002, about 86 per cent of cars presented for a technical checkup were used ones aged over 10 years; most of them were manufactured between 1984 and 1992 (Fig. 2).

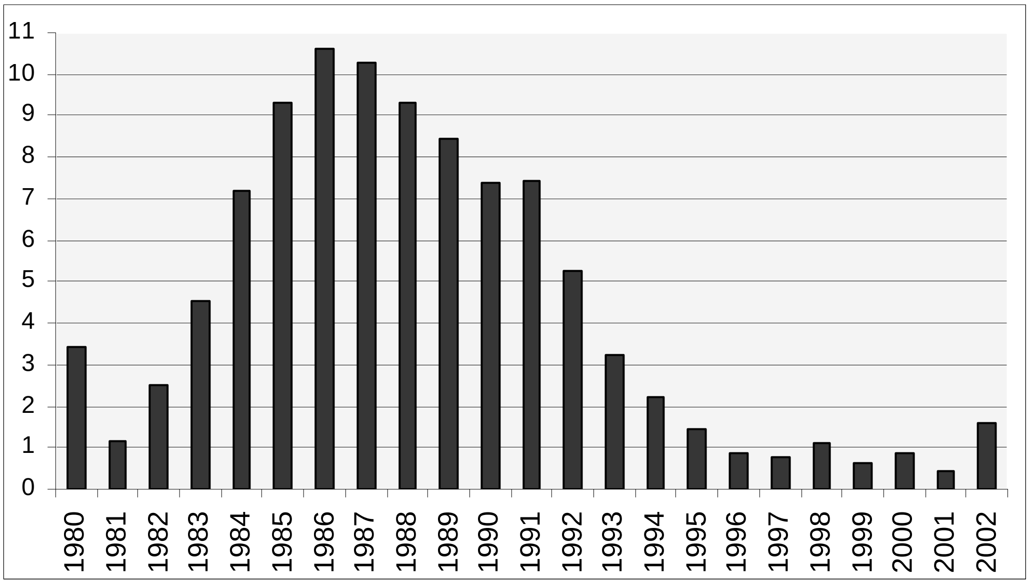


Figure 2. *Number of motor vehicles % by year of manufacture given on technical checkup in 2002*

One of the principal indicators of automobilization is number of cars per one thousand of the population. Within the last decade, this indicator in Lithuania has doubled, however, it is less than in developed European countries (Fig. 3).

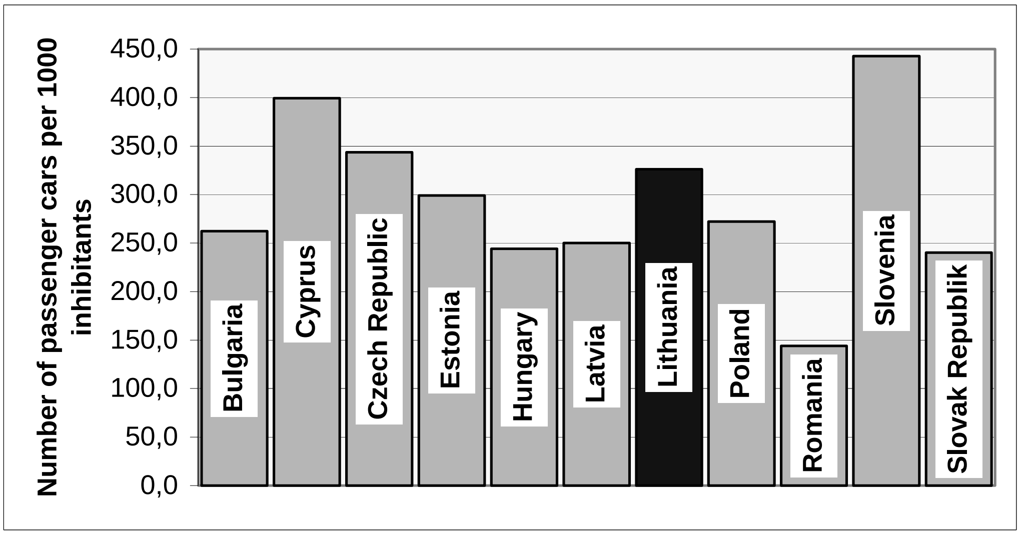


Figure 3. *Level of automobilization in the countries of Europe in 2002*

It may be seen that the number of vehicles per one thousand of the population in our country is less than in developed countries of Europe, however, the relative indicators of accident rate, i.e. the number of deaths on one million of the population or vehicles, in our country are several times worse that analogous indicators in other European countries. This fact attests not only a low level of traffic safety and emergency medical aid to injured persons, but also a poor passive security of the exploited vehicles (most old are not equipped with air cushions, ABS and other security means).

In 2001, according to statistical data, the number of deaths on one million inhabitants in road accidents in Latvia was 221, in Lithuania – 202, in Estonia – 147, in Germany – 84, in Norway – 56.

Vehicles are one of the main sources of pollution of the environment in Lithuania. They discharge over 80% of all pollutants to the environment – about 200 chemical combinations; most of them are dangerous for human health. It was approximately calculated that road vehicles in Lithuania discharge to the atmosphere about 500 thousand tons of pollutes annually. The atmosphere is polluted both by products of combustion exhausted by engines (65% of all pollutants) and gases of engine's carters that are more toxic than the products of combustion as well as evaporating hydrocarbons from engine's feeding system. The products of wear of tyres and brake shoes are also harmful for human health: these small particles mix with air, get into respiratory organs of a human and irritate them [1].

3. The Methods of Renovation of the Fleet of Vehicles

European and other states implement programmes on renovation of their fleets of vehicles. Such programmes were and are implemented in Greece (1991-1993), Hungary (since 1993 up to present time), Denmark (1994-1995), Spain (since 1994 to present time), France (1994-1996), Ireland (1995-1997), Norway (1996) and Italy (1997-1998).

The main objects of the programmes on a renovation of fleets of vehicles are the following:

- to stimulate the national industry of vehicles and the national economy by extension of manufacture and sales of new vehicles;
- to improve safety of vehicles by using newer, more safe vehicles;
- to reduce volumes of pollutants discharged by vehicles.

The objects related environmental protection were attained in all programmes. The economic objects encouraging the national industry of manufacture of vehicles were realized as well. At present, the number of deaths on road accidents became less in most European states, in spite of increase of the annual run. However, it is practically impossible to identify what improvements of vehicles should be attributed to causes of this reduction together with other traffic safety measures realized by governments.

Greece was the first European country where the programme of a renovation of vehicle fleet was implemented. The object of the programme was encouragement of use of vehicles with catalysts and improvement of air quality in towns. The excise tax was reduced by 40-60% as a compensation on purchasing a new vehicle. The additional condition was giving back a vehicle aged over 10 years. Other dues were reduced for each owner who acquired a new vehicle with a catalyst. The programme involved all Greece and was completed in 1993.

In Hungary, an implementation of the renovation programme was started in 1993. In Budapest, the started programme strove for elimination of cars and trucks with two-cycle engines (such as Trabant, Wartburg, Barkas). The owners having chosen any of vehicles sponsored by the Government were awarded with premiums in amount about USD500. The alternative was a coverage of costs of tickets for public transport for a year. The programme was applied all over the country.

In Denmark, an implementation of such programme was started in 1994. The amount of compensation for each replacement of a vehicle aged over 10 years was about USD1'000. The programme was implemented up to 1995. 100'000 vehicle were renovated. It was calculated that discharge of HC and nitrogen oxides reduced about 0.8%.

In Ireland, an implementation of the programme was started in 1995. The amount of compensation for each replacement of a vehicle aged over 10 years was about USD1'600. The programme was implemented up to 1997. 60'000 vehicles were renovated. Most of the replaced vehicles were aged about 10-12 years.

The programmes of renovation of vehicle fleet affect the environment in two expectable ways.

The first of them is positive: they may cause a reduction of pollution of the atmosphere. The second is negative: the average age of vehicles becomes shorter and – in case of repetition of the programmes – consumption of energy and materials increases. Because differences of environmental safety indicators between a part of old vehicles and most new vehicles are considerable, the positive effect is fixed in many implemented programmes.

An implementation of a programme of renovation of vehicle fleet depends on the economical conditions that affect a conduct of vehicle owners as well as on the expenses of the programme. The structure of taxation dependently on ownership and use of vehicles is the most important element of the whole economical part of the project that, in its turn, depends on the number of vehicles in the fleet.

On an implementation of programmes of renovation of fleets of vehicles in EU States Members, it is simultaneously foreseen who will be responsible for an utilization of aged vehicles, i.e. who will be obliged to allot funds for this purpose:

- In Spain, France, Italy, Greece, Germany, Austria, Belgium, Holland and Luxembourg, the manufacturers/exporters are fully responsible for coverage of expenses of utilization of a vehicle unfit for exploitation;
- In Sweden, the manufacturers/exporters are partially responsible for financing of utilization of a vehicle unfit for exploitation;
- In Denmark, a full financial responsibility is assumed by owner of the vehicle through the system of insurance of vehicles (at present, discussions on a full conformity of such order to the requirements of the Directive on Vehicles Unfit for Exploitation take place);
- In other EU States Members, the system of financing is not clear yet.

4. Conclusions

In course of performance of the analysis of the structure of the fleet of motor vehicles in Lithuania and examination of the harmful impact of road transport upon human health and environment, it was found that the main cause of such harmful impact is a poor technical condition of the fleet of vehicles that mostly depends on the age of vehicles (the average age of the vehicles of the fleet is about 15 years). This is confirmed by the experience of foreign states and the references as well.

In order to reduce the harmful impact of road transport upon the environment and society and to ensure traffic safety, it is necessary to renovate the fleet of motor vehicles in Lithuania.

One of the methods of a renovation of the fleet is a regulation of import of used vehicles, however, such regulation should be carried out, taking into account the economical and social conditions of the country. In addition, it should be taken into consideration that vehicles are not produced in Lithuania [1].

It is mentioned in the EU Directive that manufacturers of vehicles shall be obliged to ensure processing of old vehicles by own forces or to conclude agreements with specialized enterprises. Because manufacturers of vehicles never took care about such problems earlier, it is expected that an introduction of the new order will require considerable expenses and force to increase the prices of new vehicles.

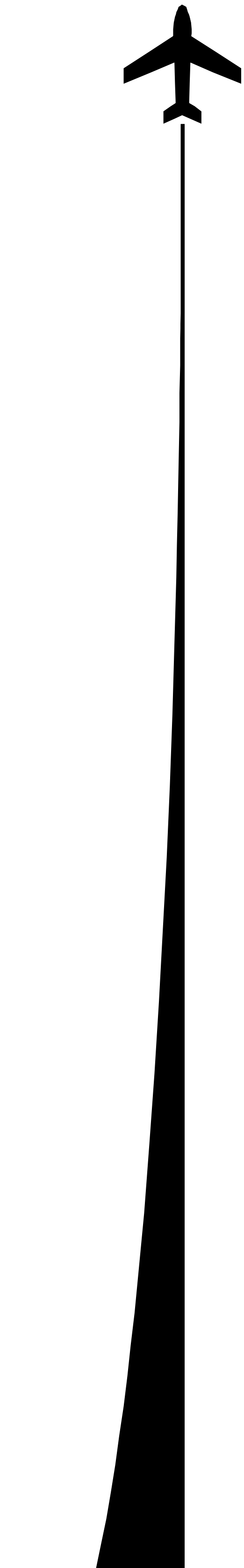
The Ministry of Environment that is officially responsible for a development of the programme of utilization of old vehicles in our country had received an assistance from Brussels. The representatives of the Ministry work in cooperation with the Association of Vehicle Businessmen and some specialists from the Environmental Protection Institute. The principal provision – to try to ensure assuming of the utilization expenses by the manufacturers of vehicles, not by tax payers of Lithuania or representatives of manufacturers employed here. The latter should be responsible for vehicles imported by them, but not for tens or even hundreds thousands of old vehicles that were brought to the country in other ways, not giving any inform to them.

The country needs such economical and legal-legislative policy that would ensure seeking of each vehicle user to acquire a newer vehicle. For a renovation of the fleet of motor vehicles in Lithuania, the following main measures should be taken:

- to improve the system of assessment of conformity of vehicles and their components;
- to differentiate taxes dependently on the age of the vehicles, the existing means and equipment of making the products of combustion harmless in the vehicles and so on;
- to stimulate an acquisition of new and little used vehicles through the system of taxes;
- to improve the methods of assessment of pollutants by assessment of improved construction of new vehicles;
- to ensure control of quality of imported spare parts and exploitation materials;
- to regulate import of used vehicles and their spare parts to Lithuania;
- to develop technical standards for enterprises engaged in vehicle maintenance and to ensure an observance of such standards in order to improve the quality of works;
- to develop a system of utilization of non-exploited vehicles and their details in Lithuania that should encourage to cross off old vehicles from the Register and to acquire new ones.
- Centers of trade with vehicles and their maintenance should extend trade with new and little used vehicles and buy up old vehicles.

References

1. Lietuvos automobilių parko atnaujinimo koncepcija / A. Pikūnas, V. Valiūnas, S. Nagurnas, I. Norvilas. Vilnius, 2001. 57 p.
2. Naudoti lengvieji automobiliai Lietuvoje: techninių trūkumų analizė. UAB “Transporto studijos”, 2001. 219 p.



Session 2

Statistical Applications

PARAMETER ESTIMATION AND HYPOTHESES TESTING FOR NONHOMOGENEOUS POISSON PROCESS: PART 2. NUMERICAL EXAMPLES

I. B. Frenkel[1*], I. B. Gertsbakh[2] and L. V. Khvatskin[1]

[1] Industrial Engineering and Management Department,
Negev Academic College of Engineering
Beer-Sheva, 84100, Israel, E-mail: iliaf@nace.ac.il/ khvat@nace.ac.il
** Corresponding author. Tel.: +972-8-6475676, Fax: +972-8-6475643*
E-mail address: iliaf@nace.ac.il

[2] Department of Mathematics, Ben-Gurion University,
P. O. Box 653, Beer-Sheva, 84105, Israel,
E-mail: eliahu@cs.bgu.ac.il

Abstract

We consider data processing for a Nonhomogeneous Poisson process (NHPP) with Log-linear and Power-law intensity functions. Parameter estimation is carried out by the maximum likelihood method.

For the case of the known intensity function, testing the hypothesis that the given sample path is a realization of NHPP, can be accomplished using the fact, that under the NHPP model the mean value functions of NHPP, computed in sequence of ordered failure times, are the failure times of Homogeneous Poisson Process (HPP) with constant intensity function of one, and the intervals between events in the HPP form a sample of i.i.d. standard exponential random variables. Thus it is possible to use standard goodness-of fit tests to check the exponentiality of the process.

The computer-intensive procedure for testing the hypothesis that the given sample path belongs to NHPP without making the assumption that the intensity function is known was described in our previous article (Frenkel at al. (2003)).

In this article we describe the different goodness-of fit tests and demonstrate our method on the failure data which exist in literature and for our own failure data for the Schlosser Vibration Machine. We also demonstrate the several methods for generating families of stochastic processes with the known probabilistic structure, which includes both NHPP and not NHPP, and check how our method recognizes the underlying process. These processes were used for testing of power properties of goodness-of-fit tests.

Key words: Nonhomogeneous Poisson process, computer-intensive procedure, goodness-of-fit tests, hypotheses testing, comparison of powers of tests

1. Introduction

Nonhomogeneous Poisson Process (NHPP) is used to model reliability of repairable systems. There are many results on estimation and hypotheses testing concerning parameters of the intensity function of NHPP (Ascher and Feingold (1984), Crow (1974), Crowder et al. (1991), Gertsbakh (2000) and others).

Basing on failure data, it is necessary to develop tests that the data actually came from NHPP.

A commonly used procedure, as far as the reliability of a repairable system is concerned, is to distinct between Homogeneous Poisson Process (HPP) and NHPP. Cox and Lewis (1966) suggest the Laplace statistic and show that this test is optimal for testing NHPP with log-linear intensity function. Crowder et al. (1991) recommend the MIL-HDBK-189 (1981) test for testing NHPP with power-law intensity function.

Crow (1984) considers a goodness-of-fit (GOF) test, proving that the data came from NHPP via Cramer-von Mises statistic and obtained a table of critical values for this statistic. Park and Kim (1992) use the Kolmogorov-Smirnov, Cramer-von Mises and Anderson-Darling statistics for GOF test of a power-law NHPP in the case of failure-truncated data. They present more precise tables of critical values for these statistics.

A simple GOF test for power-law NHPP, based on the Duane plot, is introduced by Gaudoin et al. (2003). It is based on graphical analysis of log-log plot of cumulative numbers of failures versus time. The authors claim that when a plot is completed, the coefficient of determination is used as the statistic of GOF test for NHPP. However, Ascher and Feingold (1984) at page 78 note that "plotting on log-log paper will tend to linearize any function. Hence a linear plot on such paper does not necessarily indicate that an NHPP … is the most appropriate model".

Many useful tests for testing of the assumption that the underlying distribution of life is exponential were introduced by Epstein (1960). Among others, a test of exponentiality is based on some results of Hartley (1950). This test deals with the ratio of maximum value to minimum value of the time intervals between failures. Gnedenko at al. (1969) suggest, using the Hartley's test, among othe.

In our previous article (Frenkel at al. (2003) we proposed a computer-intensive procedure for testing the hypothesis that the given sample path belongs to NHPP.

We suggest three kinds of data for demonstrating our method.

1. Using the existing in literature failure data, we compare our conclusions to other authors' results. Generally the results coincide.

2. We present a numerical example of our own data to show how the above described procedure functions.

3. We generate a set of stochastic processes with known probabilistic structure. This set includes NHPP and not NHPP. And we investigate how our method recognizes the underlying process and compare the powers of different tests.

2. Hypothesis Testing

2.1. Testing Hypothesis for NHPP

The procedure for estimation of NHPP intensity function parameters is described in literature (Cox and Lewis (1966)). Two parameterizations are especially convenient for maximum likelihood estimation: the log-linear form with $\lambda(t) = e^{\alpha+\beta t}$ and Weibull or power-law form with $\lambda(t) = \alpha^{\beta}\beta t^{\beta-1}$. The maximum likelihood estimators of (α, β) are obtained as the solutions of the systems of the ML equations (Crowder et al. (1991) pp. 167 and 173).

Testing the hypothesis, that the given process is NHPP with known intensity function $\lambda(t)$ is based on the claim, that the mean value functions of NHPP, counted in sequence of ordered failure times, are the failure times of Homogeneous Poisson Process (HPP) with constant intensity function of one, and the intervals between events in the HPP form a sample of i.i.d. standard exponential random variables (Çinlar (1975), Thompson (1981), Meeker and Escobar (1998)).

Thus it is possible to use standard goodness-of fit tests to check the exponentiality of the process.

The procedure for testing the hypothesis that the given sample path belongs to NHPP without making an assumption that the intensity function is known was described in our previous article (Frenkel at al. (2003)).

We used the following statistics notations for hypothesis testing. All statistic calculations will be explain in details in the next chapter.

S_1 – Kolmogorov-Smirnov distance between the ordered sample of in-between-events intervals $(\hat{\Delta}_1^{(i)}, \hat{\Delta}_2^{(i)}, ..., \hat{\Delta}_n^{(i)})$ in transformed time and the theoretical CDF, assuming that $\Delta_i \sim Exp(1)$;

S_2 – the Laplace statistic, also based on the in-between-events intervals $(\hat{\Delta}_1^{(i)}, \hat{\Delta}_2^{(i)}, ..., \hat{\Delta}_n^{(i)})$ in transformed time (Ascher and Feingold (1984), pp. 78-79, Crowder et al. (1991), p. 169);

S_3 – the coefficient of variation (C.V.) (Gertsbakh (2000), p. 214, Ascher and Feingold (1984), p. 82). C.V. is defined as the ratio between a standard deviation and mean value of random variable.

S_4 – The Levis-Robinson Statistic (Ascher and Feingold (1984), p. 82).

S_5 – MIL-HDBK-Test Statistic (Ascher and Feingold (1984), p. 79, Crowder et al. (1991), p. 172-173, MIL-HDBK-189 (1981)).

S_6 – The Cramer-von Mises Statistic (Park and Kim (1992)).

S_7 – The Anderson-Darling Statistic (Park and Kim (1992)).

S_8 – The Hartley Statistic (Gnedenko at al. (1969), p. 252).

After carrying out the procedure, we will obtain the simulated values of all above statistics. Determine the upper and lower α - critical values for these statistics and intervals $\left[S_i(\alpha), S_i(1-\alpha)\right], i = 1, 2, ..., k.$.

For the given sample compute the values $S_1^*,\ S_2^*, ...,\ S_k^*$ of the statistics $S_1, S_2, \ldots, S_k$. Compare $S_1^*,\ S_2^*, ...,\ S_k^*$ with the upper and lower critical values.

Reject the hypothesis that the given process is the NHPP with intensity function $\hat{\lambda}(t)$, if one of the statistics $S_1^*,\ S_2^*, ...,\ S_k^*$ falls outside of the one of the intervals $\left[S_1(\alpha), S_1(1-\alpha)\right]$, $\left[S_2(\alpha), S_2(1-\alpha)\right], ..., \left[S_k(\alpha), S_k(1-\alpha)\right]$.

2.2. Statistics for Hypothesis Testing

2.2.1. The Laplace Trend Test

The Laplace's trend test is a simple and powerful test for distinguishing between Homogeneous Poisson Process, when a constant rate at which events are occurring, and an increasing rate of occurrence of such event. Laplace's Test is discussed in details in Cox and Lewis (1966), Ascher and Feingold (1984).

Consider a situation where the system is run until m failures have occurred. Under HPP assumption, the first $m-1$ arrival times, designated as T_1, T_2, …, T_{m-1} ,are the order statistics from a uniform distribution on [0, T_m] interval. The Laplace's test statistic is defined as

$$U = \frac{\dfrac{\sum_{i=1}^{m-1} T_i}{m-1} - \dfrac{T_m}{2}}{T_m \cdot \sqrt{\dfrac{1}{12(m-1)}}}.$$

The Laplace test statistic has the following interpretation. Under constant rate of event occurrence, the arrival times to fault will occur randomly around the midpoint of the length, $T_m/2$. Therefore, the sample mean of the T_i's will be approximately equal to $T_m/2$; hence the test statistic U will be small. If events are occurring more frequently towards the end of the interval, however, the sample mean will be large. If U L is larger than the z-value of the standard normal distribution, $z_{\alpha/2}$, there is evidence at a significant level α that the event occurrence indicates the increasing trend.

2.2.2. The Lewis-Robinson Test for Trend

Meeker and Escobar (1998) note that both Laplace's Test and MIL-HDBK-189 Test can give misleading results when the renewal process is not an HPP. The Lewis-Robinson statistic (Lewis and Robinson (1974)) is formed by dividing the Laplace's statistic U by the estimated coefficient of variation of the times between recurrences,

$$U_{LR} = \frac{U}{C\hat{V}}.$$

If the underlying process is an HPP, then U_{LR} is asymptotically equivalent to U since CV=1 if the times between recurrences are exponentially distributed. In large samples, U_{LR} follow approximately a standard normal distribution if the underlying process is a renewal process.

Lawless and Thiagarajah (1966) indicate that the Lewis-Robinson Test is preferable to Laplace's test as a general test of trend in point process data.

The Lewis-Robinson Test was discussed by Ascher and Faingold (1984).

2.2.3. The MIL-HDBK-189 Test

The MIL-HDBK-189 (1981) test is based on the test statistic

$$V = 2\sum_{i=1}^{m-1} \ln\left[\frac{T_m}{T_i}\right].$$

Under the null Hypothesis of an HPP, V is distributed as χ^2 , with 2(m-1) degrees of freedom. Large values of V supply evidence against null hypothesis in favor of reliability growth. Small values of V are indicative of reliability deterioration.

The MIL-HDBK-189 test was discussed by Crowder et al. (1991), Meeker and Escobar (1998) and others.

2.2.4. Cramer-von Mises Test

The failure times, T_i/T_m , i=1, 2, …, m-1 are distributed as order statistics from the uniform distribution over [0,1] (Park and Kim (1992)). Denote $\hat{U}_i = T_i/T_m$. Cramer-von Mises statistic calculates as follows:

$$W^2 = \sum_{i=1}^{n-1}\left[\hat{U}_i - \frac{2i-1}{2(n-1)}\right]^2 + \frac{1}{12(n-1)}.$$

Critical values of this goodness-of-fit statistic are calculated by Crow (1964, 1990). Park and Kim (1992) present more precise table of critical values for this statistic.

2.2.5. Anderson-Darling Test

Anderson-Darling statistic is calculated as follows (Park and Kim, 1992):

$$A^2 = -\frac{\sum_{i=1}^{n-1}(2i-1)\left[\ln\hat{U}_i + \ln(1-\hat{U}_{n-i})\right]}{n-1} - n + 1.$$

Critical values of this goodness-of-fit statistic are calculated by Park and Kim (1992).

2.2.6. Hartley Test

This test of exponentiality is based on some results of Hartley [12]. The test uses ratio of maximum value to minimum value of the time intervals between failures. Gnedenko at al. [11] suggest, among other tests, using the Hartley's test. In our case the Hartley's Test is as follows:

$$h(n) = \frac{\max\{\Delta_i\}}{\min\{\Delta_i\}}.$$

The null hypothesis will be rejected if $h(n) > h_{1-\alpha}(n)$. The critical values for this statistic are represented in Gnedenko at al. (1969) and for big n values (n>12) may be calculated using Monte-Carlo Simulation.

3. Testing the NHPP Hypothesis on Experimental and Simulated Data

3.1. Hypothesis Testing on Literature Failure Data

Let us consider the well-known real-life failure data and compare different authors' conclusions with results, gained via our procedure.

3.1.1. Proschan Data on Air Conditioning System

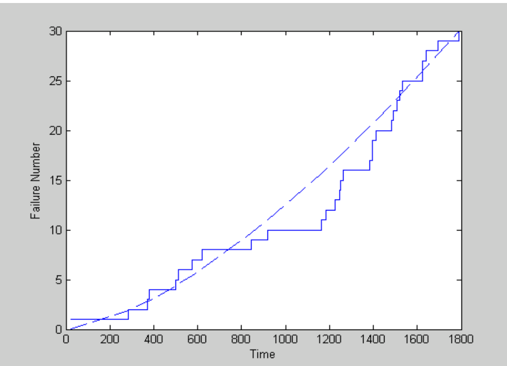


***Figure 1**. The Plot of the step function of the failure numbers and corresponding estimated intensity function against time for the air conditioning system*

Let us illustrate our methodology using data on the time intervals between successive failures of the air conditioning system of the Boeing 720 jet series number 7912 (Proschan (1963)). This data was analyzed by many researchers. Park and Kim (1992) define 3 statistics (Cramer-von Mises, Kolmogorov-Smirnov and Anderson-Darling) for the goodness-of-fit (GOF) test of this process. The authors claim, that failure data came indeed from NHPP with power-law intensity function. Gaudoin et al. (2003) introduce the graphical analysis for GOF test of the power-law process and received similar results. The same result was achived by Muralidharan (2001), Asher and Hansen (1998) and others.

We came to the similar conclusion (Table 1). All test statistic values fall inside the corresponding [0.05, 0.95] simulated intervals for all of our statistics. Therefore, we would claim that the data don't contradict the NHPP with power-law intensity function. On Figure 1 is shown the Plot of the step function of the failure numbers and corresponding estimated intensity function against time for the air conditioning system

Table 1. Comparison of Test Statistics with the Corresponding Simulated Critical Values for the Air Conditioning System

Test Statistic	Test Statistic Values	Corresponding Simulated Intervals	
		0.05	0.95
Kolmogorov-Smirnov Statistic	0.169	0.050	0.171
Laplace Statistic	1.069	-0.350	1.226
CV Statistic	1.140	0.732	1.239
Levis-Robinson Statistic	0.983	-0.352	1.323
MIL-HDBK-Statistic	60.000	59.263	60.600
Cramer-von Mises Statistic	0.140	0.029	0.228
Anderson-Darling Statistic	0.794	0.227	1.350
Hartley Statistic	196.197	34.288	2213.6

3.1.2. USS Halfbeak Engine Failures Data

Crowder et al. (1991) at the page 173 gives the data on failures of an engine of USS Halfbeak. The data were fitted using the Log-Linear and Power-Law intensity functions with the total number of observed events n=71. Using the Laplace test statistic and MIL-HDBK test statistic the authors express doubts that the data set comes from a NHPP.

Our tests reveal the following: Using the Power-Law intensity function four of eight statistics fall *outside* the corresponding [0.05, 0.95] simulated intervals for these statistics (Table 2). Using Log-Linear intensity function all our criteria do not contradict NHPP hypothesis. Our conclusion is that NHPP hypothesis is doubtful. On the Figure 2 is shown the Plot of the step function of the failure numbers and corresponding estimated intensity function against time for the engine of USS Halfbeak

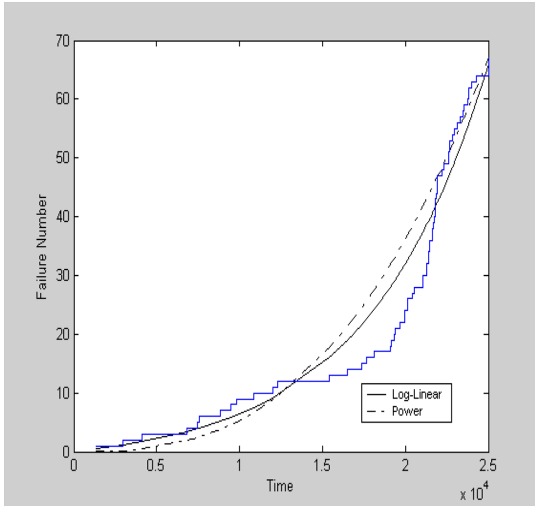


Figure 2. *The Plot of the step function of the failure numbers and corresponding estimated intensity function against time for the engine of USS Halfbeak*

Table 2. Comparison of Test Statistics with the Corresponding Simulated Critical Values for the engine of USS Halfbeak

Test Statistic	Log-Linear Intensity Function			Power-Law Intensity Function		
	Test Statistic Values	Corresponding Simulated Intervals		Test Statistic Values	Corresponding Simulated Intervals	
		0.05	0.95		0.05	0.95
Kolmogorov-Smirnov Statistic	0.150	0.061	0.202	**0.204**	0.055	0.129
Laplace Statistic	0.783	-2.149	2.559	**1.904**	-0.593	1.123
CV Statistic	1.228	0.832	1.312	1.34	0.81	1.149
Levis-Robinson Statistic	0.638	-2.090	2.543	**1.421**	-0.620	1.127
MIL-HDBK-Statistic	135.1	105.8	190.3	142.0	141.57	142.39
Cramer-von Mises Statistic	0.375	0.040	0.949	**0.680**	0.029	0.231
Anderson-Darling Statistic	1.808	0.301	4.913	**3.222**	0.225	1.373
Hartley Statistic	1130.5	113.2	5734.3	1463.4	105.9	6909.0

3.1.4. Schlosser Vibration Machine Failures Data

The data presented in Table 2 summarize the time-intervals in operating hours between failures of Schlosser Vibration Machine collected from the operation reports from 1999 to 2002 at the Yeroham Construction Materials Facility (Israel). The machine was observed to be operated for 16309 hours and 27 failures were identified.

Table 3. Schlosser Vibration Machine Failure Data

240	2352	888	88	456
4032	168	768	268	24
288	480	336	84	120
1224	1400	528	86	
624	408	72	96	
552	528	96	103	

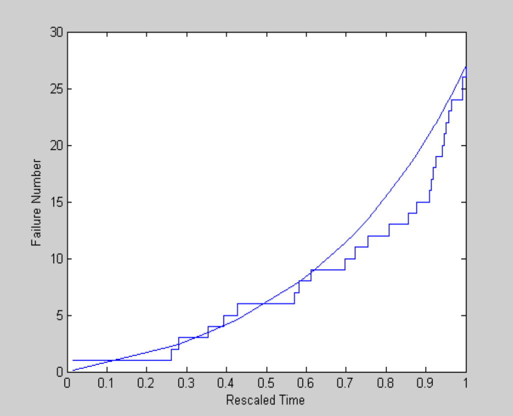


Figure 3. *The Plot of the step function of the failure numbers and corresponding estimated intensity function against rescaled time for the Schlosser Vibration Machine*

The estimated intensity function is assumed to be log-linear. Figure 3 shows the step function of the failure numbers against time in operation hours and corresponding estimated intensity function. We rescaled the failure time by a factor 16309. Hence α=1.7992 and β=2.4979.

To gain this failure data we used our method. According to Table 4 all test statistic values fall inside the corresponding simulated intervals. Therefore, we would claim that the data belong from NHPP with log-linear intensity function.

Table 4. Comparison of Test Statistics with Corresponding Simulated Intervals for the Schlosser Vibration Machine

Test Statistic	Test Statistic Values	Corresponding Simulated Intervals	
		0.05	0.95
Kolmogorov-Smirnov Test Statistic	0.075	0.070	0.507
Laplace Test Statistic	1.614	-2.180	3.002
CV Statistic	0.870	0.729	1.384
Levis-Robinson Test Statistic	1.854	-1.878	2.890
MIL-HDBK-Test Statistic	50.389	31.398	89.614
Cramer-von Mises Test Statistic	0.175	0.042	1.055
Anderson-Darling Test Statistic	0.984	0.320	5.555
Hartley Test Statistic	33.632	31.200	2221.4

3.3. Testing the Data with Known Structure

So far, we applied our testing procedure to the realizations of random processes with unknown structure. In some cases we may only suppose, using some physical considerations that the random path we have analyzed comes from NHPP. In order to show how our testing procedure may distinguishes between NHPP and not NHPP, we must simulate random processes with <u>known</u> probabilistic structure.

We suggest three methods to generate the random processes with known probabilistic structure. The first method is to fill up the random process with elements of two different known structures: exponential random variables and slightly disturbed constant value according the probability. The second method is to form the elements of random process as "mixture" of an exponential random variables and a constant value. And the third one is to slightly disturb the NHPP using non-homogeneous deterministic process.

3.3.1. The first method

Consider two urns with random numbers Ξ^1 and Ξ^2. The first urn consists of random numbers $\xi^1 \sim Exp(1)$. The second urn consists of the random numbers $\xi^2 \sim U(1-\varepsilon, 1+\varepsilon)$, where $\varepsilon = 0.01$ for example.

Make the following procedure.

1. Set the choice probability $\alpha = 0$.

2. Generate the random number $\beta \sim U(0,1)$.

3. If $\beta < 1-\alpha$, then draw a random number ξ from the first urn, otherwise, draw a random number from the second urn.

4. Repeat the process *n* times.

5. Set $\alpha = \alpha + 0.01$.

6. If $\alpha \le 1$, then go to 2.

We received a set of renewal processes $\xi_1, \xi_2, ..., \xi_n$. For $\alpha = 0$, we have a NHPP (in fact, a HPP) with $\lambda(t) = 1$ and for $\alpha = 1$, we have an "almost" deterministic sequence, which in some sense is opposite to NHPP. Using our procedure on this set of renewal processes we calculated the power of described above tests. The main results are shown in the Figures 4 and 5. Tests that not shown, are not sensitive for such processes.

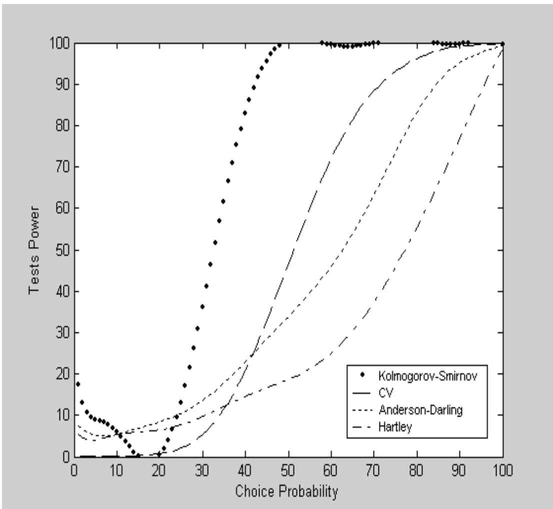


Figure 4. *Comparison of Tests Power with Log-Linear Function Recognition*

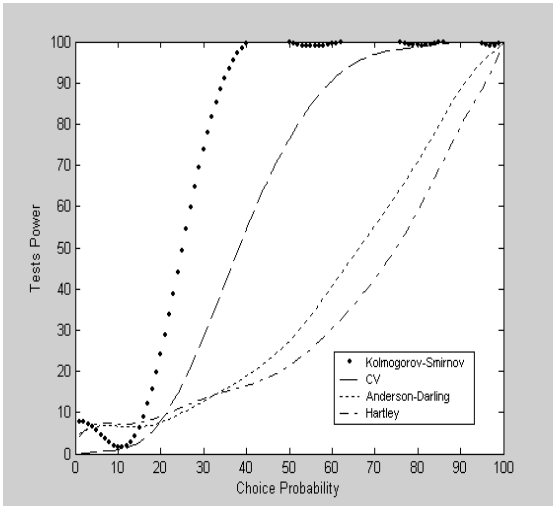


Figure 5*. Comparison of Tests Power with Power Function Recognition*

3.3.2. The second method

Consider two same urns with random numbers Ξ^1 and Ξ^2. The first urn consists of random numbers $\xi^1 \sim Exp(1)$. The second urn consists of the random numbers $\xi^2 \sim U(1-\varepsilon, 1+\varepsilon)$, where $\varepsilon = 0.01$ for example.

Set the α. Let us take the random number ξ^1 from the first urn and ξ^2 from the second. Make the linear combination

$$\xi = (1-\alpha)\xi^1 + \alpha\xi^2 .$$

Repeat the procedure *n* times. Receive the sequence of random inter invents intervals $\xi_1, \xi_2, ..., \xi_n$. Changing the parameter α from 0 to 1, we receive collection of renewal processes, where the first process for $\alpha = 0$ is a renewal process with standard exponentially distributed inter event times (NHPP with $\lambda(t) \equiv 1$,in fact, a HPP) and the last process for $\alpha = 1$ is a deterministic sequence. Using our procedure on this set of renewal processes we calculated the power of described above tests. The main results are shown in the Figures 6 and 7. Tests that not shown, are not sensitive for such processes.

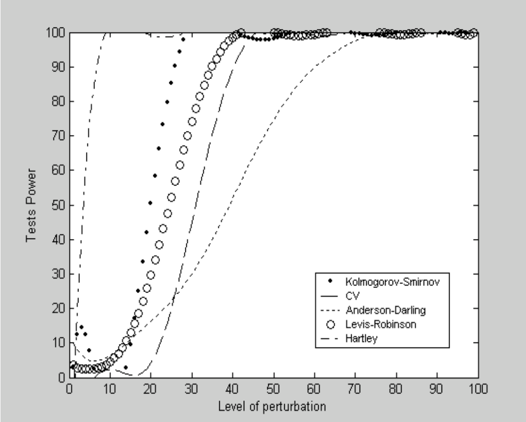


Figure 6. *Comparison of Tests Power with Log-Linear Function Recognition*

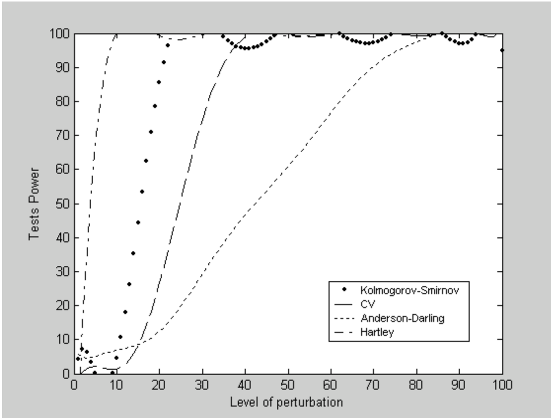


Figure 7. *Comparison of Tests Power with Power Function Recognition*

The sequence of purely exponential random variables is recognized by our tests as a HPP. Slightly perturbed sequence, e.g. $\alpha = 0.1 \div 0.2$, of exponential random variables are recognized by our tests as being HPP. Other processes are recognized by our tests as not being HPP.

3.3.3. The third method

Let us generate n evens of NHPP with intensity function $\lambda(t) = t$ using the reverse transformation of Exp(1) distributed sequence of random numbers.

1. Generate random numbers $x_i \sim Exp(1),\ \ i = 1, 2, ..., n$
2. Compute $y_1 = x_1,\ y_2 = x_1 + x_2\ ,...,\ y_{100} = \sum_{i=1}^{100} x_i$
3. Define $\upsilon \equiv \lambda(\upsilon)$
4. Let $y_k = \int_0^{t_k} \upsilon d\upsilon = \frac{t_k^2}{2},\ \ k = 1, 2, ..., n$
5. Compute $t_k = \sqrt{2y_k},\ \ k = 1, 2, ..., n$

In our generated NHPP the events occur at the time moments

$$t_1 = \sqrt{2y_1}\ ,...,\ t_k = \sqrt{2y_k}\ ,...,\ t_n = \sqrt{2y_n}$$

and intervals between events are equal to

$$\delta_1^1 = t_1,\ \delta_2^1 = t_2 - t_1,\ ...\ ,\delta_k^1 = t_k - t_{k-1},\ ...\ ,\delta_n^1 = t_n - t_{n-1}.$$

Check this NHPP using our procedure with Power intensity function. All test statistics fall inside the critical intervals. It does not contradict the NHPP assumption. On Figure 8 is shown the step function of the failure numbers and corresponding estimated intensity function. Estimated parameter β is equal to 2.

Let us generate the following sequence of time intervals:

$$t_1 = \sqrt{2},\ t_2 = \sqrt{4},\ ...\ ,\ t_n = \sqrt{2n}\ .$$

Intervals between events are

$$\delta_1^2 = \sqrt{2},\ \delta_2^2 = \sqrt{4} - \sqrt{2},\ ...\ ,\ \delta_n^2 = \sqrt{2n} - \sqrt{2(n-1)}\ .$$

In transformed time

$$\Delta_k = \int_0^{\sqrt{2k}} \upsilon d\upsilon = \frac{\upsilon^2}{2}\Bigg|_0^{\sqrt{2k}} = \frac{2k}{2} = k$$

and inter events intervals in transformed time exactly equals to 1. Check this sequence of time intervals using our procedure with Power intensity function. Estimated parameter β is equal to 2. Seven test statistics fall outside the critical intervals. It contradicts the NHPP assumption. On Figure 9 is shown the step function of the failure numbers and corresponding estimated intensity function.

Now let us mix two sequences of these time intervals.

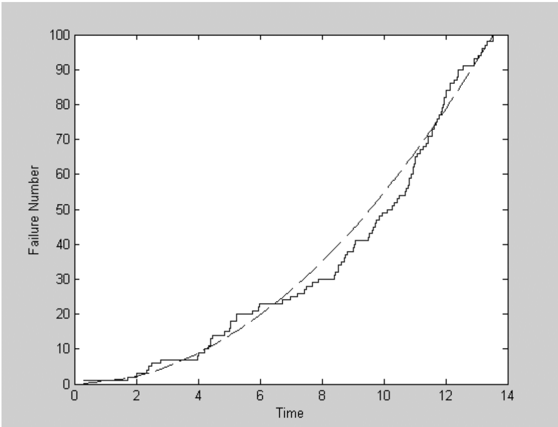


Figure 8. *The Plot of the generated NHPP step function of the failure numbers and corresponding estimated intensity function against time*

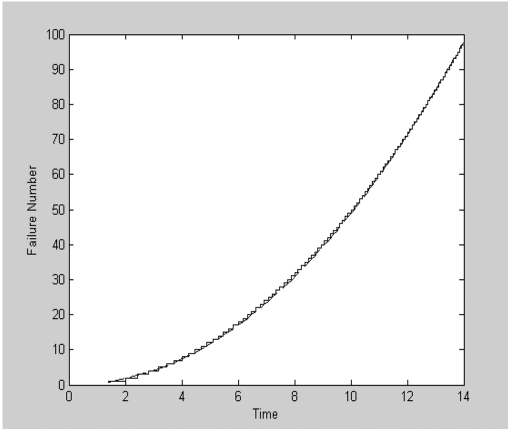


Figure 9*. The Plot of the step function of the failure numbers and corresponding estimated intensity function against time*

$$\delta_1^* = (1-\alpha)\delta_1^1 + \alpha\delta_1^2$$
$$\dots\dots\dots\dots\dots\dots$$
$$\delta_k^* = (1-\alpha)\delta_k^1 + \alpha\delta_k^2$$
$$\dots\dots\dots\dots\dots\dots$$
$$\delta_n^* = (1-\alpha)\delta_n^1 + \alpha\delta_n^2$$

If $\alpha = 0$, the sequence $\delta_1^*, \delta_2^*, \dots, \delta_n^*$ is NHPP. If $\alpha = 1$, the sequence $\delta_1^*, \delta_2^*, \dots, \delta_n^*$ is non-homogeneous deterministic process. Changing α from 0 to 1, we receive the set of perturbed processes from NHPP to deterministic process. Using procedure of power testing we received results shown on Figure 10.

The sequence of purely exponential random variables is recognized by our tests as a HPP. Slightly perturbed sequence, e.g. α = 0.1 ÷ 0.2, of exponential random variables are recognized by our tests as being HPP. Other processes are recognized by our tests as not being HPP.

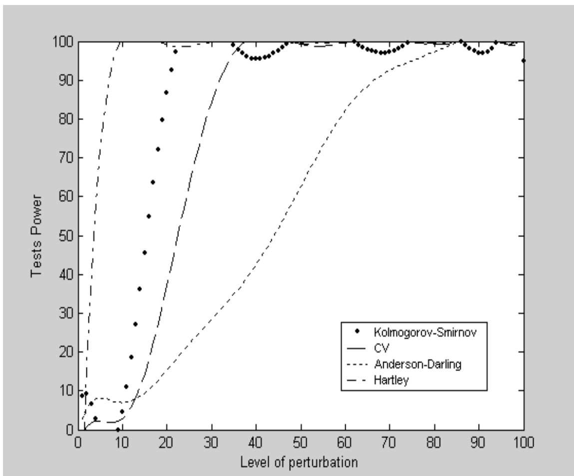


Figure 10. *Comparison of Tests Power with Power Function Recognition*

4. Conclusion

The test method proposed in this article is based on time transformation of the given process and the estimation of intensity function using the Maximum Likelihood Method. Critical values of the statistics were obtained by a repeated sampling of NHPP with estimated intensity function. Computer-intensive procedure allows us to suppose that the given sample path while testing the hypothesis belongs to NHPP.

We checked the procedure on different reliability data sets. We used failure data existing in literature and compared our conclusions to authors' results. In most cases there was the coincidence of results. Also we presented our own numerical example and showed how the above described procedure functions. Our procedure was verified on a set of the stochastic processes with known probabilistic structures, which includes HPP and a deterministic renewal process. It was demonstrated how our procedure identifies the underlying process.

The main weakness of our method is the fact that we worked only with NHPP having Log-linear and Power-law intensity functions, because there are efficient and easy to apply ML Method for estimation of the parameters. It would be desirable the develop numerical efficient methods to fitting the empirical cumulative failure rate to a wider class of functions, together with the corresponding methods of parameter estimation.

The above numerical results demonstrate the performance of our method to test the Nonhomogeneous Poisson Process hypothesis versus an alternative belong to a set of random processes obtained as addition of a "noise" to a pure NHPP. The best of our knowledge, there are no experimental data on similar test performance. As soon as they will become, we will compare the performance of our procedure versus the alternative.

Acknowledgment

This research was supported by research grant from the Negev Academic College of Engineering (NHPP), Beer-Sheva, Israel.

Reference

1. Akaike, H. (1974), A new Look at the Statistical Model Identification, IEEE Transactions on Automatic Control, AC-19(6), pp.716-723.
2. Anderson, T. W. and Darling, D. A. (1952), Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Process, The Annals of Mathematical Statistics, 23(2), pp. 193-212.
3. Anderson, T. W. and Darling, D. A. (1954), A Test of Goodness of Fit, Journal of the American Statistical Association, 49(268), pp. 765-769.
4. Ansel, J., Archibsald, T., Dagpunar, J and Thomas, L. (2001), Estimating Asset Life of Major Equipment, Proceeding of ESREL2001.
5. Ansell, J. B. and Phillips, M. J. (1989), Practical problems in the Statistical Analysis of Reliability Data, Applied Statistics, 38(2), pp. 205-247.
6. Arkin, B. L. and Leemis, L. M. (2000), Nonparametric Estimation of the Cumulative Intensity Function for a Nonhomogeneous Poisson Process from Overlapping Realization, Management Science, 46(7), pp. 989-998.
7. Ascher, H. E. and Feingold, H. (1984), Repairable Systems Reliability. Marcel Dekker, New York.
8. Ascher, H. E. and Hansen, C. K., (1998), Spurious Exponentiality Observed When Incorrectly Fitting a Distribution to Nonstationary Data. IEEE Transactions on Reliability, 47(4), pp. 451-459.
9. Bain, L. J., Engelhardt, M. And Wright, F. T. (1985), Testing for an Increasing Trend in the Intensity of a Poisson Process: a Power Study, Journal of the American Statistical Association, 80. pp. 419-422.
10. Baker, R. D. (1996), Some new tests of the power-law process, Technometrics, 38, pp. 256-265.

11. Bassin, W. M. (1973), A Bayesian Optimal Overhaul Interval Model for the Weibul Restoration Process Case, Journal of the American Statistical Association, 68(343), pp. 575-578.
12. Bergman, B. and Klefsjö, B. (1984), The Total Time to Test Concept and Its Use in Reliability Theory, Oper. Research, 32(3), pp. 596-606.
13. Bhattacharjee, M. (2001), Inference for Non-Homogeneous Poisson Processes: Models for Repairable Systems, ISI Session, Seul, August, 2001, pp. 1-16.
14. Black, S. E. and Rigdon, S. E. (1996), Statistical Inference for a Modulated Power Law Process, Journal of Quality Technology, 28(1), pp. 81-90.
15. Bracquemond, C., Chevalier, M. and Gaudoin, O. (2001), Modeling the Aging of Non Repairable Systems in Discrete Time, Proceedings of ESREL2001.
16. Çinlar, E. (1975), *Introduction to Stochastic Processes*, Prentice-Hall, New-Jersey.
17. Coetzee, J. L. (1997), The Role of NHPP Models in the Practical Analysis of Maintenance Failure Data, Reliability Engineering and System Safety, 56, pp. 161-168.
18. Cox, D. R. and Lewis, P. A. W. (1966), *The Statistical Analysis of Series of Events,* Chapman and Hall, London.
19. Crow, L.H. (1974), Reliability Analysis for Complex, Repairable Systems, Reliability and Biometry, pp. 379-410.
20. Crow, L.H. (1990), Evaluating the Reliability of Repairable Systems, Proceedings of Annual Reliability and Maintainability Symposium, pp. 275-279.
21. Crowder, M. J., Kimber, A. C., Smith, R. L. and Sweeting, T. J. (1991), *Statistical Analysis of Reliability Data*, Chapman and Hall, New York.
22. Dohi, T., Takeita, K. and Osaki, S. (2000), Graphical Methods for Determining/Estimating Optimal Repair-Limit Replacement Systems, International Journal of Reliability, Quality and Safety Engineering, 7(1), pp. 43-60.
23. Epstein, B. (1960), Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential. Part I, Technometrics, 2(1), pp. 83-101.
24. Epstein, B. (1960), Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential. Part II, Technometrics, 2(2), pp. 167-183.
25. Frenkel, I. B., Gertsbakh, I. B. and Khvatskin, L. V. (2003), Parameter Estimation and Hypotheses Testing for Nonhomogeneous Poisson Process, Transport and Telecommunication, 4(2), pp. 9-17.
26. Gaudoin, O. (1992), Optimal Properties of the Laplace Trend Test for Software-Reliability Models, IEEE Transactions on Reliability, 41(4), pp. 525-532.
27. Gaudoin, O., Yang, B. and Xie, Min. (2003), A Simple Goodness-of-Fit Test for the Power-Law Process, Based on the Duane Plot, IEEE Transactions on Reliability, 52(1), pp. 69-74.
28. Gertsbakh, I. B. (2000), *Reliability Theory with Applications to Preventive Maintenance*. Springer Verlag.
29. Gnedenko, B. V, Belyaev, Yu. K. and Solovyev, A. D. (1969), Mathematical Methods in Reliability Theory. Academic Press, New York.
30. Gokhale, S. S. and Trivedi, K. S. (1996), Log-Logistic Software Reliability Growth Model, Proceedings of ISSRE1996.
31. Guess, F., Hollander, M. and Proschan, F. (1986), Testing Exponentiality Versus a Trend Change in Mean Residual Life, The Annals of Statistics, 14(4),pp. 1388-1398.
32. Hartley, H. O. (1950), The maximum F-ratio as a short-cut test for heterogeneity of variance, Biometrica, 37, pp.308-312.
33. Jani, P. N., Shanubhogue, A. and Muralidharan, K. (1997), Some Tests for a Poisson Process and Their Power with Respect to Trend in the Intensity, Commun. Statist. – Theory Meth., 26(11), pp. 2671-2688.
34. Kaminsky, F. C. and Rumpf, D. L. (1977), Simulating Nonstationary Poisson Processes: a Comparison of Alternatives Including the Correct Approach, Simulation, 7, pp. 17-20.
35. Kao, E. P. C. and Chang, S.-L. (1988), Modeling Time-Dependent Arrivals to Service Systems, Management Science, 34(11), pp. 1369-1379.
36. Klefsjö, B. and Kumar, U. (1992), Goodness-of-fit Tests for the Power-Law Process Based on the TTT-Plot, IEEE Transactions on Reliability, 41(4), pp. 593-598.

37. Klein, R. W. and Roberts, S. D. (1984), A Time-varying Poisson Arrival Process Generator, Simulation, 10, pp. 193-195.
38. Kuhl, M. E. and Bhairgond, P. S. (2000), Nonparametric Estimation of Nonhomogeneous Poisson Processes using Wavelets, Proceedings of the 2000 Winter Simulation Conference, pp. 562-571.
39. Kuhl, M. E., Damerdji, H. and Wilson, J. R. (1998), Least Square Estimation of Nonhomogeneous Poisson Processes, Proceedings of the 1998 Winter Simulation Conference, pp. 637-645.
40. Kuhl, M. E. and Lim, S. E. (1999), Sensitivity Analysis of Simulation Output to Parameters of Nonhomogeneous Poisson Processes, Proceedings of the 1999 Winter Simulation Conference, pp. 462-465.
41. Kuhl, M. E., Wilson, J. R. and Jonson, M. A. (1997), Estimating and Simulating Poisson Processes Having Trends or Multiple Periodicities, IIE Transactions, 29, pp. 201-211.
42. Kuhl, M. E. and Wilson, J. R. (2001), Modeling and Simulating Poisson Processes Having Trends or Nontrigonometric cyclic effects, European Journal of Operational Research, 133, pp. 566-582.
43. Kvaløy, J. T. and Lindqvist, B. H. (1998), TTT-Based Tests for Trend in Repairable System Data, Reliability Engineering and System Safety, 60, pp. 13-28.
44. Kvaløy, J. T., Lindqvist, B. H. and Malmedal, H. (2001), A Statistical Test for Monotonic and Non-Monotonic trend in Repairable Systems, Proceedings of ESREL2001.
45. Lawless, J. F. and Thiagarajah, K. (1996), A point-process model incorporating renewals and time trend, with application to repairable systems, Technometrics, 38, pp. 131-138.
46. Lee, S., Wilson, J. R. and Crawford, M. M. (1001), Modeling and Simulation of a Nonhomogeneous Poisson Process having Cyclic Behavior, Communication of Statistics – Simulation, 20(2&3), pp. 777-809.
47. Leemis, L. W. (1988), Input Modeling, Proceedings of the 1998 Winter Simulatiom Conference, pp. 15-22.
48. Leemis, L. W. (1991), Nonparametric Estimation of the Cumulative Intensity Function for Nonhomogeneous Poisson Process, Management Science, 37(7), pp. 886-900.
49. Lewis, P. A. W. and Robinson, D. W. (1974), Testing for a Monotone Trend in a Modulated Renewal Process, Reliability and Biometry, pp. 163-182.
50. Lewis, P. A. and Shelder, G. S. (1976), Statistical Analysis of Non-stationary Series of Events in a Data Base Systems, IBM J. Res. Development, pp. 465-482.
51. Lewis, P. A. and Shelder, G. S. (1976), Simulation of Nonhomogeneous Poisson Process with Log Linear Rate Function, Biometrica, 63(3), pp. 501-505.
52. Lewis, P. A. and Shelder, G. S. (1979), Simulation of Nonhomogeneous Poisson Process by Thinning, Naval Research Logistics Quarterly, 26, pp. 403-413.
53. Lilliefors, H. W. (1969), On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown, Journal of the American Statistical Association, 64, pp. 387—389..
54. Lindqvist, B. H. (1997), Statistical Modeling and Analysis of Repairable Systems, Proceedings of the 1st International Conference on Mathematical Methods in Reliability, Bucharest, Romania, September, 16-19, 1997, (MMR'97).
55. Lomnicki, Z. A. (1966), A Note on the Weibull Renewal Process, Biometrica, 53(3 and 4), pp. 375-381.
56. López-Swgovia, L., Villaseñor-Alva, J. A. and Vaquera-Huerta, H. (2002), Two Goodness of Fit Tests for Non Homogeneous Poisson Processes, Agrociencia, 36(3), pp. 703-712.
57. MacLean, C. J. (1974), Estimation and Testing of an Exponential Polinomial Rate Function with the Nonstationary Poisson Process, Biometrica, 61(1), pp. 81-85.
58. Mann, H. B. (1945), Nonparametric Tests Against Trend, Econometrica, 13(3), pp. 245-259.
59. Massey, W. A., Parker, G. A. and Whirtt, W. Estimating the Parameters of a Nonhomogeneous Poisson Process with Linear Rate, Telecommunication Systems, 5, pp. 361-388.
60. Meeker, W. D. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data,* Wiley and Sons.

61. MIL-HDBK-189 (1981), Reliability Growth Management. HQ, US Army Communications Research and Development Command. ATTN: DRDCORT, Fort Monmouth, New Jersey.
62. Moller, S. K. (1976), the Rasch-Weibull Process, Scandinavian Journal of Statistics, 3, pp. 107-115.
63. Muralidharan, K., (2001), On testing of Parameters in Modulated Power Law Process, Applied Stochastic Models in Business and Industry, 17, pp. 331-343.
64. Nechval, N. A., Vasermanis, E. K., Rozevskis, U., Rozite, K. and Nechval, K. N. (2003), Characterization Theorems and Goodness-of-Fit Testing on Smal Data Samples, Transport and Telecommunication, 4(2), pp. 92-100.
65. Park, W. J. and Kim, Y. G., (1992), Goodness-of-Fit Tests for the Power-Law Process, IEEE Transactions on Reliability, 41(1), pp. 107-111.
66. Park, W. J. and Seon, M., (1994), More Goodness-of-Fit Tests for the Power-Law Process, IEEE Transactions on Reliability, 43, pp. 275-278.
67. Proschan, F. (1963), Theoretical Explanation of Observed Decreasing Failure Rate, Technometrics, 5(3), pp. 375-383.
68. Pulcini, G. (2001), A Bounded Intensity Process for the Reliability of Repairable Equipment, Journal of Quality Technology, 33(4), pp. 480-492.
69. Rigdon, S. E. (1989), Testing Goodness-of-Fit for the Power Law Process, Commun. Statist. – Theory Meth., 18(12), pp. 4665-4676.
70. Rigdon, S. E. and Basu, A. P. (1989), The Power Law Process: A Model for the Reliability of Repairable Systems, Journal of Quality Technology, 21(4), pp.251-260.
71. Rigdon, S. E. and Basu, A. P. (1990), The Effect of Assuming a Homogeneous Poisson Process When the True Process is a Power Law Process, Journal of Quality Technology, 21(4), pp.111-117.
72. Saldanha, P. C., de Simone, E., A. and Frutuoso e Melo, P., F. (2001), An Application of Non-Homogeneous Poisson Processes to the Reliability Analysis of Service Water Pump, Nuclear Engineering and Design, 210, pp.125-133.
73. Schmee, J., Gladstein, D. and Nelson, W. (1985), Confidence Limits for Parameters of a Normal Distribution from Singly Censored Samples, Using Maximum Likelihood, Technometrics, 27(2), pp. 119-128.
74. Serfling, R. J. (1974), Research in Point Process, with Applications in Reliability and Biometry, Reliability and Biometry, pp. 109-127.
75. Shin, I., Lim, T. J. and Lie, C. H., (1996), Estimation Parameters of Intensity Function and maintenance Effect for Repairable Unit. Reliability Engineering and System Safety, 54, pp. 1-10.
76. Schneidewind, N. F. (1993), Software Reliability Model with Optimal Selection of Failure Data, IEEE Transactions on Software Engineering, 19(11), pp. 1095-1104.
77. Thompson, W. A. Jr. (1981), On the Foundations of Reliability, Technometrics, 23(1), pp. 1-13.
78. Villaseñor-Alva, J. A., and Díaz-Carreño, M. A. (2003), Non-Parametric Tests for Non-Homogeneous Poisson Processes, Agrociencia, 37(1), pp. 21-31.
79. Wang, K. S., Po, H. J., Hsu, F. S. and Lui, C. S. (2002), Analysis of Equivalent Dynamic Reliability With Repairs Under Partial Information, Reliability Engineering and System Safety, 76, pp. 29-42.
80. Weckman, G. R., Shell, R. L. and Marvel, J. H. (2001), Modeling the Reliability of Repairable Systems in the Aviation Industry, Computers and Industrial Engineering, 40, pp. 51-63.
81. Xie, M., Yang, B. and Gaudoin, O. (2001), Regression Goodness-of-Fit Test for Software Reliability Model Validation, Proceedings of ISSRE2001.
82. Yañez, M., Joglar, F. and Nodarres, M. (2002), Generalized Renewal Process for Analysis of Repairable Systems with limited Failure Experience. Reliability Engineering and System Safety, 77, pp. 167-180.
83. Zeljkoviæ, V. and Begenišiæ, N. (2001), Software Reliability – Different Model Analyses, Proceeding of TELFOR2001, Telecommunication Forum, November 20-22, 2001, Beograd.
84. Zhao, M. and Xie, M. (1992), On the Log-Power NHPP Software Reliability Model, Proceedings of the 3[d] Int. Symp. On Software Reliability Engineering, North Carolina, USA, IEEE Comp.-Society Press, 7-10 Oct. 1992, North Carolina, USA, pp. 14-22.

COMBINATORIAL AND PROBABILISTIC PROPERTIES OF LOMONOSOV's "TURNIP"

I. Gertsbakh[1], Y. Shpungin[2]

[1] *Ben Gurion University of the Negev, Beersheva, P.O.Box 653, Israel*
e-mail: eliahu@cs.bgu.ac.il

[2] *Negev Academic College of Engineering, Beersheva, 84100, Israel*
e-mail: yosefs@nace.ac.il

Abstract

This paper is a review of a very efficient combinatorial approach for Monte Carlo (MC) estimation of networks reliability parameters. It is well known that the main problem of the Crude Monte Carlo and various versions of it is unbounded growth of the relative error. This new and very promising approach was first suggested by M. Lomonosov in 1974 [1] and developed later in a series of works [2,3,4,5,6]. The method eliminates the rare event problem, has several computational advantages and is very useful for calculating different reliability parameters such as the stationary network availability and so-called stationary down-up transition rate.

Key words: terminal connectivity; network reliability; Monte Carlo; closure; trajectory; transition rate

1. Introduction

Networks and their reliability is a subject of great interest. Many researches in various directions take place in this field. This paper presents one of the most useful approaches for estimating reliability parameters of static and dynamic networks. The principal idea of this approach was first suggested in [1].

1.1. What is a Network?

The common feature for all networks is that they consist of *vertices* and *edges*. There exist many types of networks varying in their performance definitions and therefore with different concepts of reliability. For our purposes, we define a network in the following manner:

By network *N=(V, E, T)* we denote an undirected graph with node-set *V* (vertices) , edge-set *E* and set *T* of special nodes named *terminals* (see Fig. 1.a). *T* is a subset of *V*.

We define nodes as *reliable* and edges as *not reliable*. In other words, nodes never fail and edge $e \in E$ fails (is being erased) with probability $q(e)$. A *state* of a network is a set of *non-erased* edges. By the definition, the network is in UP state if all its terminals are connected, i.e. if between each pair of terminals there exists a path that connects them. Otherwise we say that a network is in the DOWN state. Two most popular cases of the operational criteria for networks are the s-t connectivity and all-terminal connectivity. Fig. 1.a presents a network *N* with two terminals: *s* and *t*. In Fig. 1.b we see a DOWN state of *N*: {2,3,5,7}. Fig. 1.c shows an UP state of *N*: {4,7}.

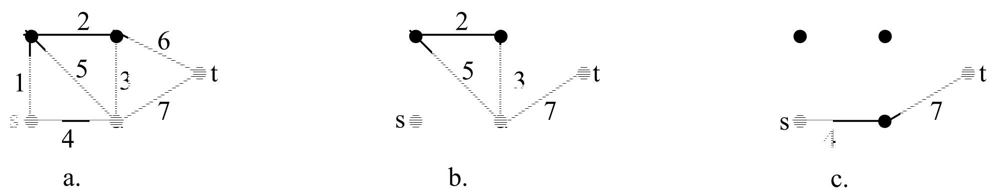


Figure 1. *The s-t terminal Network and its DOWN and UP states.*
Bold points are the vertices. Terminals are denoted by letters s,t.

The common static network problem is the calculation of its *Reliability*, i.e. the probability that a network is in the UP state. Note that the straightforward calculation of reliability is very time consuming, even for moderate size networks. Also the well known path-set and/or cut-set methods demand computation time which grow exponentially with the network size. Computational difficulties in the direct calculation of static network reliability stimulated great interest in developing Monte Carlo methods.

We do not intend to review all the approaches in this direction. The greatest part of these works is related to so-called Crude Monte Carlo (CMC) or its modifications. The main deficiency of these methods is that they are very inefficient in two extreme cases: highly reliable and highly unreliable networks (the so called rare event phenomenon). Note that the highly reliable networks have great interest in many important applications.

In this paper we present a very efficient MC approach suggested in [1] and developed in [2,3,4,5,6]. This method is especially useful for the extreme cases mentioned above. It *in principle* eliminates the rare event phenomenon and has, in addition, some useful properties for its applications in static and dynamic networks. (If edge *e* has lifetime $\xi(e)$, $e \in E$, we call such network *dynamic*).

1.2. Reliability and Crude Monte Carlo

Let us begin with a very simple network shown on Fig. 2. Suppose that the operational criterion is all-terminal connectivity and the edge probabilities of being up are equal: $p(1)=p(2)=p$.

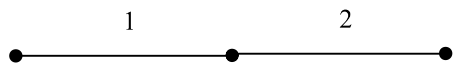


***Figure 2**. A series network*

Clearly that the only UP state here is the state $S=\{1,2\}$ and the reliability of the network $R(N) = p^2$. Suppose now that we want to estimate this reliability using CMC. Denote by X the random variable which characterizes the state of the network:

$$X = \begin{cases} 1, & \text{if the state is UP} \\ 0, & \text{otherwise} \end{cases}.$$

Note that $p(X = 1) = p^2$ and $E[X] = p^2$. The procedure of CMC is the following

(i) put $\hat{R} := 0$;

(ii) simulate the states of edges 1 and 2 with the given probability p and determine X;

(iii) $\hat{R} := \hat{R} + X$;

(iv) repeat the above operations M times;

(v) the estimate of the R is $\hat{R} := \hat{R}$ /M.

Clearly that $\hat{R}$ is an unbiased estimator of R.

The deficiency of this procedure lies in the simulation of edges' states. If the edges are highly reliable, say $p \to 1$, we will almost never observe the down states of the edges. This means that the estimated value of R may be 1! Obviously, this result is not satisfactory.

Let us look at this situation more formally. One of the measures of the MC methods is the relative error, which in our case equals

$$\delta_{CMC} = \frac{\sqrt{Var(1-\hat{R})}}{1-R} = \sqrt{\frac{Var(X)}{(1-p^2)^2 \cdot M}} = \sqrt{\frac{p^2}{1-p^2}} \cdot \frac{1}{\sqrt{M}}. \tag{1}$$

From this formula we see that when $p \to 1$, then for fixed M the value of the relative error tends to *infinity*, and this is the main obstacle for using CMC.

1.3. The CMC as an "*Urn*" Scheme

To explain the advantages of the Lomonosov's idea, let us look at the CMC from a more general point of view. Consider an *urn U* with a large number of balls *b* in it. Suppose that each ball *b* is marked with some value *z*(*b*), and we want to calculate the sum of *z*(*b*) over *b* in *U*:

$$Z = \sum_{b \in U} z(b). \tag{2}$$

This completely corresponds to computation of network reliability. For this case, the balls *b* are the states, and *z*(*b*) are defined as 0 for any DOWN state and as the probability of the state if the state belongs to the set UP. Therefore, *Z* becomes the reliability of the network. Because of a very large number of balls in *U*, we usually cannot compute exactly the whole sum, and we are forced to estimate *Z*. To convert the expression in (2) into MC scheme, we introduce the probability distribution *p(b)* on *U*. Ball *b* will be drawn with probability *p*(*b*). Then one may express the sum *Z* in the following form:

$$Z = \sum_{b \in U} p(b) \cdot \frac{z(b)}{p(b)} = E[Y(b)], \tag{3}$$

where $Y(b) = \dfrac{z(b)}{p(b)}$. (Note that values of *Y*(*b*) are 0 or 1 for the states DOWN and UP, respectively). The main idea of this representation is that the sum *Z* is expressed as an *expectation* of some random variable.

Now the CMC scheme for evaluating $Z = E[Y]$ can be stated as follows:

- draw balls b_i *M* times from the urn U, with probability $p(b_i)$;

  $$\hat{Y} = \frac{\sum_{i=1}^{M} Y(b_i)}{M}$$

- calculate, which is an *unbiased* estimator of *Z*.

As in the example in (1.2), it is easy to get the general form for the relative error:

$$\delta_{CMC} = \sqrt{\frac{R}{1-R}} \cdot \frac{1}{\sqrt{M}}. \tag{4}$$

Again we see the unbounded growth of the relative error for highly reliable networks. Let us emphasize once more that it is the main deficiency of the CMC. The reason for this lies in the above described urn scheme. Namely, each ball, which represents the state, is drawn from the urn with the probability of that state. Therefore the growth of the relative error is caused by a bad definition of probability measure *p*(*b*) on network states.

Various improvements of CMC have been suggested in order to reduce the unbounded growth of the δ_{CMC}. The approach suggested in [2] is referred to in this paper as Lomonosov's turnip.

This is a *different urn* scheme, which eliminates the rare event phenomenon and is very efficient for computing the reliability of highly reliable networks.

2. The Idea of the Turnip

The turnip is constructed using three basic ideas. The first one is introducing some *artificial* random process associated with each edge. The second one is defining the *balls b* in the *urn* scheme as certain *trajectories* which do not depend on the probabilities of the network states. The third one is using a special combinatorial operation (*closure*) which considerably accelerates the computation procedure.

2.1. Parameterization and *Artificial* Creation Process

Introduce for each edge an *artificial* creation process (see Fig. 3). Initially, at t=0, every edge e is down. Edge is "born" independently of others and remains *up* forever. The random moment $\xi(e)$ of edge "birth" is exponentially distributed:

$$P\{\xi(e) \le t\} = 1 - \exp(-\lambda(e) \cdot t), \quad e \in E. \tag{5}$$

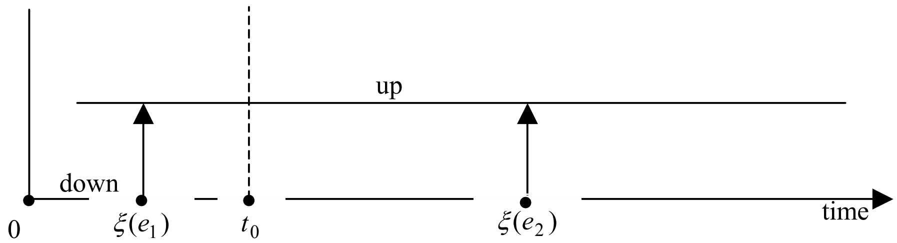


Figure 3. *Independent appearance of edges in time. At the instant t_0, edge e_1 does exist while e_2 does not*

Fix now an *arbitrary* moment t_0 (see Fig. 3), in particular, $t_0 = 1$. Chose for each edge *e*, the "birth" rate $\lambda(e)$ so that the following condition holds:

$$P\{\xi(e) > t_0\} = \exp(-\lambda(e) \cdot t_0) = q(e). \tag{6}$$

This means that in this *dynamic* process, the probability of being *down* at t_0 for each edge *coincides* with the *static down*-probability $q(e)$. At t=0 our network is DOWN. At some random moment $\xi(N)$ the network becomes UP (and remains in the UP forever). Let us denote by $P\{\xi(N) \le t_0\}$ the probability that at moment t_0 in the creation process the whole network N is UP. Then we have:

$$R(N) = P\{\xi(N) \le t_0\}. \tag{7}$$

In words: the static probability $R(N)$ that the network is UP coincides with the probability that in the creation process, the state becomes UP before t_0.

2.2. Closure

One of the important advantages of our approach is the essential use of *combinatorial* features of the network, namely the operation of *closure*. We give an intuitive explanation for closure. For more precise definition see [2].

Suppose that the edges in the creation process appear randomly and independently. In Fig. 4.a a network with four nodes and five edges is given. We are going to use this network in the following examples for two operational criteria: all terminal connectivity and *s-t* connectivity.

Suppose that first edges 1 and 2 are born. At this moment the nodes associated with edge 3 are already connected by a *path* formed by edges 1 and 2. Therefore edge 3 does not affect the connectivity of the component {1, 2, 3} and becomes *irrelevant* to the further evolution of the network.

For each state $S \subseteq E$ of the network we define a *closure* as a subset of edges in *E* so that it equals to the *union* of *S* and all irrelevant edges (for given operational criterion). For example, in Fig. 4.b the set $\sigma_{21} = \{1,2,3\}$ is a closure for each of the following states: {1,2}, {1,3}, {2,3}. Let us carry out the closure operation within each connected component of the network. Then we call the collection of all such components a *super-state.* (A single node is also a closed component).

2.3. Turnip as Evolution Process with Closure

On Fig. 4.b we present the turnip as the creation process defined in 2.1 together with the closure operation defined in 2.2. As an example, we consider a network *N* with four nodes and five edges. The operational criterion is the all-terminal connectivity.

Suppose that initially all edges are in the *DOWN* state, i.e. the initial *super-state* is a collection of one-node components. The zero-level of the turnip is the set *V* without any edge (no edge was born). The first level shows all possible evolution results from the zero level which appear as a result of a birth of a single edge. There are 5 such super-states. The second level of the turnip shows what happens when a second edge is born. We distinguish 6 such super-states $\sigma_{21}, \sigma_{22}, \sigma_{23}, \sigma_{24}, \sigma_{25}, \sigma_{26}$. It is important to stress that these super-states are shown together with the relevant closure. For example, suppose that after edge 1, edge 3 is born. Then edge 2 can be added to the existing edges, and the corresponding super-state is σ_{21}. Similarly, the same σ_{21} appears if 2 is born after 1.

The last level consists of one super-state containing one component. Note that in the case of all-terminal connectivity, only at this level the corresponding state is the UP-state. In the case of the *s-t* connectivity the UP states may be on each level. For example, consider the network on Fig. 4.c, which has the terminals *s* and *t*. We see that the UP states are situated on levels 2 and 3 (The corresponding super-states on Fig. 4.b are double-circled).

An important feature of the turnip is that simulating the transitions from one super-state to another is very easy. Compute for example the probability $P(\sigma_{11} \to \sigma_{21})$ of the transition from σ_{11} to σ_{21}. Suppose that the birth rates for edge *k* is $\lambda(k),\ k = 1,...,5$. Then, by the well known fact, we have:

$$P(\sigma_{11} \to \sigma_{21}) = \frac{\lambda(2)+\lambda(3)}{\lambda(2)+\lambda(3)+\lambda(4)+\lambda(5)} = \frac{\lambda(\sigma_{11})-\lambda(\sigma_{21})}{\lambda(\sigma_{11})}, \qquad (8)$$

where $\lambda(\sigma_{ij})$ denotes the total birth rate for the super-state σ_{ij}. Indeed, the transition $\sigma_{11} \to \sigma_{21}$ takes place if and only if edge 2 or 3 is born out of four possibilities of the birth of edges 2, 3, 4 and 5.

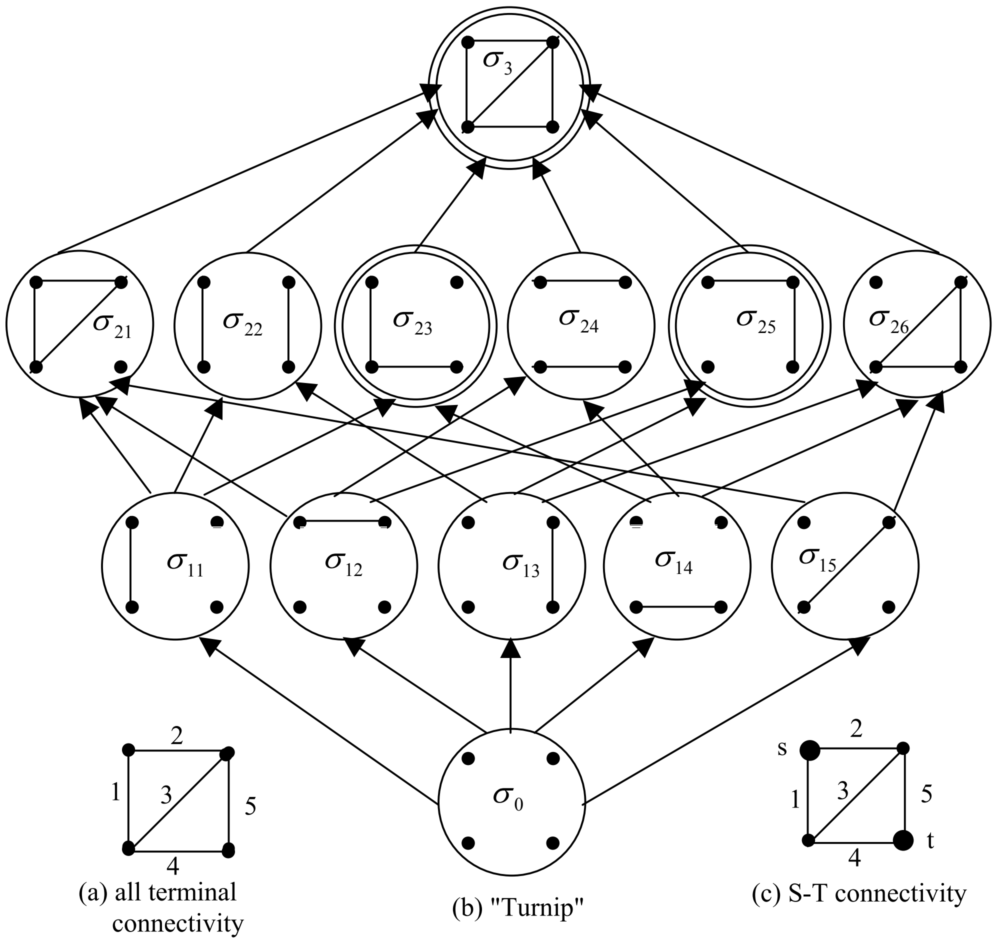


Figure 4. *The "Turnip" for two cases of connectivity*

Actually, the "turnip" describes an artificial creation process with closure, and the probability that the network in this process will be UP at some moment t_0, coincides with the corresponding static probability of the network. For more precise description of the process see [2].

Consider a random process $\sigma(t)$ whose states are the super-states of the above described network evolution process. (For example $\sigma(t=0)=\sigma_0$). We already mentioned that each state corresponds to some super-state. The following claim was proved in [2].

Claim 1. (i) $\sigma(t)$ is a Markov process;

(ii) The time spent by $\sigma(t)$ in a particular super-state σ' is distributed as $\exp(\lambda(\sigma'))$;

(iii) Let σ'' be the direct successor of σ'. Then the transition $\sigma' \to \sigma''$ takes place with the probability $P(\sigma' \to \sigma'') = (\lambda(\sigma') - \lambda(\sigma''))/\lambda(\sigma')$.

Remark. Each successor σ'' of σ' is obtained by merging exactly two components of σ'. The process $\sigma(t)$ is called in [2] the Merging Process (MP).

Define now a new term, *trajectory*, which plays the central role in the Merging Process (and therefore also in the corresponding Monte Carlo scheme). A trajectory is a sequence $u=(\sigma_0,\sigma_1,...,\sigma_r)$ of super-states such that σ_0 is the trivial initial super-state, each σ_i is the direct successor of σ_{i-1}, and r is the first i such that σ_i is UP. For example, the sequence $(\sigma_0,\sigma_{11},\sigma_{23},\sigma_3)$ on the Fig. 4.b is a trajectory.

Now, in terms of the trajectories, the network is UP at moment t if there exists at least one trajectory which reaches the UP state before t. It is easy to calculate the probability $p(u)$ for a trajectory $u = (\sigma_0, ..., \sigma_r)$:

$$p(u) = \prod_{i=0}^{r-1} P(\sigma_i \to \sigma_{i+1}). \tag{9}$$

Suppose that the evolution goes along the trajectory u. Denote by $P(t|u)$ the probability that the UP-state (i.e., σ_r) will be reached before time t *given* that it goes along u. By property (ii) of the claim 1, our MP is sitting in each super-state σ_j an exponentially distributed random time $\tau(\sigma_j)$ and due to the Markovian property, the total evolution time along the trajectory u is a sum of the respective exponential random variables. More formally,

$$P(t \mid u) = P(\tau(\sigma_0) + \tau(\sigma_1) + ... + \tau(\sigma_{r-1}) < t \mid u = (\sigma_0, ..., \sigma_r)). \tag{10}$$

Note that given u, this quantity can be computed directly as a convolution of respective exponents.

Now, the probability that at moment t the MP process reaches UP equals

$$R(N) = P\{\xi(N) \le t\} = \sum_{u \in U} p(u) P(t \mid u), \tag{11}$$

where U is the set of all trajectories. The latter expression has the form of an expectation as in (3) and determines in fact, the simulation strategy for its evaluation:

(i) put $\hat{R} := 0$;

(ii) *generate* trajectory u from the trivial super-state to the super-state in the UP, with the probability (9);

(iii) *calculate* $\hat{R} := \hat{R} + P(t \mid u)$;

(iv) repeat the (ii) and (iii) M times;

Then the $\dfrac{\hat{R}}{M}$ is an unbiased estimator of $R(N)$.

In terms of the *urn scheme* (see 1.3) the trajectory is a *ball* and it is drawn with probability (9) which *does not depend* on the edge probabilities. This, at least intuitively, indicates that in the MP scheme the rare event phenomenon does not exist. The following claim was proved in [3].

Claim 2. For a given number of nodes and a given operational criterion the coefficient of variation δ^2_{MP} is bounded uniformly for all $t \in [0, \infty)$ and all λ-vectors satisfying $\max \lambda(\cdot) / \min \lambda(\cdot) \le C$ (for any C).

Claim 2 assures (through the Tshebyshev inequality), that for any given number of nodes n and rates $\lambda(e)$, and for any positive ε, δ, there exists a sample size N such that for all $0 \le t < \infty$, we have

$$1 - \varepsilon < \frac{\hat{R}(N)}{R(N)} < 1 + \varepsilon \tag{12}$$

with probability at least $1 - \delta$. Note that this property does not hold for the CMC.

3. Applications of Turnip

3.1. Availability $A_N(t)$

The static reliability *R(N)* also expresses the equilibrium instant availability of *N*. Suppose that the life of each edge *e* is described by an alternating renewal process. Assume that for this process there exist mean *up* and *down* periods. Consider the process *E(t)*- the set of all edges from *E* which are UP at moment *t*. Then the availability $A_N(t)$ as $t \to \infty$ of the network *N* equals the static probability *R*(*N*), with edge *down*-probability given by $q(e) = \dfrac{\lambda(e)}{\lambda(e)+\mu(e)}$, where $\lambda(e)$ is the failure rate and $\mu(e)$ is the repair rate. For more details see Gnedenko et al [7].

3.2. The Mean Period T_{UP} and the Transition Rate $\Phi(N)$

One of important characteristics of the *dynamic* network is the mean equilibrium *up* and *down* periods of the network, T_{UP} *and* T_{DOWN} respectively. Let us define T_{UP} (T_{DOWN} is similar). Consider the sequence $\{U_i\}, i = 1,...,k$ of *UP* periods. Then, it may be proved that, as $k \to \infty$, the sequence $\{U_k\}$ converges in probability to limit U with $E[U] = \lim_{k\to\infty} E[U_k]$. Denote $T_{UP} = E[U]$ (see [7]). Computing the values T_{UP} and T_{DOWN} is a very difficult task. One can compute them using the MP. At first, note that the following formulas are valid ([7]):

$$R(N) = \frac{T_{UP}}{T_{UP} + T_{DOWN}}, \tag{13}$$

$$\Phi(N) = \frac{1}{T_{UP} + T_{DOWN}}, \tag{14}$$

where $\Phi(N)$ is the so-called stationary transition rate (see its definition below). Clearly, computing $\Phi(N)$ opens the way for computing T_{UP}. The turnip effectively helps in doing so.

Let us now describe the notion of $\Phi(N)$, which is itself an interesting characteristic of the renewable network. By definition, $\Phi(N)$ is the transition rate from DOWN to UP. In fact, a transition DOWN$\to$UP may take place only from so-called *border* states. A border state is such particular DOWN state of the network which may be transformed in an UP state by adding a single edge. Denote *BD* the set of all border states. For any border state *S*, denote by S^+ the set of all edges *e* such that $S \cup e \in UP$. Consider, for example, the state *S*={3, 4} for the network on Fig. 4.a. This state is presented by super-state σ_{26} on Fig. 4.b. Clearly, S is a border state, and $S^+ = \{1,\ 2\}$.

Now we are ready to define the stationary transition rate more formally ([7]):

$$\Phi(N) = \sum_{S \in BD} P(S) \cdot \mu(S^+), \tag{15}$$

where *P(S)* is the probability of the border state *S* and $\mu(S^+) = \sum_{e \in S^+} \mu(e)$.

3.3. Estimating $\Phi(N)$

The number of border states in network is usually very large. Therefore we can only estimate the transition rate $\Phi(N)$. Let us explain here shortly the MC scheme for estimating $\Phi(N)$.

We consider the all-terminal criterion only. The case of s-t connectivity is more complicated *technically,* but not in principle. The difference between two criteria lies in the fact that for the all-terminal connectivity, all border states are located on the (r-1)-st level of the turnip, where r is the highest level. All states of this level contain exactly two connected components. In the case of the s-t connectivity, the border states may be located at each level of the turnip and may consist of any number of components. For example, we see on Fig. 4.b that for the s-t connectivity the state S={5} (which is associated with the super-state σ_{13}) is a border state and it consists of 3 components.

Consider now the sum for $\Phi(N)$ in (15), which is represented in the static form. We adopt the same parameterization which was defined in (2.1) and consider the Merging Process. So, we replace in the (15) the states by the super-states. Formally, after regrouping terms, we arrive at the following expression:

$$\Phi(N) = \sum_{\sigma \in BD} P(\sigma) \cdot \mu(\sigma^{+}). \tag{16}$$

As it was explained earlier, the static probability $P(\sigma)$ may be interpreted as the probability that the Merging Process is in the super-state σ at moment t_0. Emphasize that we say "at moment t_0" but not "before the moment t_0". The reason for this is that in comparison with the UP-state, the border state (super-state) is not absorbing, i.e. our process may "jump" to the next (UP) state. Denote by $\xi(\sigma)$ the random time needed for the MP process to reach the super-state σ. We denoted earlier by $\xi(N)$ the random time for reaching the UP state. We have now for the $P(\sigma)$ the following formula:

$$P(\sigma) = P\{\xi(\sigma) \le t_0\} - P\{\xi(N) \le t_0\}. \tag{17}$$

The latter expression means that the MP enters the border state σ before t_0 and remains there at least up to t_0. Introducing now the trajectories (in the same manner as it was in (2.3)), we get the following formula:

$$\Phi(N) = \sum_{u \in U} p(u)(P\{\xi(\sigma(u)) \le t_0 \mid u\} - P\{\xi(N) \le t_0 \mid u\}) \cdot \mu(\sigma(u)^{+}). \tag{18}$$

We can do it because in the case of all-terminal connectivity, each trajectory u determines the unique border state $\sigma(u)$ (this does not take place for the s-t connectivity). So the sum in (18) has the form of (11) and therefore allows using the Monte Carlo scheme on turnip in the manner similar to (2.3): we simulate the trajectories with the probabilities $p(u)$ defined by (9) and compute the probabilities in brackets using convolutions (10).

4. Conclusions

As mentioned, the above described approach was first suggested in [1] and developed later in the works [2-6]. It does a very successful use of the network combinatorics to develop a highly efficient Monte Carlo scheme. This approach has various aspects and problems: combinatorial algorithms on graphs, various computational problems (for example, calculating long convolutions), efficiency of the method in some cases, etc. All these aspects were examined in referreed works in exact and formal manner. Our first purpose was to explain the idea of the approach itself. We had no intention of reviewing formally complicated problems and technicalities.

Another purpose was to describe briefly the main advantages of the turnip, some of which were mentioned in this paper. Note that numerous simulation results given in [2, 3, 6] confirm the high performance of the turnip approach. The advantages are the following:

(a) The MP method (turnip) *eliminates in principle* the rare event phenomenon;

(b) The relative error of the method is bounded uniformly with respect to time;

(c) A single simulation *run* (a trajectory over the turnip) serves for as many values of t as needed, which means that in one Monte Carlo trial we can calculate the reliabilities of networks with same topologies and *different* vectors of edge down-probabilities;

(d) The turnip is very computationally efficient, especially for dense networks;

(e) For certain combinatorial cases of the networks, the turnip allows to carry out the algorithm complexity analysis;

(f) The "turnip" approach has many applications: estimation of T_{UP} and T_{DOWN}, Availability, differentiation formula for reliability (the gradient), etc.

(g) The approach is applicable to evaluating non-static (dynamic) reliability measures;

(h) The approach may be generalized from networks to binary monotonic systems.

We consider this approach to be very promising and further investigation in this field may yield many results in various directions.

References

1. Lomonosov, M.V., *Bernoulli scheme with closure*, Problems of Information Transmission (USSR), 10, 1974, 73-81.
2. T. Elperin, I. Gertsbakh, and M. Lomonosov, *Network reliability estimation using graph evolution models*, IEEE Transactions on Reliability, R-40, 1991, 572-581.
3. T. Elperin, I. Gertsbakh, and M. Lomonosov, *An Evolution Model for Monte Carlo Estimation of equilibrium Network Renewal parameters*, Probability in the Engineering and Informational Sciences, 6, 1992, 457-469.
4. M. Lomonosov, *On Monte Carlo estimates in network reliability*, Probab.Eng. Inf. Sci, 8, 1994, 245-264.
5. M. Lomonosov, Y. Shpungin, *Combinatorics of reliability Monte Carlo*, Algorithms&Complexity, 1999, 329-343.
6. Shpungin Y., *Combinatorial and Computational Aspects of Monte Carlo Estimation of Network Reliability*, PhD Thesis, Dept. of Mathematics and Computer Science, Ben-Gurion University of the Negev, 1997.
7. Gnedenko, B.V., ed., *Mathematical Aspects of Reliability Theory*, Moscow, "Radio I Sviaz" (in Russia), 1983.

Session 3

Computing Mathematics

EXACT ANALYTICAL SOLUTION AND SOME OF ITS APPROXIMATIONS FOR TWO-DIMENSIONAL SYSTEMS WITH RECTANGULAR FIN

Andris Buikis[1], Margarita Buike[1] and Sharif Guseinov[1, 2]

[1] Institute of Mathematics, Latvian Academy of Sciences and Latvian University,
1 Akademijas Sq., Riga LV-1524, Latvia
Ph: (+371)- 7225674, Fax (+371)-7227520, E-mail: buikis@latnet.lv, sharif@one.lv

[2] Transport and Telecommunication Institute
1 Lomonosova str., Riga LV-1019 Latvia
Ph.: (+371)-7100598, Fax (+371)-7100660, sharif@one.lv

Abstract

The paper describes a two-dimensional heat transfer in the element of a periodic system with a rectangular fin by steady-state conditions. Firstly we solve the system of ordinary differential equations, which is obtained from the two-dimensional problem statement by the method of conservative averaging, and we receive an approximate solution. Further the well known one-dimensional statement of problem is examined. Finally we receive the exact solution of two-dimensional problem by using Green functions for two rectangles (the wall and the fin). This solution is obtained in form of Fredholm's integral equation of the second kind.

Key words: Extended surface, Heat exchange, Analytical solutions

Introduction

Systems with fins play very important role in various branches of engineering, including aircraft and aerospace [1], [2]. Cooling and heating systems with fins are widely exploited in various technical facilities. Therefore the modelling of heat transfer in these systems is very actual problem. In our previous publications we have constructed appropriate two-dimensional solution of this problem [3]-[5]. In this paper we construct the exact analytical solution. Traditionally this problem is solved in the one-dimensional approximation [6], [7].

1. The Statement of the Two-Dimensional Problem

In this paper we consider a periodical system with a rectangular fin whose elements, in dimensionless arguments, are a wall $\{x \in [0, \delta], y \in [0,1]\}$ and a fin $\{x \in [\delta, \delta + l], y \in [0, b]\}$. The steady-state heat process in such system is described as follows [3]-[5] (here $U_0(x, y)$ $(U(x, y))$ is a dimensionless temperature of the wall (fin)).

The heat transfer for the wall is given by the equation:

$$\frac{\partial^2 U_0}{\partial x^2} + \frac{\partial^2 U_0}{\partial y^2} = 0, \quad x \in (0, \delta), y \in (0,1), \tag{1}$$

for the fin – by the equation:

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0, \quad x \in (\delta, \delta + l), y \in (0, b). \tag{2}$$

On the left side of the wall the heat exchange with the environment at the dimensionless temperature $\Theta = 1$ is given:

$$\frac{\partial U_0}{\partial x} + \beta_0^0 (1 - U_0) = 0, \quad x = 0, y \in (0,1). \tag{3}$$

On the right side of the wall the heat exchange with the environment at the dimensionless temperature $\Theta = 0$ is given:

$$\frac{\partial U_0}{\partial x} + \beta_0 U_0 = 0, \quad x = \delta, y \in (b,1). \tag{4}$$

The heat exchange with the same environment is given on the side and on the tip of the fin:

$$\frac{\partial U}{\partial y} + \beta U = 0, \quad x \in (\delta, \delta + l), y = b, \tag{5}$$

$$\frac{\partial U}{\partial x} + \beta U = 0, \quad x = \delta + l, y \in (0,b). \tag{6}$$

We have the symmetry conditions on the top and the bottom of the wall and on the centre line of the fin:

$$\left.\frac{\partial U_0}{\partial y}\right|_{y=0} = \left.\frac{\partial U_0}{\partial y}\right|_{y=1} = 0, \quad x \in (0, \delta), \tag{7}$$

$$\left.\frac{\partial U}{\partial y}\right|_{y=0} = 0, \quad x \in (\delta, \delta + l). \tag{8}$$

On the line between the wall and the fin $x = \delta$ the conjugation conditions are given:

$$U_0\big|_{x=\delta-0} = U\big|_{x=\delta+0}, \tag{9_1}$$

$$\beta \left.\frac{\partial U_0}{\partial x}\right|_{x=\delta-0} = \beta_0 \left.\frac{\partial U}{\partial x}\right|_{x=\delta+0}. \tag{9_2}$$

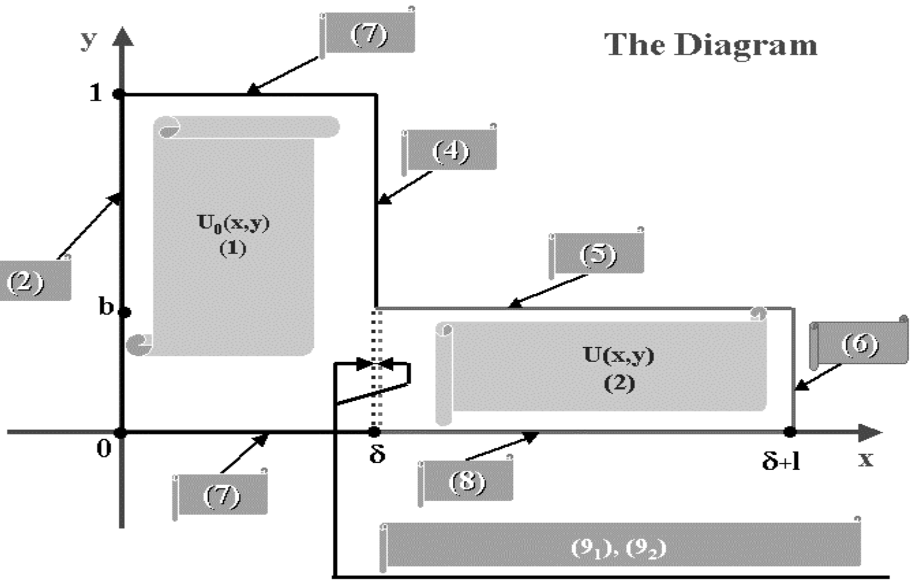
In the mathematical formulation (1)-(9_2) we use following dimensionless parameters:

$$\beta_0^0 = \frac{h_0(B+R)}{k_0}, \quad \beta_0 = \frac{h\ (B+R)}{k_0}, \quad \beta = \frac{h(B+R)}{k}, \tag{10}$$

$$\delta = \frac{\Delta}{B+R}, \quad l = \frac{L}{B+R}, \quad b = \frac{B}{B+R},$$

where $k(k_0)$- heat conductivity coefficient for the fin (wall), $h(h_0)$- heat exchange coefficient for the fin (wall), $c(c_0)$ - specific heat of the fin (wall), 2B – width of the fin, L – its length, Δ- width of the wall, 2R – distance between two fins.

Geometrically, the statement of the (1)-(9_2) problem shows up as follows:



2. The Approximate Analytical Solution by the Conservative Averaging Method

Continuing our previous publications [3]-[5] we approximate the solution in the fin in the form

$$U(x,y) = f_0(x) + (e^{\rho y} - 1) f_1(x) + (1 - e^{-\rho y}) f_2(x), \quad \rho = b^{-1}, \tag{11}$$

where $f_i(x),\ i = 0,1,2$ are three unknown functions.

First of all, we introduce the integral average value of function $U(x,y)$ in the y - direction:

$$u(x) = \rho \int_0^b U(x,y) dy. \tag{12}$$

Then, as in our papers [4], [5] we can represent the solution for the fin in the form:

$$U(x,y) = u(x)\Phi(y), \tag{13}$$

where

$$\Phi(y) = \frac{\sinh(1) + \beta b[\cosh(1) - \cosh(\rho y)]}{\sinh(1) + \beta b[\cosh(1) - \sinh(1)]}. \tag{14}$$

The function $u(x)$ is the solution of the ordinary differential equation ($\mu^2 = \beta\rho\Phi(b)$):

$$\frac{d^2u}{dx^2} - \mu^2 u = 0, \quad x \in (\delta, \delta + l) \tag{15}$$

with the boundary condition at $x = \delta + l$:

$$\frac{du}{dx} + \beta u = 0. \tag{16}$$

The solution of equation (15) with boundary condition (16) has the form:

$$u(x) = C_1(\mu_1 \exp(\mu x) + \exp(-\mu x)), \ \mu_1 = \frac{\mu - \beta}{\mu + \beta}\exp(-2\mu(\delta + l)) \tag{17}$$

with one free constant C_1. We will determine this constant later from the conjunction of the solutions for the two parts of the wall and the fin.

For the wall we express the exponential approximation in the x - direction in the form:

$$U_0(x,y) = g_0(y) + [e^{d(\delta-x)} - 1]g_1(y) + [1 - e^{d(x-\delta)}]g_2(y) \tag{18}$$

with $d = \delta^{-1}$ and we introduce the integral average value again, but of the function $U_0(x,y)$ and now in the x - direction:

$$u_0(y) = d\int_0^{\delta} U_0(x,y)dx\,. \tag{19}$$

The boundary condition (3) and equality (19) allow us to exclude $g_1(y)$ and $g_2(y)$ from representation (18):

$$g_i(y) = (-1)^i[-a_i g_0(y) + b_i u_0(y) - d_i]\,.$$

Here (see [4], [5]) $a_i = \dfrac{A_i}{K_1},\ b_i = \dfrac{B_i}{K_1},\ d_i = \dfrac{D_i}{K_1},\ i = 1{,}2$,

$$K_1 = e^{-1}[2 + \beta_0^0\delta(e-1)(3-e)],$$

$$A_1 = e^{-1}(1 + \beta_0^0\delta(e-2)),\ A_2 = e + \beta_0^0\delta,$$

$$B_1 = e^{-1}(1 + \beta_0^0\delta(e-1)),\ B_2 = e\ \ + \beta_0^0\delta(e-1)),$$

$$D_1 = e^{-1}\beta_0^0\delta,\ D_2 = \beta_0^0\delta(e-2).$$

Expression (18) for $U_0(x,y)$ can be rewritten in following form now:

$$\begin{aligned} U_0(x,y) &= [1 + (e^{d(\delta-x)} - 1)a_1 - (1 - e^{d(x-\delta)})a_2]g_0(y) + \\ &[(1 - e^{d(x-\delta)})b_2 - (e^{d(\delta-x)} - 1)b_1]u_0(y) + (e^{d(\delta-x)} - 1)d_1 - (1 - e^{d(x-\delta)})d_{2.} \end{aligned} \tag{20}$$

For the upper part of the wall b<y<1 we use the boundary condition (4) to get the connection between functions $g_0(y)$ and $u_0(y)$:

$$g_0(y) = b_0 u_0(y) - d_0\,. \tag{21}$$

The coefficients b_0 and d_0 are given by the expressions:

$$b_0 = \frac{B_0}{K_0},\ d_0 = \frac{D_0}{K_0},\ K_0 = A_2 - A_1 + \beta_0\delta K_1,\ B_0 = B_2 - B_1,\ D_0 = D_2 - D_1\,.$$

Thus we have expressed the solution for upper part of the wall trough the function $u_0(y)$:

$$\begin{aligned} U_0(x,y) &= [b_0 + (a_1 b_0 - b_1)(e^{d(\delta-x)} - 1) + (b_2 - a_2 b_0)(1 - e^{d(x-\delta)})]u_0(y) - \\ &d_0 + (d_1 - a_1 d_0)(e^{d(\delta-x)} - 1) + (a_2 d_0 - d_2)(1 - e^{d(x-\delta)}). \end{aligned} \tag{22}$$

In this representation function $u_0(y)$ is the solution of the ordinary differential equation

$$\frac{d^2u_0}{dy^2} - k^2u_0 + \Theta_2 = 0, \quad b < y < 1 \tag{23}$$

with boundary condition at y=1:

$$\frac{du_0}{dy} = 0. \tag{24}$$

Parameters k^2 and Θ_2 are given by expressions:

$$k^2 = 2\delta^{-2}K_0^{-1}[(\beta_0 + \beta_0^0)\delta\sinh(1) + 2\beta_0\beta_0^0\delta^2(\cosh(1)-1)],$$

$$\Theta_2 = \delta^{-2}[(d_1 - a_1 d_0)(e-1) + (d_2 - a_2 d_0)(1-e^{-1})].$$

The solution of the problem (23), (24) is

$$u_0(y) = C_2\cosh(k(1-y)) + U_2, U_2 = k^{-2}\Theta_2 \tag{25}$$

with a free constant C_2.

For the lower part of the wall we use conjugation conditions $(9_1), (9_2)$. We have from expressions (13) and (20):

$$g_0(y) = u(\delta)\Phi(y),$$

$$\left.\frac{\partial U_0}{\partial x}\right|_{x=\delta-0} = \frac{\beta_0}{\beta}\Phi(y)\left.\frac{du}{dx}\right|_{x=\delta+0}.$$

It means that we have following representation for the lower part of the wall:

$$\begin{gathered} U_0(x,y) = [1 + (e^{d(\delta-x)} - 1)a_1 - (1 - e^{d(x-\delta)})a_2]u(\delta)\Phi(y) + \\ [(1-e^{d(x-\delta)})b_2 - (e^{d(\delta-x)} - 1)b_1]u_0(y) + (e^{d(\delta-x)} - 1)d_1 - (1 - e^{d(x-\delta)})d_2. \end{gathered} \tag{26}$$

Finally we have got non-homogenous ordinary differential equation of second order:

$$\frac{du_0}{dy^2} - \lambda^2 u_0 + C_1\Theta_{3,0} - C_1\Theta_{3,1}\cosh(\rho y) + D_3 = 0, \quad 0 < y < b \tag{27}$$

with constant coefficients:

$$\lambda^2 = \frac{\beta_0^0(e-1)^2}{e\delta K_1}, \; D_3 = \frac{2\lambda^2}{(e-1)^2}, \; \Theta_{3,i} = \Theta_3\Phi_i, \; i = 0,1,$$

$$\Theta_3 = \frac{(A_1 e - A_2 e^{-1})(1+\mu_1)}{\delta^2 K_1} - \frac{\beta_0\mu}{\beta\delta}(1-\mu_1),$$

$$\Phi_1 = 1/(\sinh(1)/(\beta b) + \cosh(1) - \sinh(1)), \Phi_0 = (\sinh(1)/(\beta b) + \cosh(1))\Phi_1$$

and boundary condition at $y = 0$:

$$\frac{du_0}{dy} = 0. \tag{28}$$

The solution of ordinary differential equation (27) with boundary condition (28) for $\rho \neq \lambda$ is:

$$u_0(y) = C_3\cosh(\lambda y) + C_1(U_3 + U_4\cosh(\rho y)) + d_3 \tag{29}$$

and

$$u_0(y) = C_3 \cosh(\lambda y) + C_1(U_3 + \Theta_{3,1} \frac{y \sinh(\rho y)}{2\rho})) + d_3 \tag{29'}$$

for $\rho = \lambda$.

Here:

$$U_3 = \frac{\Theta_{3,0}}{\lambda^2}, U_4 = \frac{\Theta_{3,1}}{\rho^2 - \lambda^2}, d_3 = \frac{D_3}{\lambda^2}.$$

For the determination of three free constants $C_i, i = 1,2,3$ we make three natural assumptions; two of them are for the function $u_0(x)$ in point $y = b$:

$$u_0(b-0,t) = u_0(b+0,t), \quad \left.\frac{du_0}{dy}\right|_{y=b-0} = \left.\frac{du_0}{dy}\right|_{y=b+0}. \tag{30}$$

and the third requirement is continuity of temperature of the fin and the upper part of the wall in the corner $\{x = \delta, y = b\}$:

$$u(\delta)\Phi(b) = b_0 u_0(b) - d_0. \tag{31}$$

These conditions give three linear algebraic equations for the determination the three free constants [4], [5]. The second and the third equations are given for the case $\rho \neq \lambda$ only:

$$\begin{gathered} C_1(\mu_1 \exp(\mu\delta) + \exp(-\mu\delta))\Phi(b) - C_2 b_0 \cosh(k(1-b)) = U_2 b_0 - d_0, \\ C_2 \cosh(k(1-b)) - C_1[U_3 + U_4 \cosh(1)] - C_3 \cosh(\lambda b) = d_3 - U_2, \\ C_1 U_4 \rho \sinh(1) + C_2 k \sinh(k(1-b)) + C_3 \lambda \sinh(\lambda b) = 0. \end{gathered} \tag{32}$$

It is easy task to solve this system of linear algebraic equations. It means we have solved the two-dimensional problem (1)-(9_2): for the given constants $C_i, i = 1,2,3$ uniquely are determined the functions $u(x)$ and $u_0(x)$; then from equations (13), (22), (26) is possible to calculate the two-dimensional temperature fields in the wall and the fin.

3. The One-Dimensional Statement of Problem

A typical statement of approximate one-dimensional model for the periodical system with rectangular fin is following [6], [7]:

$$\frac{d^2 U_0(x)}{dx^2} = 0, \quad x \in (0, \delta), \tag{33}$$

$$b \frac{d^2 U(x)}{dx^2} - \beta U(x) = 0, \quad x \in (\delta, \delta + l), \tag{34}$$

$$\frac{dU_0}{dx} + \beta_0^0 (1 - U_0) = 0, \quad x = 0, \tag{35}$$

$$\frac{dU}{dx} + \beta U = 0, \quad x = \delta + l, \tag{36}$$

$$U_0\big|_{x=\delta-0} = U\big|_{x=\delta+0}, \tag{37}$$

$$\beta\left[\frac{dU_0}{dx}+\beta_0(1-b)U_0\right]\Bigg|_{x=\delta-0}=\beta_0 b\frac{dU}{dx}\Bigg|_{x=\delta+0}. \tag{38}$$

It is easy to see, that this one-dimensional statement follows from our two- dimensional formulation (1)-(9_2), if we assume independence of both temperatures – functions $U(x,y)=U(x), U_0(x,y)=U(x)$ from second argument y.

4. The Exact Solution of Two-Dimensional Problem

We will seek this solution by the method of Green function. To find the exact solution of the problem (1)-(9_2) we will represent this problem geometrically decomposed in two rectangles: one of them is the wall, the second is the fin. We will start with the wall. The new boundary condition on the right side of the wall can be written in the following form (this form arise from the boundary condition (4) and the combination of the conjugation conditions (9_1), (9_2)):

$$\frac{\partial U_0(x,y)}{\partial x}\Bigg|_{x=\delta-0}+\tilde{\beta}(y)U_0(x,y)\Big|_{x=\delta-0}=F_0(y), \tag{39}$$

$$\tilde{\beta}(y)=\begin{cases}\dfrac{1}{\beta} & if \quad 0<y<b,\\ \beta_0 & if \quad b<y<1,\end{cases} \tag{40}$$

$$F_0(y)=\begin{cases}\dfrac{\beta_0}{\beta}\dfrac{\partial U(x,y)}{\partial x}\Bigg|_{x=\delta+0}+\dfrac{1}{\beta}U(x,y)\Bigg|_{x=\delta+0} & if \quad 0<y<b,\\ 0 & if \quad b<y<1.\end{cases}$$

The solution of this problem for the wall has well known form:

$$U_0(x,y)=\beta_0^0\int_0^b G_1^0(x,y;0,\eta)d\eta+\beta_0^0\int_b^1 G_2^0(x,y;0,\eta)d\eta+$$
$$+\int_0^b\left[\frac{\beta_0}{\beta}\frac{\partial U(x,y)}{\partial x}\Bigg|_{x=\delta+0}+\frac{1}{\beta}U(x,y)\Bigg|_{x=\delta+0}\right]G_1^0(x,y;\delta,\eta)d\eta, \tag{41}$$

where the functions $G_1^0(x,y;\xi,\eta)$ and $G_2^0(x,y;\xi,\eta)$ are Green's functions, and they have the following expressions:

$$G_1^0(x,y;\xi,\eta)=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\left\{(\tilde{\mu}_n)^2\tilde{C}_n^{(1)}(x,\xi)-(\beta_0^0)^2\tilde{C}_n^{(2)}(x,\xi)+2\beta_0^0\sin\tilde{\mu}_n(x+\xi)\right\}}{(\beta_0^0)^2(1+\delta+\beta)+(\tilde{\mu}_n\beta\beta_0^0)^2\left(\delta+\dfrac{1}{\beta_0^0}\right)+(\tilde{\mu}_n)^2\beta^2\left(\delta+\beta+(\tilde{\mu}_n\beta)^2\right)}\times$$
$$\times\frac{\left\{1+(\tilde{\mu}_n\beta)^2\right\}\cdot\left\{\cos\pi m(y+\eta)+\cos\pi m(y-\eta)\right\}}{\tilde{\mu}_n^2+(\pi m)^2},$$

$$\tilde{C}_n^{(1)}(x,\xi)=\cos\tilde{\mu}_n(x+\xi)+\cos\tilde{\mu}_n(x-\xi),$$
$$\tilde{C}_n^{(2)}(x,\xi)=\cos\tilde{\mu}_n(x+\xi)-\cos\tilde{\mu}_n(x-\xi),$$

$$G_1^0\left(x,y;\xi,\eta\right)=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\left\{\left(\tilde{\tilde{\mu}}_n\right)^2\tilde{\tilde{C}}_n^{(1)}(x,\xi)-\left(\beta_0^0\right)^2\tilde{\tilde{C}}_n^{(2)}(x,\xi)+2\beta_0^0\sin\tilde{\tilde{\mu}}_n(x+\xi)\right\}}{\left(\beta_0^0\right)^2\left(\beta_0+\delta\tilde{\tilde{\mu}}_n^{\,2}+\beta_0^{\,2}\right)+\tilde{\tilde{\mu}}_n^{\,2}\left(\beta_0+\delta\tilde{\tilde{\mu}}_n^{\,2}+\beta_0^{\,2}\delta+\left(\beta_0^0\right)^2\right)+\delta\left(\beta_0\beta_0^0\right)^2}\times$$
$$\times\frac{\left\{\tilde{\tilde{\mu}}_n^{\,2}+\beta_0^{\,2}\right\}\cdot\left\{\cos\pi m(y+\eta)+\cos\pi m(y-\eta)\right\}}{\tilde{\tilde{\mu}}_n^2+\left(\pi m\right)^2},$$

$$\tilde{\tilde{C}}_n^{(1)}(x,\xi)=\cos\tilde{\tilde{\mu}}_n(x+\xi)+\cos\tilde{\tilde{\mu}}_n(x-\xi),$$
$$\tilde{\tilde{C}}_n^{(2)}(x,\xi)=\cos\tilde{\tilde{\mu}}_n(x+\xi)-\cos\tilde{\tilde{\mu}}_n(x-\xi).$$

Here $\tilde{\mu}=\tilde{\mu}_n\left(n\in N\right)$ and $\tilde{\tilde{\mu}}=\tilde{\tilde{\mu}}_n\left(n\in N\right)$ are the positive roots of following transcendental equations accordingly:

$$\tilde{\mu}=\frac{1}{\delta}\cdot arctg\left(\frac{1+\beta\beta_0^0}{\tilde{\mu}^2\beta-\beta_0^0}\cdot\tilde{\mu}\right),\ \tilde{\tilde{\mu}}=\frac{1}{\delta}\cdot arctg\left(\frac{\beta_0+\beta_0^0}{\tilde{\tilde{\mu}}^2-\beta_0\beta_0^0}\cdot\tilde{\tilde{\mu}}\right)$$

Now we will consider the similar problem for the fin. The new boundary condition on the left side of the fin can be written in the similarly to the (39) form:

$$\left.\frac{\partial U(x,y)}{\partial x}\right|_{x=\delta+0}-\frac{1}{\beta_0}U(x,y)\bigg|_{x=\delta+0}=\frac{\beta}{\beta_0}\left.\frac{\partial U_0(x,y)}{\partial x}\right|_{x=\delta-0}-\frac{1}{\beta_0}U_0(x,y)\bigg|_{x=\delta-0},\quad 0<y<b.$$

The solution of this problem for the fin has well known form too:

$$U(x,y)=\int_0^b\left[\frac{1}{\beta}U_0(\xi,\eta)\big|_{\xi=\delta-0}-\frac{\beta}{\beta_0}\left.\frac{\partial U_0(\xi,\eta)}{\partial\xi}\right|_{\xi=\delta-0}\right]G(x,y;\delta,\eta)d\eta,\tag{42}$$

where the function $G(x,y;\delta,\eta)$ is the Green function having following expression:

$$G\left(x,y;\xi,\eta\right)=\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\left\{\left(\beta_0\cdot\mu_n\right)^2C_n^{(1)}(x,\xi)-C_n^{(2)}(x,\xi)-2\cdot\sin\mu_n(x+\xi-2\delta)\right\}}{\beta_0^{\,2}\cdot\left[\beta\mu_n^{\,2}+l\left(\beta\mu_n\right)^2+l\mu_n^{\,4}\right]+\mu_n^{\,2}\cdot\left[l-\beta_0\right]+\beta^2\left[l-\beta_0+\dfrac{1}{\beta}\right]}\times$$
$$\times\frac{\left\{\beta^2+\mu_n^{\,2}\right\}\cdot\left\{\cos\omega_m(y+\eta)+\cos\omega_m(y-\eta)\right\}\cdot\left\{\beta^2+\omega_m^{\,2}\right\}}{\left\{\beta+b\left[\beta^2+\omega_m^{\,2}\right]\right\}\cdot\left\{\mu_n^{\,2}+\omega_m^{\,2}\right\}},$$

$$C_n^{(1)}(x,\xi)=\cos\mu_n(x+\xi-2\delta)+\cos\mu_n(x-\xi),$$
$$C_n^{(2)}(x,\xi)=\cos\mu_n(x+\xi-2\delta)-\cos\mu_n(x-\xi).$$

Here $\mu=\mu_n$ $\left(n=1,2,3...\right)$ and $\omega=\omega_m$ $\left(m=1,2,3...\right)$ are positive roots of the following transcendental equations accordingly:

$$\mu=\frac{1}{l}\cdot arctg\left(\frac{\beta\beta_0-1}{\mu^2\beta_0+\beta}\cdot\mu\right),\omega=\frac{1}{b}\cdot arctg\left(\frac{\beta}{\omega}\right).$$

Thus with the formula (41) we represent the solution $U_0(x,y)$ for the wall through the solution $U(x,y)$ for the fin. For the solution $U(x,y)$ we have a similar formula (42). Finally we have following Fredholm's integral equation of the second kind from these formulae:

$$\left[U_0(x,y)-\frac{\partial U_0(x,y)}{\partial x}\right]\Bigg|_{x=\delta-0}=\left[f(x,y)-\frac{\partial f(x,y)}{\partial x}\right]\Bigg|_{x=\delta-0}+$$

$$\int_0^b\Bigg[\left\{g_1(x,y;\eta)-\frac{\partial g_1(x,y;\eta)}{\partial x}\right\}\Bigg|_{x=\delta-0}\left\{g_2(x,y;\eta)-\frac{\partial g_2(x,y;\eta)}{\partial x}\right\}\Bigg|_{x=\delta-0}+$$

$$\left\{g_1(x,y;\eta)-\frac{\partial g_2(x,y;\eta)}{\partial x}\right\}\Bigg|_{x=\delta-0}\left\{g_2(x,y;\eta)-\frac{\partial g_1(x,y;\eta)}{\partial x}\right\}\Bigg|_{x=\delta-0}\Bigg]\times \tag{43}$$

$$\times\left[U_0(\xi,\eta)-\frac{\partial U_0(\xi,\eta)}{\partial \xi}\right]\Bigg|_{\xi=\delta-0}d\eta\ .$$

In the formula (43) we have introduced following designations:

$$g_1(x,y;\eta)=\int_0^b G_1^0(x,y;\delta,\zeta)\left[\frac{\beta_0}{\beta}\frac{\partial G(\xi,\zeta;\delta,\eta)}{\partial\xi}\Bigg|_{\xi=\delta+0}+\frac{1}{\beta^2}G(\xi,\zeta;\delta,\eta)\Bigg|_{x=\delta+0}\right]d\zeta\ ,$$

$$g_2(x,y;\eta)=\int_0^b G_1^0(x,y;\delta,\zeta)\left[\frac{\partial G(\xi,\zeta;\delta,\eta)}{\partial\xi}\Bigg|_{\xi=\delta+0}+\frac{1}{\beta_0}G(\xi,\zeta;\delta,\eta)\Bigg|_{x=\delta+0}\right]d\zeta\ ,$$

$$f(x,y)=\beta_0^0\int_0^b G_1^0(x,y;0,\zeta)d\zeta+\beta_0^0\int_b^1 G_2^0(x,y;0,\zeta)d\zeta\ .$$

The solution of equation (43) together with the formulae (41) and (42) gives the desired exact solution of two-dimensional problem (1)-(9_2).

References

1. Kern,D.Q, Kraus,A.D., Extended Surface Heat Transfer. – McGraw-Hill Book Company. 1972.
2. Manzoor,M., Heat Flow Through Extended Surface Heat Exchangers. Springer-Verlag: Berlin and New York, 1984.
3. Buikis,A., Two-dimensional solution for heat transfer in a regular fin assembly. Latvian Journal of Physics and Technical Sciences., No. 5, pp. 51-58, 1995.
4. Buikis,A., Buike,M., Closed Two-dimensional Solution for Heat Transfer in a Periodical System with a Fin. Proceedings of the Latvian Academy of Sciences. Section B, Vol.52, Nr.5, pp.218-222, 1998.
5. M.Buike. "Simulation of steady-state heat process for the rectangular fin- containing system", Mathematical Modelling and Analysis, 1999, vol. 4, pp. 33-43.
6. P.J.Heggs and P.R.Stones. "Effects of dimensions on the heat flow rate through extended surface", Trans. ASME Journal of Heat Transfer, **102**, 1980, 180-182.
7. A.S.Wood, G.E.Tupholme, M.I.H.Bhatti and P.J.Heggs. "Performance indicators for steady-state heat transfer through fin assemblies", Trans. ASME Journal of Heat Transfer, **118**, 1996, 310-316.

OPTIMISING FLIGHT ROUTE FOR CONFORMING TO THE PLANNED TIME OF ARRIVAL

Dalia Ziliene, Jonas Stankunas

Vilnius Gediminas Technical University
Antanas Gustaitis Aviation Institute
Address: 30 Rodunes kelias, Vilnius,LT – 2038, Lithuania
Phone: +370 5 2745808, Fax: +370 5 2745808 E-mail: D.Ziliene@one.lt E-mail: just@ai.vtu.lt

The work deals, in some detail, with a trajectory of flight, which is calculated by using the route plan, i.e. the parameters are identified, elimination of which enables to set the most optimal trajectory of flight, and remodeling can be made in accordance with the changing situation. The goal of this data processing is to make a maximally precise forecast of the arrival time at the airport of destination, which would significantly smooth the functions of the air traffic control in tacking the issues of space use and regulating intensive flows of aircraft, also give the precise information to land services on the arrival of aircraft, which would, in their own turn, be able to regulate flows of passengers in the airports better.

Key words : air traffic intensity, flight pats, flight profile, time of arrival

Introduction

Air transport is one of the fastest and safest ways of transportation. Yet air transport intensity is constantly growing with time alongside with increasing demand for air transportation services. Each year millions of aircrafts fly through European airspace. Air traffic intensity reaches up to over 100 aircrafts per hour in busier air traffic control regions and one air traffic controller manages about 20 planes simultaneously. In order to regulate flight intensity, airplane route should be optimised to ensure that the planned time of arrival is followed. The study herein analyses the impact of air traffic intensity growth on reasons of airplane time of arrival deviation from the planned time.

Prognoses on Air Traffic Intensity

Analysis of statistical data on passenger and cargo flow on international airlines for the past five years and future prognoses show the growth of the said rates. Based on that, air traffic intensity in Europe also increases. Figure 1 shows the percentage of minimum and maximum growth of forecasted air traffic intensity given by several European air traffic control centres for the years 1999 till 2005. [1]

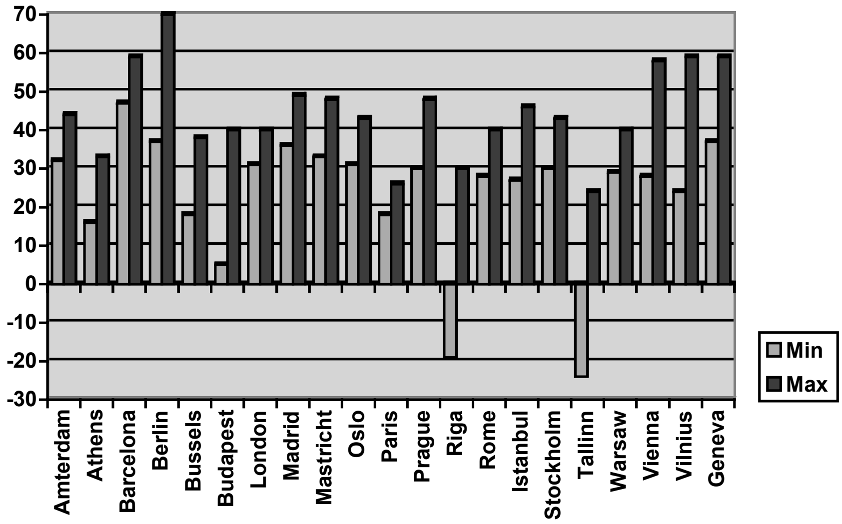


Figure 1. *Graphical representation of forecasted air traffic intensity expressed in percentage for the years 1999 – 2005 given by several European air traffic control centers*

Forecast of air traffic intensity for the years 1999 – 2005 for major European cities are shown in Figure 2. [1]

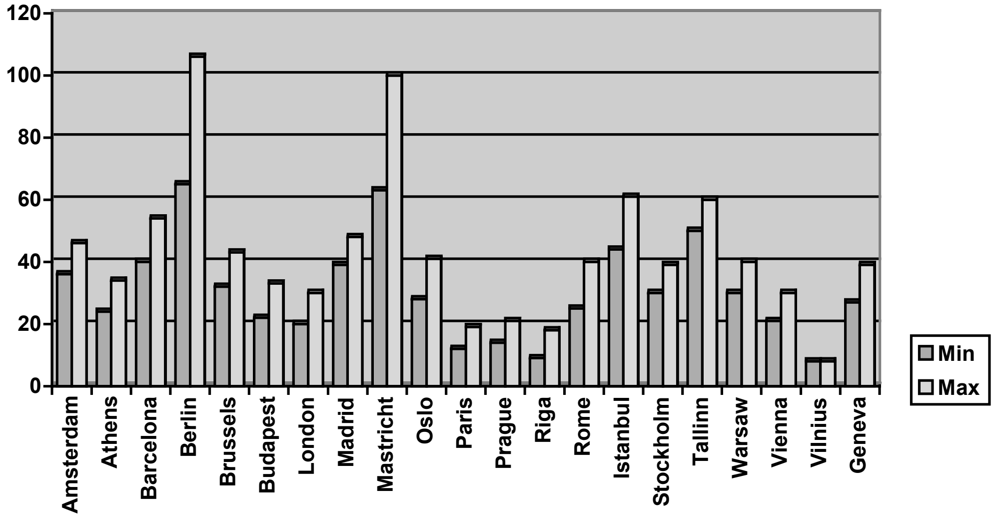


***Figure 2**. Graphical representation of air traffic intensity expressed in percentage for the years 1999 – 2005 for major European cities*

From charts shown in Figure 1 and 2 we can see that air traffic intensity growth for air traffic control centres and airports differs. It can be explained by the fact that forecasted intensity is influenced by various factors such as:

- Development of aerodrome and airway network;
- Aircraft types;
- Flight schedules;
- Degree of automation of air traffic control processes;
- Technical equipment used for flight handling;
- Prospects of development of certain country regions;
- Growth of Gross Domestic Product, etc.

Evaluation of all the above factors is a rather complicated task, which can be resolved with enough precision only by using methods of mathematical modelling. Such forecasts could be used for making decisions on development of the system regarding issues of airspace prospects and of air traffic flow management.

It is necessary to take into account that with growing air traffic intensity, delays appear in great air traffic intensity regions. Predicted ATFM (Air Traffic Flow Management) delays in Europe, expressed on average for a delayed flight, minutes on route and in an airport are shown in Figures 3 and 4. [1]

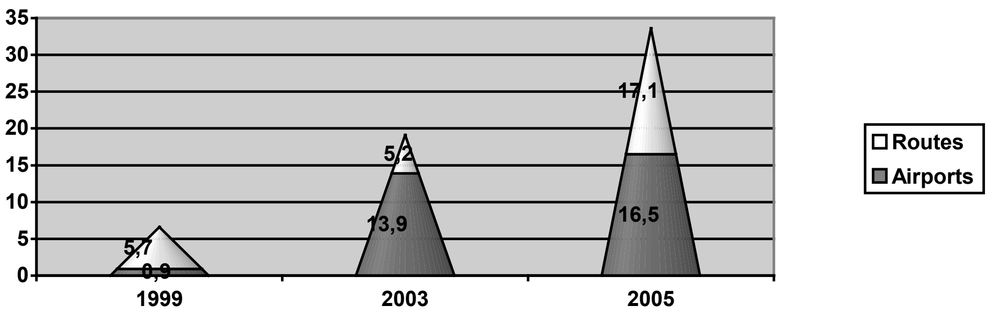


***Figure 3**. Forecasted ATFM delays in Europe: aircraft flights are operated according to routes – average growth of delays is expressed in percentage*

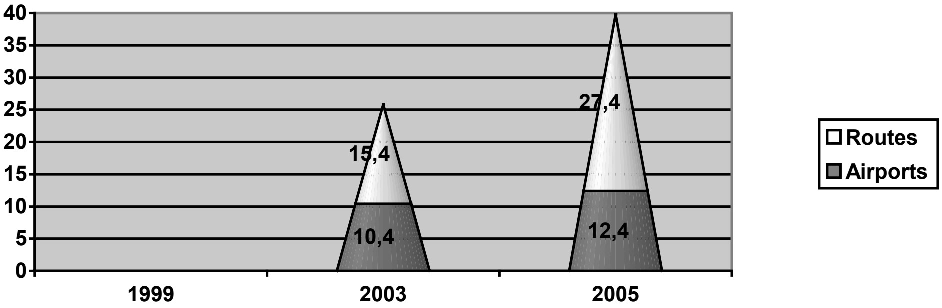


Figure 4. Forecasted ATFM delays in Europe: aircraft flights are operated on non-route basis (cutting flight track) – average growth of delays is expressed in percentage

As seen from Figures 3 and 4, forecasted longer delays, which are unfavourable both to air traffic control and airlines, and cannot bring comfort to passengers and cargo receivers and as a consequence become financially detrimental to airlines.

Therefore, to avoid air traffic delays on routes and airports, great work is being done to increase air traffic throughput.

Airplane Flight Paths

One of feasible ways to increase air traffic throughput dealing with flight flow planning is to consolidate factors influencing airplane flight and employ a possibility to measure time of plane flight on the chosen route as precisely as possible, which would enable planning and correcting flight flows in airspace and destination airport.

For this purpose airplane flight, i.e. its flight path and related elements must be examined. Airplane flight path could be divided the following way in respect of time [2]:

- 3D (dimension) path is airplane flight in airspace changing its vertical and horizontal position. That is the most often used and practised path concept.
- 3,5D (dimension) path is airplane flight in airspace from its take-off to its landing, including fixed time at certain points of the route. This concept of a path is actually implemented, when the system can present precise and exact time of flying above a certain point regarded as a control point.
- 4D (dimension) path is airplane flight in airspace from its take-off to landing, while flight information is monitored and time is evaluated during the entire flight.

Current and future path is comprehended solidly and is evaluated with the airplane day flight plan, which should have additional information that is not included in day or scheduled flight plans. It is real time information such as take-off – let-down profiles, day's speed and flight level, special chosen route for the day, flight restrictions – those can be restrictions related to air traffic control such as airplane flow distribution, activated temporary, separate territories, restrictions of air traffic control or airports, weather conditions. This additional information could help the general system to more precisely estimate airplane flight over the air traffic control centres on the route and time of arrival to the destination airport. Such huge amount of additional information, which may change with time, can correct airplane flight path, therefore the path must be monitored and verified in real time.

Airplane Flight Profile

Airplane flight profile is also related to its path. It is drawn by airplane following its route. The said profile is not permanent as it can vary with time. In most cases airplane flight profile is calculated based on flight plan information, i.e. by using the given flight levels and the route specified. See Figure 5.

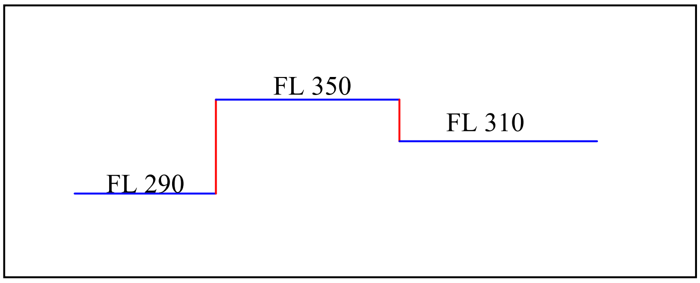


Figure 5. *Airplane stepped flight profile*

The above representation of airplane flight profile is called stepped profile [2] as no reckoning is made of vertical climb or descent, which is important in respect of time. In order to precisely calculate the time of airplane's climb or descent, airplane flight profile must conform to the real flight path as shown under Figures 6 and 7.

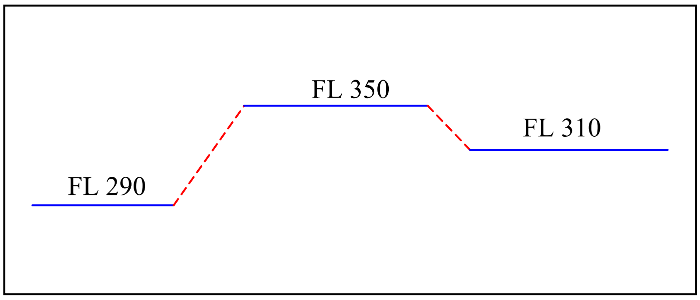


Figure 6. *Airplane flight profile*

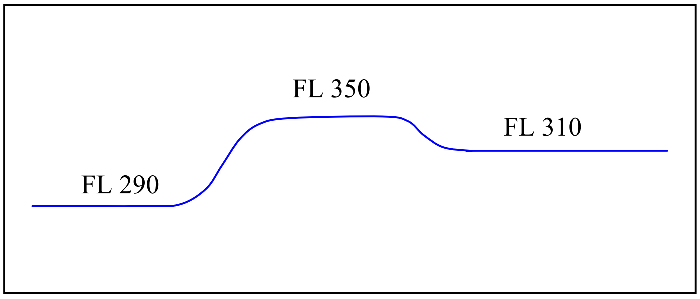


Figure 7. *Airplane real flight profile*

The above flight profile representation gives an opportunity to calculate a more precise airplane flight over air traffic control centres and the time of arrival to the destination airport.

Airplane Flight Speed

Results of calculating precise airplane flight time are influenced by airplane flight speed as well. To estimate speed, airplane flight information is used, which has the route and cruising speed indicated, which is regarded invariable and corresponding to airplane technical parameters. Therefore scheduled airplane flight time biases emerge. In order to estimate airplane flight time as precisely as possible, the following information must be available: forecasted wind speed and direction in airplane chosen flight route, real airplane flight speed and airplane flight speed upon changing flight profile, i.e. during airplane climb or descent.

Thus a more precise calculation on airplane time of arrival at the destination airport could be done by using the following formula:

$$t = D_1 / W + D_2 / V_p; \quad (1)$$

t – time of arrival,
D_1 – airplane chosen route not including climb and descent,
D_2 – distance covered by airplane during climb and descent,
W – airplane speed in regards to the ground (speed of the trip),
V_p – airplane horizontal speed component upon changing flight profile.

Airplane speed in regards to the ground can be calculated by using navigation speed triangle:

$$W = V \cos\alpha + U \cos\varepsilon ; \quad (2), [3]$$

W – airplane speed in regards of the ground (speed of the trip),
V – airplane speed in point of air mass,
U – speed of the wind,
α – deviation angle,
ε – vector's angle between the wind speed component and airplane's speed in respect of the ground component.

By combining formulae (1) and (2) we get formula (3), which enables us to calculate airplane's time of arrival at the destination airport.

$$t = D_1 / (V \cos\alpha + U \cos\varepsilon) + D_2 / V_p ; \quad (3)$$

With the help of formula (3), for the sake of an example we can analyse times of arrival of aircrafts B747 and B767 (Figure 8) at the destination airport following the chosen route. In order to make it easier to compare aircrafts' times of arrival, for both types of aircrafts chosen route 4000 kilometres and speed of the wind are regarded the same.

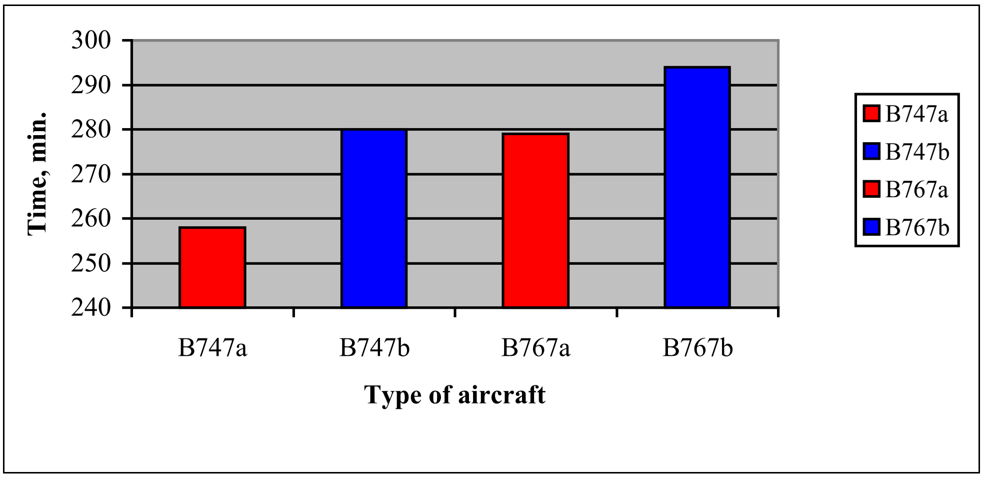


***Figure 8.** Calculated times of arrival of aircrafts B747 and B767 at destination airports*

$B747_a$ and $B767_a$ columns show the calculated time of aircrafts' arrival at the destination airport applying measurements of route distance and cruising speed. $B747_b$ and $B767_b$ columns show the calculated time of aircrafts' arrival at the destination airport employing formula (3). As seen from the chart, aircraft B747 times of arrival at the destination airport differ by 22 minutes, while the times of aircraft B767 – by 15 minutes. No doubt such a vast difference in times underlies huge delays either by following the route or at the destination airport itself. Deviations of 15-20 min. are very wide considering that airplanes land in an airport at an interval of 2-3 min.

After a more detailed analysis of formula (3) we can see that its components have a great impact on airplane arrival time at the destination airport. Big influence is also predetermined by forecast speed and direction of wind as well as aircraft's ability to, with respect to its technical parameters, take the necessary cruising flight level, i.e. aircraft's speed upon changing flight profile.

Thus, after entering additional information into day flight plans, it is possible to more precisely calculate airplane time of arrival at the destination airport, and that provides an opportunity to more accurately scheduled air traffic flows and manage airplane intensity in the air as well as in airports.

Conclusions

1. As air traffic intensity is growing, it is necessary to stiffen the conformity of real time of aircraft arrival to the flight plan of day. Advanced evaluation of real flight conditions' influence on aircraft flight time and related correction of day flight plans become inevitable.

2. As air traffic intensity is growing, it is necessary to include exact time of airplane flight over air traffic control centres on route and time of arrival at the destination airport in day flight plans.

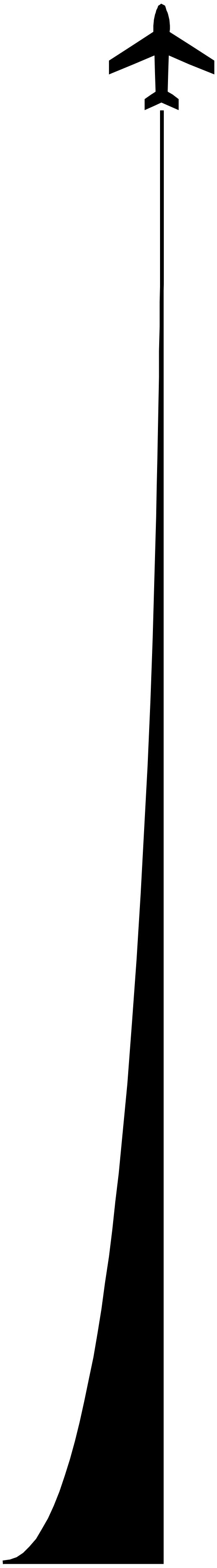
3. In order to exact the planned time of aircraft arrival it is necessary day flights plans have to be filed with real results of flight path calculation, which would assess aircraft climb and descent profiles, speeds of climb and descent, forecasted direction and speed of the wind in the chosen route and real aircraft speed.

References

1. Dalichampt M., Mahlich S. Medium – term Capacity Shortfalls 2003–2005 Inkluding national and supra-national Capacity Enhancement Plans. Eurocontrol Experimental Centre. 1999.

2. DPS-AADP ATC Advanced Data Processing 9-13 December 2002, Training Documentation.

3. Русол В.А., Киселев В.Ф., Крылов Г.О. и др. Справочник пилота и штурмана гражданской авиаций / Под ред. И.Ф.Васина.. Транспорт, 1988; 203 с.



Session 4

Management and Economics of Transport

IMPACT OF ELECTRONIC COMMERCE ON FORMING COMPETITIVE ADVANTAGES

Daiva Novikoviene

Vilnius Gediminas Technical University,
Department of Transport Management
Plytinės 27, LT-2016, Vilnius, Lithuania
Ph: (+370) 52744780. E-mail: daiva.novikoviene@post.5ci.lt

Abstract

The article analyses the impact of electronic commerce on forming competitive advantages in terms of the added value chain. The researches done by the following authors were reviewed and conclusions were made. The matrix of four alternative possibilities suggested by Rockart and Scott Morton was reviewed, as well as three main ways in which electronic commerce may be used to form competitive advantages by distinguishing direct benefit to the customers and to the company. Five overlapping Internet development stages in business distinguished by M.Porter were analysed. The matrix was provided on the importance of the organisation development and electronic commerce for the company activity. The following issues were analysed: Is it possible to use electronic commerce in seeking to make significant changes in the company's business processes so that the company could have a competitive advantage? Could the company focus on electronic commerce in order to improve its position on the market? Is it better to aim efforts at internal improvements in the same way as the company's activity is performed at the moment?

Key words: Electronic commerce, Competition, Competitive advantages, Internet, Market

1. Contribution of Different Authors in Analysing the Impact of Electronic Commerce on Forming Competitive Advantages of the Economic Entities

The analysis of the added value chain, with the aim of forming competitive advantages, received more attention from various scientists than the Porter model of competitive analysis. In the opinion of M.Porter (2001), Rockart (1984), Scott Morton (1984), Ives (1984), Learmonth (1984), in order to establish the fields where a company could apply electronic commerce to gain competitive advantages, we must analyse the company's added value chain and search for the possibilities to apply the EC in each stage of the cost value chain.

Seeking to describe the potential opportunities and competitive advantages arising due to the use of electronic commerce, Rockart and Scott Morton (1984, 1990) analysed the added value chain.

2. Impact of Electronic Commerce on Forming Competitive Advantages in Terms of the Added Value Chain

According to Rockart and Scott Morton (1990), the research of the stages of the added value chain is one of the methods to use the possibilities of electronic commerce more efficiently and to successfully develop the company's business. However, in practice the result of such research is appearing of new efficient technologies and their successful application in the company's business. When evaluating this approach of the authors, it must be remarked that the information services/technologies unit of the company must review its activity limits and actively participate in the business process, and maintain contacts with the customers and suppliers.

Electronic commerce may be applied in establishing relations with the customers and starting new kinds of business, e.g. managing information on the insurance indemnity which must be paid to the consumer by the insurance companies.

A thorough analysis of the added value chain allows to identify the areas where using of electronic commerce can give the highest benefit. When making such analysis, the manager seeking to establish the critical points where electronic commerce can be best applied must make a thorough analysis of all steps in the business process.

According to Rockart and Scott Morton (1984), in analysing possible use of electronic commerce in respect of the added value chain it is most important to answer the following questions:

- Is it possible to use electronic commerce in seeking to make significant changes in the business process of the company so that the company can have a competitive advantage?
- Could the company focus on electronic commerce in order to improve its position on the market? Would it be better to direct the efforts to internal improvements in the same way as the company's activity is being carried out at the moment?

The first question is very important since in certain industry sectors quite a number of possibilities exist for using electronic commerce in order to provide revolutionary new products able to change the industry sector substantially or to change the present attitude to production, purchasing, etc. This huge competitive "jump" can become very important in the company's activity. If such jump is not acceptable to one company, it can be used by others. Another alternative is that if it comes out that such an opportunity cannot be implemented, then the company should be focused on improving of the present business by using electronic commerce.

The second question makes a special emphasis on two methods: by improving the company's position on the market to improve the main internal operations of the company by reducing the expenses or improving the services (minimisation of the costs or differentiation of the products). At the moment most of the attention is drawn on using of the technology in order to improve the company's position on the market. Nevertheless, most companies have an opportunity to improve their main internal operations by using information technologies.
To summarise the use of electronic commerce in order to get competitive advantages it can be stated that the answers to these questions offer a matrix of four alternative possibilities.

	Position on the market	Internal operations
Structural changes	1	2
Traditional products and services	3	4

Figure 1. *Matrix of using electronic commerce to gain strategic advantages (Rockart and Scott Morton, 1984)*

As the matrix shows, possibilities 2 and 4 are related with the use of electronic commerce to form competitive advantages in separate stages of the added value creation, while possibilities 1 and 3 are related with using of electronic commerce to increase competitiveness in respect of the industry sector.

Generalising the Rockart and Scott Morton's approach (1984), we can state that the authors distinguished the major methods how electronic commerce can be used to form competitive advantages:

- To improve each value adding function
- To increase the costs of selecting customers and suppliers
- To create new fields of activity.

Rockart and Scott Morton (1984) were the first to try to define the possibilities of using electronic commerce to form competitive advantages. B.Ives and G.P.Learmonth (1984) further developed the research started by Rockart and Scott Morton (1984). In their works B.Ives and G.P.Learmonth (1984) analysed the model of thirteen stage life cycle of resources seeking to establish how the company can gain competitive advantages. Although the authors did not use the term "added value chain" in their works, however the model of the resources life cycle used by them is very closely related with the added value chain.

3. Using of Electronic Commerce as an Instrument for Gaining Competitive Advantages

The analysis of the added value chain or life cycle of the resources aimed at achieving the operational efficiency and functional efficiency, is closely related with the performance strategy of the companies. In the opinion of B.Ives and G.P.Learmonth (1984), the ways of using EC for gaining competitive advantages are shown in the table.

<table>
<tr><td></td><td></td><td>Use</td><td>Examples</td></tr>
<tr><td rowspan="8">Using electronic commerce in seeking competitive advantage</td><td rowspan="4">Direct benefit to the customers</td><td>Offering services</td><td>Purchasing at home by electronic means</td></tr>
<tr><td>Offering new products</td><td>"Mobile" sales of insurance policies</td></tr>
<tr><td>Creating of the distribution channel</td><td>Regional bank automates able to pay amounts</td></tr>
<tr><td>Offering other elements</td><td>Electronic medicine</td></tr>
<tr><td rowspan="4">Direct benefit to the company</td><td>Providing information (knowledge)</td><td>Integrating of external and internal marketing information</td></tr>
<tr><td>Reducing costs of the product</td><td>Integrated production and distribution systems with feedback and control</td></tr>
<tr><td>Reducing costs of the services</td><td>Capital for changing the work force</td></tr>
<tr><td>Organisational change</td><td>Automation of the office functions</td></tr>
</table>

Figure 2. *Using of electronic commerce as an instrument for gaining competitive advantages (Ives and Learmonth, 1984, Notowidigdo 1984)*

4. Importance of Electronic Commerce in the Company

In order to see how information technologies, including electronic commerce and Internet, influence the added value chain, we must evaluate the historic perspective of their use. Development of the Internet in business may be evaluated by expressing it in five overlapping stages each of which originated from the previous generation due to the progress in technology (M.Porter, 2001).

The earliest use of the Internet and electronic commerce in transport companies allowed to automate separate operations such as accepting of orders and accounting. Another stage is characterised by the automation of separate higher activities and fulfilment of additional functions such as human resources management, trading operations and product development of transportation services. The third stage accelerated by the Internet covers integration of various actions, e.g. joining of trading actions with order processing. Many actions were related due to such measures, like customer relations management and company's resources planning systems. The fourth stage, which is just starting, provides an opportunity to integrate the added value chain and the whole system of value, i.e. the complex of the value chains and the whole industry sector, including suppliers, distribution channels, and customers.

In the future fifth stage the commerce based on information technologies and Internet will be used not only to maintain relations in the value system of various actions and participants but also to optimise its operation in real time. The decisions will be made basing on the information from the joint actions and common organisations. The production tasks will be shared in accordance with the companies' capacities and reserves which can be received from suppliers. Since the first implementation phases of the fifth stage will cover a relatively simple optimisation of the resources supply, production, logistics and service operations, a more complete optimisation will in its turn cover product development, too.

The summary of the stages of applying electronic commerce in the company provided by M. Porter (2001) is shown in the figure.

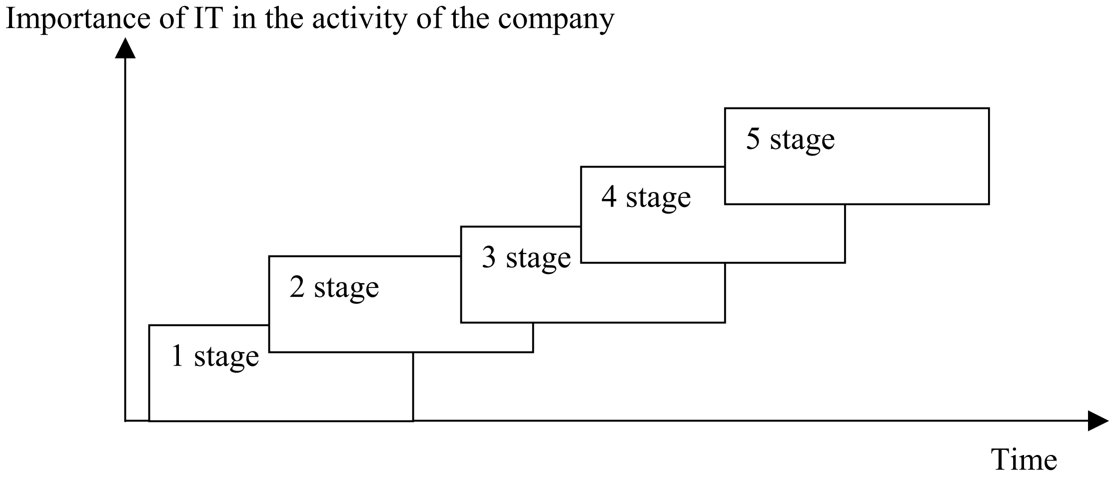


Figure 3. *The growth of importance of electronic commerce in the company's activity time-wise (M.Porter, 2001)*

Generalising of the importance of electronic commerce to the company time-wise allows to form a matrix, which evaluates the development of the organisation and the importance of electronic commerce for the company's activity. According to the reasoning provided by M.Porter, the company starting to apply electronic commerce in its activity is included in field 1 of the matrix and only after passing five stages it gets to field 4. However, the companies the managers of which think that electronic commerce has a strategic importance to the company's activity, get into field 3 by attempting to implement both the first and the second stage in their activity at the same time.

		Organisation	Development
		Young	Mature
EC importance	Little	1	2
In company's activity	Great	3	4

Figure 4. *Matrix of the importance of organisation development and electronic commerce in the company's activity*

5. Conclusions

The question whether electronic commerce is a competitive instrument in business remains important to many company managers. For many years technical experts have been trying to find out whether the computer systems of the company correspond to its needs. Unfortunately, these experts did not realise the business needs in terms of competition, as they did not realise that electronic commerce based on information technologies is part of the competitive business field. Despite of this main lack of the strategic direction, many companies, including transport sector, use electronic commerce to achieve their competitive advantage.

References

1. Porter, M. Competitive Strategy, The Free Press, 1982
2. Porter, M., Miller, V., "How Information Gives You Competitive Advantage", Harvard Business Review, July-August 1985
3. Scott Morton, M.S., "Strategy Formulation Methodologies and IT", in Kling, R.(Ed.), Computerization and Controversy, Academic Press, 1997.

DISTRIBUTION OF MARKET SHARES AMONG COMPETITORS IN LITHUANIAN TRANSPORT SECTOR

Jonas Lazauskas

Vilnius Gediminas Technical Univercity,
Department of Transport Management
Plytinės 27, LT-2016, Vilnius, Lithuania
Ph: (+370) 52744780. E-mail: jonlaz@ti.vtu.lt

Abstract

The article analyses the intensity of competition among Lithuanian companies in the transport sector and its evaluating. Intensity of competition is defined by using Herfindal's –Heshman's index. Distribution of market shares is defined among carriers of passengers by taxi cars, air transport companies, cargo carrying companies, and forwarding companies. The nature of the market share distribution is evaluated which shows the concentration degree of production and services in the Lithuanian transport sector. The market monopolisation level in the Lithuanian transport sector is defined.

Key words: Competition. Transport company. Services. Competition intensity. Market share

1. Introduction

During the recent decade increase of competition has been practically observed in the whole world. Not so long ago it was absent in many countries and sectors of economy. The market was secured and its dominating positions were clearly defined. Even in those areas where there was competition, it was not very strong. The increase of competition was suppressed by direct interference of the state.

Severe competition pushes away weak non-efficient companies and concentrates production or service providing in large companies. The tension of competition is conditioned by the increasing number of participants in international trading and service providing operations, increasing variety of products and services, constant improvement of technologies, etc. Very much influence on rapid and hardly predictable changes in this field has the progress of science and technology. It is not by chance that innovation of a product or a service today is becoming one of the basic factors of competitiveness. The increase of competition in its turn speeds up the characteristic processes: rapid renewing of its forms and methods, active creating of new competitive products, services and technologies, search for new outlets.

Competitiveness of a product or service is mostly conditioned by the ability to satisfy the consumer's needs: price, quality acceptable to the consumer, accessibility of a product or service, awareness to the consumer. Each of these factors is conditioned by many other factors. E.g.: the price is influenced by the production costs (which in their turn are conditioned by the working efficiency, etc.), distribution, advertising costs and other promotion costs. The price depends on the management and the pricing strategy of the company – it may be based on costs or value. The quality depends on the production technologies, the materials used, and on work organization.

The need for large financial resources when implementing a new production or providing new services is a critical factor restricting appearance of new companies on the market. On the one hand, it is financially difficult to establish a company, and on the other, circulating assets must be increased, obligations to the bank must be fulfilled, etc. This makes entering to the market harder.

With competitive products or services many companies still cannot efficiently use this advantage as they do not have experience in using the marketing complex: flexible policy on the product range and prices, effective sales stimulation methods, etc. The conjunctive situation is more and more burdened every year by extending of the market margins and entering of foreign companies with large experience in the conditions of strong competition into the market.

The factors having influence on the company's competitiveness are management, financial resources, staff qualification, modern technologies, convenient location of the company, experience on the market, flexibility in reacting to changes on the market, etc. Many Lithuanian businessmen very well realize the importance of management, finance, and implementation of modern technologies, however improving of staff qualification and training is a new process in Lithuania and far from many manages of the companies realize its importance.

Infrastructure of business services is a source of a competitive advantage and its absence or low level may be an important drawback of competitiveness. This is especially important to the transport, logistics, and telecommunications system. In Lithuania this general economic infrastructure is relatively well developed and is improving. Development of the transport service sector will depend on many circumstances. Good geographical location is only one important precondition for development, however the regulations for border crossing and transit control are also of no less importance as well as the investments of the state in this infrastructure. The state must privatize the services which, as the world's experience shows, are operating more efficiently when being private. The best quality and most rational services may be more rapidly ensured only by strong international companies. Certainly, this does not apply to the services which may be provided by small and medium companies. In this case it would be most perspective to use common services of foreign and local companies since the biggest effect is achieved this way.

Many small companies see competition as copying large competitors. This gives them self-confidence. However, following others means loosing any advantage. Absence of the competitive advantage is the straightest way to bankruptcy. Some companies that have an advantage do not take any measures to maintain it. Possession of a competitive advantage is to be seen as an accomplished fact or achieved goal and then another one must be searched for. The essence of the competitive struggle is not in the actions against the company but in taking over the competitors using the services of that company.

2. Distribution of Market Shares Among Competitors and Intensity of Competition

One of the most important characteristics of the competitive struggle is the degree of competitors' resistance in fighting for consumers and new gaps in the market. Due to the difficulties in direct evaluating of interaction of the competitive environment factors, the indirect measurement of competitive intensity is possible. It is based on evaluation of the after-effects of the actually controlled relationships. The results of the analysis allowed us to distinguish three aggregated actions showing the intensity of competition. The nature of the market share distribution among the competitor, market expansion rate, and market profitability must also be ascribed to such factors.

To evaluate the nature of the market share distribution the ratio is used showing the concentration degree of production and services in the sector. It allows defining the monopolization level on the market and is the opposite value to the competition intensity. The ratio used consists of four parts

(CR_4 – Concentration Ratio):

$$CR_4 = \frac{QR_1 + QR_2 + QR_3 + QR_4}{QR},$$

where QR_i is the sales amounts of the i[th] company;

QR – total sales amounts of the planned product range, LTL thou.; $QR_1 = \max\{QR_i\}, i = 1 \div n$;

$QR_2 = \max\{QR_i \setminus QR_1\}, i = 1 \div (n-1)$;

$QR_3 = \max\{QR_i \setminus QR_1, QR_2\}, i = 1 \div (n-2)$;

$QR_4 = \max\{QR_i \setminus QR_1, QR_2, QR_3\}, i = 1 \div (n-3)$.

Where n is the total number of companies selling certain products or services.

In other words, CR_4 shows the total market share of the first four companies. Those companies sell most products or services on the whole analyzed market. This ratio is used by the Department of Justice, USA to define the competition level on the trading market. If CR_4 is higher than 0,7 (75 per cent), then restrictions on company mergers are introduced, since such market is becoming a monopoly.

At the moment concentration ratios in the USA and France are calculated to 4,8,20,50,100 leading companies on the market, while in Germany, England, and Canada this is done to 3,6,10 companies. In the middle of the eighties, the eastern European countries started to use CR, as well as partially Hungary and Poland.

An important drawback of the concentration ratio is that it does not reflect the distribution of various market shares among the competitors. E.g., CR is the same and equals to 0,8 in two totally different market situations: 1) one company controls 77 per cent of the market, while other 23 companies have 1 per cent each; 2) 5 equivalent companies control 20 percent of the market each.

This drawback in a way is eliminated by the quadrate sum of the competitors' market shares – Herfindal's index:

$$Y_h = \sum D_i^2 \text{ or } Y_h = 10000 \sum D_i^2,$$

where Y_h is Herfindal's index ($0 < Y_h \leq 1$)

$$D_i = QR_i \setminus QR, \; i = 1, \ldots, n$$

D_i is the share of the i[th] company in all sales amount of the planned product range.

Herfindal's index increases when increasing concentration in the market and is equal to 1 in case of pure monopoly. When 100 equivalent companies are operating on the market, then $Y_h = 0{,}01$.

Since 1984 this index has been used in practice with certain Hershman's specifications in the USA government anti-monopoly activity. The specifications are related with the number of companies and parts used in formula 1. In defining Herfindal-Hershman's index the quadrate sums of the shares are defined only for the first 50 larger companies operating on the analyzed market. In case the value is higher than 0,18 then the competition intensity is low and the market concentration is high. Therefore, the government must interfere and monopolize the situation on the market. In case in this situation the companies merger leads to $Y_h = 0{,}005$ (50 points), then it is prohibited by the laws.

In order to define the market monopolization level among the road cargo carriers of the Republic of Lithuania, a research was made which showed that the market concentration level ratio CR_4 in 2001 stood at 0,385 and in 2002 it was 0,428, which shows low monopolization level among the Lithuanian cargo carriers. Having used Herfindal-Hershman's index during the research the quadrate sums of the shares are defined only for the first 50 larger companies operating in the analyzed market. If the value is higher than 0,18 then the competition intensity is low and the market concentration is high. The research showed that among 50 largest road cargo carriers of the Republic of Lithuania the values of Herfindal-Hershman's index in 2001 stood at 0,0658, and in 2002 – 0,0715. This shows that the competition intensity is high and the market concentration in low. When analyzing the cargo carrying companies of Lithuania it was defined that the cargo carrying related activity share of the analyzed companies amounts to 50-100 per cent.

Having made an equivalent research of the cargo forwarding companies of the Republic of Lithuania, it was established that the market concentration level ratio CR_4 in 2001 amounted to 0,426 and in 2002 it was 0,489, which shows low monopolization level among the Lithuanian cargo forwarding companies.

The research showed that among 50 largest cargo forwarding companies of the Republic of Lithuania the values of Herfindal-Hershman's index in 2001 stood at 0,1151, and in 2002 – 0,1088. This shows that the competition intensity is high and the market concentration in low. When analyzing the cargo forwarding companies of Lithuania it was defined that the cargo forwarding related activity share of the analyzed companies amounts to 60-100 per cent.

Having made an equivalent research of the taxi companies of the Republic of Lithuania, it was established that the market concentration level ratio CR_4 in 2001 amounted to 0,686 and in 2002 it was 0,729, which shows high monopolization level among the Lithuanian taxi companies. The research showed that among 50 largest taxi companies of the Republic of Lithuania the values of Herfindal-Hershman's index in 2001 stood at 0,1513, and in 2002 – 0,1555. This shows that the competition intensity is close to the limit when the market concentration is considered to be high.

A totally different situation is among air transport companies. 2 Lithuanian companies and 6 representatives of foreign companies in Lithuania are dominating on the market. Having calculated the market concentration level ratio, which in 2001 amounted to 0,871 and in 2002 – 0,849, we can state that there is a high monopolization level among air transport companies in Lithuania, and the market concentration is high. This is also illustrated by the values of Herfindal-Hershman's index, which in 2001 stood at 0,3788 and in 2002 – 0,3806.

As the formula shows, Y_h does not evaluate the companies' rank. This drawback is eliminated by Rozenblut's index (Y_r), which is calculated by considering the running number of the companies received having divided by rank from the maximum to the minimum (i):

$$Y_r = \frac{1}{2\sum\left(i * D_i\right) - 1}; i = 1,...,n .$$

Sometimes the entropy ratio (E) is used in evaluating the distribution of the shares. In this case the measurement of the share is done not by ranks but by making the share logarithms:

$$E = \sum D_i * \ln D_i; i = 1,...,n .$$

In principle, a totally different method giving equivalent results is the evaluating of the qualitative integration of Lorence's diagram – Ginny's ratio (G):

$$U_D = 1 - \frac{\sigma(D)}{D_{vid}} \text{ or } U_D = 1 - \frac{\sqrt{\left(1/n\right)\sum\left(D_i - D_{vid.}\right)^2}}{D_{vid}}; \text{ i = 1,...,n} .$$

Where U_d – competition intensity measured by evaluation of the similarity of the competitors' shares;

$\dfrac{\sigma(D)}{D_{vid}}$ – variation ratio of the competitors' market share (D_i);

$\sigma(D)$ – average quadrate deviation D_i;

D_{av} – arithmetic D_i average;

N – number of companies on the market.

It can be seen that the arithmetical average of the shares depends on the number of companies (n):

$$D_{vid} = \frac{1}{n} .$$

Therefore, $U_D = 1 - n\sqrt{(1/n)\sum_i \left(D_i - (1/n)\right)^2}$, i = 1,...,n.

3. Market Development and Competition Intensity Rates

U_D ratio is necessary but not enough since it does not evaluate the nature of the market development. We are talking about the characteristics of dynamic demand and supply, which are disclosed by the growth rates of the sales and services amounts. The speeded-up development of the market even in the situation of equal competition may eliminate many contradictions among the companies on their satisfaction in the increase rates. High growth rates, increasing demand and supply eliminate many problems, including competition. This happens because the market share for the company increases not on account of the competitors but due to the increased number of consumers or increased purchasing power. In such case the competition intensity falls.

On the other hand, such increase cannot go on forever. Due to many objective and subjective reasons, any market undergoes stagnation, depression or a small positional growth when the sales amounts increase due to the decreasing position of the competitors. In such case the competition intensity increases significantly. Such fact must be included into the complex evaluation of the competition intensity. The main difficulty is several meanings of the marginal values of the growth rates. The competition intensity in minimal (growth rates higher than 100 per cent) or maximal (growth rates lower than 100 per cent). Most situations describing the dynamics of certain product or service markets may be limited by two marginal values of growth rates in the sales amounts: 70 per cent and 140 per cent. In this range the following values can exist of the competition intensity evaluating the growth rates of the sales amounts in the analyzed market:

$$U_{TA} = 1 - \frac{T_A - 70}{140 - 70} = \frac{140 - T_A}{70} ,$$

where T_A is the annual growth rate of the sales amounts in the analyzed market without considering the inflation, %.

General scheme for defining U_{TA}:

$$\text{If } \begin{Bmatrix} T_A \geq 140\% \\ 70\% < T_A < 140\% \\ T_A = 70\% \end{Bmatrix}, \text{ then } \begin{Bmatrix} \Rightarrow 0 \\ U_{TA} = \left(140 - T_A\right)/70 \\ \Rightarrow 1 \end{Bmatrix}.$$

It is also important that when T_a is lower than 70 per cent (market change), then the competition intensity becomes lower. The given situation cannot be transferred to the operating market. It means that the products are not sold or services are not provided, or the market has experienced a strong economic shock. These cases must be analyzed separately.

Another important economic factor showing the competition intensity is the market profitability ratio (R_R), which is defined by the relation of general profit to sales amounts:

$$R_R = \frac{P}{Q} \ .$$

High market profitability is when the demand exceeds the supply. The situation allows the company to fulfill the goals without suppressing the competitors' interests. When the profitability gets lower, the situation becomes more inverted.

This tendency has an objective basis and is mentioned in the researches of many economists. A number of microeconomics theories use Lerner's ratio (*L*) to evaluate dominating of the company on the market.

$$L = \frac{K - MC}{K} \text{ or } L = -\frac{1}{E_K} \ ,$$

K – product price;
MC – marginal costs;
E_k – demand elasticity by price.

The ratio provided shows the possibility of the seller or service provider to make influence on the price of the product. The higher is Lerner's ratio, the less is the company dependent on the competitors, suppliers, consumers, etc. Having integrated this expression by all companies on the market, we would come to the aforementioned market profitability ratio. This way, regardless of the market profitability, R_R shows the company's competitive activity and the company's "freedom" degree in gaining profit. The higher is R_R, the lower is the pressure of the competitive environment, lower competition intensity. This conclusion can be expressed by the formula:

$$U_R = 1 - \frac{P}{Q} = 1 - R_R \ ,$$

U_R – competition intensity ratio evaluating the market profitability level.

When profitability is higher than 100 per cent, the value of U_R is closer to 0, and when the business is loss making, the U_R value is closer to 1.

General scheme for defining U_R:

$$\text{If } \begin{Bmatrix} R_R \geq 1 \\ 0 \leq R_R < 1 \\ R_R \leq 0 \end{Bmatrix}, \text{ then } \begin{Bmatrix} \Rightarrow 0 \\ U_R = 1 - R_R \\ \Rightarrow 1 \end{Bmatrix}.$$

It is important to mention that the factors provided cannot fully evaluate the impact of the competitive environment to the competition intensity. Speaking about U_D, U_{TA}, U_R, the attention was not focused on technologic innovations, product modification, change in market strategy, and other moments able to make significant changes in the nature of the competitive struggle. This was not done accidentally. After detailed analysis, changes in the competitive environment are reflected from the dynamics of the competitors, market shares, market profitability, and growth rates. This is confirmed by the competition practice and special competition researches.

Analysis by the competition intensity characteristics is convenient when analyzing the competition intensity in separate markets (its segments) and evaluating the attraction of the market. Beside the comparison base, such analysis allows to specify the analysis results of separate elements of the competitive environment.

The ratios U_D, U_{TA}, U_R can be generalized in the following way:

$$U_k = \sqrt[3]{U_D * U_{TA} * U_R} \ ,$$

U_k – generalized competition intensity ratio, $0 \le U_k \le 1$.

This way the aggregated, general evaluation of the company's competitive environment is received. Its main purpose is to measure the competition intensity. It would, however, be wrong to think that the whole situation of the analyzed activity can be described by the given characteristics. Like any other complicated process, the competitive struggle also requires a system of ratios. By U_k it is impossible to establish the conditions for receiving certain advantages over the competitors, since there is only indirect relation with the results of the concrete measures. This fact requires detailed evaluation in analyzing the competitors' activity.

4. Conclusions

1. There are over 3000 companies in the road cargo carriers market of the Republic of Lithuania, the competition intensity is high, and the market concentration is low.

2. The monopolization level among the cargo forwarding companies of the Republic of Lithuania is a low, i.e. the competition intensity is high, and the market concentration is low.

3. The research has shown that the competition intensity among the Lithuanian taxi companies is approaching the limit when competition is considered to be high.

4. The monopolization level among air transport companies in Lithuania is high and the market concentration is high too. Two Lithuanian air transport companies are dominating on the market.

References

1. Porter, M. On Competition. Harvard Business Press, 1998, 495p.
2. Yakovlev, A. Statistical estimation of production monopolization level. Moscow. Statistics manual, 1990. 46p. (in russian)
3. Azoev, G. Competition: analysis, strategy, and practice. Moscow: Economic and marketing center, 1996. 208p. (in russian)
4. Azoev, G. Competitors activity analysis. Moscow. GAU. 1995. 80p. (in russian)

SOCIAL AND ECONOMIC PRINCIPLES OF DEVELOPING AN INTEGRATED PUBLIC TRANSPORT NETWORK IN VILNIUS

Daiva Griskeviciene, Algirdas Griskevicius

Vilnius Gediminas Technical University, Transport Management Department,
Plytines g. 27, LT-2016 Vilnius, Lithuania,
E-mail: daiva1@centras.lt

Abstract

In the present paper, the problems of offer and demand for public transport services as well as satisfaction of the needs of potential passengers and their attitudes towards the available routes and transport services and their expansion have been considered. Transport services, their quality and amount, network of routes, schedules, tariffs and fares have been analysed from social and economic perspectives. The most urgent transport problems are encountered in newly built residential districts which emerged in the last decade. Therefore, a survey based on questionnaires distribution among the residents of these districts with the aim to learn their opinion as the potential public transport users about the above issues was made. The data obtained was then analysed from social and economic perspectives.

Key words: public transport; routes; demand; supply; suburban; service quality

1. Introduction

With the aim to investigate the present situation of the changing public transport market in 2001–2002 there were carried out the social-economic investigations. In this research work, the problems of offer and demand for public transport services as well as satisfaction of the needs of potential passengers and their attitudes towards the available routes and transport services and their expansion have been considered. Transport services, their quality and amount, network of routes, schedules, tariffs and fares have been analysed from social and economic perspectives. The most urgent transport problems are encountered in newly built residential districts which emerged in the last decade. Therefore, a survey based on questionnaires distribution among the residents of these districts with the aim to learn their opinion as the potential public transport users about the above issues was made. The data obtained was then analysed from social and economic perspectives.

2. Supply and Demand

The key market economy principle is the correspondence of the public transport service market supply to the demand of passengers. The transport service supply analysis confirms the insufficient capacity of the public transport system.

1. In Vilnius, 57.8 % passengers using public transport are carried by trolley-buses. Buses are mainly used on suburban routes to carry passengers to the downtown. They service 39.3 % passengers, with 15 % passengers using taxi buses and 14 % – fixed route taxis (Fig. 1).

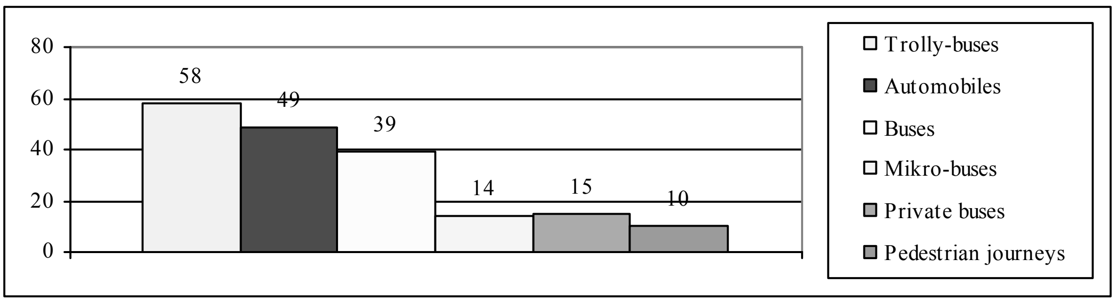


Figure 1. *The choice of transport facilities in the outskirts and in the suburbs of Vilnius in 2001–2002 (%)*

2. Because of the high rate of motorization, public transport has lost a part of passengers. In 2001, the degree of motorization of the population reached 304.0 auto/1000 inhabitants, with 0.8 automobile per household. The motorization boom in Lithuania was caused by the flood of old relatively cheap Western – made cars on the market (Fig. 2).

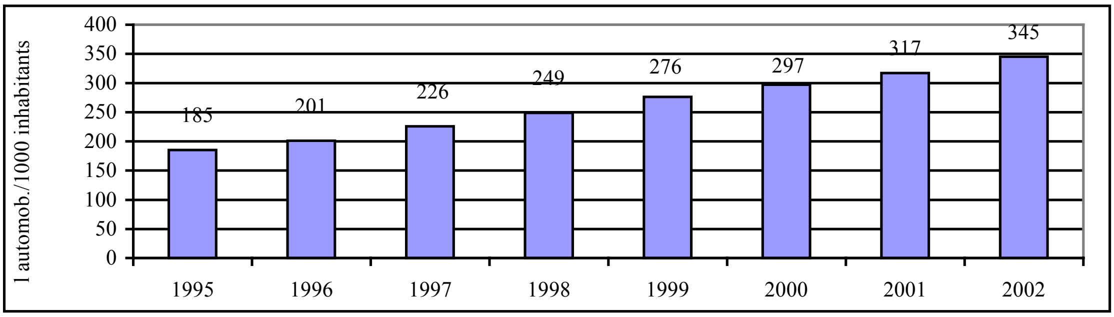


Figure 2*. The number of automobiles in Vilnius, automob./1000 inhabitants*

The mobility of the inhabitants increased not only on account of commuters but because people began to travel more at the weekends and in their leisure time. Even seasonal travel fluctuations became not so sharp. Now these fluctuations correspond to the variations of expenditures for cultural and recreational purposes.

3. Firms providing public transport services tend to maintain currently achieved level. However, the trolley-buses and buses exploited can now satisfy only the existing demand for transportation. For renovation and upgrading of public transport facilities the subsidies from Vilnius municipality are needed. New buses and trolley-buses are needed to serve the suburbanites and integrate the suburban routes into a unified urban transport network.

4. The municipality is responsible for providing quality transport services for urban inhabitants. It announces bids for providing transport services on the existing routes and gives licenses for using minibuses on the routes serviced by other transport facilities. In announcing bids, the municipality should set the criteria to evaluate quality and quantity of the services offered.

5. The problems of suburban transport have not been paid due attention yet. New routes for public transport are not being laid, passengers should still make changes, the network of streets has not been completed, and not all suburban roads have been turned into the city streets. These unsolved integration problems prevent rational allocation of funds of Vilnius region and the city into transport network development. The currently used tariffs and fares do not meet the requirements raised to passengers' transportation. Vehicle operators supported by the municipality tend to increase the tariffs. A survey has shown that the users of transport services are interested in maintaining low tariffs as long as possible because they believe that public transport should be partially state-supported. A discount system seems most acceptable for them. It may be accounted for the economic situation and the purchasing power of the inhabitants as well as social spectrum of potential passengers embracing groups of state-supported people, unemployed family members and students for whom any price rise is painful. As a result, they simply stop using public transport.

6. The priorities of passengers include better transport services and an integrated transport network of urban and suburban routes. In the questionnaires, customers emphasized the importance of maintaining schedules with short intervals for fixed-route transport, as well as fast and comfortable journey. The survey conducted in Vilnius revealed the main factors influencing the selection of a particular transport facility.

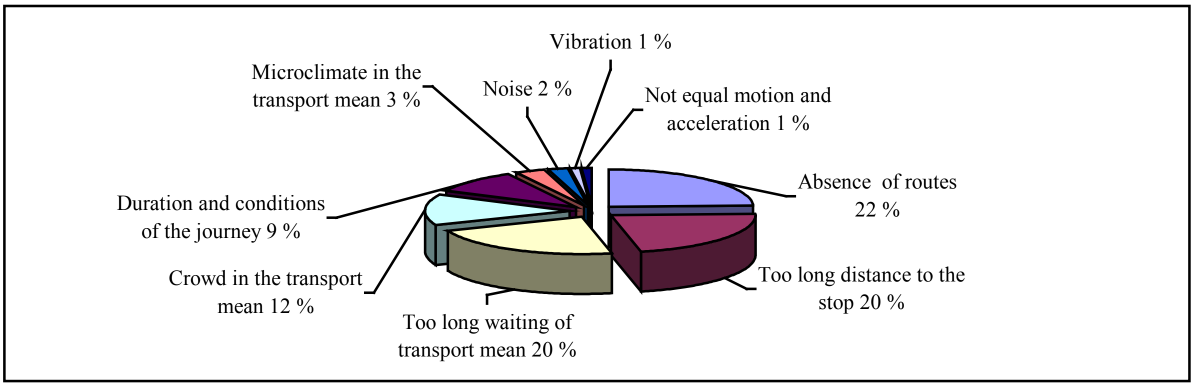


Figure 3. *Conditions when preference is given to automobile*

The preference of an automobile is conditioned by the lack of the route (22 %), long way to a bus (trolley-bus) stop (20 %), large intervals in the schedule (20 %), non-regular public transport service (3 %), time and conditions of journey (9 %), need for making changes and incompatibility of traffic schedules (9 %), etc. A unified network of urban and suburban transport providing services of the same quality to all inhabitants is needed.

3. Integration of Public Transport

The research made has shown that public transport system of Vilnius should be developed along three main lines [1, 2]:

1. Improving the quality of the available fixed-route public transport services;

The passengers give the priority to higher quality transport services and the extension of the existing routes. They emphasize such aspects as density, speed and comfort of fixed-route traffic. The users have to adapt to changes, therefore they prefer long-term stable schemes of transport routes and tariffs.

2. Integrating the suburban transport into the urban transport network

To make public transport more easily available for the inhabitants, an integrated urban transport system including suburban transport should be created. The inhabitants of the near suburbs would like to have stable routes and transport for getting to the workplace and for other purposes which should be the same as in the city in terms of time, space and tariffs.

To achieve this end, the fleet of transport facilities providing the above services should be extended.

Routes lying outside the city limits and the municipality's responsibility should be financed by the Vilnius region authorities. The tariffs on these routes should be revised in the context of the integration into the urban transport system.

3. Developing more advanced tariffs and ticket systems

Passengers are concerned about stable and low tariffs believing that public transport should be partially supported by the state. They support the allowances embracing a large number of people belonging to various social layers. The economic situation of the country and the impact of the priorities of the former planned economy account for this.

Tickets for journeys by public transport facilities are simple to print, to use and to change, if necessary. Mechanical ticket stamping machines are simple, cheap and easily maintained or repaired. However, a system of ticketing is not flexible, being difficult for keeping account and control.

The development of a harmonized public transport and its integration should be performed along three main lines:

- integration of public transport services includes coordinated exploitation and market of services, interaction of various transport facilities and services, coordinated selling of tickets and easily available information about the services;
- integrated development of all transport facilities implying their more effective use and coordination as well as the use of inter-modal transport and easy transfer from private to public transport;
- integration of transport development into other branches of economy.

These measures comply with major strategic goals of harmonized public transport development in Lithuanian cities [3, 4]:

1) to promote the use of public transport by urban inhabitants;
2) to raise the standards of public transport services provided in Lithuanian towns in compliance with the quantitative and qualitative standards of the developed EU states;
3) to extend the urban public transport network so that it could satisfy the needs of the particular territory;
4) to develop a mechanism of passenger transport system financing on the fixed routes which could be competitive in the market;
5) to authorize municipalities to be responsible for public transport services, their provision, planning, financing, expansion, quality and control;
6) to legitimize the responsibility of transport operators to satisfy the needs of passengers;
7) to support a competition between companies providing public transport services;
8) to take into account social principles in developing public transport and fixing the tariffs;
9) to start the integration process of urban and suburban public transport;
10) to introduce a well-thought (balanced) tariffs of public transport services;
11) to update the system of ticketing;
12) to introduce more flexible and varied system of tickets.

An urban pubic transport system is not sufficiently flexible. The response of the passengers to the novelties introduced is also often delayed, requiring some time for analysis and adaptation. Therefore, timely information is very important.

In the market, the spirit of uncertainty prevails in the expectation of the consequences of the integration processes in the area of public services. The users are expecting the improvements in public transportation and higher quality services. However, they are afraid of possible rise of prices which would affect the price of transport tickets.

4. Conclusions

1. A system of public transport is closely connected with the territorial planning of the city. If the urban territory is being expanded, transport system should also be extended not to become an obstacle for urban territory expansion.
2. Public transport system of Vilnius should be developed along three major lines:
   - Raising quality standards of the available fixed-route transport;
   - Promoting the integration of suburban transport into the urban transport system and providing services of higher quality;
   - Upgrading ticketing and tariffs;

3. A major purpose of the urban transport system is the integration of various transport facilities to provide a free choice of the particular means of transport for passengers.

4. To achieve more frequent use of public transport by urban inhabitants, a financing mechanism of passenger transportation should be developed which could be competitive in the market.

References

[1] P. Juskevicius. *Planning of the urban transport systems*. (Vilnius: Technika, 1995).

[2] P. Juskevicius. *Harmonizing the expansion of cities and their transport systems*. (VGTU, Vilnius: 2003).

[3] M. Burinskiene, G. Paliulis. The analysis and perspectives of public transport // *Development of cities and roads*. (Vilnius: Technika, 2000).

[4] J. Butkevicius. Passenger transportation. (Vilnius: Technika, 2002).

Session 5

Reliability of Mechanical and Aviation Systems

THE DESIGN OF LOW SPEED WING SECTION FOR SAFETY AND PERFORMANCE

Eduardas Lasauskas

Gediminas Technical University, Vilnius, Lithuania
Rodunios kelias 30, LT-2038 Vilnius.
E-mail: edulas@ai.vtu.lt

Abstract

The R. Eppler Airfoil Program System is used for the multipoint design of wing sections. The goal of the design is the low sensitiveness to leading edge contamination and low drag. The XFOIL and MSES codes of M. Drela are used for the analysis. The examples of wing sections, calculated and measured results are presented.

Key words: wing section, design, calculation, experiment

Introduction

The airfoil lift-to-drag ratio is rarely the only desirable feature of an airfoil. The airfoil high lift, climb parameter, thickness, pitching moment, stall characteristics, and sensitivity to the leading edge contamination or roughness are all important factors, among others, that must each be weighed separately when one considers selecting or designing an airfoil. This study focuses on the factors most related to low drag and low sensitivity to the surface contamination of airfoils at medium Reynolds Numbers (about one million). Only single-element airfoils are considered in the current work.

The well-known method for low drag of airfoils is a long range of laminar boundary layer on the upper and lower surfaces. At the medium Reynolds numbers the transition from the laminar to the turbulent boundary layer occurs through laminar separation bubble mechanism. In this case it is very important the exact prediction of transition location during design process of airfoil. Several codes use several transition prediction methods. In this work was used the individual frequencies e^n method implemented in the two codes MSES [1] and BUBBE [2]. This method is most exact but requires more computational resources.

The early laminar airfoils were very sensitive to the leading edge contamination: an abrupt decrease of lift and accompanying increase of drag. The modern laminar airfoils do not have this disadvantage, because at the maximum lift of the smooth airfoil the free transition on the upper side occurs near the leading edge [3]. In this work an additional criterion was used: no or very small turbulent separation area at the trailing edge on the airfoil with forced transition near the leading edge in the low drag bucket of the airfoil with free transition. This additional criterion minimises growth of the drag when transition is forced near the leading edge.

The research, airfoil design methods and two airfoil examples are presented in this work.

Experiments

The experiments were carried out at the sailplane-manufacturing factory AB “Sportine Aviacija” in cooperation with the Central Aerohydrodynamic Institute (TsAGI). The goal of these experiments was the better understanding of flow processes at medium Reynolds numbers and better interpretation of calculation results. The experiments were carried out in the free flight and in laminar wind tunnel of TsAGI.

AB “Sportine Aviacija” equipped the flight test laboratory on the two-seater metal sailplane L-13 Blanik for TsAGI. The test part of the airfoil was mount on the wing. Pressure distribution was measured in three sections. The loss of the impulse was measured in the wake for drag estimation. The boundary layer characteristics were measured on another wing. The registration was performed using photo camera and liquid manometers in the fuselage.

Later the own flight laboratory SL-2P was developed at AB Sportine Aviacija" (Fig.1). The equipment on the top was similar to the one of the Super Janus DLR, Germany. The lift and pitch moment of the airfoil was estimated from the pressure distribution. The drag was estimated from the wake measurement. The liquid manometers and photo camera were in the right fuselage. The pilot and experimenter were sitting in the left fuselage.

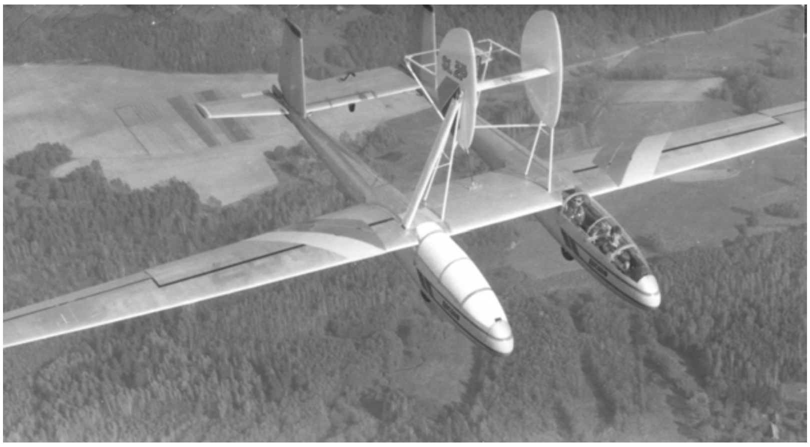


Figure 1. *Flying airfoil testbed SL-2P*

The same airfoil model was tested with the flight laboratory and also in the wind tunnel of TsAGI [4]. The comparison of measurements from free flight and wind tunnel is presented in Fig. 2. The congruence of this comparison is very good.

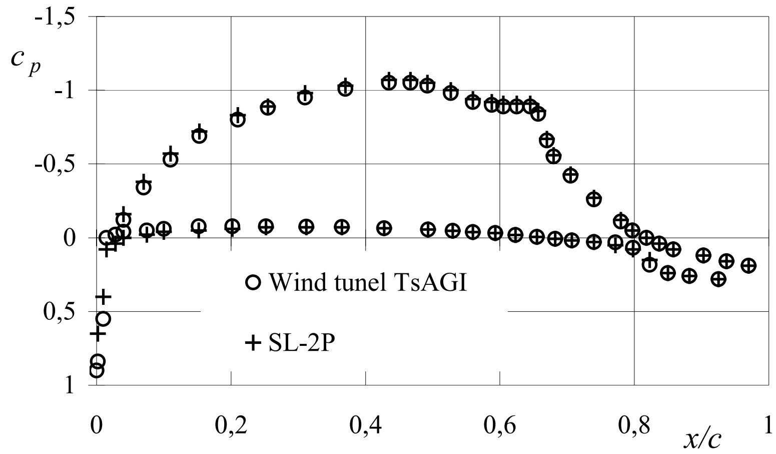


Figure 2. *Measured pressure distribution of airfoil FX67-K-170 at* $\alpha = 0$ *deg and* Re = $0,8 \times 10^6$

The pressure distribution on the upper side of the airfoil between 65% and 70% clearly shows the existence of laminar separation bubble (Fig. 2). Laminar separation bubbles occur, if the laminar boundary layer separates from the surface. Due to the destabilizing effects of the shear layers velocity profiles with an inflection point the laminar shear layer rapidly becomes unstable and transition occurs. Next the turbulent shear layer reattaches, thus forming a bubble. The laminar part of the bubble is characterised by nearly constant pressure and the turbulent part by a very steep pressure rise. The size of the laminar separation bubble is related to the stability of the laminar boundary layer, which in turn depends on Reynolds number and velocity gradient. By increasing the Reynolds number transition is moving forward and the size of the bubble decreases.

The Corner in the pressure distribution corresponds to the laminar-turbulent transition. It is very important, that the transition location is the same from the free light experiment and from wind tunnel experiment.

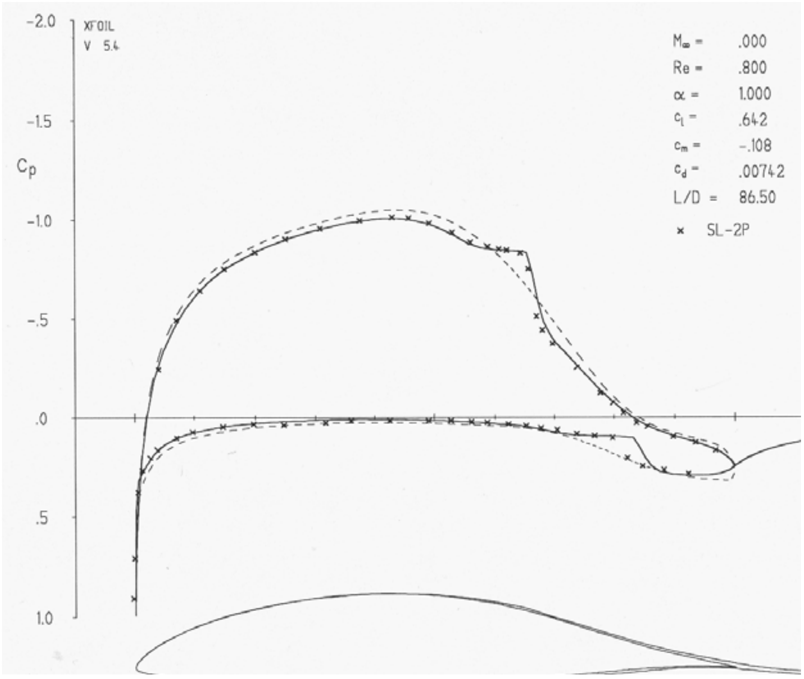


Figure 3. *Measured and calculated pressure distribution of airfoil FX67-K-170 at* α *= 1 deg and* Re *= 0,8 × 10*^{6}

Fig. 3 shows the comparison of measured and calculated (XFOIL code [5]) pressure distribution of airfoil FX67-K-170 at angle of attack $\alpha = 1$ deg and Reynolds number $Re = 0{,}8 \times 10^6$. The congruence is also very good except the position of laminar-turbulent transition on the lower side.

Airfoil Design

Using the experience from experiments some wing airfoils were designed through the several design and analysis codes: the R. Eppler code [6], XFOIL [5] and MSES [1]. First, R. Eppler multipoint conformal mapping method was used for the design. The new airfoil that appeared to meet the performance objectives was then analysed using Eppler analysis method and then XFOIL code.

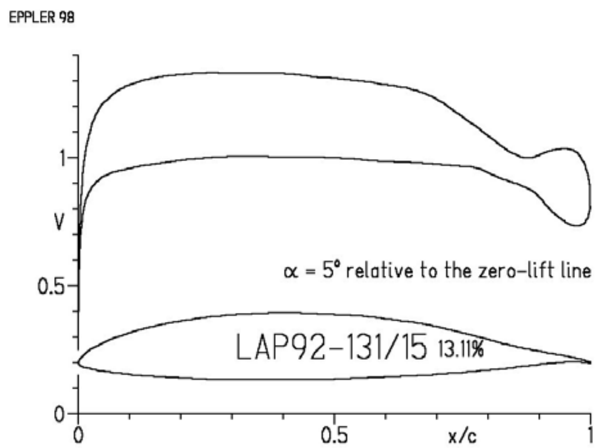


Figure 4*. Airfoil shape and inviscid pressure distribution at* α *= 5 deg relative to the zero lift line*

If at any point the candidate airfoil failed to meet the design goals, the experience gained was used to redesign the airfoil to more closely match the desired performance objectives. This iterative process was continued until a successful airfoil was designed.

Fig. 4 shows one example: a flapped wing section for the high performance sailplane and the pressure distribution with the flap in zero position. Laminar-turbulent transition on the lower side is fixed at 78%, laminar-turbulent transition on the upper side is free.

The laminar-turbulent transition on the upper side was predicted using several methods: R. Eppler empirical criterion [7], simplified e^n envelope method (XFOIL, MSES) and individual frequencies e^n method (MSES).

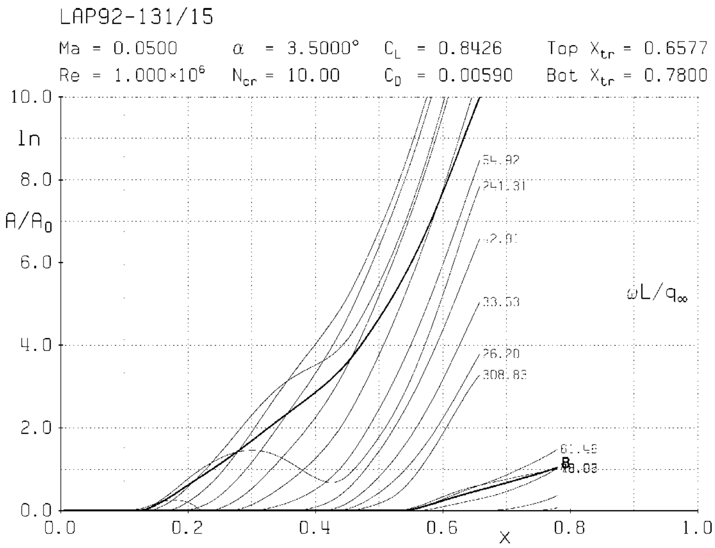


Figure 5. *The ratio of amplification in the laminar boundary layer of airfoil LAP92-131/15 at* α *= 3,5 and* Re *=* 1×10^6

Fig. 5 shows the comparison of the e^n envelope method and individual frequencies e^n method for the airfoil LAP92-131/15 at $\alpha = 3{,}5$ and $Re = 1 \times 10^6$. The standard transition prediction in XFOIL and MSES is performed using the envelope e^n method. Thick lines correspond to the envelope of disturbance amplification ratio. Thin lines correspond to different frequencies of disturbance. The user of XFOIL and MSES should prescribe the critical value n_{krit}. This is always a problem. The information about the ratio of amplification of separate frequencies helps in this case.

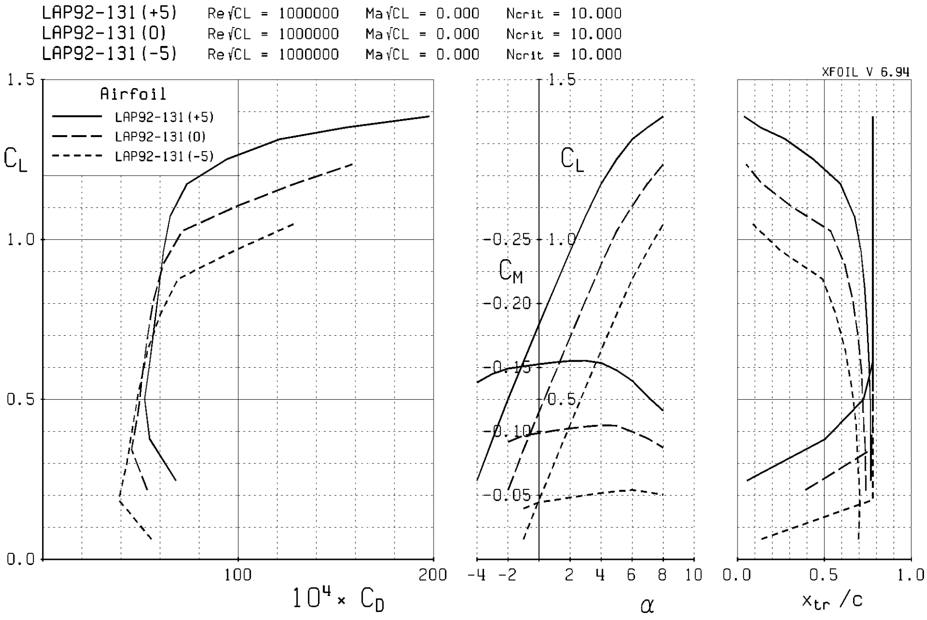


Figure 6. *Characteristics of the airfoil LAP92-131/15*

Fig. 6 shows the data of the airfoil at the different flap positions. In the left part the Cl-Cd curves are presented, the middle part shows the lift curves, the right part shows transition location versus lift coefficient. The free transition on the upper side is near the leading edge at the maximum lift. It means that maximum lift is not sensitive to the leading edge contamination.

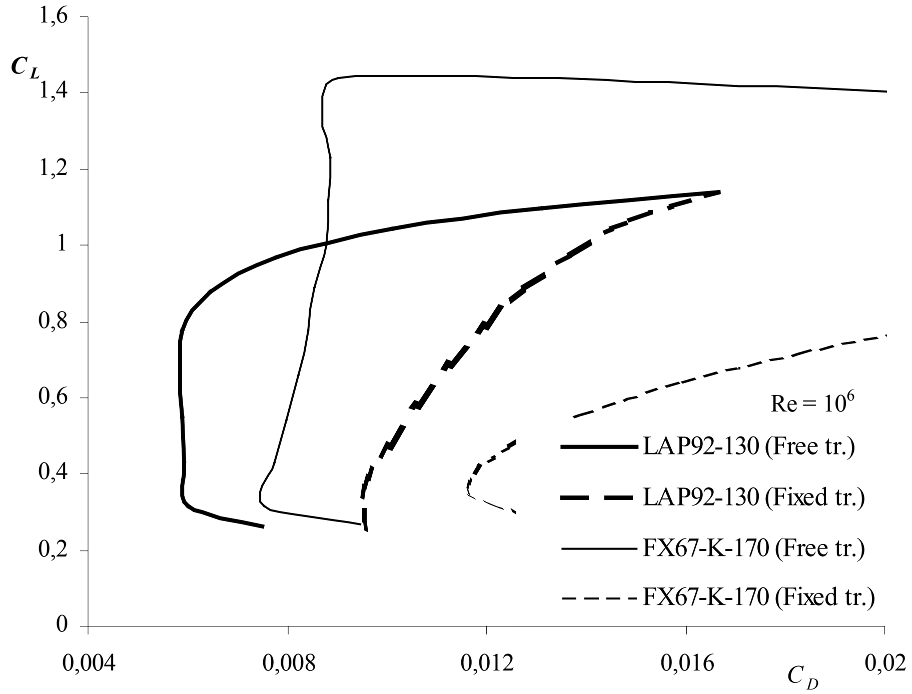


Figure 7. *Polar comparison of the airfoils LAP92-131/15 and FX67-K-170 at* Re=10^6. *Solid lines: free transition on the upper side. Dashed lines: fixed transition at 0,05 of chord on the upper side*

Fig. 7 shows polar comparison of the airfoils LAP92-131/15 and FX67-K-170 at Re=10^6. Solid lines correspond to free transition on the upper side. Dashed lines correspond to fixed transition at 0,05 of chord on the upper side. The drag of the airfoil LAP92-131/15 is less in both cases. The maximum lift of airfoil LAP92-131/15 with free transition is less than of airfoil FX67-K-170. But if the transition is fixed at 0,05 of chord on the upper side, the maximum lift of airfoil LAP92-131/15 remains the same and maximum lift of airfoil FX67-K-170 drops down.

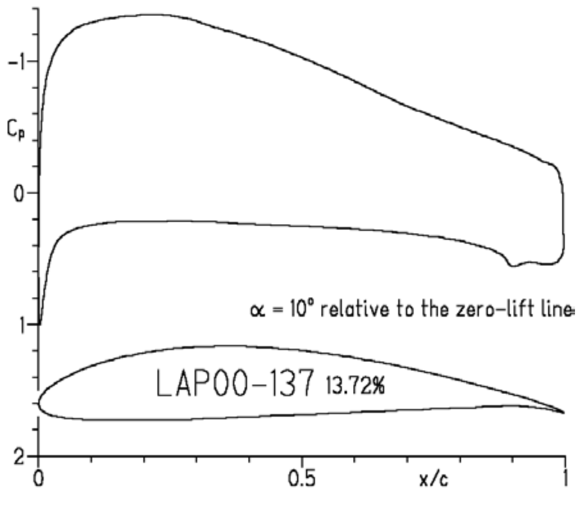


Figure 8. *Airfoil shape and inviscid pressure distribution at* α *= 10 deg relative to the zero lift line*

Second example is airfoil LAP00-137 for primary glider application [8]. Fig.8 shows the airfoil form and pressure distribution. Fig. 9 shows the CL-CD curves of this airfoil.

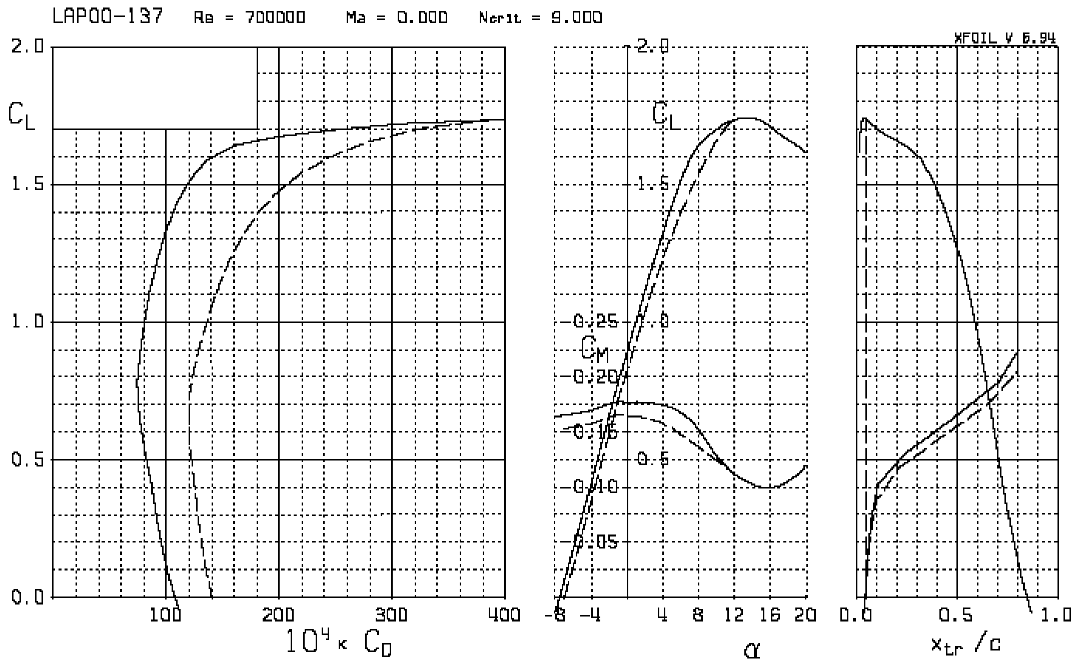


Figure 9. *Calculated characteristics of the airfoil LAP00-137 at* Re=10^6. *Solid lines correspond to free transition on the upper side. Dashed lines: fixed transition at 0,05 of chord on the upper side*

The maximum lift of airfoil LAP00-137 is not sensitive to leading edge contamination. The shape of this airfoil is convex, except 0.1 of chord on the upper side at the trailing edge. The convex shape of airfoil is better for the homebuilding.

Conclusions

It is shown, that the usage of exact method for transition prediction allows achieving less drag of airfoils at medium Reynolds numbers. An additional criterion for low area of separated turbulent boundary layer near trailing edge on the disturbed airfoil warrants small rise of drag. The presented examples show that low drag and low sensitivity to the contamination is simultaneously possible.

References

1. M. Drela, Design and optimization method for multi-element airfoils, *AIAA Paper 93-0969*, 1993.
2. P. Dini, M.S. Selig and M.D. Maughmer, A Simplified Linear-Stability Transition Prediction Method for Separated Boundary Layers, *AIAA Journal*, 1992, v.30, No 8, 953-1961.
3. L.M.M. Boermans and A. van Garrel, Design and Windtunnel Test Results of a Flapped Laminar Flow Airfoil for High-Performance Sailplane Applications, *Technical Soaring*, 21, No1, 1997, 11-17.
4. E.Lasauskas and M.W.Ryshonkow, *Methodical Flow Investigation of two airfoils FX67–K–170 and PP-17 Using Flying Test Laboratory SL-2P* (Moscow: TsAGI Report No. 8319, 1988).
5. M. Drela, XFOIL: An Analysis and Design System for Low Reynolds Number Airfoils, *Proceedings of the Conference Notre Dame*, *Indiana, June 5-7, 1989*. (New York: Springer-Verlag, 1989), 1-12.
6. R. Eppler, *Airfoil Design and Data* (Berlin: Springer-Verlag, 1992).
7. R. Eppler, An Empirical Criterion for Laminar-to-Turbulent Boundary-Layer Transition, *Technical Soaring*, 23(2), 1999, 34-42.
8. E. Lasauskas, Sensitiveness to Leading Edge Contamination of Wing Sections for Primary Glider Application, *Aviation*, 5, 2000, 61-64.

INFLUENCE OF CORROSION ON THE REQUIRED NUMBER OF AN AIRFRAME INSPECTION

M. Shujauddin Wahab, Yu. M. Paramonov

Riga Technical University, Aviation Institute, Riga, Latvia
Lomonosova 1, Riga, LV-1019, Latvia
Tel.: (+371)-7089950, Fax: (+371)-7089990, E-mail: shujawahab@hotmail.com

Abstract

The behavior of fuselage splice joints containing multi-site fatigue damage and corrosion is investigated. The objectives of this paper is to develop probabilistic analysis methodology that allows to calculate the probability of fatigue failure as function of specified interval between inspections and to make comparison of the cases when there is and there is not corrosion. As initial information the results of fatigue test of specimens of special type are used [1]. By the use of this information parameters of the stochastic damage growth model were estimated. The probability of failure at the specified number of inspections at the fixed specified life was calculated by the use of Monte Carlo method and related formulas. It was shown that if the specified life is very large then relatively large probability of failure is nearly the same for both cases when there is and there is not corrosion. But if the specified life is relatively small then corrosion increases the probability of failure at the same number of inspection very significantly.

Key words: Fatigue Crack Growth, Corrosion Medium, Failure Probability and Interval Between Inspections

1. Introduction

Fatigue crack growth analysis in the presence of corrosion is an important subject as shown in Figure1 because it can degrade the structural integrity and damage tolerance of fatigue critical structural components in aging aircrafts. Multiple site fatigue damage (MSD) in a longitudinal skin splice has been recognized as a major airworthiness problem. It had a very significant influence in Aloha B-737 incident in 1988.

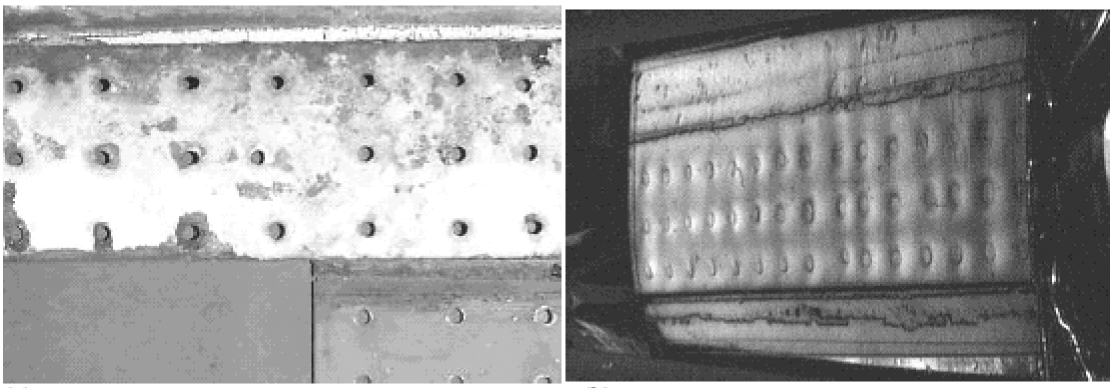

(a) (b)

Figure 1. *Illustration of a corroded longitudinal fuselage splice from a retired 727: (a) white corrosion product on faying surface, (b) corrosion pillowing detected by D Sight*

For fleet management it is important to know the effects of corrosion in normal service on the durability and damage tolerance (DADT) characteristics of the fuselage. The DADT characteristic of any structure are defined by the crack initiation and growth patterns, the critical crack scenarios that could develop and the number of load cycles it takes for cracks to become detectable and then grow to a critical condition.

2. Test Program

The MSD concept is illustrated by the generic lap splice version of the specimen shown in Figure 2. A finite element model of the loop stress distribution in specimens is also shown. The concept is the use of bonded side straps to simulate the load transfer from cracked areas to surrounding structure that occurs on aircraft. The specimen shown is a 25.4 cm (10 in) wide version designed to be representative of the longitudinal fuselage splices in some narrow body transport aircraft. The splice in the generic specimen comprises two sheets of 1.0 mm (0.04 in) thick 2024-T3 Alclad held by three rows of 4 mm (5/32 in) diameter 20177-T4 rivets (MS20426AD5-5) without adhesive, paint or sealant. The rivet geometry results in a knife-edge countersink.

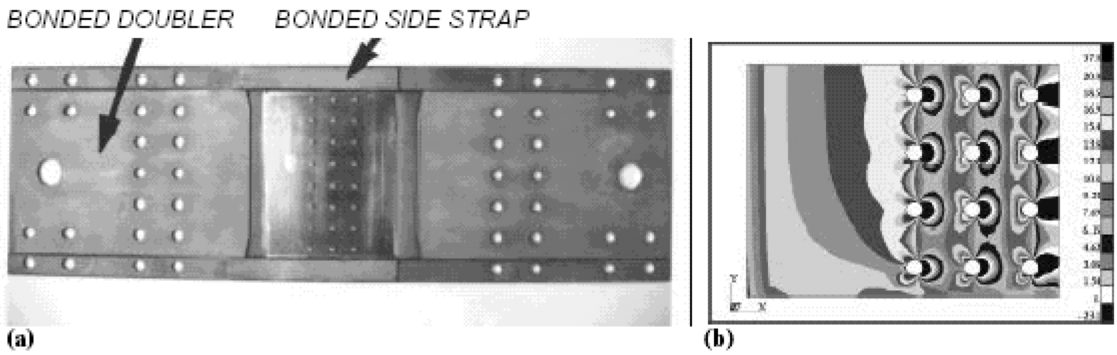


Figure 2. *Illustration of MSD specimen (a) bonded doubler, (b) with a hoop stress distribution at faying surface by finite element prediction*

The average cycle number for the final failure for the corroded specimens is 207640 cycles. As shown in Figure 3, the corrosion damage in this MSD specimen (average thickness loss of between 5% and 6%) was compared with the damage in a section of splice from a Boeing 727 aircraft, shown in Figure 1, which was naturally corroded to a comparable level during 48,665 flights over 24 years. So 1 flight is approximately equivalent to 4.266 cycles.

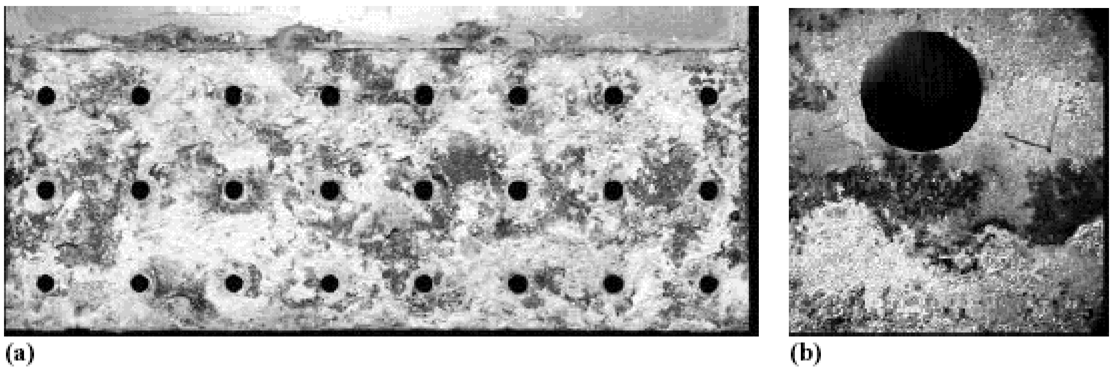


Figure 3. *MSD specimen corroded to 5% to 6% average thickness loss: (a) countersunk sheet with corrosion product still in place; (b) close-up near hole with corrosion product removed*

The combination of corrosion and fatigue assumes that corrosion/fatigue interactions occurs only in the context of pre-existing corrosion and in a dry splice. This is a reasonable approximation for two reasons. First, teardown of aircraft splices and evidence indicates that substantial corrosion often exists without any associated fatigue cracking. Second, the highest in-service loads occur when any moisture in the splice is likely to have frozen.

There are altogether nine MSD specimens out of which five are non-corroded and four are relatively heavily corroded. They all are fatigue tested. These specimens are listed in Table 1 along with their respective fatigue life at visible crack detection, first link up and final failure.

Table 1. Fatigue life of MSD Specimen

Specimen #		Fatigue Life (Cycles)		
		1st observed	1st Linkup	Final failure
Non-corroded	Cgc-f38	387500	491711	501933
	Cgc-f46	314000	398908	403718
	Cgc-f51	304001	381378	392591
	Cgc-f60	290000	368650	378754
	Cgc-f61	368500	473397	481353
			Average Final Failure	431670
Corroded to 5%-6% level	Cgc-cf34	160001		222450
	Cgc-cf43	144000		189074
	Cgc-cf45	104107		177129
	Cgc-cf58	142000		241909
			Average Final Failure	207640

3. Failure Characteristics

A significant difference was noticed in the behavior of the MSD specimen with and without corrosion. The visible crack were observed to start in different scenarios and there were distinct differences in load cycles to first observed cracks, which are shown in Table 1. The five non-corroded specimens showed visible cracks at between 2.9 and 3.88 * 10^5 cycles and failed at between 3.79 and 5.02 * 10^5 cycles. The statistical dispersion of visible crack detection and growth damage accumulation is large, which is a typical phenomenon of MSD specimens. The load cycles to visible crack detection of the non-corroded specimens represented 70% to 80% of their total fatigue life and similar behavior was observed in the corroded specimen. The observed reduction due to corrosion in the mean cycles to visible crack detection was 59% for the specimens corroded to the 5% to 6% level.

In non-corroded specimens the crack grew with increasing load cycles from the central holes outward forming a pattern of multi-site damage as shown in Figure 4. Changes in gross failure modes were observed in the corroded specimens with the 5% to 6% level. The two dominant failure modes in corroded specimen are: (i) non-uniform MSD – one crack developed from only one site – at the rivet locations in the upper row and (ii) fatigue cracking at one or more sites in the inner (driven) sheet 5.08 to 7.62 mm (0.2 to 0.3 in) below the lower rivet row.

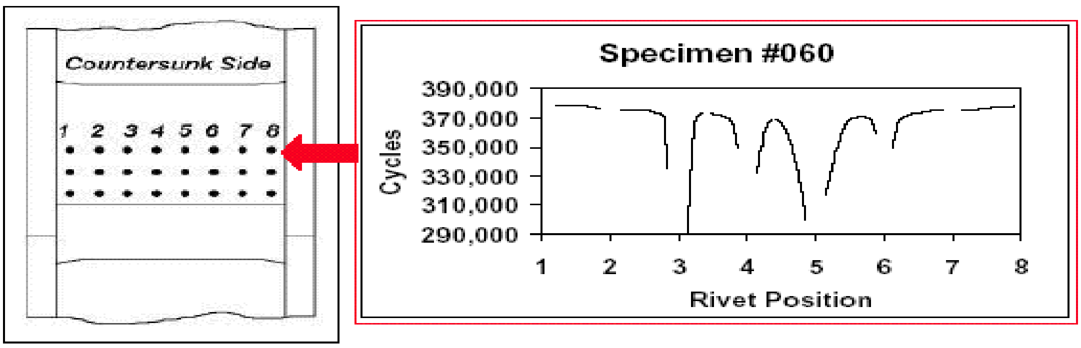


Figure 4. *Typical MSD growth pattern in a non-corroded specimen*

MSD tends to develop in clusters within the boundaries of a frame-bay. Similarly the linkup of MSD and the formation of a lead crack also tend to occur initially within a frame-bay for this curves that record the sum of all individual crack lengths at any given time. The crack length at a rivet hole is measured from the edge of the drilled hole. For cracks that developed away from the rivet rows, as in some corroded specimens the aggregate crack length is taken as the total tip to tip crack length. Where there were several such cracks in a specimen in an interacting MSD formation, overlapping cracks were regarded as linked cracks. The test data for the crack growth history of the two specimen groups are shown in Figure 5-6. In the non-corroded specimens, first linkup occurred at an aggregate crack length of about 50.8 mm (2 in). Subsequent crack growth was relatively fast and produced a pronounced knee in the growth curve. On the other hand in the corroded specimens the overall crack growth rate was relatively stable during the whole growth period similar to the growth progression of a single crack.

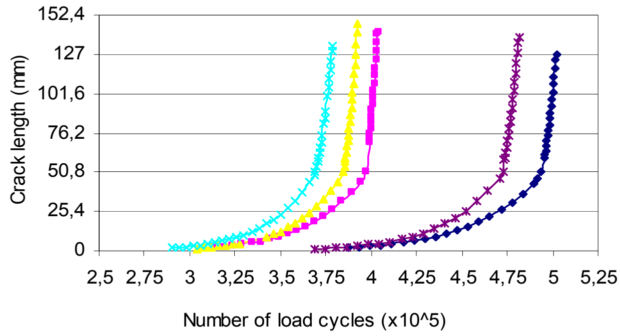


Figure 5. *Crack growth history data of non-corroded specimens*

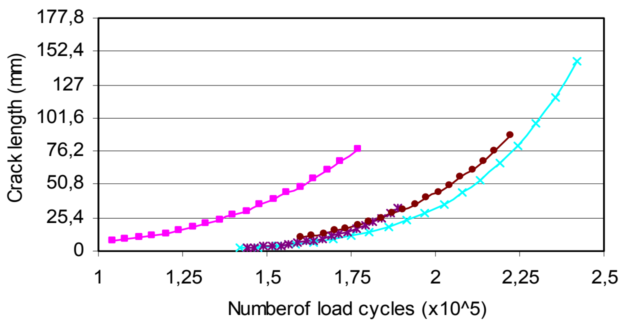


Figure 6. *Crack growth history data of corroded specimens at 5% - 6% level*

With above observation the total service life of a specimen is divided into two or three stages. For non-corroded specimens, the total fatigue life, N_t, is divided into three parts: life to visible cracks or visible damage starting life, N_s, growth life before linkup, N_{g1}, and growth life after linkup, N_{g2}, that follow $N_t = N_s + N_{g1} + N_{g2}$. For the corroded specimens to the 5% to 6% level, a single stage with growth life N_s, is used for the whole growth period because of their relatively stable growth behavior and the total fatigue life is $N_t = N_s + N_g$. The visible damage starting life is the number of load cycles at which the first crack was observed and the total life of a specimen is when the final failure occurred. The growth life is the difference between the total life and the damage starting life $N_g = N_t - N_s$.

In a modern transport aircraft, the critical length of a single longitudinal crack in a fuselage skin is typically in excess of the frame bays about 10.16 cm (40 in). Crack growth rates are high when a lead crack reaches a length of several inches. The presence of MSD in adjacent frame-bays could reduce the critical length of the lead crack. Therefore first the splice is considered to have failed when the first linkup occurs at which the length of the aggregate lead crack $\mathbf{a}$, reaches a specific value $\mathbf{a_{lk}}$. Second the splice is considered to have failed when the aggregate crack length reaches a critical value $\mathbf{a_{cr}}$. The specific crack length for linkup and the critical crack length for final failure are taken from the mean values of the crack length obtained fro the test data corresponding to the linkup and failure life, respectively. For corroded specimen, only the final failure is considered.

4. Damage Starting Life and Stochastic Growth Model

4.1. Curve-Fitting of Test Data

The test data are fitted by the help of growth function, expressed as:

$$a = C_1 * N^{C_2} \tag{1}$$

Where **a** is the aggregate crack length, **N** is the number of load cycles, $\mathbf{C_1}$ and $\mathbf{C_2}$ are two constants taken from the Table 2.

Table 2. Constants in fitted growth curves

Specimen #		Growth stage 1		Growth stage 2	
		C_1	C_2	C_1	C_2
Non-corroded	Cgc-f38	1.39e-9	13.18	1.90e-40	57.69
	Cgc-f46	2.16e-8	13.29	1.21e-35	58.84
	Cgc-f51	5.59e-9	14.62	1.97e-30	51.30
	Cgc-f60	2.48e-8	13.93	6.60e-22	37.86
	Cgc-f61	1.70e-10	14.92	4.52e-35	51.40
Corroded to 5% - 6% level	Cgc-cf34	0.0194	6.48		
	Cgc-cf43	0.0028	9.65		
	Cgc-cf45	0.2412	4.44		
	Cgc-cf58	0.0057	7.82		

4.2. Determination of Fatigue Crack Growth Function Parameters

It is assumed that fatigue crack rate of some items of airframe is defined by Paris's formula [2]:

$$\frac{da}{dt} = C\left(\Delta K\right)^m = Q * a^{m/2} \tag{2}$$

a – size of a crack
t – service time

Where

$$Q = C\left(\lambda\left(\sigma_{\max} - \sigma_{\min}\right)\sqrt{\pi}\right)^m \tag{3}$$

C,m – crack growth function parameters
λ - takes into account the width of the panel, influence of stringers
σ_{max} , σ_{min} – maximum and minimum stress in a flight

The solution of this differential equation for m≠2 is:

$$a(t) = a(o) \Big/ \left(1 - \mu\left(a(o)\right)^{\mu} Qt\right)^{1/\mu} \tag{4}$$

Where, $\mu = m/2 - 1$ (5)

μ – depends on the material characteristics

$$a(o) = \left(1 \Big/ \left(\mu Q t_n + a_n^{-\mu}\right)\right)^{1/\mu} \tag{6}$$

a(0) – equivalent beginning size of a crack

Residual strength

$$\sigma_p(t) = K_c \big/ \lambda\sqrt{\pi a(t)} \tag{7}$$

K_c – critical value of stress intensity factor

Lets investigate the results of data processing for crack growth during fatigue experiments. Transferring to logarithm scale and putting new designation, we will accept following form of description of crack growth speed.

$$y = \ln(da/dt) = \ln Q + m/2*\ln a = b_0 + b_1*x \tag{8}$$

Where, $b_0 = \ln Q$, $b_1 = m/2$ (9)

Then sequence $\{(a_i, t_i), I=1,\dots, n\}$ transmits to sequence (y_i , x_i), $I = \{1,\dots,n-1\}$, where:

$$x_i = \ln((a_{i+1} + a_i)/2); \tag{10}$$

$$y_i = \ln(\Delta a/\Delta t)_I = \ln((a_{i+1} - a_i)/ (t_{i+1} - t_i)), \tag{11}$$

Parameters b_0, b_1 can be found using least – squares method by formulas:

$$b_1 = \frac{\left(\overline{xy} - \overline{x}\cdot\overline{y}\right)}{\left(\overline{x}^2 - \overline{(x)}^2\right)} \tag{12}$$

$$b_1 = y-b_1*x \tag{13}$$

Where: (x_i,y_i), $I=\{1,..,n\}$

Using the evaluation of parameters b_0,b_1 can be found estimations:

$$Q = \exp(b_0),\ m = 2b_1,\ C = Q/(\lambda(\sigma_{max} - \sigma_{min})*\sqrt{\pi})m \tag{14}$$

After finding the estimation parameters Q, m, a(0) for both the non-corroded and corroded specimens, we get their statistical analysis with the help of Excel.

Examples of Relevant curves a(t) and $\sigma_p(t)$ are shown in Figures 7-8.

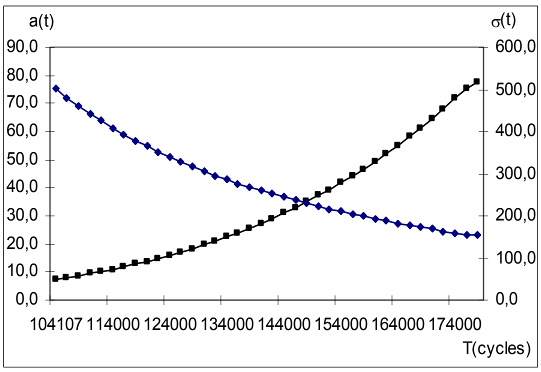


Figure 7. *Fatigue crack growth and loss of residual strength (FCG&RS) for corroded specimen.*

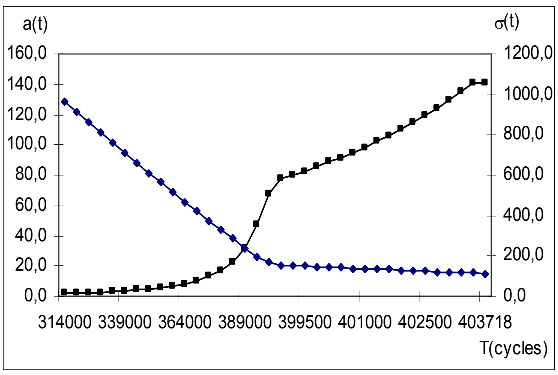


Figure 8. *Fatigue crack growth and loss of of residual strength (FCG&RS) for non-corroded specimen.*

Results of processing four fatigue crack growth data for corroded specimens at 5%-6% level are given in Table 3.

Table 3. Fatigue crack growth parameters for corroded specimens at 5%-6% level

Serial No.	Specimen #	μ	b_o = ln Q	a(0)
1.	Cgc-cf34	-0.15424	-9.75393	2.37E-19
2.	Cgc-cf43	-0.10353	-9.52048	1.48E-27
3.	Cgc-cf45	-0.22511	-9.61464	4.51E-13
4.	Cgc-cf58	-0.12758	-9.70537	1.63E-18
Average		-0.15262	-9.6486	1.13E-13
Standard Deviation		0.052581	0.103096	2.26E-13

Results of processing five fatigue crack growth data for non-corroded specimens are given in Table 4.

Table 4. Fatigue crack growth parameters for non-corroded specimens

Serial No.	Specimen #	μ	b_o = ln Q	a(0)
1.	Cgc-f38	0.432273	-11.2323	0.079004
2.	Cgc-f46	0.44934	-11.0738	0.091935
3.	Cgc-f51	0.334914	-10.6355	0.02707
4.	Cgc-f60	0.249426	-10.5026	0.014186
5.	Cgc-f61	0.322694	-10.7252	0.018577
Average		0.35773	-10.8339	0.046155
Standard Deviation		0.082805	0.307091	0.036475

4.3. Simulation of a Process of Fatigue Crack Inspection

It is assumed that some inspection technology is characterized by two values: a_d and w_i; a_d - the minimum size of a detectable crack and w – is interpreted as probability that the earlier scheduled inspection will be made with required accuracy. Service time when crack becomes detectable t_d and service time to fatigue failure t_f are defined below:

$$t_d = \frac{C_d}{Q} \qquad t_f = \frac{C_f}{Q} \tag{15}$$

we consider, that t_d and t_f are functions of random variable Q.

C_d – constant for both non-corroded and corroded specimens.

$$C_d = \frac{1 - \left(\frac{a(o)}{a_d}\right)^{\mu}}{\mu\left(a(o)\right)^{\mu}} \tag{16}$$

where a_d is equal to 20 mm.

C_f – constant for non-corroded specimens

$$C_f = \frac{1 - \left(\frac{a(o)}{\frac{K_c^{\,2}}{\sigma_{\max}^{\,2}\pi}}\right)^{\mu}}{\mu a(o)^{\mu}} \qquad \mu \text{ - positive} \tag{17}$$

C_f – constant for corroded specimens

$$C_f = \left[\frac{\left(\left(K_c/\sigma_{\max}\right)^2 * 1/\pi\right)^{\gamma} - \left(a(o)\right)^{\gamma}}{\gamma}\right] \qquad \mu \text{ - negative.} \tag{18}$$

Bar chart of crack undetectable and crack detectable time periods (CUCDTP) in both cases are shown in Figures 9-10.

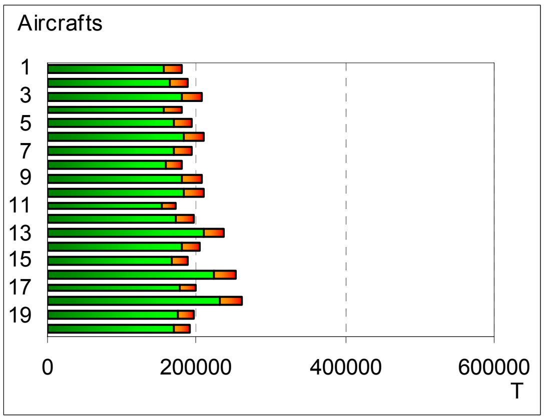


Figure 9. *Bar chart of CUCDTP for the corroded specimens at 5% and 6% level.*

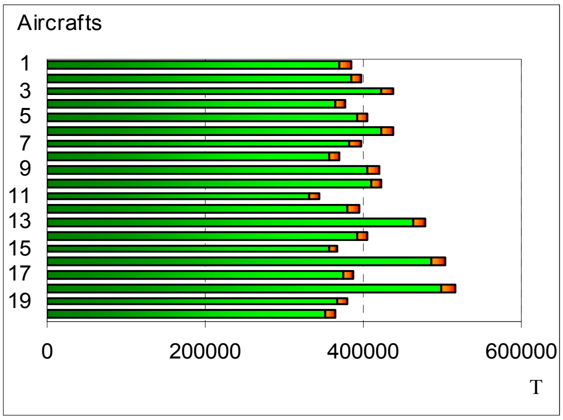


Figure 10. *Bar chart of CUCDTP for the non-corroded specimens.*

4.4. Interval between Inspection and Estimation Fatigue Failure Probability

$$\Delta = \frac{t_{SL}}{(n+1)} \tag{19}$$

where t_{SL} – specified life of an aircraft
n – number of inspections.

If we use Monte Carlo method then the failure probability in the interval $(t_n, t_p)_j$ with r_j inspections on the j-th airplane is defined by formula:

$$\hat{p}_{f1j} = (1-w)^{r_j} \tag{20}$$

w – is a probability that planned inspection will be made with required accuracy

Estimation of mean P_f for N airplanes:

$$\hat{p}_f = \frac{1}{N}\sum_{j=1}^{N} P_{f1j} \tag{21}$$

Relevant curves $P_f = P_f(\Delta)$ are shown in Figures 11-14

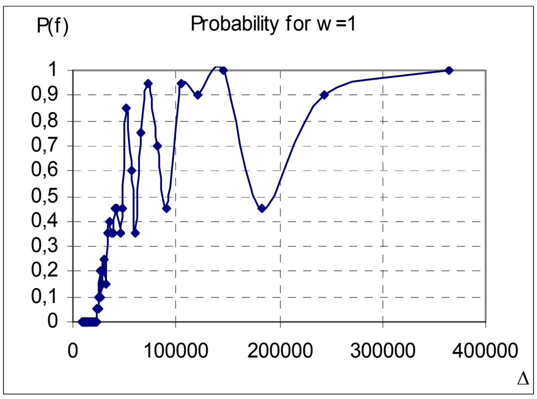


Figure 11. *Failure probability for the corroded speciemn with specified life = 730000 cycles*

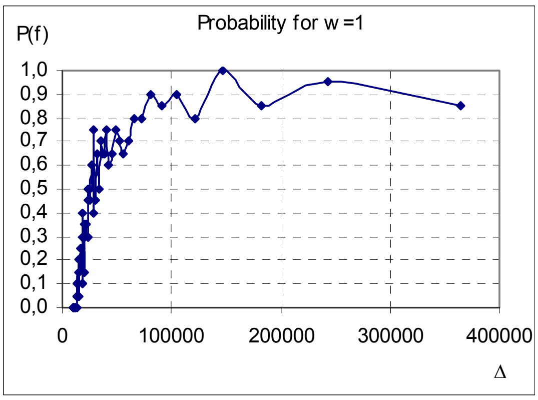


Figure 12. *Failure probability for the non-corroded speciemn with specified life = 730000 cycles*

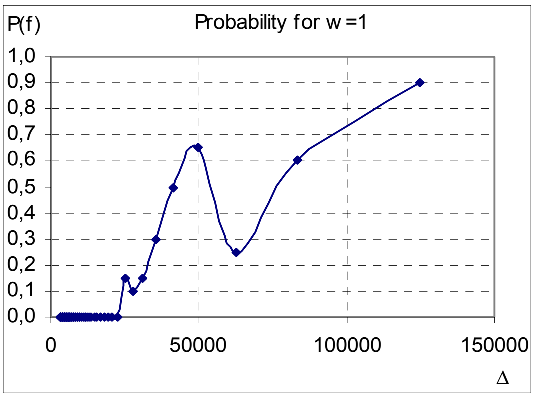


Figure 13. *Failure probability for the corroded speciemn with specified life = 250000 cycles*

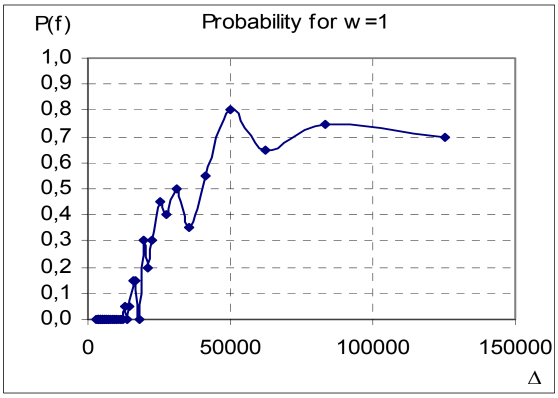


Figure 14. *Failure probability for the non-corroded speciemn with specified life = 250000 cycles*

Failure can happen in any interval $(0,t_1)$, (t_1, t_2), …, (t_{n-1}, t_n), (t_n, t_{n+1}). The probability that the visual fatigue crack appears in the interval and conditional probability that simultaneously the critical time T_f will be less than t_i [3]:

$$p_i^* = P\left(\left\{t_{i-1} < T_d < t_i\right\} \cap \left\{t_{i-1} < T_f < t_i\right\}\right) = \tag{22}$$

$$p_i^* = P\left(\left\{t_{i-1} < \frac{C_d}{Q} < t_i\right\} \cap \left\{t_{i-1} < \frac{C_f}{Q} < t_i\right\}\right) =$$

$$p_i^* = P\left(\max\left(\frac{C_f}{t_i}, \frac{C_d}{t_i}\right) < Q < \min\left(\frac{C_d}{t_{i-1}}, \frac{C_f}{t_{i-1}}\right)\right) = \tag{23}$$

$$p_i^* = P\left(\frac{C_f}{t_i} < Q < \frac{C_d}{t_{i-1}}\right), \text{ if } \frac{C_d}{t_{i-1}} \le \frac{C_f}{t_i}.$$

So in this case

$$p_i^* = \begin{cases} 0, if\ \dfrac{C_d}{t_{i-1}} \le \dfrac{C_f}{t_i} \\ \pi_i, if\ \dfrac{C_d}{t_{i-1}} > \dfrac{C_f}{t_i} \end{cases} \tag{24}$$

$$\pi_i = F_Q\left(\frac{C_d}{t_{i-1}}\right) - F_Q\frac{C_f}{t_i} = \tag{25}$$

$$\pi_i = \phi\left(\frac{\ln\left(C_d/t_{i-1}\right) - \theta_0}{\theta_1}\right) - \phi\left(\frac{\ln\left(C_f/t_i\right) - \theta_0}{\theta_1}\right)$$

In which θ_0, θ_1 are mean and standard deviation of lnQ. We suppose the normal distribution of lnQ. If we assume that the α is random variable also and use the denotation P_i^* as a function of α then in order to get the mean value of P_i^* we have to do integration.

$$p_i^* = \int_0^\infty p_i^*(X)\,dF_\alpha(X) \tag{26}$$

And finally the failure probability is defined in the following way:

$$p^{*}=\sum_{i=1}^{n}p_{i}^{*} \tag{27}$$

After we have got function $P_f = P_f(n)$, we can choose n (number of inspections) for allowable failure probability as shown in Figure 15-18.

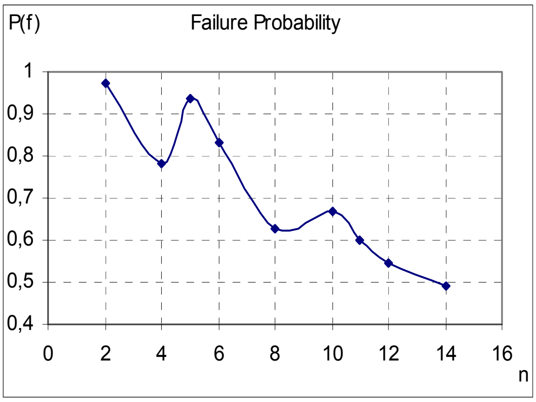


Figure 15. *Failure probability for the corroded speciemn with specified life = 730000 cycles*

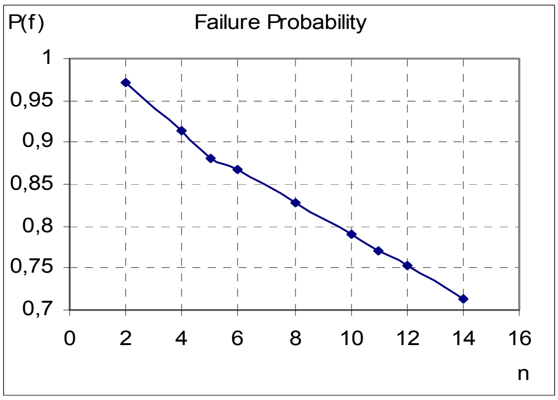


Figure 16*. Failure probability for the non-corroded speciemn with specified life = 730000 cycles*

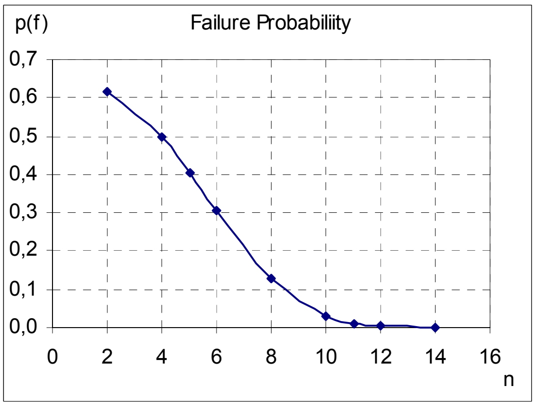


Figure 17. *Failure probability for the corroded speciemn with specified life = 250000 cycles*

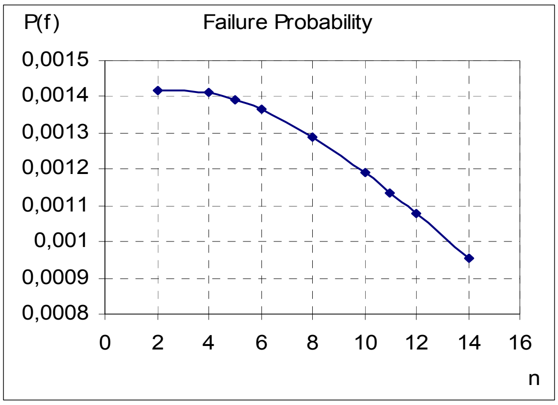


Figure 18*. Failure probability for the non-corroded speciemn with specified life = 250000 cycles*

In the following Table 5., we observe the probability of failure (Pf) for different specified lives (SL). We see that for very large SL (171120 flights) 14 inspections (n=14) give comparable value of Pf and in this case there are not significant difference between corroded and non-corroded specimens. But the difference is very significant for relatively small SL (58603 flights): for 8 inspections Pf is equal to 0.127 for corroded and Pf is equal to 0.00129 for non-corroded specimens.

Table 5. Influence of corrosion on the required number of inspections

Cycles	Flights	Corroded Specimens		Non – Corroded Specimens	
		P_f	n	P_f	n
730000	171120	0.49	14	0.71	14
250000	58603	0.127	8	0.00129	8

5. Conclusions

1. If the specified life is 730000 cycles (171120 flights) then probability of failure is unacceptably large for both cases when there is and there is not corrosion even for inspection number n=14.

2. If the specified life is relatively small that is 250000 cycles (58603 flights) than corrosion increases the probability of failure at the same number of inspections very significantly.

References

1. Dr. Y. Xiong, Mr. G. Eastaugh and Dr. G. Shi, Probabilistic failure analysis of fuselage splice joints with multiple site fatigue damage and corrosion, www.nrc.ca/iar/pdfs/aiaaj99_paper.pdf.
2. Ю. М. Парамонов и В. П. Кузнецов, Живучесть авиационных конструкций. Оценка эффективности эксплуатации по состоянию, МУ по выполнению самостоятельной работы, Рига, 1990, 24с.
3. Yu. M. Paramonov and R. Bharati, Mathematical Models of the Process of Airframe Fatigue Problem Solution for the purpose of Automated System Development, Transactions of RAU, Mechanical Development, 1997, 123-126.

APPLICATION OF STATISTICAL METHODS FOR QUALITY EVALUATION OF ASPHALT CONCRETE MIXTURE PRODUCTION

Henrikas Sivilevičius

Vilnius Gediminas Technical University, Faculty of Transport Engineering,
Department of Transport Technological Equipment
Plytines Street 27-303, LT-2016 Vilnius, Lithuania,
Tel. +370 5 2744785, Fax (+370 5) 2745060, E-mail: henrikas@ti.vtu.lt

Abstract

Mathematical models revealing the ideal structural homogeneity of asphalt concrete, shown by uniform distribution of its mineral components (coarse aggregate, fine aggregate, filler) and bitumen, are presented. The asphalt concrete structure is proposed to be investigated as a whole of three levels (asphalt binder, asphalt mastic, asphalt concrete mixture) of incomplete, optimal and smallest structural elements of the structure, consisting of grains of various diameters as well as adhesive films of oriented and volumetric bitumen. It is demonstrated that it is not always possible to produce an asphalt concrete mixture conforming to the project composition and standard physical-mechanical parameters due to technological errors. Mathematical models, showing technological factors and their impact on the quantity of components in the mixture, were constructed based on the factual homogeneity of asphalt concrete mixture identified by statistical methods. The additive model consists of variance, which depend on the stability of the used aggregate grading, their dosing errors and the quality of mixing. Average variance of mineral components' quantities in the produced asphalt concrete mixture as well as the impact in per cent of comprising technological factors were calculated from experimental data. The most important trends of technology upgrading, the practical application of which may considerably improve the quality of asphalt concrete mixture production were substantiated.

Keywords: asphalt concrete mixture, production technology, homogeneity, mathematical simulation

1. Introduction

In European and countries of other continents as well as Lithuania mixtures produced with organic binder for laid elastic road pavement have been used widely. The quality of asphalt concrete and other bitumen mixtures is not always high; therefore, the service life of pavement set up of these mixtures is short-term. If the dependability and service life of the roads of national significance were increased, the efficiency of the road transport system would be improved and construction, reconstruction, repair and maintenance works' costs would be reduced.

When integrating into the EU and accessing NATO, the Lithuanian road network and its technical parameters should be appoximated to the world standards and comply not only with the current needs but with the rapidly increasing number of cars as well as fast, safe, and convenient international communication level. Therefore, the conditions of the roads in the country shall be improved, and their asphalt pavement shall conform to the quality indicators meeting the increasing traffic of heavy vehicles and cars.

Studies were conducted to identify how asphalt concrete functions in the road pavement when it is influenced by climatic conditions and loads of vehicle wheels, how characteristics of materials used for its production and structure change. The most important findings were published in the United Kingdom [1, 2], USA [3], Canada [4], Russia [5,6], as well as other countries.

Lithuanian researchers have been working in the field of increasing the durability of asphalt concrete pavement as well [7–9]. Data received from research show that characteristics, structure, and insufficient durability of asphalt concrete depend on the quality of the asphalt concrete mixture used. Therefore, thorough investigations have been carried out. The strength of asphalt concrete is influenced not only by external and internal factors, road pavement base structure, the thickness of layers, but also by the mixture characteristics which are formed when producing it in a plant.

The influence of stochastic parameters of the process of asphalt concrete production on the product quality in the mixing plant of the old type was studied by S. Rokas [10] and V. A. Borisov [11] in 1975–77. These researchers as well as their disciples identified the actual values of materials characteristics and technological parameters of asphalt concrete mixture (ACM) production and its composition as well as deviations from physical-mechanical indices. They proved that the values of deviations shall be reduced and actual deviations shall be taken into account when handling the technological process of ACM. However, the technological causes influencing the quality of ACM production due to the complexity of special experiments requiring to interfere with the process of production and produce some products of a bad quality as well as a lack of theoretical knowledge have not been studied thoroughly yet.

Designers and manufacturers of asphalt concrete mixing plants (ACMP) shall know how old and new equipment functions (if the products produced in them suit for road pavement construction). The construction of new computerized ACMP differs from the older equipment; however, the most technological processes applying in them have the same kinetics and are the same in essence. Different systems of parameters' control and handling, the effectiveness of which depends on the operators' work, are used in them. The latest research shows [12,13] that the technological process of ACM production should be improved since it means improving the quality of the product.

2. The Condition of Asphalt Concrete Paved Roads and Causes Influencing it

Rapid increase of the number, mass, driving speed of vehicles and stricter requirements of road transport negative impact (accident rate, air pollution) improves not only the road network but also road technical parameters. In the last 40 years, the length of Lithuanian roads has increased from 38 700 km in 1960 to 76 600 km in 2002. The specific weight of paved roads has increased even three times: from 30,7% to 91,3%. However, the length of state roads did not increase last year and made up 21 312,6 km. The specific weight of roads with improved pavement made up 56,7%.

In Lithuania as well as other countries in the world road pavement is set up from hot mixtures produced with organic binder, which enables to pave and compact strong conglomerant single or double-layer system. The Lithuanian roads of national significance are paved with asphalt concrete: 56,18%, cement concrete: 0,40%, gravel: 43,31% and cobble-stone: 0,11%. Such ratio in per cent of the road pavement distribution shows that ACM has become the most important material in their construction. To implement the gravel road paving programme, to reconstruct, regenerate, repair and maintain exploited roads, each year 2-3 million tons of ACM has to be produced, which costs approximately 350-500 million Litas.

The data of research carried out by the Transport and Road Research Institute each year shows that the road strength index of Lithuanian main and national roads I_{st} has decreased by 12% on average from 1992 to 2000. The main roads deteriorate slowlier (I_{st} decreased from 1,00 to 0,93) than national roads (I_{st} decreased from 1,00 to 0,85), and this dynamics is approved by the research carried out by us.

The rapidly decreasing road strength index and the increasing pavement defects (cracks, plastic deformations, defects due to bitumen, surface wear) are influenced not only by the increasing traffic of heavy vehicles, but also by the insufficient quality of ACM used for paving. The types of technological equipment of ACM production used at the end of 40s, at the end of 50s and the beginning of 50s, have consuderably changed. Asphalt concrete mixture plant construction has changed; they have become more efficient and universal, the adjusted mixture production technology has improved, and separate dosing (batching) operations and the whole production process has been automated and later computerized.

ACM quality assessment investigators noted that the selected optimal materials' mass ratio selected in laboratory research is not complied with in large-scale production. It is suggested to control and manage ACM quality through the application of methods of mathematical statistics [14] enabling to verify its characteristics more precisely. However, technological causes influencing on the variance of the produced component quantity and the impact of factors were not investigated thouroughly.

3. Models of Asphalt Concrete and its Mixture‘s Ideal Structural Homogeneity

When analyzing the homogeneity of ACM, it is divided into the smallest structural elements (SSE). The smallest structural element ABM_{SSE} of asphalt binding material (ABM) of the first level is made up of one particle of fillers $\left(F_1\right)$ with protective films of oriented $\left(B_{oF}\right)$ and volumetric $\left(B_{vF}\right)$ bitumen:

$$ABM_{SSE} = F_1 + B_{oF} + B_{vF} . \tag{1}$$

The smallest structural element AM_{SSE} of the second level asphalt mastic (AM) is made up of one fine aggregate particle $\left(FA_1\right)$ with the protective films of oriented $\left(B_{oFA}\right)$ and volumetric $\left(B_{vFA}\right)$ bitumen and evenly laid out related n_{opt} ABM_{SSE}:

$$AM_{SSE} = FA_1 + B_{oFA} + B_{vFA} + \sum_{i=1}^{n_{opt}} ABM_{SSE\ i} . \tag{2}$$

The smallest structural element ACM_{SSE} of the third level asphalt concrete mixture (ACM) is made up of one coarse aggregate particle $\left(CA_1\right)$ with protective films of oriented $\left(B_{oCA}\right)$ and volumetric $\left(B_{vCA}\right)$ bitumen and evenly laid out related $n_{opt}^{'}$ AM_{SSE}:

$$ACM_{SSE} = CA_1 + B_{oCA} + B_{vCA} + \sum_{i=1}^{n_{opt}^{'}} AM_{SSE\ i} . \tag{3}$$

Having arranged SSE of all three levels next to each other technologically when their number approximates to infinity, we obtain ABM, AM and ACM, the structure of which is demonstrated through the following mathematical models:

$$ABM = \sum_{j=1}^{\infty}\left(F_1 + B_{oF} + B_{vF}\right)_j , \tag{4}$$

$$AM = \sum_{j=1}^{\infty}\left[\left(FA_1 + B_{oFA} + B_{vFA}\right) + \sum_{i=1}^{n_{opt}} ABM_{SSE\ i}\right]_j , \tag{5}$$

$$ACM = \sum_{j=1}^{\infty}\left\{\left(CA_1 + B_{oCA} + B_{vCA}\right) + \sum_{i=1}^{n_{opt}^{'}} AM_{SSE\ i}\right\}_j . \tag{6}$$

If the quantity of SSE component of any level is less than the average mass (volume) specified in the composition design, such SSE is incomplete and, if the quantity is larger, it is satiated. In the produced, the quantity of ACM incomplete and satiated ABM, AM and ACM, SSE shall be as little as possible since it is the most important technological process strategy of ACM production, enabling the road pavement to use homogeneous production maximally.

4. Statistical Criteria of Asphalt Concrete Mixture Production Quality

The charateristics of produced ACM are influenced by the characteristics of used materials, the quality of composition design and technological factors of ACMP production. The parameters of ACM production technological factors shall enable to produce ACM, the composition and physical-mechanical indicators of which shall not differ from the project and standard values. The quality of ACM production shows how to meet this requirement. ACM production quality is the wholeness of ACM production technological process characteristics which influence on the compliance of the process parameters and its results with the set requirements.

Physical-mechanical indicators show the optimum, lay-out rationallity, adhesion, interaction of hard, liquid and gas phases as well as a capability to resist climatic factors and transport loads of the produced and used ACM components‘ quantity. Some of their values are standard and have two (lower T_L and upper T_U) permitted values (residual porosity T_{bit}, Marshall flow PM), only one lower T_L (Marshall stability SM, the ratio of stability and flow TM, compaction coefficient K_{st}) or upper T_U (swelling H - non-standardized).

Deviations from ACM composition (the quantity of components in it) or physical-mechanical indicators from optimal values are shown by the role of population mean μ_0 in terms of design and norms as well as width $2t\sigma_0$ of its separate values‘ dispersion around μ_0 ($t-$ coefficient of significance, depending on the accepted probability and the number of experiments, σ_0 – population standard deviation in the given period).

The data obtained from our and other authors‘ research showed that ACM component quantity and physical-mechanical indicators distribute according to the normal (Gauss) law. When improving ACM quality according to the indicator of normal distribution of its characteristics during the minimal short time (from t_0 to t_1), population mean μ_0 is replaced by value μ_1, and value σ_0 of its dispresion width shall be reduced to σ_1 (Fig. 1). Such purposeful correction of ACM production technological process enables to reduce the total number of separate values of the indicator not complying with the requirements in the norms, which is aimed at through research measures.

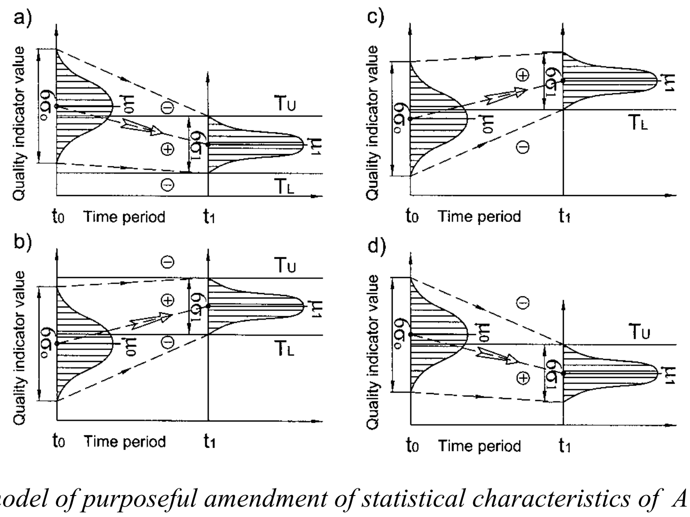


***Figure 1**. The model of purposeful amendment of statistical characteristics of AC and its mixture Quality indicator value when technical specifications specify:*
a, b – upper T_U and lower T_L tolerance limits; c – lower limit T_L; d – upper limit T_U

5. Classical Technology Model of Asphalt Concrete Mixture Production

ACM is produced according to different technological schemes applied in ACMP, which impact on the heterogeneous connection of initial materials and intermediate products, distribution and segregation of theirparticles, due to which the final product with different characteristics is obtained. When producing it following the technological scheme (Fig. 2), which has become traditional, initial cold mineral materials (ICMM) are dosed uniformly, their cold mixture (CM) is dried, heated and dust is removed from air and gas (MRD), hot mixture (HM) is sieved into fractions of 3-5 (Fr1, ..., Fr5), its MRD and cold imported fillers (IF) dosed discretely, a mixture of finally dosed materials (FDMM) is mixed for a short time (GDMM), a batch of dosed bitumen is added (B) and ACM is mixed. ACM components (F, FA, CA, B) redistribute in the technological production process by changing the mass in per cent and and its stability in materials and mixtures.

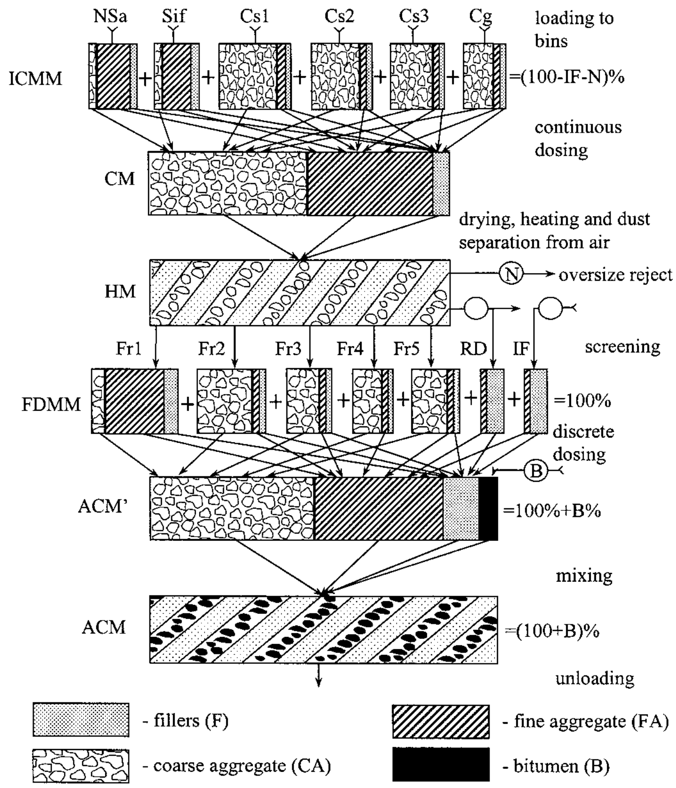


Figure 2. *The model of ACM components' variance in mineral materials and mixtures when producing it in asphalt concrete mixing plant according to the traditional technology*

Empiric data obtained through the study of initial mineral materials, intermediate mixtures and materials' grading used in seven ACMP showed that finally dosed hot fractions are contaminated not only with by-particles but they are heterogeneous due to the segregation in bin's sections. The main function of screening through ACMP technological sieves in most cases does not give a positive effect. Hot fraction (HFr) of 0-5 mm is contaminated with CA particles larger than 5 mm on average, $\mu_{ij} = 5,8\,\%$ (from 0,9% to 13,7%) and F finer than 0,071 mm particles $\mu_{ij} = 9,8\,\%$ (from 6,2% to 13,4%). In hot fraction of 5-15 mm, the quantity of particles finer than 5 mm is $\mu_{ij} = 16,8\,\%$ on average (from 4,6 % to 32,8%). In particles finer than 5 mm, the quantity is $\mu_{ij} = 1,4\,\%$ F on average (from 0,2% to 3,0 %), the actual quantity of which shall be known when adjusting dosing units.

Having used the data from 337 samples of HFr of 0-5 mm in 336 samples of HFr of 5-15 mm and 144 samples of cold IF grading, strong curved-linear correlation dependence is identified between sample mean $\overline{X}$ and standard deviation S of full siftings through control screens. Having approximated empiric data, the following regression equations are written: hot mineral materials of 0-5 mm fraction

$$\hat{S} = \sqrt{6,07 \cdot 10^{-6}\, \overline{X}^{2,478} \left(100 - \overline{X}\right)^{1,698}}\ ; \tag{7}$$

- hot mineral materials of 5-15 mm fraction

$$\hat{S} = \sqrt{7,28 \cdot 10^{-5}\, \overline{X}^{1,928} \left(100 - \overline{X}\right)^{1,481}}\ ; \tag{8}$$

- cold imported filler

$$\hat{S} = \sqrt{9,84 \cdot 10^{-6}\, \overline{X}^{2,072} \left(100 - \overline{X}\right)^{1,308}}\ . \tag{9}$$

Regression curves drawn according to them (Fig. 3. Fr 0-5 mm) show that the greatest variance (destability) is typical of particles with the diameter of 50% mass of the total quantity of particles of mineral materials or mixture. This $\overline{X} = 50\ \%$ screening through control screens enables to compare grading homogeneity of any aggregates: the larger $S_{50}\ \%$ is, the more heterogeneous material is.

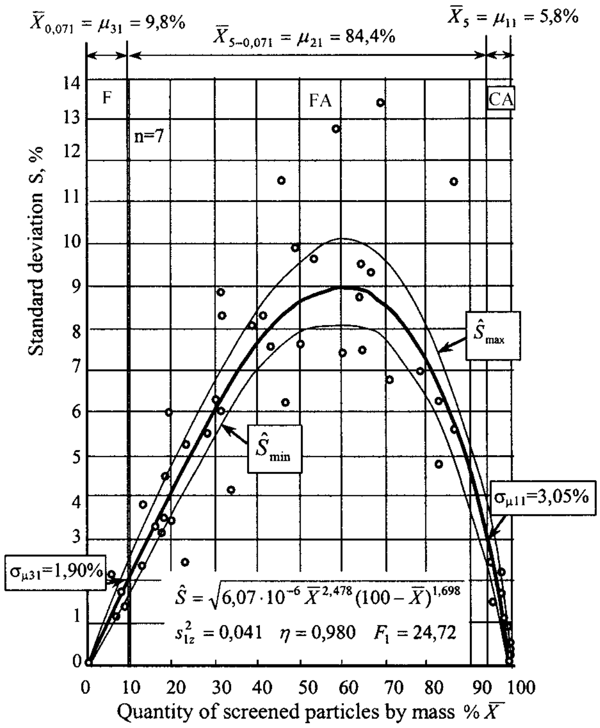


Figure 3. *Dependence of standard deviation S of HMM finest fractions of 0-5 mm siftings' mass screened through technological sieves of seven ACMP on arithmetic mean* $\overline{X}$

The spatial distribution model (Fig. 4) of HFr 0-5 mm probabilities density dependence on $\overline{X}$ and $\hat{S}$ shows the same interrelation. If j – mineral materials' $\hat{S} = f\left(\overline{X}\right)$ is known and $\overline{X} = \mu_{ij}$ is accepted, standard deviation $\sigma_{\mu ij}$ of any i diamater particle can be identified from regression equations or graphically (any mineral component i of ACM: F, FA, CA).

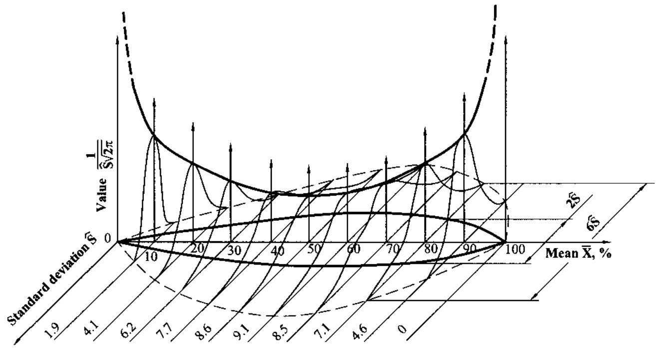


Figure 4. *The spatial model of probabilities' density function of the finest hot fraction of 0-5 mm full siftings screened through control screens of ACMP technological sieves*

6. Statistical Study of Factors Influencing ACM Homogeneity

Complete population variance σ_i^2 of mineral component quantity in ACM lot (party) is calculated according to the following formula

$$\sigma_i^2 = \sigma_{Mi}^2 + \sigma_{Ei}^2 + \sigma_{Ti}^2, \tag{10}$$

here σ_{Mi}^2 – ACM i mineral component quantity variance, which depends on the sample selection methodology, %, σ_{Ei}^2 – variance, which depends on the sample testing errors, %; σ_{Ti}^2 – variance, which depends on ACM production ACMP technological operations' accuracy and stability, % .

Variance σ_{Ti}^2 value mostly depends on the grading stability of mineral materials (HFr, cold IF and RD) used for producing ACM (value $\sigma_{g.s.ij}^2$) and dosing errors (value $\sigma_{d.e.ij}^2$) and the quality of mixing them in the batch mixer (value $\sigma_{b.h.i}^2$). Its actual value in periodic ACMP producing ACM according to the traditional technology is calculated from the additive model (Fig. 5):

$$\sigma_{Ti}^2 = \sum_{j=1}^{m} \sigma_{g.s.ij}^2 + \sum_{j=1}^{m} \sigma_{d.e.ij}^2 + \sigma_{b.h.i}^2, \qquad i = 1,...,k,\ j = 1,...,m \tag{11}$$

or extended mathematical model

$$\sigma_{Ti}^2 = \sum_{j=1}^{m} \left(\frac{q_j \sigma_{\mu ij}}{Q_{mp}} \right)_j^2 + \sum_{j=1}^{m} \left(\frac{q_j \mu_{ij} V_{qj}}{100 Q_{mp}} \right)_j^2 + \left(\sigma_{hio} e^{-Mt} \right)^2, \tag{12}$$

here q_j – average factual mass of j mineral material dose, kg; $\sigma_{\mu ij}$ – mass % of quantity standard deviation of mineral component i quantity in j mineral material; μ_{ij} – mass % of population mean of i mineral component quantity in j mineral material; V_{qj} – j mineral material dosing error the numerical value of which is equal to the variation coefficient of its dose mass, %; Q_{mp} – average mass of batch ACM mineral part, kg.; σ_{hio} – quanity of i mineral component in ACM sample standard deviation when starting the mixing process, mass %; e – base of Napierian logarithm; M – coefficient showing the characteristics of the batch mixer; t – duration of mixing the materials in a batch mixer, s.

Having inserted average values of most frequently used arguments $q_j, \sigma_{\mu ij}, \mu_{ij}, V_{qj}$ and Q_{mp} in the mathematical model (12), the impact in per cent of its factors on the produced ACM mineral part heterogeneity is estimated (Table 1).

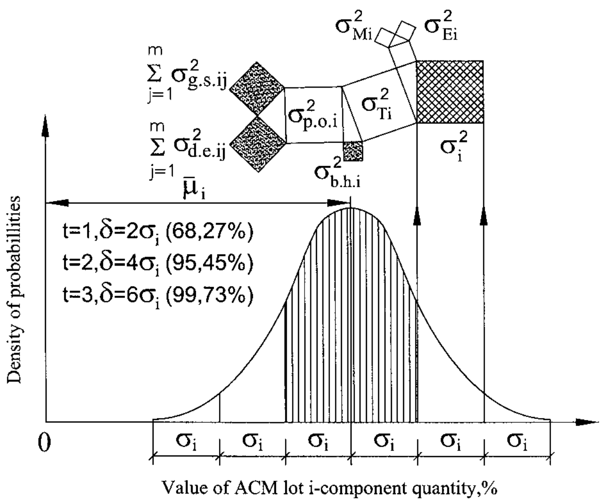


Figure 5. *Model of conjugation of standard deviation σ_i and variance σ_i^2 of asphalt concrete mixture lot i-mineral component (F, FA, CA) quantity*

Table 1. Technological factors making up calculated values of variance of ACM components' quantity and their impact in per cent

Mineral component of asphalt concrete mixture	Calculated values of separate production factors standard deviations σ_{Ti}, variance σ^2_{Ti} and their part in percent						The calculated influence of all mixture production factors on the values of mineral component quantity stability $\sigma_{Ti}/\sigma^2_{Ti}$ %		Variance of production errors σ^2_{Tif} obtained from experimental data in %
	Due to instability of finally sieved mineral materials grading			Due to dosing errors of finally sieved mineral materials					
	Hot fraction 0-5 mm $\sigma^{Fr0-5}_{g.s.}$	Hot fraction 5-15 mm $\sigma^{Fr5-15}_{g.s.}$	Cold imported filler $\sigma^{IF}_{g.s.}$	Hot fraction 0-5 mm $\sigma^{Fr0-5}_{d.e.}$	Hot fraction 5-15 mm $\sigma^{Fr5-15}_{d.e.}$	Cold imported filler $\sigma^{IF}_{d.e.}$			
Coarse aggregate (particles larger than 5 mm)	1,45/2,10	1,64/2,69	0/0	0,11/0,01	2,55/6,50	0/0	3,36/11,30		9,86–17,53
	18,7	23,7	0	0,1	57,5	0	-	100	100
Fine aggregate (particles from 0,071 to 5 mm)	1,72/2,96	1,66/2,74	0,09/0,01	1,53/2,34	0,47/0,22	0,16/0,026	2,88/8,30		9,04–17,60
	35,7	33,1	0,1	28,2	2,6	0,3	-	100	100
Fillers (particles smaller than 0,071 mm)	0,91/0,82	0,17/0,03	0,09/0,01	0,18/0,03	0,04/0,01	0,74/0,541	1,20/1,44		0,60–1,79
	57,3	2,1	0,5	2,2	0,1	37,8	-	100	100

To identify the influence of actual values of technological factors parameters of ACM production on the stability of of its mineral part components quantity, values of summands making up variance σ^2_{Ti} were calculated. Values of variables mostly occurring in practice were used, which were obtained from studies carried out earlier [15,16].

When adequacy of model (12) was tested, the following values depending on concrete technological conditions were used:

- ACM under production the diameter of its largest particles is 16 mm (e g 0/16 S-V) in sampling action ACMP sieving the mixture of hot mineral materials into hot fraction of 0-5 mm, 5-15 mm and 15-35 mm;
- Two finally dosed hot fractions (Fr 0-5 mm and Fr 5-15 mm) weighed with cold imported filler in one bin of the dosing equipment (IF): at Fr 0-5 mm dose by adding Fr 5-15 mm dose, and, in the end, IF dose;
- Mass of mineral part of ACM batch is $Q_{mp} = 610$ kg, bitumen dose is equal to $q_B = 40$ kg; mass of ACM batch is $Q = 650$ kg;
- Equal portions of each hot fracion are used (291 kg dose Fr 0-5 mm: 47,7% and 292 kg dose Fr 5-15 mm 47,9% and 27 kg dose of IF: 4,4%);
- Errors V_{qj} of materials dosing and quantities μ_{ij} of components (crushed stone, sand, and imported filler) in dosed mineral materials and their average standard deviations $\sigma_{\mu ij}$ are taken from [15].

Values of standard deviations σ_{Ti} and variance σ_{Ti}^2 of crushed stone, sand, and imported filler quantity in asphalt concrete mixture, depending on technological factors of its production, calculated from mathematical model (12) are presented in Table 1. When compared with actual values σ_{Tif}^2 obtained during the experiment, the calculated values of variance $\sigma_{Ti\,f}^2$ comply with actual values: quantity of coarse aggregate $\sigma_{T1}^2 = 11,30\,\%$, and $\sigma_{T1f}^2 = 9,86 - 17,53\,\%$; quantity of fine aggregate $\sigma_{T2}^2 = 8,30\,\%$, and $\sigma_{T2f}^2 = 9,04 - 17,60\,\%$; quantity of filler $\sigma_{T3}^2 = 1,44\,\%$, and $\sigma_{T3f}^2 = 0,60 - 1,79\%$. Obtained data enables to state that the constructed model (12) is dependable and shows the influence of separate factors on the stability of mineral part of ACM.

7. Conclusions

1. The influence of separate factors on the stability of asphalt concrete mixture as well as the sequence of error transmission presented in Table 1 show that variations of coarse aggregate (CA) quantity in it depend on dosing errors of hot fraction 5-15 mm (57,5 %), stability of its grading (23,7 %) as well as stability of hot fraction 0-5 mm grading (18,7 %). They hardly depend on dosing errors of fraction 0-5 mm (0,1 %). Errors of dosing and stability of grading of imported filler does not influence on the variation CA quantity in asphalt concrete mixture.

2. Stability of fine aggregate (FA) quantity in the produced asphalt concrete mixture mostly depends on the stability of grading of fraction 0-5 mm and fraction 5-15 mm (35,7 % and 33,1 %, respectively), dosing errors of fraction 0-5 mm (28,2 %). It hardly depends on dosing errors of fraction 5-15 mm (2,6 %) and almost does not depend on dosing errors and stability of grading of imported filler (0,3 % and 0,1 %, respectively).

3. Stability of grading of fraction 0-5 mm (57,3 %) and dosing errors of imported filler (37,8 %) have greatest influence on the variation of filler (F) quantity in asphalt concrete mixture. It hardly depends on dosing errors of fraction 0-5 mm (2,2 %) and stability of grading of fraction 5-15 mm (2,1 %) and almost does not depend on the stability of grading of imported filler (0,5 %) and dosing errors of fraction 5-15 mm (0,1 %).

4. When values of mathematical model (12) are changed, values of variance σ_{Ti}^2 and the influence in per cent of summands making it up change as well. When increasing homogeneity of asphalt concrete mixture composition, the parameters of those technological factors due to which values of variance σ_{Ti}^2 increase shall be improved at first.

References

1. *Bituminous mixtures in road construction* / Edited by Robert N. Hunter. London: Thomas Telford, 1997. 441 p.
2. Jones Lee. Recent developments in coating plant technology. *Quarry Management*, 1986, 13, No 10, p. 25-30.
3. G Li., Y Li., J. B Metcalf., S.-S.Pang Elastic modulus prediction of asphalt concrete. *Journal of Materials in Civil Engineering.* 1999, Aug., Vol 11, No 3, p. 236-241.
4. M.T. Obaidat, H.R. Al-Masaeid, F.Gharaibeh, An innovative digital image analysis approach to quantify the percentage of voids in mineral aggregates of bituminous mixtures. *Canadian Journal of Civil Engineering.* Vol 25, No 6, 1998, p.1041-1049.
5. G. N.Kirjukhin, V. M. Jumashev, M. B.Sokalskaja, Laying of the uniform top dressing. *Science and engineering in road building* (Наука и техника в дорожной отрасли). 1998, No 1 (4), p. 13-15 (in Russian).
6. G. S. Bakhtrakh, Model of estimation of the non-rigid pavement service life. *Science and engineering in road building* (Наука и техника в дорожной отрасли), 2002, No 2 (21), p. 17-20 (in Russian).
7. K.Petkevičius, H. Sivilevičius, The substantiation of demanded service properties of highway asphalt concrete covers. *Transport Engineering* (Transportas). 2000, Vol XV, No 4, p. 184-195 (in Lithuanian).
8. V.Puodžiukas, A.Pakalnis, Determination of seasonal factors for structural condition values of bituminous pavements. *Journal of Civil Engineering and Management.* 2002, Vol VIII, No 2, p. 133-137.
9. D.Čygas, A.Laurinavičius, The main problems of using asphalt concrete mixtures for pavements of automobile roads in town streets. *City development and road. Supplement of journal „Civil Engineering“* (Miesto plėtra ir keliai. Mokslo žurnalo „Statyba“ priedas). 2000, p. 20-26 (in Lithuanian).
10. S.Rokas, *Statistical control of quality in road construction.* (Статистический контроль качества в дорожном строительстве). Moscow: Transport, 1977. 152 p. (in Russian).
11. V.A.Borisov, *Technological accuracy asphalt concrete plants and methods of its increase.* (Технологическая точность асфальтобетонных заводов и методы ее повышения) Saratov: Izdatelstvo Saratovskogo universiteta, 1975. 159 p. (in Russian).
12. I. F.Bunkin, *Automation of asphalt concrete production management.* Summary of the research report presented for habilitation (Автореферат диссертации на соискание ученой степени доктора технических наук). Moscow.: 2002, 38 p. (in Russian).
13. H. Sivilevičius, *The quality improvement system of asfalt concretemixture production technological process.* Summary of the research report presented of habilitation . Vilnius: Technika, 2003, 37 p.
14. D. C. Montgomery, *Introduction to Statitical Quality Control*. New York, Chishester, Brisbane, Toronto, Singapore: John Wiley and Sons, Inc. 1997. 667 p.
15. H. Sivilevičius, Theoretical principles and experimental data to identify stability of asphalt concrete components in finally dosed mineral materials. *Transport*, 2002, Vol XVII, No 1, p. 19-29 (in Lithuanian).
16. H. Sivilevičius. Influence of homogeneity of mineral materials‘ grading and dosing errors on the stability of asphalt concrete mixture composition // *Journal of Civil Engineering and Management*. 2003, Vol IX, No1, p. 25-35.

THE EFFECT OF VARIABLE RIGIDITY OF HAULAGE ROPE ON THE VIBRATION RATE AND RELIABILITY OF "TRAILED VIBRATION ROLLER – HAULAGE ROPE" SYSTEM

Vilius Bartulis

Vilnius Gediminas Technical University, Faculty of Transport Engineering,
Department of Transport Technological Equipment
Plytines Street 27-303, LT-2016 Vilnius, Lithuania,
Tel. +370 5 2744794, Fax (+370 5) 2744800, E-mail: vilius@ti.vtu.lt

Abstract

If a slope is being compacted by a vibration roller trailed by means of a haulage rope, then the vibrations of the roller are transferred not only to the slope surface but to the haulage rope as well. To avoid resonant vibration with the natural oscillations of the system, the proper characteristics of the roller forces should be chosen. The article aims to determine the effect of the roller's parameters on its oscillations along the generating line of the slope.

Keywords: reliability of vibration rollers, soil compaction by vibration

One of the systems used in slope strengthening (compacting) consists of a trailed vibration roller connected by a haulage rope to the main machine at the top of the slope provided with a winch. If a roller with non-directional circular vibrations (which is suitable for such operations) is used, the vibration of the roller is transmitted to a haulage rope and then to a basic machine. The larger the amplitude of roller vibration along the haulage rope, the higher the dynamic effect of the vibration roller on a basic machine and the higher the possibility of irregular vibration of the roller resulting in poor soil compaction.

The oscillations along the generating line of the slope, caused by the force exerted by a trailed vibration roller compacting the slope are usually dampened due to the rigidity and the internal friction of the haulage rope of the winch. When the roll running along the generating line of the slope vibrates, it rises off the ground (partially or completely) at every revolution of a shaft, which makes it possible to assume that the rolling friction may be neglected. Therefore, in the present paper only the interaction between the roller and a haulage rope of the winch is considered in analysing the vibration of the roll along the slope.

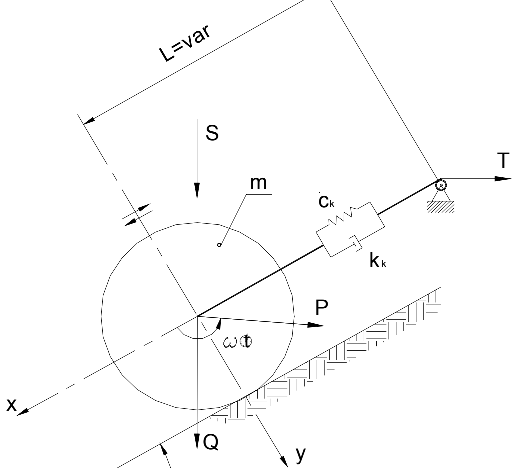


Figure 1. *A design scheme of a vibration roller trailed with a haulage rope on the slope*

A design scheme presenting the above interaction is given in Fig. 1. The vibration of a roll along the axis x may be described by an equation:

$$x''m + x'k + xc = P\cos\omega t + (Q+S)\sin\alpha\,, \quad (1)$$

here m – mass of the vibrating parts of the roller; k – internal friction reflecting the damping characteristics of the rope; c – rope elasticity; P – centrifugal force of the roller causing the vibration; Q – gravitational force of vibrating parts of the roller; S – static force of roll ballasting; α – slope angle (with respect to horizon); x – movement of the roller along the haulage rope; ω – angular frequency of forced vibration of the roll; t – duration of vibration at the particular moment of time (ωt – phase angle, determining the direction of centrifugal force P at the moment).

The natural angular vibration frequency ω_0 of the system "vibration roller – haulage rope" depends on the varying rope rigidity c (due to slowly changing length of the rope plumb line L – Fig. 2) and on the steepness α of the slope on which the roller works:

$$\omega_0 = \sqrt{c/m - (k/2m)^2} \quad (2)$$

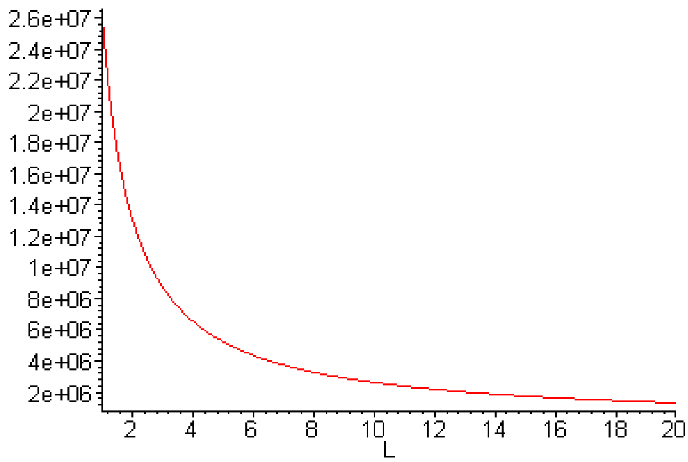


Figure 2. *Variation of rope rigidity c, depending on the length of the plumb line of the rope L*

Since we are considering the oscillations in the plane, making angle α with the horizon, while the mass vibrating in this direction is expressed as $m \sin\alpha$, then the equation (2) takes the form:

$$\omega_0 = \sqrt{c/m\sin\alpha - (k/2m\sin\alpha)^2}\,. \quad (3)$$

The vibration frequency f expressed in Hz is obtained from the angular frequency ω by using a well-known expression as follows:

$$f = \omega/2\pi\,. \quad (4)$$

As one can see in Fig. 3, high natural oscillation frequency of the considered "roller-rope" system, reaching 140 Hz is found only in a horizontal plane, when the length of the rope plumb line is minimal. The influence of the slope angle decreases when the length of the rope plumb line increases.

When the roller remotes from the basic machine, i.e. when the length of the rope plumb line increases, natural vibration frequency of the system is sharply decreased, which corresponds to the change in rope rigidity (Fig. 2). The most effective frequency of soil compaction is about 30…50 Hz [1,2]. With the particular values of slope angle and rope plumb line length, these values of the forced frequency are in a dangerous zone of the "roller-rope" system's resonance. This entails the increase of oscillation frequency along the haulage rope. The forced frequency is considered in the following range [3]:

$$0,5 f_0 \le f \le 1,5 f_0 . \tag{5}$$

The amplitude of forced oscillations may be expressed in the following way:

$$X_a = P / \sqrt{\left(C - \omega^2 m \sin\alpha\right)^2 + \omega^2 k^2} . \tag{6}$$

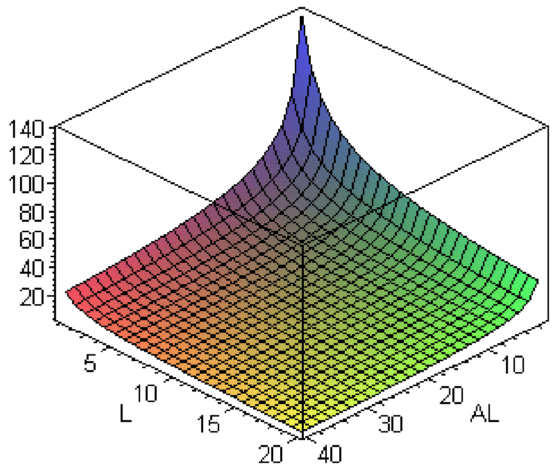
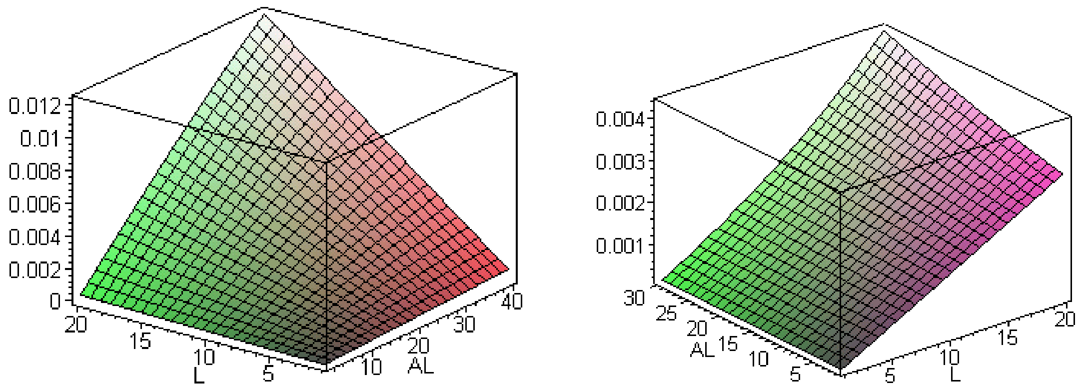


Figure 3. *The dependence of natural frequency of roller oscillations f_0 on the angle of the compacted surface f the slope α (AL) and the length of the rope plumb line L*

It is graphically represented in Fig. 4a. The influence of the slope angle gets stronger when the length of the plumb line increases, however, the latter factor is decisive.



a b

Figure 4. *The influence of slope angle AL and length of the rope plumb line L on the amplitude of forced vibration of the roller along the haulage rope x_a (a), and rope deformation under the action of static gravitational force of the roller x_c (b)*

When the roller is some meters away from the pulling winch, the amplitude of roll oscillations along the haulage rope may reach several millimeters. The more the amplitude of forced oscillations exceeds rope deformation due to the action of static gravitational force of the roller along the axis *x*, the higher is the amplitude of roll oscillations. This, in turn, depends on the relationship between force P causing oscillations and the static force $Q\sin\alpha$ [4]. The influence of AL and L on the rope deformation X_c under the action of the static force is similar to dynamic impact but is less prominent (Fig. 4b). The relationship between rope deformation X_a under the action of the dynamic force P and X_c – caused by the static gravitational force of the roller $Q\sin\alpha$, depending on the same factors, is given in Fig. 5. One can see that the vibration mode of the roller along the generating line of the slope expressed as the relationship of deformations X_a / X_c, largely depends on the slope angle (AL). Therefore, in determining force characteristics of the vibration roller for slope compaction, the impact of the slope angle and resolution of forces into components causing the changes in the dynamic characteristics of the roller should be taken into account.

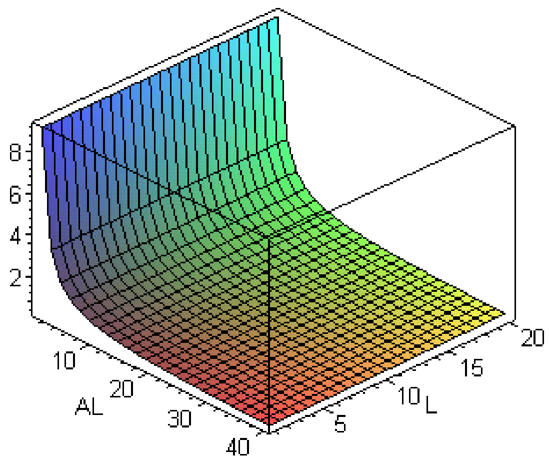


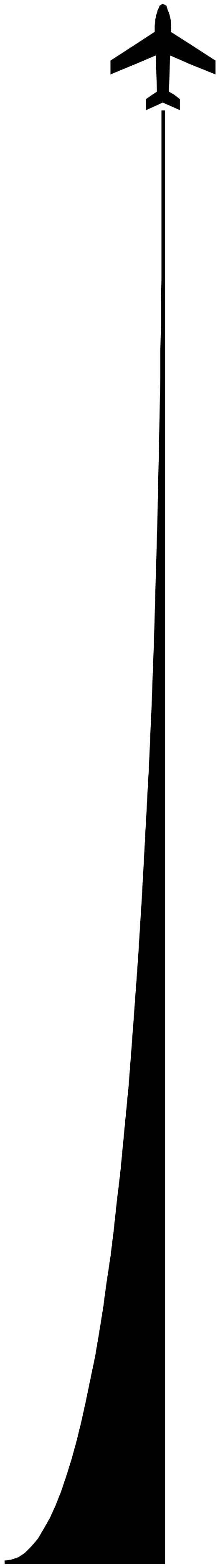
Figure 5. *The relationship between the amplitude of the forced roller oscillations along the generating line of the slope and the static rope deformation under the action of the static force* (X_a / X_c) *for varying slope angle AL and the angle of rope plumb line L*

Conclusions

Under certain conditions, the trailed vibration roller of circular oscillations operating on the slope and connected to a pulling winch by a rope may cause a resonance of the roller oscillations along the haulage rope. To avoid the resonance and, therefore, irregular vibration of the roller reducing the reliability of the system, the forces relating to the operation of the roller should be properly determined. The relationships offered in the present investigation which take into account slope angle and the length of the plumb line of the haulage rope may help to achieve this.

References

1. G. Arquie, G. Morel. *Le compactage*. Paris, Eyrolles, 1988.
2. Хархута Н. Я., Васильев Ю. М. *Прочность, устойчивость и уплотнение грунтов земляного полотна автомобильных дорог.* М.: Транспорт, 1975.
3. Пановко Я. Г. *Введение в теорию механических колебаний*. М., Наука, 1991.
4. В. Бартулис. Влияние угла откоса на выбор машины для уплотения наклонной поверхности // The Fourth International Conference *"Heavy Machinery HM 2002"*,A11-A13.



Session 6

Efficiency of Electronics Systems and Devices

ANALYSIS OF DIFFERENCE SCHEMES IN MODELING OF GYROTRON EQUATION

O. Dumbrajs, H. Kalis and A.Reinfelds

Helsinki University of Technology, Espoo, FIN-002150, Finland
Institute of Mathematics of Latvian Academy of Sciences and University of Latvia,
Akadēmijas laukums 1, Rīga, LV-1524, Latvia

Abstract

In this paper we will examine in detail different in the time two and three layer finite-difference schemes (Richardson's, Simpson's, Richtmayer-Morton's) that arises in solving single mode gyrotron equation. It is initial-boundary value problem for Schrödinger type partial differential equation with complex value boundary conditions of third kind. For stability analysis we will use the uniform grid in the space.

1. Formulation of Problem

We have the following initial-boundary value problem for Schrödinger type partial differential equation [1]

$$\begin{cases} i\dfrac{\partial f}{\partial t} = \dfrac{\partial^2 f}{\partial x^2} + \partial f + F \\ \quad f(t,0) = 0, \qquad \dfrac{\partial f(t,L)}{\partial x} = -i\gamma f(t,L),\, f(0,x) = f_0(x), \end{cases} \tag{1}$$

where $x \in (0,L)$ – is the space coordinate, $t > 0$ – is time, $\delta, \gamma > 0$ – are real parameters, $i = \sqrt{-1}$, $f(t,x)$ – is the unknown function, $f_0(x)$, $F(t,x)$ – are given complex functions.

Denote by $\omega_h = \{x_j = jh,\, j = \overline{1, N-1},\, Nh = L\}$ a uniform homogeneous spatial grid for x and by $\omega_\tau = \{t_n = n\tau,\, n-1,2,\ldots\}$ a temporal grid for t (h and τ are corresponding step-lengths). Substitute the continuous function $f = f(t,x)$ by the discrete grid function $y = y(t,x)$, $t \in \omega_\tau$, $x \in \omega_h$ with values $y(t_n, x_j) \equiv y_j^n$.

The derivatives of equation (1) approximate using the following finite-difference expressions:

1. the second order approximation in space

$$\frac{\partial^2 f(t_n, x_j)}{\partial x^2} = \Lambda f_j^n + r_1,$$

$$\Lambda f_j^n \equiv \frac{f(t_n, x_{j+1}) - 2f(t_n, x_j) + f(t_n, x_{j-1})}{h^2}$$

$$r_1 = -\frac{h^2}{12}\frac{\partial^4 f(t_n, \varsigma_1)}{\partial x^4},\ \varsigma_1 \in [x_{j-1}, x_{j+1}];$$

2. the first order approximation in time (for two level schemes)

$$\frac{\partial f\left(t_n,x_j\right)}{\partial t}=\frac{f\left(t_{n+1},x_j\right)-f\left(t_n,x_j\right)}{\tau}+r_2,$$

$$r_2-\frac{\tau}{2}\frac{\partial^2 f\left(\theta_2,x_j\right)}{\partial t^2},\ \theta_2\in\left[t_n,t_{n+1}\right];$$

3. the second order approximation in time (for three level schemes)

$$\frac{\partial f\left(t_n,x_j\right)}{\partial t}=\frac{f\left(t_{n+1},x_j\right)-f\left(t_{n-1},x_j\right)}{2\tau}+r_3,$$

$$r_3-\frac{\tau^2}{6}\frac{\partial^3 f\left(\theta_3,x_j\right)}{\partial t^3},\ \theta_3\in\left[t_{n-1},t_{n+1}\right].$$

The derivatives of boundary conditions (1) approximate using the following expressions:

1. the first order approximation in space

$$\frac{\partial f\left(t_n,L\right)}{\partial x}=\frac{f\left(t_n,x_N\right)-f\left(t_n,x_{N-1}\right)}{h}+r_4,$$

$$r_4=-\frac{h}{2}\frac{\partial^2 f\left(t_n,\varsigma_2\right)}{\partial x^2},\ \varsigma_2\in\left[x_{N-1},x_N\right],\ x_N=L,\ x_{N-1}=L-h;$$

2. the second order approximation in space

$$\frac{\partial f\left(t_n,L\right)}{\partial x}=\frac{3f\left(t_n,x_N\right)-4f\left(t_n,x_{N-1}\right)+f\left(t_n,x_{N-2}\right)}{2h}+r_5,$$

$$r_5=-\frac{h^2}{3}\frac{\partial^3 f\left(t_n,\varsigma_3\right)}{\partial x^3},\ \varsigma_3\in\left[x_{N-2},x_N\right],\ x_{N-2}=L-2h.$$

Hence, ignoring the remained terms r_4 and r_5, we have two approximation forms for boundary conditions in fixed time moment t_n:

$$y_N^n=Cy_{N-1}^n \tag{2}$$

$$y_N^n=\tilde{C}_1 y_{N-1}^n+\tilde{C}_2 y_{N-2}^n, \tag{3}$$

where $C=\left(1+i\gamma h\right)^{-1}$, $\tilde{C}_1=2\left(1.5+i\gamma h\right)^{-1}$, $\tilde{C}_2=-0.5\left(1.5+i\gamma h\right)^{-1}$.

Then ignoring the remained term r_1 we obtain 3-point difference expression, that approximate the second order derivative in the point $\left(t_n,x_j\right)$

$$\Lambda y_j^n\equiv\frac{\left(y_{j+1}^n-2y_j^n+y_{j-1}^n\right)}{h^2}. \tag{4}$$

To study the stability of the discrete problems (finite-difference schemes) we rewrite the homogeneous difference equations with respect to the difference $z_j^n=y_j^n-f\left(x_j,t_n\right)$ in the matrix operator form

$$z^{n+1}=Gz^n, \tag{5}$$

where G is the transition operator with the eigenvalues ρ_k and eigenfunctions $g^{(k)}\left(x_j\right)$, $k=\overline{1,N-1}$ of the difference operator Λ. The solution of difference equations corresponding to values z_j^n can be found in the form

$$z_j^n=\left(\rho_k\right)^n g^{(k)}\left(x_j\right),\ k=\overline{1,N-1}. \tag{6}$$

Using (2) the corresponding eigenvalues λ_k of difference operator Λ can be obtained from the three-point finite-difference scheme

$$\begin{cases}\Lambda g_j^{(k)}+\lambda_k g_j^{(k)}=0, & j=\overline{1,N-1}\\ g_0^{(k)}=0, & g_N^{(k)}=Cg_{N-1}^{(k)}\end{cases} \tag{7}$$

For the boundary condition (3) in (7) we introduce the expression $g_N^{(k)}=C_1 g_{N-1}^{(k)}+C_2 g_{N-1}^{(k)}$.

Now the solution of (7) can be written as

$$g_j^{(k)}=E_k\sin\left(q_k x_j\right), \tag{8}$$

where E_k are arbitrary constants, $1-\dfrac{\lambda_k h^2}{2}=\cos\left(q_k h\right)$. From boundary conditions (2) follows that the complex parameter q_k is determined by the complex transcendental equation

$$\sin\left(q_k L\right)=C\sin\left(q_k\left(L-h\right)\right), \tag{9}$$

where the parameter $q_k=a_k+ib_k$, $k=\overline{1,N-1}$ has complex values. Therefore

$$\lambda_k=\frac{2}{h^2}\left(1-\cos\left(q_k h\right)\right)=A_k+iB_k, \tag{10}$$

where $A_k=2h^{-2}\left(1-\cos\left(a_k h\right)\cosh\left(b_k h\right)\right)$, $B_k=2h^{-2}\sin\left(a_k h\right)\sinh\left(b_k h\right)$, $k=\overline{1,N-1}$.

If $\gamma=\infty$ (boundary conditions of first kind $g_N^{(k)}=0$), then $C=0$ and $q_k=\dfrac{k\pi}{L}$ (real numbers),

$$\lambda_k=\frac{4}{h^2}\sin^2\frac{k\pi h}{2L} \tag{11}$$

and $g_j^{(k)}=\sqrt{\dfrac{2}{L}}\sin\dfrac{k\pi x_j}{2L}$, $k=\overline{1,N-1}$ [2].

Calculations with the help of MAPLE for $L=15$, $\gamma=2$, $h=0{,}1$ are shown in Table (1), where a_k and b_k are solutions of (9) and $\dfrac{\pi\left(k-1\right)}{L}<a_k<\dfrac{\pi k}{L}$, $0<b_k<1$.

Table 1. The discrete values of q_k , $\tilde{\alpha}_k$

k	a_k	b_k	A_k	B_k
1	0.2092	0.0070	0.0437	0.0029
2	0.4183	0.0141	0.1748	0.0118
3	0.6273	0.0216	0.3929	0.0271
4	0.8360	0.0296	0.6976	0.0494
5	1.0442	0.0383	1.0879	0.0798
6	1.2515	0.0481	1.5618	0.1201
7	1.4569	0.0595	2.1154	0.1728
8	1.6587	0.0725	2.7396	0.2393

In Fig. (1) we show the first 50 eigenvalues q_k for $h = 0,02$.

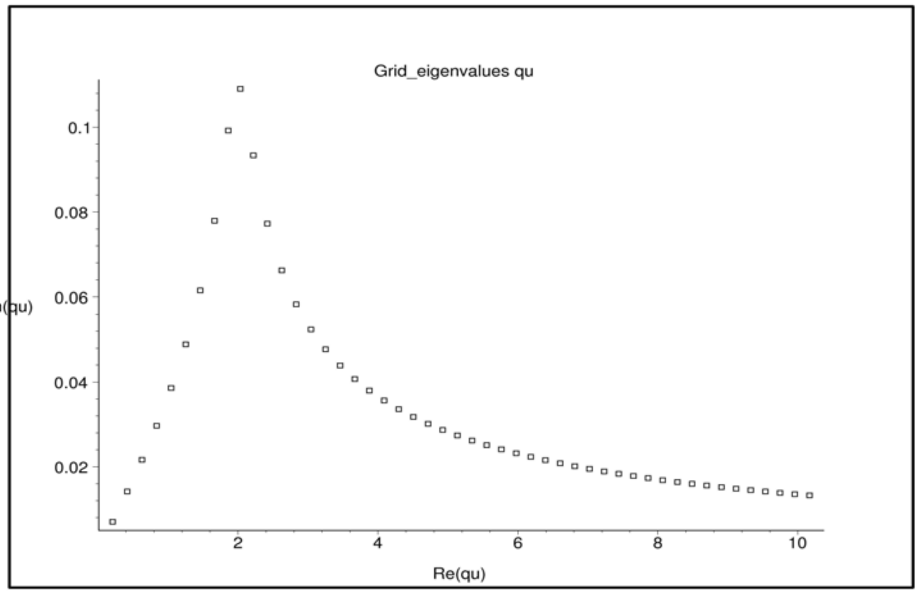


***Figure 1**. Eigenvalues of the discrete problem* q_k , $N = 750$

It is seen that, if (a_k, b_k) is a solution of these systems, then also $(-a_k, -b_k)$ is a solution. The values of the coefficients A_k, B_k do not change and it is sufficient to consider only $a_k > 0$. If simultaneously $b_k > 0$, then also $B_k > 0$. Calculations with the help of "MAPLE" show that positive variables a_k correspond to positive variables b_k i.e. $B_k > 0$. If the parameter $\gamma < 0$, then it can be easily seen that positive a_k correspond to negative b_k .

Now we study in detail the stability and precision of different schemes for $|\rho_k| \le 1$, $k = \overline{1, N-1}$.

2. Two Level Finite-Difference Scheme with Parameter σ

Substituting differences (2) – (4) into the problem (1), we obtain a two-layer finite-difference scheme with weight $\sigma \in [0,1]$

$$\begin{cases} \dfrac{y_j^{n+1} - y_j^n}{\tau} = -i\left(\sigma\Lambda^* y_j^{n+1} + (1-\sigma)\Lambda^* y_j^n + F_j^\sigma\right), j = \overline{1, N-1} \\ y_0^{n+1} = 0,\ y_N^{n+1} = Cy_{N-1}^{n+1} \end{cases}, \tag{12}$$

where $F_j^\sigma = F\left(t_n + \sigma\tau, x_j\right)$, $\Lambda^* = \Lambda + \delta$. For the truncation error

$$\psi_j^n \equiv \frac{f_j^{n+1} - f_j^n}{\tau} + i\left(\sigma\Lambda^* f_j^{n+1} + (1-\sigma)\Lambda^* f_j^n + F_j^\sigma\right)$$

we get $\psi_j^n = O\left(\tau^\alpha + h^2\right)$, where $\alpha = 1$ if $\sigma \neq 0.5$ and $\alpha = 2$ if $\sigma = 0.5$.

Boundary conditions are approximated only to the first order. To obtain second order, one has use expression (3)

For error z_j^n (6) from (12) follows the expression

$$\rho_k = \frac{1 + i\tau(1-\sigma)\lambda_k^*}{1 - i\tau\sigma\lambda_k^*}, \tag{13}$$

where $\lambda_k^* = \lambda_k - \delta$.

If λ_k are real numbers, e.g., in the case of the first kind boundary conditions $\left(\gamma = \infty,\ z_N^{n+1} = 0\right)$ $q_k = \dfrac{k\pi}{L}$, then from the stability condition [2]

$$|\rho_k|^2 = \left(1 + \tau^2(1-\sigma)^2\left(\lambda_k^*\right)^2\right)\left(1 + \tau^2\sigma^2\left(\lambda_k^*\right)^2\right)^{-1} \leq 1,$$

it follows that

$$\sigma \geq \frac{1}{2} \tag{14}$$

independent of τ. Similar problem for Schrödinger type differential equation was investigated in [3].

Taking the boundary condition of the third kind in the form $z_N^{n+1} = Cz_{N-1}^{n+1}$ and determining the complex parameter $q_k = a_k + ib_k$, we find the complex values $\lambda_k^* = A_k^* + iB_k$, where $A_k^* = A_k - \delta$, $B_k > 0$.

Then from

$$|\rho_k|^2 = \frac{\left(1 - \tau(1-\sigma)B_k\right)^2 + \left(A_k^*\right)^2\tau^2(1-\sigma)^2}{\left(1 + \tau\sigma B_k\right)^2 + \left(A_k^*\right)^2\tau^2\sigma^2} \leq 1$$

follows that

$$-2B_k + \tau(1-2\sigma)\left(\left(A_k^*\right)^2 + B_k^2\right) \leq 0,$$

the inequality (14) holds and the difference scheme (12) is stable.

3. Richardson's Three Level Finite-Difference Scheme

The difference equations of three level Richardson's scheme are in the following form

$$\frac{y_j^{n+1} - y_j^{n-1}}{2\tau} = -i\left(\Lambda * y_j^n + F_j^n\right). \tag{15}$$

Corresponding truncation error is $\psi_j^n = O\left(\tau^2 + h^2\right)$. For real heat transfer equations ($-i$ is replaced by 1) this scheme is absolutely unstable [2]. The transfer modulus ρ_k can be obtained from the following quadratic equation

$$\rho_k^2 - 2i\tau\lambda_k^*\rho_k - 1 = 0\,. \tag{16}$$

For the stability of the difference scheme it is necessary that the all roots of (16) satisfy the inequality $\left|\rho_k\right| \le 1$ (the identity is valid only for the simple roots of (16)). For the complex quadratic equation

$$a\rho^2 + b\rho + c = 0 \tag{17}$$

this inequalities holds, if [4]

$$\left|\bar{b}a - \bar{c}b\right| \le |a|^2 - |c|^2\,,\ |b| < 2|a|\,, \tag{18}$$

where $\bar{b}$, $\bar{c}$ are the complex conjugate values of b, c.

If $\lambda_k^* \in R$, then from (18) follows that

$$\tau\left|\lambda_k^*\right| < 1\,,\quad \tau < \frac{h^2}{4 - h^2\delta}\,. \tag{19}$$

If $\lambda_k^* \in C$, then from (18) follows that

$$B_k = 0\,,\quad \tau < \frac{1}{\max\left|A_k^*\right|}\,, \tag{20}$$

which is equivalent to (19).

Therefore, if $B_k > 0$, then Richardson's scheme for equation (1) is absolutely unstable and conditionally stable by (19) only in the limit case $\gamma = \infty$ $\left(B_k = 0\right)$ for the boundary conditions of the first kind.

4. Simpson's Three Level Finite-Difference Scheme

The difference equations of three level Simpson's scheme is in the following form

$$\frac{y_j^{n+1} - y_j^{n-1}}{2\tau} = -i\left(\left(\Lambda * y_j^{n+1} + 4\Lambda * y_j^n + \Lambda * y_j^{n-1}\right)/6 + F_j^n\right). \tag{21}$$

The truncation error is $\psi_j^n = O\left(\tau^4 + h^2\right)$. For real heat transfer equations this scheme is absolutely unstable [2]. The transfer modulus ρ_k can be obtained from the following quadratic equation

$$\left(1-\frac{i\tau\lambda_k^*}{3}\right)\rho_k^2-\frac{4i\tau\lambda_k^*}{3}\rho_k-\left(1+\frac{i\tau\lambda_k^*}{3}\right)=0. \qquad (22)$$

If $\lambda_k^* \in R$, then from (18) follows that

$$\left(\tau\lambda_k^*\right)^2<3,\ \ \tau<\sqrt{3}\,\frac{h^2}{4-h^2\delta}. \qquad (23)$$

If $\lambda_k^* \in C$, then from (18) follows that

$$B_k=0,\ \ \tau<\frac{\sqrt{3}}{\max\left|A_k^*\right|}, \qquad (24)$$

which is equivalent to (23).

Therefore, if $B_k>0$, then Simpson's scheme also for equation (1) is absolutely unstable and conditionally stable by (23)) only if $\gamma=\infty$ $\left(B_k=0\right)$ for the boundary conditions of the first kind.

5. Richtmayer-Morton's Three Level Finite-Difference Scheme

The difference equations of three level Richtmayer-Morton's scheme is in the following form

$$1.5\frac{y_j^{n+1}-y_j^n}{\tau}-0.5\frac{y_j^n-y_j^{n-1}}{\tau}=-i\left(\Lambda * y_j^{n+1}+F_j^{n+1}\right). \qquad (25)$$

The truncation error is $\psi_j^{n+1}=\tau^2+h^2$. For real heat transfer equations this scheme is absolutely stable [2]. The transfer modulus ρ_k can be obtained from the following quadratic equation

$$\left(1.5-i\tau\lambda_k^*\right)\rho_k^2+2\rho_k+0.5=0. \qquad (26)$$

If $\lambda_k^* \in R$, then from (18) follows that this difference scheme is absolutely stable $\left(\tau<\infty\right)$.

If $\lambda_k^* \in C$, then from (18) follows also that this difference scheme is absolutely stable, if $B_k>0$.

6. General Three Level Finite-Difference Schemes

For fixed grid point x_j in the space we can considered 1-D three level approximation in the time of the function $u=u\left(t\right)$ in following form:

$$\frac{Au\left(t_{n+1}\right)+Bu\left(t_n\right)+Cu\left(t_{n-1}\right)}{\tau}=A_1u'\left(t_{n+1}\right)+B_1u'\left(t_n\right)+C_1u'\left(t_{n-1}\right)+\tau_n, \qquad (27)$$

where $\tau_n=\tau^{\alpha-1}\dfrac{u^{(\alpha)}\left(\varsigma_n\right)}{\alpha!}C_0$ is the error term, $\varsigma\in\left[t_{n-1},t_{n+1}\right]$, A,B,C,A_1,B_1,C_1,C_0, $\alpha\ge 3$ are the unknown coefficients, $u'=\dfrac{d}{dt}$, $u'\left(t_n\right)=-i\left(\Lambda * y_j^n+F_j^n\right)$.

The unknown coefficients can be obtained from the condition, that the expression (27) is valid for polynomials with highest degree. To this end we consider normalized function $\tilde{u}(\tilde{t})$, where $\tilde{t} = \frac{t - t_n}{\tau}$. Then from (27) follows the expression

$$A\tilde{u}(1) + B\tilde{u}(0) + C\tilde{u}(-1) = A_1\tilde{u}'(1) + B_1\tilde{u}'(0) + C_1\tilde{u}'(-1) + C_0\frac{\tilde{u}^{(\alpha)}(\tilde{\varsigma})}{\alpha!} \qquad (28)$$

where $\tilde{u}' = \frac{d\tilde{u}}{d\tilde{t}}$, $\tilde{\varsigma} \in [-1,1]$. Using the elementary functions $\tilde{u}(\ddot{t}) = \tilde{t}^k$, $k = \overline{0,\alpha}$, we obtain the unknown coefficients from the following system of linear algebraic equations:

$$A + B0^k + C(-1)^k = k\left(A_1 + B_1 0^{k-1} + C_1(-1)^{k-1}\right), \qquad (29)$$

where $0^0 = 1$, $k = \overline{0,\alpha - 1}$.

If $k = \alpha$, then we obtain the equation for coefficient C_0

$$C_0 = A + C(-1)^a - \alpha\left(A_1 + C_1(-1)^{\alpha-1}\right). \qquad (30)$$

For $\alpha = 5$ from (29), (30) can be obtained the Simpson finite difference equations (21) with

$$A = \frac{1}{2},\ B = 0,\ C = -\frac{1}{2},\ A_1 = \frac{1}{6},\ B_1 = \frac{4}{6},\ C_1 = \frac{1}{6},\ C_0 = -\frac{4}{6}.$$

For $\alpha = 3$ can be obtained the Richtmayer-Morton finite difference equations (25) with

$$A = \frac{3}{2},\ B = -2,\ C = \frac{1}{2},\ A_1 = 1,\ B_1 = C_1 = 0,\ C_0 = -2.$$

For $\alpha = 4$ can be obtained the finite difference equations, depending of two parameters A, C, in the form

$$\begin{cases} B = -(A + C),\ A_1 = \frac{1}{12}(5A + C),\ B_1 = \frac{2}{3}(A - C), \\ C_1 = -\frac{1}{12}(A + 5C),\ C_0 = -(A + C) \end{cases}. \qquad (31)$$

If $C = -A$, then we obtain (21).

The stability investigations of general difference equations (27) $(r_n = 0)$ lead to the following characteristic equations

$$(A - \mu_k A_1)\rho_k^2 + (B - \mu_k B_1)\rho_k + (C - \mu_k C_1) = 0,$$

where $\mu_k = i\tau\lambda_k^*$.

From (18) follows inequalities:

1) if $\lambda_k^* \in R$, then

$$\begin{cases} \left|(B + \mu_k B_1)(A - \mu_k A_1) - (C + \mu_k C_1)(B - \mu_k B_1)\right| \le \\ \qquad A^2 - (\mu_k A_1)^2 - C^2 + (\mu_k C_1)^2, \\ B^2 - (\mu_k B_1)^2 < 4\left(A^2 - (\mu_k A_1)^2\right); \end{cases} \qquad (32)$$

2) if $\lambda_k^* \in C$, then $\left(\mu_k^* = i\tau A_k^*\right)$

$$\begin{cases} \left|\left(B+\mu_k^* B_1+\tau B_1 B_k\right)\left(A-\mu_k^* A_1+\tau A_1 B_k\right)-\left(C+\mu_k^* C_1+\tau C_1 B_k\right)\left(B-\mu_k^* B_1+\tau B_1 B_k\right)\right| \leq \\ \qquad \left(A+\tau A_1 B_k\right)^2-\left(\mu_k^* A_1\right)^2-\left(C+\tau C_1 B_k\right)^2+\left(\mu_k^* C_1\right)^2, \\ \left(B+\tau A_1 B_k\right)^2-\left(\mu_k^* B_1\right)^2<4\left(\left(A+\tau A_1 B_k\right)^2-\left(\mu_k^* A_1\right)^2\right). \end{cases} \tag{33}$$

From (31) by $A = C = 1$ it follows, that

$$A = 1,\ C_1 = -\frac{1}{2},\ A_1 = \frac{1}{2},\ C = 1,\ B_1 = 0,\ B = -2,\ C_0 = -2,$$

and inequalities (32), (33) holds.

Therefore we obtain the following three level finite difference equations

$$\frac{y_j^{n+1}-y_j^n}{\tau}-\frac{y_j^n-y_j^{n-1}}{\tau}=-i\left(0.5\left(\Lambda^* y_j^{n+1}-\Lambda^* y_j^{n-1}+F_j^{n+1}-F_j^{n-1}\right)\right) \tag{34}$$

with the truncation error

$$\psi_j^n = -\frac{\tau^3}{12}\frac{\partial^4 f\left(\theta_n, x_j\right)}{\partial t^4}+\frac{ih^2}{12}\frac{\partial^4 f\left(t_n, \varsigma_j\right)}{\partial x^4},$$

where $\varsigma_j \in \left[x_{j-1}, x_{j+1}\right]$, $\theta_n \in \left[t_n, t_{t+1}\right]$.

References

[1] M.I. Airila, O. Dumbrajs, A. Reinfelds, and U. Strautiņš. Nonstationary oscillations in gyrotrons. Phys. Plasmas, 8 (10), 4808–4812, 2001.

[2] A.A. Samarski. Theory of difference schemes. Moscow, Nauka, 1977 (in Russian).

[3] M. Radžiunas, F. Ivanauskas. The stability conditions of finite difference schemes for Schrödinger, Kuramoto-Tsuzuki and heat equations. Mathematical modeling and analysis, 3, 177–194, 1998.

[4] A. Štikonas. The root condition for the third order polynomials. Mathematical modeling and analysis, 4, 163–173, 1999.

HIGH-SELECTIVE DIGITAL FILTERS WITH OPTIMIZED LINEARITY OF PHASE CHARACTERISTIC

Takhir Mamirov

Transport and Telecommunication Institute
Lomonosova Str1, Riga, LV 1019, Latvia.
Tel.: (+371)-7100650. Fax: (+371)-7100660. E-mail: mamirov_t@yahoo.com

The method of increase of phase characteristic linearity of high-selective filters is offered. The modelling results are introduced.

Key words: transfer function with complex coefficients, linear phase characteristic

One of the important additional requirements at designing supernarrow-band filters frequently is a high linearity of phase-frequency characteristic. Especially sharply this problem arises at designing high-selective filters. Unfortunately, the regular approach at definition of the best realization meeting these requirements does not exist yet; therefore the problem is reduced to optimization selection its parameters at imitating or natural modelling [1].

In the work the method of filters' phase characteristics linearization by a choice of parameters of transfer function with complex coefficients is offered.

As base, we shall consider a transfer function of the prototype of the elliptic filter possessing the maximal selectivity with other equal conditions:

$$H(p) = H_0 \prod_{i=1}^{N/2} \frac{p^2 + A_i}{p^2 + B_i p + C_i} \qquad \text{- odd order} \tag{1}$$

$$H(p) = \frac{H_0}{p+G} \prod_{i=1}^{N-1/2} \frac{p^2 + A_i}{p^2 + B_i p + C_i} \qquad \text{- even order} \tag{2}$$

Let $p = k\dfrac{1-z^{-1}}{1+z^{-1}}$, so transfer function of separate biquad in Z-plane:

$$H_i(z^{-1}) = h_{0i} \frac{(k^2 + A_i) + 2(A_i - k^2)z^{-1} + (k^2 + A_i)z^{-2}}{1 + 2h_{0i}(C_i - k^2)z^{-1} + h_{0i}(k^2 - B_i k + C_i)z^{-2}}, \tag{3}$$

where $h_{0i} = \dfrac{1}{k^2 + B_i k + C_i}$, $k = ctg\dfrac{\pi}{2}\tilde{\Omega}_1$.

The coefficients of transfer function are real. Let's consider, for example, frequency characteristics of the elliptic filter at the following requirements:

1. cut-off frequency, $\tilde{\Omega}_1 = 0.2$;
2. passband attenuation, a_{max} = 0.1 dB;
3. stopband attenuation, a_{min} = 60 dB.

These requirements are carried out by the filter of 6-th order.

The result is shown on picture 1. Thus the maximal deviation of a phase in a passband reaches size about 40-50 grad.

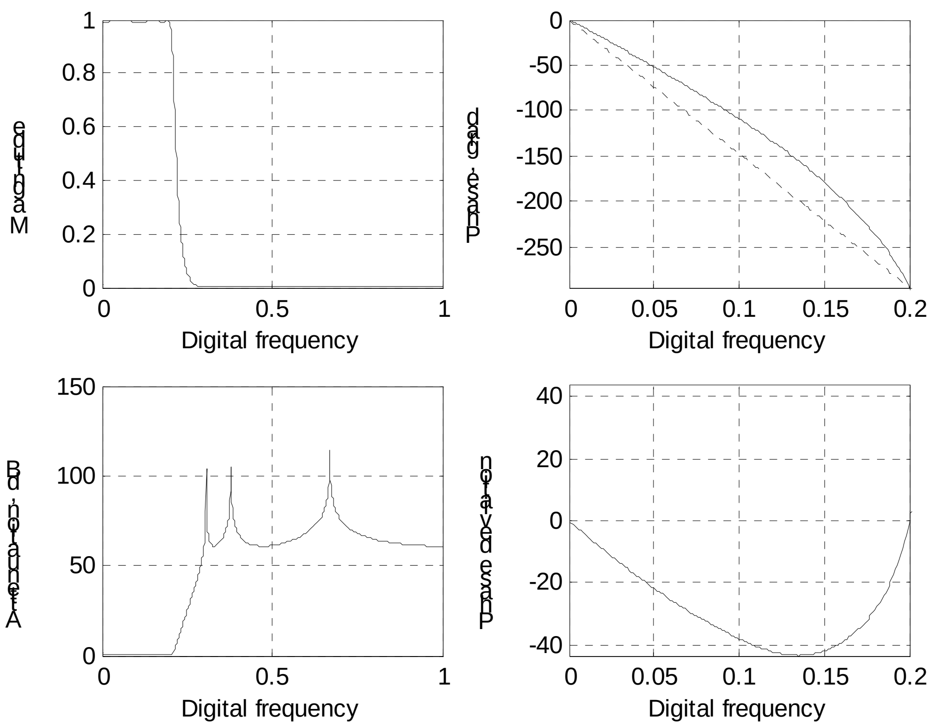


Figure 1. *Frequency characteristics of elliptic filter*

The absence of analytical methods for designing supernarrow-band digital filters with a linear phase forces to address to numerical methods of the decision.

As the possible approach to phase linearization, instead of traditional transfer function (3) it is offered to use the modified function:

$$H_i(z^{-1}) = \dot{h}_{0i}\,\frac{(k_1^{\,2}+A_i+\alpha_1+j\beta_1)+2(A_i-k_1^{\,2}+\alpha_2+j\beta_2)z^{-1}+(k_1^{\,2}+A_i+\alpha_3+j\beta_3)z^{-2}}{1+2\dot{h}_{0i}(C-k_1^{\,2}+\alpha_5+j\beta_5)z^{-1}+\dot{h}_{0i}(k_1^{\,2}-Bk_1+C+\alpha_6+j\beta_6)z^{-2}} \qquad (4)$$

where $\dot{h}_{0i} = h_{0i} + (\alpha_4 + j\beta_4)$, $k_1 = k + \alpha_7$.

The parameter k_1 defines the cut-off frequency of modified amplitude-frequency characteristic. It’s possible to show, that at minimization of phase deviation in a passband the cut-off frequency increases.

All multipliers of modified biquad generally are complex. It complicates the realization; however there are approaches to the decision of this problem [2].

The analysis of the synthesized filter we shall lead as follows:

1. We shall weaken requirements to cut-off frequency, control frequency and passband attenuation of the modelled filter, keeping the order.
2. We shall set the requirement to the maximal deviation of a phase in a passband (at least, much less to traditional deviation).
3. Selecting coefficients in (4), we optimize a phase a numerical method (for example, Nelder-Mead), keeping the control above amplitude-frequency characteristic.
4. If the received maximal deviation of a phase in a passband does not meet the requirement, should set other initial approach and repeat item 3-4.

As a result of iterative optimization the following result is received: deviation of a phase in a passband - about 10^{-1}-10^{-2} grad. Thus the resulting mistake depends on rigidity of requirements to control frequency. Varying control frequency in allowable limits, it is possible to reduce considerably phase deviation in a passband.

On fig. 2 frequency characteristics after two iterations of global search are shown.

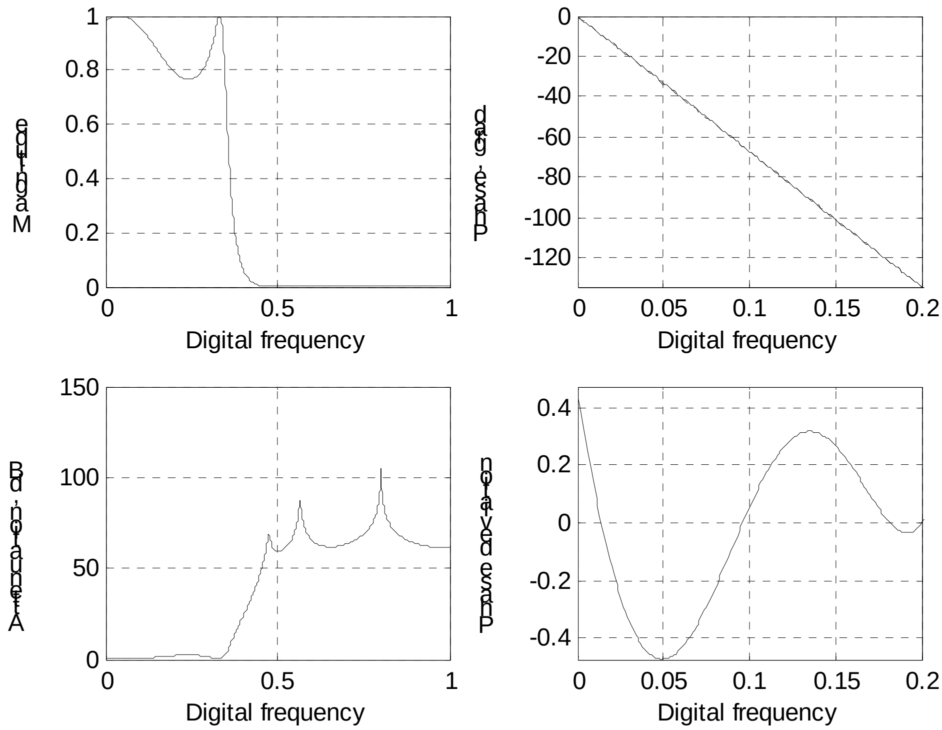


Figure 2. *Frequency characteristics of synthesized filter*

At optimization of a phase deviation in a working band it is necessary to supervise required attenuation on control frequency and higher. Thus cut-off frequency can be changed in limits in which safety of the order is guaranteed.

Conclusion

One of possible approaches to a problem of increase of phase characteristic linearity of high-selective filters at insignificant easing requirements to other parameters is offered.

References

1. Оппенгейм А.В., Шафер Р.В. «Цифровая обработка сигналов». Связь, Москва, 1979.
2. Мамиров Т., «Цифровые фазовые звенья с комплексными умножителями», Transport and Telecommunication, ISSN 1407-6160, стр. 34-38.
3. Еремеев В.П., Гуменюк А., Мамиров Т. Новые структуры малочувствительных цифровых фильтров, Computer Modelling & New Technologies, ISSN 1407-5806, стр. 132-137.
4. Антонью А. Цифровые фильтры: анализ и проектирование, Москва: Радио и связь, 1983, 320 с.
5. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов, Москва: Мир, 1978, 848 с.

Poster session

THE DISTRIBUTION OF CONDITIONAL ON REALISED VOLATILITY FINANCIAL RETURNS

***Max V. Moldovan*[1], *Andrew C. Worthington*[2], *Nicholas A. Nechval*[3], *Helen Higgs*[4]**

[1] *ESH Group Ltd.,*
6-450 Ezermalas Street, Riga, Latvia, LV-1016
Ph: (+371) 9 82 69 66. Fax: (+371) 7 518 519. E-mail: max@esh.lv

[2] *School of Economics and Finance, Queensland University of Technology,*
2 George Street, Brisbane, Australia, Qld 4001
Ph: (+61) 07 3864 2658. Fax: (+61) 07 3864 1500. E-mail: a.worthington@qut.edu.au

[3] *Mathematical Statistics Department, University of Latvia,*
19 Raina Blvd, Riga, Latvia, LV-1050
Ph: (+371) 7034701. Fax: (+371) 7034702. E-mail: nechval@junik.lv

[4] *School of Economics and Finance, Queensland University of Technology,*
2 George Street, Brisbane, Australia, Qld 4001
Ph: (+61) 07 3864 4252. Fax: (+61) 07 3864 1500. E-mail: h.higgs@qut.edu.au

Abstract

The following study explores the stochastic properties of financial time series. Special attention is devoted to the unusual two-peaked shape of the empirically observed distribution of the conditional on realised volatility financial returns. The performed analysis leads to the conclusion that the conditional on realised volatility returns are distributed with the specific previously undocumented distribution. The probability density that represents this distribution is derived, characterised and applied for improving the forecasts of Value at Risk.

Key words: Realised volatility, High-frequency data, Value at Risk, Two-component effect, Monte Carlo simulation

1. Introduction

A distribution of financial returns plays a central role in financial econometrics and its applications. It is important to distinguish between unconditional and conditional return distributions. The unconditional distribution reflects stochastic characteristics of returns observed in the markets. Since unconditional returns can be directly observed, they have been well studied. The situation is more complicated in the case of conditional returns. The term conditional generally refers to conditioning returns on underlying volatility. Since financial volatility is not directly observable, volatility estimates are used for conditioning. Realised volatility is an unbiased, consistent and highly efficient estimator of actual financial volatility. In this study realised volatility estimates are applied to conditioning financial return series. Despite the theoretical expectations, it has been found that conditional on realised volatility returns are not standard normally distributed but follow the specific previously undocumented two-peaked distribution. The probability density that represents this distribution is derived and applied to forecasting the future return probability density quintiles, or so-called Value at Risk.

The paper is structured as follows. Section 2 presents the traditional return decomposition and introduces the realised volatility estimator. In Section 3 the unusual previously undocumented distribution of conditional on realised volatility return series is documented and probability density that represents this distribution suggested. Section 4 demonstrates how the new probability density can be used in the financial model with application to Value at Risk forecasting. Section 5 summarises the results of the study.

2. Return Decomposition and Realised Volatility

We assume that daily return generating process can be expressed as follows.

$$r_t = \sigma_t z_t \tag{1}$$

where $r_t = \ln(P_t) - \ln(P_{t-1})$, P_t is a last price observed over day t, and $z_t \sim N(0,1)$ and σ_t are a stochastic component and standard deviation of return over day t respectively.

This view on the return generating process relies on the Efficient Market Hypothesis exposed by Fama [1] and widely accepted among academics and practitioners (see Campbell et al. [2]). Following this, the conditional on volatility returns $z_t = r_t / \sigma_t$ are expected to be standard normally distributed: $z_t \sim N(0,1)$. For such conditioning it is required to estimate volatility over each day t. The recent developments in financial econometrics allow estimating the variability of financial returns using more frequently sampled intra-period data in place of traditional multi-period datasets. Andersen and Bollerslev [3] and Andersen et al. [4] related intra-period accumulation of squared returns to the mathematical concept of quadratic variation and proposed the realised volatility estimator. Realised volatility is an unbiased, consistent and efficient volatility estimator that does not requires information outside an estimation interval. Realised volatility in day t based on k intraday return observations is expressed as follows.

$$v_{t,k} = \sum_{n=1}^{k} x_{t,n}^2 \tag{2}$$

where $v_{t,k}$ is a realised volatility estimate computed by summing squared intraday returns $x_{t,n}^2$ that sampled k times over equally spaced time intervals within day t.

The especially useful property of the realised volatility estimator is consistency: $v_{t,k} \to \sigma_t^2$ as $k \to \infty$. However, the value of k should not be the highest available in order to avoid the domination of the microstructure bias, though high enough to minimise the measurement error. For selection of the optimal sampling frequency Andersen et al. [5] suggested the volatility signature plot approach, which is illustrated in Appendix 2 on the example of data described in Appendix 1. It has been found that the optimal sampling frequency for series considered in the study is $k = 8$. Hence, the realised volatility series $v_{t,8}$ has been computed for return conditioning.

3. The Two-Component Effect and J Distribution

The conditional return series has been computed as follows.

$$z_t = r_t / \sigma_{t,8} \tag{3}$$

where the log return r_t over day t is standardised by the realised standard deviation $\sigma_{t,8} = \sqrt{v_{t,8}}$.

According the theoretical view on the return generating process given by (1), conditional returns should be approximately standard normally distributed $z_t \sim N(0,1)$. Indeed, Andersen et al. [6] performed the standardisation (3) and found that distributions of the resulted series are "remarkably close to a standard normal". Authors interpreted this result as consistent with Mixture-of-Distributions hypothesis suggested by Clark [7], completely ignoring platykurtosis, which clearly violets normality. In this study it has been found that standardised return series computed by (3) has a highly unusual two-peaked distribution.

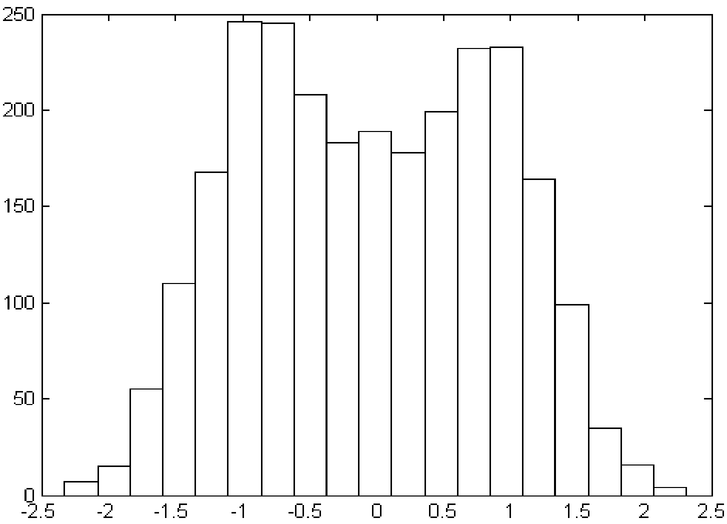


Figure 1. *Standardised by* $\sigma_{t,8}$ *AUD/USD futures returns*

This empirical effect has been entitled the two-component effect. After the careful examination it has been found that the two-component effect arrises due to the special case of heteroscedasticity in intraday returns. Specifically, the simple standard deviation of returns over the first intraday period is more than three times larger than standard deviations of the rest of intraday returns. The table below demonstrates this.

Table 1. Simple standard deviations of intraday returns

n =	1	2	3	4	5	6	7	8
s.d. of returns	0,0051	0,0015	0,0015	0,0015	0,0014	0,0012	0,0011	0,0012

In mathematical terms this situation can be expressed as follows.

$$z_t = \frac{(x_{t,1} + j_t) + x_{t,2} + \ldots + x_{t,k}}{\sqrt{(x_{t,1} + j_t)^2 + x_{t,2}^2 + \ldots + x_{t,k}^2}} \qquad (4)$$

where $x_{t,n} \sim N(0, \sigma_t^2 / k)$, $j_t \sim N(0, \sigma_{j,t}^2)$, $\sigma_{j,t}^2 > 0$ and k is the number of intraday returns within day t.

It has been found that with $k = 8$, $\sigma_t^2 = 1$ and $\sigma_{j,t}^2 = 3$ the random variable given by (4) has a distribution that is visually identical to the distribution of empirically observed standardised returns. Recognising that if $k \to \infty$ then $x_{t,n} \to 0$ and relating the homogenous variance σ_t^2 to the variance of a jump as $\sigma_{j,t}^2 = \alpha \cdot \sigma_t^2$, the function (4) can be presented as follows.

$$z_t = \frac{y_t + j_t}{\sqrt{\sigma_t^2 + j_t^2}} \qquad (5)$$

where $y_t \sim N(0, \sigma_t^2)$, $j_t \sim N(0, \alpha \cdot \sigma_t^2)$, $\alpha \geq 0$ reflects the portion of σ_t^2 in $\sigma_{j,t}^2$ and $Cov(y_t, j_t) = 0$.

Since the denominator of (5) is an unbiased and consistent estimator of the standard deviation of the numerator and the variance of y_t is directly related to the variance of j_t through the parameter α, it can be demonstrated that the distribution of z_t does not depend on the variance of y_t. Therefore (5) can be generalised as follows.

$$z_t = \frac{y_t + j_t}{\sqrt{1 + j_t^2}} \tag{6}$$

where $y_t \sim N(0,1)$, $j_t \sim N(\gamma, \alpha)$, $\alpha \geq 0$ and $Cov(y_t, j_t) = 0$.

The new function is a combination of two i.n.d. random variables one of which is standard normally distributed and the other is normally distributed with arbitrary parameters. Since the distribution of z_t given by (6) depends only on two parameters γ and α, this probability density will be referred to as $J(\gamma, \alpha)$. Figure 2 demonstrates some shapes that $J(\gamma, \alpha)$ takes with variation of parameters γ and α.

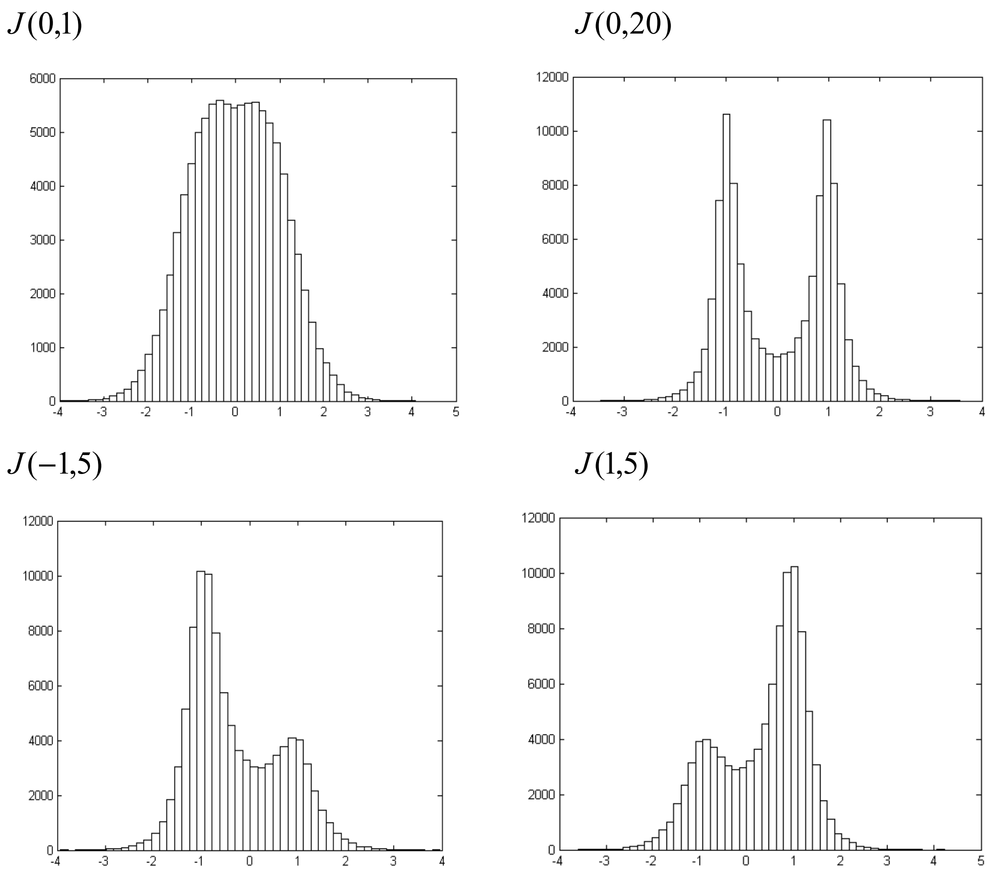


***Figure 2**. Histograms of $J(\gamma, \alpha)$ distributed random variables*

It should be noted that if $\gamma = 0$ then z_t given by (6) always have expected unconditional mean and variance parameters 0 and 1 respectively. More detailed characterisation of $J(\gamma, \alpha)$ together with visual examples and technical comments is given in Moldovan [8]. It also should be noted that since at the present moment $J(\gamma, \alpha)$ has not a probability density function, this distribution is simulated but not theoretical. However, since the random number generator is already (see [9]), J can already be applied practically. In the next section it is demonstrated how $J(\gamma, \alpha)$ can be used in the financial model for forecasting purposes.

4. Application of $J(\gamma,\alpha)$: Forecasting Probability Quintiles of Future Price Distribution

In this section is shown how $J(\gamma,\alpha)$ can be applied for improving the forecasting performance of the financial model. Specifically, it is demonstrated how taking into account the previously unnoticed distributional specifics of conditional on realised volatility returns can substantially improve forecasts of Value at Risk (VaR), which after the acceptance by the Basle Committee (Basel Committee on Banking supervision [10]) has become the most widely used risk measurement tool in the banking sector.

VaR of a financial asset or portfolio of assets is defined as the maximum loss that can be expected with a certain level of confidence over a particular interval of time. Thus, VaR is equivalent to a quintile of a future gain/losses distribution on a considered asset or portfolio of assets that corresponds to a desired confidence level. One of the most efficient ways of VaR forecasting is based on the Monte Carlo simulation technique (see Jorion [11]). This technique can be implemented by forecasting the future gain/losses distribution and selecting a required quintile. The following financial model is used for simulation.

$$P_t = P_{t-1} \cdot \exp(r_t) \tag{7a}$$

$$r_t = \sigma_{t,k} z_t \qquad z_t \sim i.i.d.N(0,1) \tag{7b}$$

$$q_t = \ln(\sigma_{t,k}) \tag{7c}$$

$$q_t = c_1 + c_2 q_{t-1} + c_3 \eta_{t-1} + \eta_t \tag{7d}$$

$$\eta_t = \xi_t \sqrt{h_t} \qquad \xi_t \sim i.i.d.N(0,1) \tag{7e}$$

$$h_t = w + a\eta_{t-1}^2 + bh_{t-1} \tag{7f}$$

where P_t and r_t are the price and log return at time t respectively, $\sigma_{t,k}$ is the realised standard deviation estimated from returns sampled k times per time interval, c_1, c_2 and c_3 are the unconditional mean, AR and MA are coefficients in (7d) respectively, η_t is the serially uncorrelated innovation, h_t is the conditional second moment of η_t. w, a and b are the unconditional mean, ARCH and GARCH coefficients in (7f) respectively.

For evaluation of the forecasting performance of the 1000 days backtesting procedure the following confidence intervals have been used (Kupiec [12]).

$$\text{95\% confidence level:} \quad 37 < X < 65 \tag{8a}$$

$$\text{99\% confidence level:} \quad 4 < X < 17 \tag{8b}$$

where X is a umber of prices given by (7a) that are less than the actually observed prices.

In words, over 1000 days we expect 50 and 10 prices, which called exceptions, to lie below the actually observed prices for 95% and 99% confidence levels respectively. However, since VaR is a stochastic measure some deviations from the expected values should be anticipated. Values that are within the intervals given by (8a) and (8b) can be attributed to a chance. Values outside these intervals will indicate the model misspecification. The results of the backtesting procedure applied for 1 to 10 days ahead VaR forecasts are shown below with * indicating the numbers of exceptions that lie outside the non-rejection regions given by (8a) and (8b).

Table 2. Numbers of exceptions of (7) with $z_t \sim N(0,1)$ in (7b)

Days ahead	1	2	3	4	5	6	7	8	9	10
VaR^{95}	54	47	47	44	42	41	33*	35*	36*	34*
VaR^{99}	4*	7	9	8	5	7	9	7	6	7

As can be seen, the model (7) is misspecified. The misspecification can be attributed to standard normally distributed z_t in (7b) that in fact found to be J distributed. After replacing $z_t \sim N(0,1)$ for $\sim J(0,14.8)$ in (7b), the following backtesting results have been obtained.

Table 3. Numbers of exceptions of (7) with $\sim J(0,14.8)$ in (7b)

Days ahead	1	2	3	4	5	6	7	8	9	10
VaR^{95}	57	51	47	47	42	43	38	38	38	36*
VaR^{99}	9	10	11	12	8	7	10	9	8	8

The model has become almost correctly specified.

5. Conclusion

In the study the distribution of conditional on realised volatility financial returns has been studied. It has been demonstrated that conditional returns are not standard normally distributed, as argued in Andersen et al. [6], but follow a specific previously undocumented distribution. The probability density that represents this distribution has been derived. Since the shape of the new density depends on two parameters it was referred to as $J(\gamma,\alpha)$. It has been demonstrated how $J(\gamma,\alpha)$ can be used in the financial model for Value at Risk forecasting. It should be noted that J does not have a probability density function and therefore is simulated but not a theoretical distribution. It should also be noted that bimodal $J(\gamma,\alpha)$ is the simplest in the J^m family that includes probability densities with any desired number of peaks *m+1*. We are confident that time will indicate additional applications as well as limitations of J^m.

References

1. E. Fama, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 1970, pp. 383-417.
2. J.Campbell, A. Lo and C. MacKinlay, The econometrics of financial markets, (Princeton University Press, New Jersey, 1997).
3. T. Andersen and T. Bollerslev, Answering the critics: yes, ARCH models do provide good volatility forecasts, National Bureau of Economic Research, 1997, Working Paper No. 6023.
4. T. Andersen, T. Bollerslev, F. Diebold and P. Labys, The distribution of realised exchange rate volatility, *Journal of the American Statistical Association* 96, 2001, pp. 42-55.
5. T. Andersen, T. Bollerslev, F. Diebold and P. Labys, Realised volatility and correlation, Working Paper, University of Pennsylvania, 1999.
6. T. Andersen, T. Bollerslev, F. Diebold and P. Labys, Exchange rates returns standardised by realised volatility are (nearly) Gaussian, *Multivariate Finance Journal* 4, 2000, pp. 159-179.
7. P. Clark, A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* 41, 1973, pp. 135-156.
8. M. Moldovan, Stochastic modelling of random variables with an application in financial risk management, Masters Thesis, School of Economics and Finance, Queensland University of Technology, 2003.
9. http://www.esh.lv/r&d/Matlab_codes/
10. Basel Committee on Banking Supervision, Overview of the amendment to the capital accord to incorporate market risks, January, 1996.
11. P. Jorion, Value at Risk, (McGraw Hill, 2001).
12. P. Kupiec, Techniques for verifying the accuracy of risk management models, *Derivatives* 2, 1995, pp. 73-84.

Appendices

Appendix 1. Data description

One main dataset is used in this study. This is the foreign futures exchange rates between the Australian dollar (AUD) and American dollar (USD). AUD/USD futures exchange rates series is obtained from Tick Data, Inc (www.tickdata.com) and covers the period from 2 January 1990 to 31 March 2000 ($T = 2586$ trading days). The duration of each trading day t is 400 minutes. Trading is open Monday to Friday, from 7.20 a.m. to 2.00 p.m. Data is recorded tick by tick and contains 234,905 observations of prices on AUD/USD contracts.

Appendix 2. Volatility signature plot

For selecting the optimal sampling frequency for realised volatility estimation the volatility signature plot approach has been used in the study. A volatility signature plot has been suggested in Andersen et al. [5] and is a graphical diagnostic, designed to provide some guidance in selecting the optimal return sampling frequency for realised volatility series computation. The problem is to find the balance between a measurement error, which is the most severe in the case of using squared returns as volatility estimates, and microstructure noise, which distorts realised volatility estimates as sampling frequency increases. AUD/USD futures exchange rates series, containing tick-by-tick data, is resampled with different frequencies. Eight sampling frequency intervals have been used: 5, 10, 20, 40, 50, 100, 200 and 400 minutes. The corresponding numbers of intraday prices are $k =$ 80, 40, 20, 10, 8, 4, 2 and 1. For return calculation, the last prices that were observed over a sampling period have been used. If no transactions occurred over a sampling period, then the latest recorded prices were taken.

Realised volatility in day t based on k intraday price observations is calculated as follows.

$$v_{t,k} = \sum_{n=1}^{k} x_{t,n}^2 \tag{A2.1}$$

As a result, eight realised volatility series $v_{1,80},\dots,v_{T,80};\dots;v_{1,1},\dots,v_{T,1}$ were obtained. To define a volatility signature plot, the average values of each of eight series were calculated.

$$\overline{v}_k = T^{-1}\sum_{t=1}^{T} v_{t,k} \tag{A2.2}$$

The average realised volatilities $\overline{v}_k$, $k = 80,40,\dots,1$ were plotted against the lengths of sampling intervals in minutes $5,10,\dots,400$. Figure A2.1 shows the result.

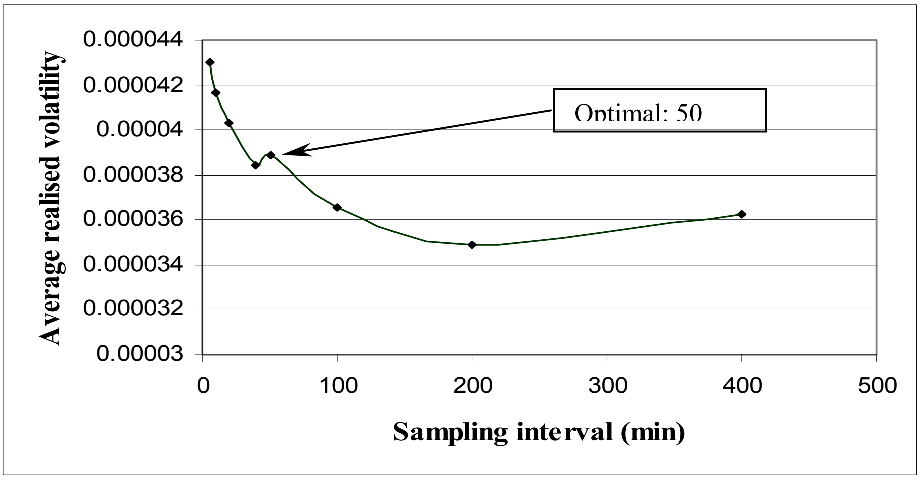


***Figure A2.1**. Realised volatility signature plot for AUD/USD series*

Remembering that microstructure effects are of the main concern in series with the highest sampling frequency, one should follow the volatility signature plot from the left to the right. Andersen et al. [5] argued that the point where the volatility signature plot is "stabilised" should be selected as optimal. In the AUD/USD series plot, this point corresponds to the 50 minute sampling interval ($\overline{v}_8$). This means that for obtaining the optimal realised volatility estimates, returns should be sampled 8 times per trading day. This sampling frequency was used in the study for estimating the realised volatility series.

THE SIMULATION-FREE SIMPLE TEST FOR NORMALITY: SOME COMPARISONS OF SIZE AND POWER

***Max V. Moldovan*[1]*, Konstantin N. Nechval* [2] *and Nicholas A. Nechval* [2]**

[1] *ESH Group Ltd.,*
6-450 Ezermalas Street, Riga, Latvia, LV-1016
Ph: (+371) 9 82 69 66. Fax: (+371) 7 518 519. E-mail: max@esh.lv

[2] *Mathematical Statistics Department, University of Latvia,*
19 Raina Blvd, Riga, Latvia, LV-1050
Ph: (+371) 7034701. Fax: (+371) 7034702. E-mail: nechval@junik.lv

Abstract

The question of whether or not a set of random variables follows a specific theoretical distribution often arises in many theoretical and applied disciplines. When true population parameters of a hypothesised distribution are known the test is known as simple. However, more often one is not able to deal with a whole population because of reasons such as the absence of access to an entire dataset or unacceptable cost of information collection. In this case, parameters have to be estimated from a sample and the test that is based on parameter estimates is known as composite. Inferences that are based on a composite test are in general less reliable than in the case with a simple test. In the following paper we suggest the analytical simple test for normality that allows performing the testing without knowing mean and variance parameters of a population. The presented test can be especially effective in the time-critical applications such as automatic control. Size and power of the presented test are given and compared with the traditional composite test for normality.

Keywords: Decision support systems; Simple test for normality; Change-point detection

1. Introduction

The question of whether or not a set of random variables follows a specific theoretical distribution often arises in many theoretical and applied disciplines. When true population parameters of a hypothesised distribution are known the test is referred to as *simple*. However, more often one does not have the ability to deal with an entire population because of reasons such as the absence of access to a complete dataset or unacceptable cost of information collection. In this case, parameters have to be estimated from a sample and the test that is based on parameter estimates is known as *composite*. Inferences that are based on a composite test are in general less reliable than in the case with a simple test. To understand the reason for that, it should be noted that estimates obtained from a sample will often differ from parameters of a population. The sample estimates are the 'best fit' parameters, or parameters that fit especially well to a sample that they are estimated from. We can say that a sample and estimates are 'relatives' and therefore composite tests are too liberal, failing to reject H_0 of a hypothesised distribution reliably enough.

The normal distribution is certainly the most widely used by academics and practitioners theoretical distribution. Some authors suggest simulation-based techniques that allow overcoming the weaknesses of composite tests for normality. For example, Dufour et al. [1] use the Monte Carlo simulation for controlling size of the several tests. Becker and Hurn [2] also used the simulation-based bootstrap technique for correcting the normality test size. Although with the power of modern computers simulation-based techniques are already not excessively computationally intensive, there are still some areas where the time consuming simulation methods are not affordable. In the following paper we suggest the analytical simple test for normality that allows performing the testing without knowing values of population mean and variance parameters. The test is based on the method first suggested by Nechval [3]. This method is originally designed for change-point detection in regression relationship and especially efficient in small samples of data.

The paper is structured as follows. In Section 2 the parameter-free transformation and the simple test for normality are presented in the explicit algorithmic form. In Section 3 power and size of the suggested test are presented and compared with the traditional composite test. Section 4 summarises the study.

2. Parameter-Free Transformation and Simple Test for Normality

The presented test is based on the technique originally suggested in Nechval [3]. This technique allows transforming a set of random variables to a slightly smaller set of random variables (less than a number of eliminated parameters) without knowing any parameters of the initial sample. The idea behind the method can be stated as follows. A stochastic process observed up to time t can be described by an invariant statistic that fully defines the character of the process with respect to known assumptions. This statistic follows a known theoretical distribution. At time t+1 the next outcome of the statistic becomes available. Since distribution of the statistic is known, the confidence interval $1-\alpha$ can be specified. If a new outcome of the statistic falls outside the confidence interval it can be concluded that the process is broken with confidence $1-\alpha$. The analytical derivation of the method is given in Moldovan et al. [4]. Here the simple test for normality, which follows directly from the method, is presented in the explicit algorithmic form.

Suppose a sample of random variables X_t ($t = 1, 2, ..., T$) is drawn from a population with an unknown stochastic character. It is required to test $H_0 : X_t \sim N(\mu, \sigma^2)$ against H_A : X_t is not $N(\mu, \sigma^2)$. The strategy that employed below transforms X_t to a set of random variables u_t ($t = 3,4...,T$) distributed on the interval $(0,1)$. Note that a resulted set u_t ($t = 3,4...,T$) is two variables less than the initial sample X_t because two parameters, mean and variance, are eliminated from consideration. $X_t \sim N(\mu, \sigma^2)$ if and only if u_t ($t = 3,4...,T$) is identically uniformly distributed $U(0,1)$. The test in the algorithmic form is presented below.

1. Remove the unconditional mean from the initial sample X_t:

$$x_t = X_t - \frac{\sum_{t=1}^{T} X_t}{T} \tag{1}$$

Now the resulted sample x_t ($t = 1,2,...,T$) has an unconditional expectation 0.

2. Take the first sub-sample x_t ($t = 1,...,n$, $n = 2$) and calculate the sum of squared residuals:

$$s\_sq_1 = \sum_{t=1}^{n} \dot{x}_t^2 , \qquad \dot{x}_t = x_t - \frac{\sum_{t=1}^{n} x_t}{n} \tag{2}$$

3. Take the second sub-sample x_t ($t = 1,...,n$, $n = 3$) and calculate the sum of squared residuals:

$$s\_sq_2 = \sum_{t=1}^{n} \dot{x}_t^2 , \qquad \dot{x}_t = x_t - \frac{\sum_{t=1}^{n} x_t}{n} \tag{3}$$

4. Calculate the first characteristic variable:

$$z\_sq_n = n - 2\left(\frac{s\_sq_2}{s\_sq_1} - 1\right),\ n = 3 \tag{4}$$

5. Repeat Step 3 with the next sub-sample x_t $(t = 1,...,n,\ n = 4)$ and calculate the next characteristic variable:

$$z\_sq_n = n - 2\left(\frac{s\_sq_3}{s\_sq_2} - 1\right),\ n = 4 \tag{5}$$

6. Repeat Step 5 until $n = T$. As a result the set z_sq_t $(t = 3,4...,T)$ of characteristic variables is obtained.

7. Transform characteristic variables to variables that lie on the interval $(0,1)$ by applying the cumulative distribution function of the central F distribution with 1 and $t-2$ $(t = 3,...,T)$ degrees of freedom to the characteristic variables z_sq_t $(t = 3,4...,T)$:

$$u_t = 1 - F_{1,t-2}(z\_sq_t),\ (t = 3,4...,T) \tag{6}$$

8. Test if the sample u_t $(t = 3,4...,T)$ is identically uniformly distributed $U(0,1)$. Any test for uniformity, such as Kolmogorov-Smirnov (K-S) test (Smirnov [5]) can be used for that. The initial sample X_t is distributed $N(\mu, \sigma^2)$ if and only if u_t $(t = 3,4...,T)$ is identically uniformly distributed $U(0,1)$. The end of the algorithm.

Note that no information about the mean and variance of the initial sample X_t has been used in the test indicating that the presented test is simple.

3. Comparisons of Size and Power

The size and power of the proposed test have been measured with application of the K-S test at Step 8 of the presented algorithm. The size of the test is based on $M = 10000$ Monte Carlo trials and reported in Table 1.

Table 1. Size of the simple test

sample size T =	**50**	**100**	**200**	**300**	**400**	**500**
α = 0,01	0,0127	0,0086	0,0084	0,0108	0,0107	0,0100
0,05	0,0560	0,0491	0,0443	0,0497	0,0498	0,0499
0,10	0,1052	0,0957	0,0978	0,1044	0,0979	0,0989

Note: α is the level of significance

As can be seen, the size of the presented test is asymptotically correct. For comparison, the size results of the ordinary K-S test for normality have been computed. Note that this alternative test is composite since mean and variance parameters must be estimated from the samples and used for testing. The size of the alternative test is reported below.

Table 2. Size of the composite test

sample size T =	**50**	**100**	**200**	**300**	**400**	**500**
α = 0,01	0,0075	0,0077	0,0079	0,0085	0,0085	0,0084
0,05	0,0395	0,0356	0,0430	0,0435	0,0422	0,0449
0,10	0,0751	0,0844	0,0860	0,0849	0,0920	0,0916

Note: α is the level of significance

It is obvious that the size of the composite test is substantially distorted due to deviations of mean and variance estimates from the true parameters of a population. In other words, the composite test is less reliable.

Next, the power of the presented simple test is measured against some common theoretical distributions and reported in Table 3.

Table 3. Power of the simple test

	Distribution						
	Uniform	$\chi^2(2)$	$t(2)$	$t(4)$	$t(6)$	$t(8)$	$t(10)$
$\alpha = 0{,}01$	0,1135	0,4596	0,8127	0,2778	0,1277	0,0713	0,0511
0,05	0,3027	0,6514	0,8919	0,4477	0,2621	0,1740	0,1474
0,10	0,4354	0,7451	0,9258	0,5411	0,3569	0,2591	0,2233

Note: Power is measured with samples of length $n = 100$

Compare with the power of the composite K-S test reported in Table 4.

Table 4. Power of the composite K-S test

	Distributions						
	Uniform	$\chi^2(2)$	$t(2)$	$t(4)$	$t(6)$	$t(8)$	$t(10)$
$\alpha = 0{,}01$	0,0035	0,5519	0,0551	0,0128	0,0103	0,0101	0,0069
0,05	0,0070	0,8995	0,2097	0,0668	0,0526	0,0186	0,0162
0,10	0,0367	0,9696	0,3540	0,1277	0,1056	0,0431	0,0424

Note: Power is measured with samples of length $n = 100$

With exemption of $\chi^2(2)$ distribution, the suggested simple test is more powerful than the alternative composite K-S test. Although there are some composite tests with greater power than the considered K-S test, it should be noted that these tests can also be used at Step 8 of the presented in Section 2 algorithm, thus increasing a relative power of the suggested simple test. Moreover, to the best of our knowledge, no simple or composite tests exist that have considerable power in detection of such hardly distinguishable from a normal distributions as Student $t(v)$ with $v \geq 4$. As can be seen from Table 3 the presented simple test for normality keeps a substantial power with these distributions.

4. Conclusion

In the paper the simple, as opposed to composite, test for normality was presented. The technical derivation of the method that the suggested test is based on is given in Moldovan et al. [4]. In this paper the eight-step explicit algorithm is given as an alternative to the mathematical representation. Size and power of the suggested simple test are reported and compared with the traditional composite K-S test. It is shown that in opposite to the composite test, the simple test has the correct size, or more reliable. Additionally, with exception of $\chi^2(2)$ distribution, the simple test is more powerful than the composite K-S test. Specifically, the presented simple test detects distributions that are only slightly deviate from a normal in the tail regions, such as Student $t(v)$ with $v \geq 4$, much better than any alternatives that we are aware of. This feature makes the suggested simple test especially useful in areas where extreme outcomes of the process are of the main concern, such as financial risk management. Finally, it should be noted that the suggested simple test, as opposed to the known simulation-based simple alternatives, is analytical and therefore can be used even in the time-critical applications, such as automatic control systems.

References

1. J.-M Dufour, A. Farhat, L. Gardiol and L. Khalaf, Simulation-based finite sample normality test in linear regressions, *Econometrics Journal* 1, 1998, pp. C154-C173.
2. R. Becker and S. Hurn, A bootstrap test for uniformity using the Gini coefficient, Working Paper, Queensland University of Technology, 2002.
3. N.A. Nechval A general method for constructing automated procedures for testing quickest detection of a change in quality control, *Computers in Industry* 10, 1988, pp. 177-183.
4. M. Moldovan, N. Nechval, E. Vasermanis, U. Rozevskis and K. Rozite, Testing for two-phase multiple regressions, in Proceedings of the International Conference on Reliability and Statistics in Transportation and Communication, 2002, pp. 73-81.
5. N. Smirnov, Table for estimating goodness of fit of empirical distributions, *Annals of Mathematical Statistics* 19, 1948, pp. 279-281.

TRANSPORT and TELECOMMUNICATION

ISSN 1407-6160

The journal of Transport and Telecommunication Institute (Riga, Latvia)
The journal is published since 2000.

Articles and review are presented in the journal in English (preferably), Russian and Latvian languages (at the option of authors).

EDITORIAL CORRESPONDENCE

Transporta un sakaru institūts (Transport and Telecommunication Institute)
Lomonosova iela 1, LV-1019, Riga, Latvia. Phone: (+371)-7100594. Fax: (+371)-7100535.
E-mail: kiv@tsi.lv, http:// www.tsi.lv

The K. Kordonsky
Charitable Foundation

The International Conference

RELIABILITY and STATISTICS in TRANSPORTATION and COMMUNICATION (RelStat'04)

15-16 September 2004. Riga, Latvia

PURPOSE

The purpose of the conference is to bring together academics and professionals from all over the world to discuss the themes of the conference:

- Theory and Applications of Reliability and Statistics
- Reliability and Safety of Transport Systems
- Rare Events and Risk Management
- Software Reliability and Testing
- Modelling and Simulation
- Intelligent Transport Systems
- Transport Logistics and Economics
- Education Programmes and Academic Research in Reliability and Statistics

DEDICATION

The Conference is devoted to the memory of Prof. Kh.Kordonsky.

OFFICIAL LANGUAGES

English and Russian will be the official languages of the Conference.

SUPPORTED BY:

Transport and Telecommunication Institute (Latvia) and
The K. Kordonsky Charitable Foundation (USA) in co-operation with:
Latvian Transport Development and Education Association (Latvia)
Latvian Academy of Science (Latvia)

SPONSORED BY

Transport and Telecommunication Institute (Latvia)
The K. Kordonsky Charitable Foundation (USA)
ALTA - Consulting Company in the Air Navigation Services (Latvia)
Latvian Operations Research Society
PAREX bank (Latvia)

HOSTED BY

Transport and Telecommunication Institute (Latvia)

SECRETARIAT

Prof. Igor Kabashkin, Latvia - Chairman
Ms. Inna Kordonsky-Frankel, USA - Co-Chairman
Mr. Nikolay Gudanets, Latvia – Secretary

DEADLINES AND REQUIREMENTS

Abstracts Submissions:	**June 1, 2004**
Camera-ready final manuscript:	**September 16, 2004**
Conference start:	**September 16, 2004**

Abstracts submitted for review should be a maximum of 600 words in length, should present a clear and concise view of the motivation of the subject, give an outline, and include a list of references.

The abstracts should reach the Secretariat before June 1, 2004. Authors should provide a maximum of five key words describing their work. Please include the full name, affiliation, address, telephone number, fax number, and e-mail address of the corresponding author.

Camera-ready documents must be handed in at the registration desk. Papers presented at the conference will be included in the conference proceedings. Instruction for papers preparing can be found on the conference WWW page: www.tsi.lv.

REGISTRATION FEE

The registration fees will be Euro 30 before the deadline, and Euro 40 after the deadline. This fee will include: hard copy of the Conference Abstracts, hard copy of the Conference Proceedings (the proceedings will be mailed to the delegates after the conference), coffee breaks, Conference Dinner.

Each registration fee might include just one paper, which presentation will be included in the conference program and published in the conference proceedings.

VENUE

Riga is the capital of the Republic of Latvia. Thanks to its geographical location, Riga has wonderful trade, cultural and tourist facilities. Whilst able to offer all the benefits of a modern city, Riga has preserved its historical charm. It's especially famous for its medieval part – Old Riga.

Old Riga still preserves many mute witnesses of bygone times. Its old narrow streets, historical monuments, organ music at one of the oldest organ halls in Europe attract guests of our city. In 1998 Old Riga was included into the UNESCO list of world cultural heritage.

ACCOMMODATION

A wide range of hotels will be at the disposal of participants of the conference and accompanying persons.

FURTHER INFORMATION

Contact:

Nikolay Gudanets, Secretary, RelStat'04
Transport and Telecommunication Institute
Lomonosova iela 1, Riga, LV-1019, Latvia
Telephone: +(371)-7100651
Fax: +(371)-7100660
E-mail: nlg@tsi.lv
WWW: www.tsi.lv