

SADRŽAJ**ORIGINALNI NAUČNI RAD**

Informatička definicija statističkog istraživanja *Branko Jiriček* 3

STRUČNI RADOVI

Obrada podataka u istraživanju o prevozu robe drumom *Jelena Kovijanić* 15

Pilot istraživanje o prevozu robe u drumskom saobraćaju u Srbiji 2008.
godine *Jelena Kovijanić, Olga Meloski Trpinac, Željko Lesić* 24

OSVRT

Savremene tendencije u lingvostatistici *Nikola Dobrić* 43

PRIKAZ NAUČNOG DOGAĐAJA

XXXV Simpozijum o operacionim istraživanjima *Dr Luka Filipović* 51

IN MEMORIAM

Prof. dr Vladislav Milošević *Prof. dr Branislav Boričić* 57

Prof. dr Vladislav Milošević *Prof. dr Ljiljana Petrović* 59

Prof. dr Mileva Žižić *Prof. dr Dubravka Pavličić* 61

Prof. dr Mileva Žižić *Prof. dr Branislav Boričić* 65

STATISTIČKA REVIJA

IZLAZI TROMESEČNO - PUBLICATION TRIMESTRIELLE - QUARTERLY

Redakcioni odbor: Conseil de rédaction: Editorial Committee:

*prof. dr Miladin Kovačević (Beograd), prof. dr Vesna Bucevska (Skoplje),
prof. dr Ksenija Dumičić (Zagreb), prof. dr Vaso Dragović (Pale), dr Milorad Kovačević
(Otava, Kanada), prof. dr Srđan Bogosavljević (Beograd), prof. dr Nahod Vuković (Beograd),
prof. dr Nada Lakić (Beograd), prof. dr Miodrag Lovrić (Beograd), prof. dr Dragana Milojić
(Beograd), mr Olga Melovski Trpinac (Beograd), prof. dr Pavle Mladenović (Beograd), mr Emilija
Nikolić Đorić (Novi Sad), prof. dr Zdravko Šolak (Novi Sad)*

Glavni i odgovorni urednik: Rédacteur en chef: Editor in Chief:

prof. dr Miladin Kovačević

Sekretar redakcije: Secrétaire de la rédaction: Editorial staff secretary:

Božidar Popović

Tehnički urednik: Redacteur technique: Technical Editor:

Ljiljana Stanić

Adresa redakcije:

Statističko društavo Srbije, 11000 Beograd, Milana Rakića 5

Žiro-račun br. 205-99554-89 kod Komercijalne banke

Tiraž: 200

SARADNICI U OVOM BROJU:

*Branko Jiriček
Jelena Kovijanić
Olga Meloski Trpinac
Željko Lesić
Nikola Dobrić
Dr Luka Filipović
Prof. dr Branislav Boričić
Prof. dr Ljiljana Petrović
Prof. dr Dubravka Pavličić*

Predato u štampu: maj 2010. godine

Izašlo iz štampe: maj 2010. godine

Štampa: Republički zavod za statistiku Srbije, Beograd, Milana Rakića 5

ORIGINALNI NAUČNI RAD

Branko Jireček
Statistical Office of the Republic of Serbia

AN IT DEFINITION OF A STATISTICAL SURVEY

ABSTRACT

Statistical Offices have been dealing with statistical surveys for a long time. Various attempts of defining surveys were made during all that time, and this is yet another one, based mostly on the IST integrated solution for processing data in a statistical survey.

IST is a meta-data driven application for processing data from statistical surveys, developed first in the former Statistical Office of Serbia and Montenegro, and later transferred to the Statistical Office of the Republic of Serbia, where it is further developed and in a rather wide use. As the meta-data that IST uses is mostly the IT definition of the survey, the idea of formalizing this definition arose, so it could be a standalone definition of a survey. This is done by transforming the traditional relational database definitions that IST uses into XML language.

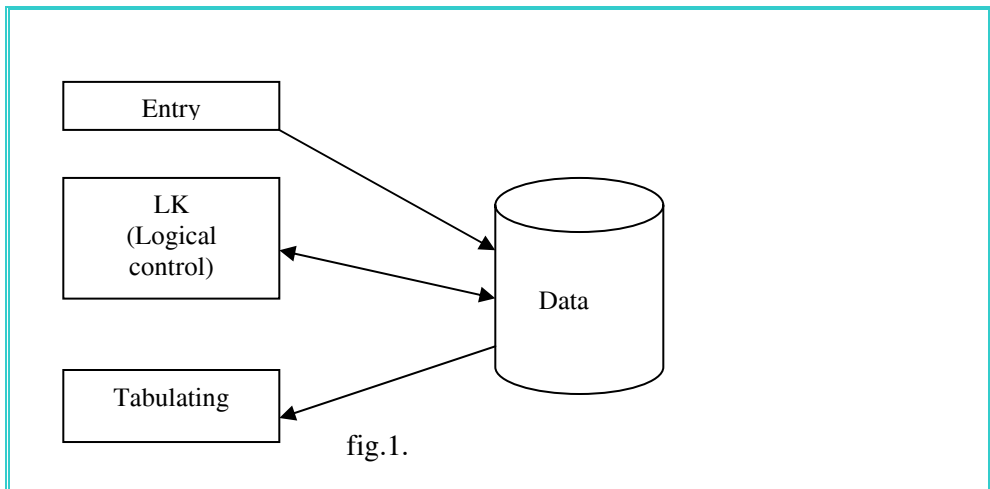
I INTRODUCTION. SHORT HISTORY OF IST.

For a long time data processing in statistical systems was (and still is) organized on a "stovepipe" principle, which results in a set of various, often incompatible solutions dealing with similar problems. A more general approach was designed and implemented in the Statistical Office of Serbia and Montenegro, which deals with processing of all the statistical data in a completely different manner. As that Office ceased to exist, this approach, together with all the program solutions, was transferred to the Statistical Office of the Republic of Serbia, where it is used and further developed. At the time being, about 50 statistical surveys are being processed this way.

I.1. The “usual” way of data processing

The most usual way of processing data of a statistical survey is a, so called, “stovepipe” principle, where the complete process is performed – from data entry to publishing the final results separately for each survey. Every processing of a survey data consists mainly of the following three phases:

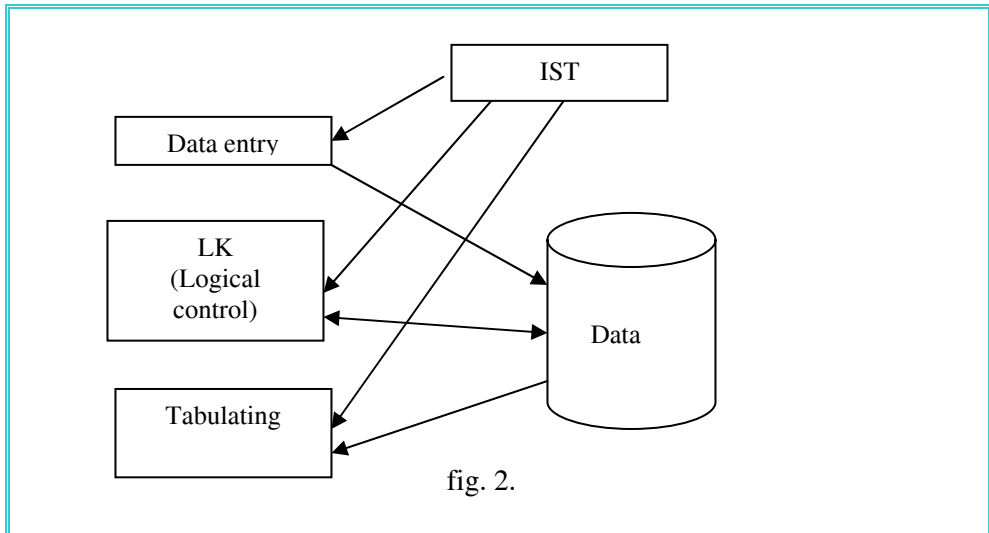
- (a) Data entry
- (b) Logical control (data “cleaning”)
- (c) Tabulating (dissemination, publishing)



The development of IT caused that various surveys had various, in that time “popular”, implementations. Depending on many factors, such as the importance of the certain survey, (un) willingness for changes etc., we had (and still have) the situation where various surveys are IT “handled” in a wide variety of ways – from surveys where data is stored in “plain” sequential files to those where they are (unfortunately) in various data bases. For a long time, IT was exclusively done on IBM mainframe platform(s). Naturally, the development of PCs caused some surveys to be “down-sized” to this platform, and, of course, we have “mixed” solutions, just to make the situation even more complicated. Therefore, we have a situation where the processing of statistical data is done on various platforms, implemented on various software tools, with data stored in various ways, separately for every survey, and even worse, for every phase, as shown in fig. 1. This is a very simplified picture, which is done deliberately, because this was the biggest issue that we found all the surveys had “in common” and from where we had to start.

I.2. “And now for something completely different”

A. Basic IST concepts



It appears to be very interesting that, although we contacted relevant people from several statistical offices aiming to solve the mentioned problems, no one seemed to be interested. Hence, we had to start “from scratch” and “on our own”. The only exception we found was EUROSTAT’s Euro Trace, but, besides it is not able to deal with most of our surveys, we also heard about it too late, since at that time we had already developed the main foundations of our idea. Nevertheless, it was very helpful to exchange experiences with people working with Euro Trace. The worst of all is the fact that during all these years we all have been writing, more or less, only three applications referring to data entry, logical control and tabulating. We felt that it was high time for these three to be finally “written” so that we would not ever (at least in the way we did until now) have to “worry” about them again. The aim is to make the whole process of “handling” statistical surveys with as little “real” programming as possible. Even better, with no “real” programming at all. Although there is strong opposition to universal applications, we had some rather successful experiences in the past. We would like to mention TABS, a universal tabulating program developed for the IBM mainframe in the Federal Statistical Office and GoDar, a data entry and logical control system for the same platform, developed in the Statistical Office of the Republic of Serbia. Although both of these solutions were outdated long time before, we were positively encouraged with the existence of this kind of tradition. It was obvious that data had to be stored in a relational database. The problem was which database (DBMS) to choose, because we already had two (IBM’s DB2 and Microsoft’s SQL-Server) in use, without any knowledge about the possible expanding of this number. Due to such a reason we had to make our solution independent of the physical implementation of the database, which, on the other hand, made the limitation of using only the most basic set of SQL commands, with absolutely no enhancements or improvements of any DBMS, which, unfortunately, excluded stored procedures, triggers and all of this more than useful stuff.

The main idea, as shown on fig. 2, was to try “organizing” the surveys as to “take care” of its own data and to process it, instead of the current situation where every survey is a “world for itself”. Technically speaking, IST shown in figure 2, is a database where all the survey relevant data, both operational (such as data description, conditions for logical control, definitions of the output tables) and descriptive (methodologies, field descriptions, etc.) is stored. The first “test bunny” was the TU-11 survey in the field of tourism statistics, that after all expected and unexpected troubles was “set up” and is presently working “on IST”.

However, there is still a lot of work on the whole idea, many things are to be improved, written and/or rewritten, but we are glad that it is brought to an operational level, which means that the whole idea was correct

B. “Inside” IST

The IST is mainly designed as a meta-data database “handler”. The whole IST database is, as mentioned above, made that way. Although it may look complicated, the whole IST database is rather simple. Here are the main tables:

- (a) **IST** Table which contains global information about particular surveys. Here we have some global information, such as the full name of the survey, database(s) where data are, periodicity of the survey, paths to the methodological texts, queries for output tables etc.
- (b) **ISTBaze** Table with the “real” connection strings, which the application uses to connect to “real” databases and data within them
- (c) **ISTTabele** List and definition of tables that the survey consists of. Here are the relations between the tables too.
- (d) **ISTPolja** Definition of the data fields of surveys with accompanying properties of them, such as the type, length, possible bounds with some consultation tables etc.
- (e) **ISTLK** Logical Control definitions, or, better to say, a list of possible errors and their definitions
- (f) **ISTTABS** List of output tables. As a matter of fact, here is the list of names of queries and possibly Excel, Word, SAS, etc. files that the application uses to make an output table.

At the moment, logical control criteria and queries for output tables are more or less limited to “pure” SQL. Although this should not have to be a real problem, it may be uncomfortable to manage some things with only “pure” SQL, so certain efforts are already made in making “functions” for some usual situations, and even a bigger effort shall be made in, probably, creating a little language that will ease this process. Since the complete solution for data entry already exists in the Statistical Office of the Republic of Serbia, this solution was included in IST, and later a whole system for automatically generated data entry was developed, so that IST “generates” input mask(s) and does all the control while data entry or editing. It seems reasonable that IST should consist of, at least, two “parts” – one for the end-user (statistician), and the other for the IT specialist who would “install” a survey “into” IST. At the moment, only the first part exists.

A special effort was made to overcome the “time consistency” problem – the fact that consulting tables, questionnaires, even the whole survey changes in time. It would be desirable if the system would work correctly through these changes. This is done with the Vod – Vdo (Valid from – Valid to) variables (fields) which are implemented in all of the main IST tables. A step more, these fields could be implemented almost everywhere

(consulting tables, even main survey tables) and IST will “recognize” them and act accordingly to the time limits set in these variables (fields).

I.3. Implementation

IST, as already mentioned, is created with the intention to work with any relational database. This also stands for the IST database. Maybe it would be better to say that it works with any database, because there is no limitation that data must be in only one database, the data from one survey can be in different (even physically different) databases, which is not so hard to expect, because this way we are planning to have unique registers, classifications and other “consulting tables” for all surveys. There was no doubt about the operating system on which IST would function. It had to be made a Microsoft Windows application, because all the “infrastructure” of the Statistical Office of the Republic of Serbia is running on that platform. There were certain doubts about how Windows will “communicate” with the IBM mainframe and, particularly, with the DB2 database, but, fortunately, this works very well. In the near future, the intention is to make the IST as a web application, and make it platform independent as well.

A. Programming language

There was much more dilemma about the programming language in which this would be implemented. Initially, the decision was made to do it in the Microsoft Access, or, better to say, in its VBA. The main reason for such a decision was that Microsoft Office is already installed on every PC in the Statistical Office, and users (statisticians) are “a bit” familiar with it, so they may be able to do some “elementary” operations in the Microsoft Access (certainly much easier than, for instance, in the DB2). The user (statistician) was supposed to choose the survey he wants (and has the right) to deal with, and either input data, logically test and clean them, or produce outputs. The logical control phase is the most interesting one, where the user was allowed either to test and change the data “locally” (on his PC) or “globally” (directly in the database where the data are stored, enabling more people to simultaneously perform this operation). And yet, all of this was abandoned when it came to the implementation in the Statistical Office of the Republic of Serbia, where the whole thing was first “translated” to the Microsoft Visual Basic 6, which was their strategic language, and later to the VB.NET. These “translations” were a good chance for “ironing up” the whole project, so it was all rather simplified from the users’ point of view, several new possibilities were added, and everything was a bit “strengthened”. The final goal is to make the IST a completely web-oriented application.

B. RDBMS

Since the IST is supposed to be basically independent of the database platform, it was interesting to see how it would work “in real life”. After the first test with TU-11, which “ran” on the Microsoft’s SQL Server, three surveys were set up in the Statistical Office of the Republic of Serbia, which are “running” on IBM’s DB2 on a Z800 mainframe. Also the IST database was “transferred” to DB2. As suspected, different databases suffer an expected problem of “incompatibility” of their SQL “dialects”. Except for some minor solutions, this problem stands. The most uncomfortable thing that had to be dealt with was the absolutely non-expected DB2’s “hostility” towards programmers – for instance, DB2 does not allow joins inside an update statement, so every update, although relatively simple,

had to be “turned” into a rather long and quite non-understandable query. The situation is even worse in creating queries for output tables - DB2 “forces” you to write rather long (and complicated☹) queries, so in the end the limit of 32K for a SQL query begins to be a problem. This certainly is not the subject of this topic, but I simply couldn’t resist not grudging and complaining a little☺. At the moment, the IST has “moved back” to the SQL Server, with about 50 surveys both on SQL Server and on DB2.

C. Samples

One of the surveys currently “running on” IST is the APG, or “The Cattle sample”, which is the survey based on a sample. Although discussions about how to generally deal with sample surveys will continue for a long time, the experience of how the IST apparatus “dealt” with this was more than positive, which also was a confirmation that the general idea was good.

I.4. The user’s view

Beside the main IST’s goal of organizing and uniformly dealing with all statistical data, one of the things that had to be achieved was simplifying the job of the end-user (statistician). So, let us see how it looks from the user’s point of view. After identification, the user has to choose the survey and the time period with which he is intending to work with. Then next step can be “cleaning” the material or making some outputs. After pushing the button for logical control, the application sets all the “errors” in the material. “Cleaning” of the material is made through a grid (grids, if there are more tables with data), where the statistician sees the material and the list of logical errors in the row on which he is positioned, and has a chance of an easy changing of the data in the grid by simply “re-typing” it over. After changing data, the logical control is done for that row, so he can interactively see the results of the change(s). There is also a possibility of “calling” the data entry part or program and change the data that way. The statistician, of course, can choose which columns and rows he wants to see in the grid(s). All that the programmer has to do is to make a list of logical errors and “define” them with a little “enriched” SQL language. No “real” programming.

Output tables can be “made” in two ways – the application enables the statistician either to make “already-made” tables or to “create” his own. “Creation” of output tables functions relatively easy – on the left is a list of all possible fields and the statistician’s job is to “put” some of them either in the resulting table’s heading, fore-column or as an expression that will be summed. The application generates a SQL query, and, after execution, the table with the results of the query which can be exported to some Office application.

Some output tables are made very often, so it seems like a good idea to make them once and just call them later. The statistician just has to choose one of the offered “already-made” tables and do with the result table whatever he wishes. “Behind the curtains”, the application “runs” a “pre-made” file with a SQL query (or more of them), exports this to Excel, possibly “runs” an Excel macro (if there is a need for it), exports all of this to the Word and, running a little Word macro, combines the “pre-made” Word table heading with the data. The result is a “ready to print” Word document. The interesting thing is that the programmer has only to write the SQL query (queries), to “prepare” the table heading and, if needed, to create an Excel macro. No “real” programming again☺.

I.5. Things yet to be done

First of all, the IST must “evolve” to a true web-oriented application. Steps are already taken, but this will surely be a rather “long run” action. Definitely, at least a little “language” will have to be made to upgrade and “unify” various SQLs and make the life of programmers much easier. A whole dissemination mechanism involving data warehouses and serious user-end description metadata would also be more than welcomed.

Although there is still a lot of work to be done, we are satisfied with the progress we have made so far. The whole system is functioning in the Statistical Office of the Republic of Serbia and it is chosen to be the future strategic tool for processing statistical data.

II. SURVEY DEFINITION

All the (meta)data you have to define in order to “put” a survey into the IST apparatus is, in fact, the definition of the survey. The thing remaining to do is just to try to “formalize” all this (meta)data. That will be done using an XML representation of the IST (meta) data and here is the .xsd definition:

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="Istrazivanje">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="Sif1st" type="xs:string" />
        <xs:element name="sifraT" type="xs:string" />
        <xs:element name="Naziv" type="xs:string" />
        <xs:element name="Metodologija" type="xs:string" />
        <xs:element name="Upitnik" type="xs:string" />
        <xs:element name="Periodika" type="xs:string" />
        <xs:element name="BrojTabela" type="xs:integer" />
        <xs:element name="TxtDir" type="xs:string" />
        <xs:element name="ProgUnos" type="xs:string" />
        <xs:element name="ISTBaze">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="BAZA" type="xs:string" />
              <xs:element name="FizBaza">
                <xs:complexType>
                  <xs:sequence>
                    <xs:element name="FizBaza" type="xs:string" />
                    <xs:element name="GCHINT" type="xs:string" />
                    <xs:element name="FMOD" type="xs:string" />
                    <xs:element name="FDateAdd" type="xs:date" />
                  </xs:sequence>
                </xs:complexType>
              </xs:element>
              <xs:element name="IzrazODBC" type="xs:string" />
              <xs:element name="IzrazOLE" type="xs:string" />
              <xs:element name="vlasnik" type="xs:string" />
              <xs:element name="prefix" type="xs:string" />
              <xs:element name="VODVDO">
                <xs:complexType>
                  <xs:sequence>
```

```

        <xs:element name="VOD" type="xs:dateTime" />
        <xs:element name="VDO" type="xs:dateTime" />
    </xs:sequence>
</xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="ABAZA" type="xs:string" />
<xs:element name="ISTTabela">
    <xs:complexType>
        <xs:sequence maxOccurs="unbounded">
            <xs:element name="Tabela" type="xs:string" />
            <xs:element name="DG" type="xs:string" />
            <xs:element name="NadTabela" type="xs:string" />
            <xs:element name="Opis" type="xs:string" />
            <xs:element name="ISTPolja">
                <xs:complexType>
                    <xs:sequence maxOccurs="unbounded">
                        <xs:element name="Polje" type="xs:string" />
                        <xs:element name="Kljuc" type="xs:string" />
                        <xs:element name="Naziv" type="xs:string" />
                        <xs:element name="RelBaza" type="xs:string" />
                        <xs:element name="RelTabela" type="xs:string" />
                        <xs:element name="RelPolje" type="xs:string" />
                        <xs:element name="VS" />
                        <xs:element name="OD" />
                        <xs:element name="DO" />
                        <xs:element name="DG" type="xs:string" />
                        <xs:element name="Opis" type="xs:string" />
                        <xs:element name="Rbr" type="xs:int" />
                        <xs:element name="tip" type="xs:string" />
                        <xs:element name="duzina" type="xs:string" />
                        <xs:element name="Forma" type="xs:string" />
                        <xs:element name="Izraz" type="xs:string" />
                        <xs:element name="VODVDO">
                            <xs:complexType>
                                <xs:sequence>
                                    <xs:element name="VOD" type="xs:dateTime" />
                                    <xs:element name="VDO" type="xs:dateTime" />
                                </xs:sequence>
                            </xs:complexType>
                        </xs:element>
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
        </xs:sequence>
    </xs:complexType>
</xs:element>
<xs:element name="ISTLK">
    <xs:complexType>
        <xs:sequence maxOccurs="unbounded">
            <xs:element name="RbrGreske" type="xs:string" />
            <xs:element name="GRESKA" type="xs:string" />
            <xs:element name="USLOV" type="xs:string" />
            <xs:element name="NazivGreske" type="xs:string" />
            <xs:element name="Akcija" type="xs:string" />
            <xs:element name="Tezina" type="xs:string" />
            <xs:element name="VODVDO">
                <xs:complexType>
                    <xs:sequence>
                        <xs:element name="VOD" type="xs:dateTime" />
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
        </xs:sequence>
    </xs:complexType>
</xs:element>

```

```

        <xs:element name="VDO" type="xs:dateTime" />
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="VODVDO">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="VOD" type="xs:dateTime" />
      <xs:element name="VDO" type="xs:dateTime" />
    </xs:sequence>
  </xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
<xs:element name="ISTTABS">
  <xs:complexType>
    <xs:sequence minOccurs="0" maxOccurs="unbounded">
      <xs:element name="RbrTabele" type="xs:integer" />
      <xs:element name="Naziv" type="xs:string" />
      <xs:element name="NazivE" type="xs:string" />
      <xs:element name="KVIRI" type="xs:string" />
      <xs:element name="UpitZaTabelu" type="xs:string" />
      <xs:element name="RIPORT" type="xs:string" />
      <xs:element name="OpisTabele" type="xs:string" />
      <xs:element name="Excel" type="xs:string" />
      <xs:element name="SAS" type="xs:string" />
      <xs:element name="excelZag" type="xs:string" />
      <xs:element name="Program" type="xs:string" />
      <xs:element name="VODVDO">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="VOD" type="xs:dateTime" />
            <xs:element name="VDO" type="xs:dateTime" />
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:schema>

```

XML files defined with this definition would be mostly useful, first of all, practically, for transporting an IST survey from one point to another (for instance, to or from a branch office). As major part of the survey is defined here (tables / record definitions, logical control, etc.) there is no reason not to include this XML file in the methodology of the survey. The only thing that is really missing is the definition of the "output" reports. However, as making a better dissemination mechanism of surveys in IST is one of primary things to be done in the near future, this will probably cause a change in the XML file, so this branch will automatically improve. The already mentioned application for IT specialists will, in fact, be made over these XML definitions of a survey and changing the XML will simply result in changing the whole survey. In that way, this XML presentation of a survey could be more understandable to a non-IT statistician, so she/he can, in a relatively easy way, make the intended changes even without the IT specialist, which would definitively prove that this whole idea is just what it was meant to be.

CONCLUSION

The IST metadata-driven statistical survey data processing tool has, for the time being, proved to be more than useful, not only in the every-day work in a Statistical Office, but also in an even more theoretical and abstract way. One of such ways is certainly the idea of formalizing the survey definition. Beyond the fact that every formalization is an achievement by itself, some interesting consequences arise almost instantly. First of all, a need for further formalization of this metadata already becomes wanted, and further on, this could lead to even more interesting analyses that would lead to the better statistical system.

ACKNOWLEDGEMENTS

Microsoft Windows, Microsoft Office, Visual Basic, VBA, VB.NET, Microsoft Access, Microsoft Excel, Microsoft Word, SQL Server are registered trademarks of Microsoft Corporation

DB2, Z800 are registered names of International Business Machines Corporation (IBM)

REFERENCES

- [1] Branko Jireček: Development of IST system for statistical data processing; supporting paper for Meeting on the Management of Statistical Information Systems (MSIS 2007), Geneva, 8-10 May 2007.

REZIME

Osnovna delatnost Statističkih Zavoda su statistička istraživanja, tako da postoje mnogi pokušaji, kako formalnog tako i manje formalnog, definisanja istraživanja. Ovo je još jedan takav pokušaj, zasnovan na IST integrisanom rešenju za obradu podataka statističkih istraživanja.

IST je programski paket zasnovan na meta-data definicijama, prvobitno napravljen u Zavodu za statistiku Srbije i Crne Gore, a trenutno u upotrebi u Republičkom zavodu za statistiku Srbije. Transformacijom meta –podataka organizovanih u klasičnoj relacionoj bazi podataka koje IST koristi u odgovarajući XML format se i dobija definicija (i to pre svega vezana za IT) statističkog istraživanja.

STRUČNI RADOVI

Jelena Kovijanic
Statistical Office of the Republic of Serbia, Serbia
e-mail: jkovijanic@stat.gov.rs

DATA EDITING AND IMPUTATION IN THE SURVEY ON ROAD FREIGHT TRANSPORT IN SERBIA

ABSTRACT

This paper focuses on data editing and imputation processes in the pilot Survey on Road Freight Transport conducted in Serbia during two weeks in 2008. It reviews errors in data from this survey and methods which were used to correct them. Analysis of these methods were performed in order to assess their efficiency and to discuss future methods for preventing data entry errors and correcting field errors.

1. INTRODUCTION

The pilot Survey on Road Freight Transport was conducted in Serbia during two weeks in 2008. Data on road performances of road freight vehicles in and outside the transport sector of the national economy were collected. These data provide information such as origin and destination of transport, weight of transported goods, type of goods, performed tonne-kilometres, type of packaging, category of dangerous goods, degree of utilization of vehicles, axle configuration, age of vehicles etc.

The survey methodology has been defined according to European Commission Regulation 1172/98 on statistical returns in respect of the carriage of goods by road. French statistics has provided expert consultancy concerning the methodology, design of questionnaire, data editing and estimation procedures.

The survey on road freight transport is complex in many respects. All the phases from the definition of the goals, preparation of the frame, sample design, design of the questionnaire, management and organization of the filed work, data editing, imputation, estimation, to the dissemination of tables demand careful preparation and checking. When it is launched it will be carried all through the year. For each week data will be collected from a sub sample of vehicles. The processing and dissemination tables will be on quarterly basis. Therefore, the survey needs to be very well planned because there will not be much space for unexpected problems and slippage from the defined time table.

The pilot survey was conducted to test the survey methodology, instruments and data processing and estimation. Many problems have been encountered and analysed in order to eliminate them from the future survey.

This paper examines data editing methods used in the Pilot survey: the data checks performed before, during and after the phase of data entry, which were used to identify invalid and inconsistent data. Also special attention is given to logical and mathematical edits performed in order to detect inconsistencies between variables. Data editing methods are analysed in order to make suggestions for future improvements.

2. DATA EDITING PROCEDURES

2.1. Introduction to data editing procedures performed in the survey

The checks performed in the survey involve checks of:

- Concepts and definitions (nature and type of variable, allowed values)
- Logical relationships (if... then...else)
- Mathematical relationships (e.g. $v1+v2*v3=v4$)

Errors were localized using two types of procedures:

- Use of “human intervention” ;
- Use of software to identify errors.

Data editing was performed in several steps:

1. Each questionnaire was examined by pool of experts and errors and inconsistencies were treated, prior to data entry. Both localization and correction were performed manually.
2. Automatic check of concepts and definitions was performed by data entry software.
3. Finally, automatic error localization in Access was done, according to defined logical and mathematical relationships, after all data from the two weeks of pilot survey had been entered. Corrections and imputations were performed according to the predefined set of rules.

2.2. Manual treatment of questionnaires and preparation for data entry

In the first stage of data editing each questionnaire was examined, and interpreted. Certain values were coded. The data that have complicated structure were firstly analyzed by staff in the Department of Statistics of Transport and Communications.

The problem is that data are collected about journeys, and journeys can be quite complicated. For example, a hauler sets off empty from his depot, picks up a first load, travels to a second pick-up point where he takes on a second load, travels to a factory where he unloads both loads and then travels empty back to base.

Here is a specific example that illustrates the potential complications that can arise. A lorry is loaded at the depot with 20 tonnes of goods. It travels 30 kilometres to A where it unloads 10 tonnes, then 15 kilometres to B where it unloads 5 tonnes, then 10 kilometres to C where it unloads the rest (5 tonnes) and travels empty 10 kilometres back to the depot (see Appendix 2).

The main statistics that the road freight survey is attempting to estimate are *tonne-kilometres* and *tonnes lifted*. In the above, tonne kilometres are easily calculated, as follows:

$$\text{Tonne-kilometres} = (20t * 30\text{km}) + (10t * 15\text{km}) + (5t * 10\text{km}) = 800\text{tkm}$$

So, for the purposes of calculating total *tonne-kilometres*, each stage can be taken separately. However total *tonnes lifted* for the journey above is just 20 tonnes. The total distance travelled is one loaded journey of 55 kilometres and one empty journey of 10 kilometres. Notice that in this case, the total distance travelled has to be calculated as the sum of the stages travelled. The calculation tonne-kilometres divided by tonnes lifted does not give the total distance travelled, as it would for a one-stop journey. In the above case, tonne-kilometres divided by tonnes would be 40, not the 55 kilometres of the total multiple stop journey.

Let's look at another example where a lorry is loaded at the depot with 20 tonnes of goods, it travels to A where 20 tonnes are off-loaded and 10 tonnes loaded, then it travels to B where it is fully unloaded and where 5 tonnes are loaded and taken to C and unloaded. In this case, the tonne-kilometres and distance travelled would be the same as in the previous example but there would be four journeys recorded and the total *tonnes lifted* would be 35 tonnes.

Although the respondents were given a detailed instruction and asked to give each journey an appropriate code for the type of journey (1-one stop journeys, 2-two to four stop journeys, 3-collection/distribution journeys with more than 5 stops, 4-empty journeys) and restructure the type 2 journeys into consignments, their responses had to be checked by experts due to complexity of the problem. The example of restructuring type 2 journeys into consignments is given in the Appendix 2 of this paper.

Therefore, in this stage the staff in the Department of Statistics of Transport and Communications checked the details provided by the hauler on the questionnaire, in order to ensure that the journeys were properly restructured and assigned the correct "type of journey" code. In this way main statistics of tonnes and tonne-kilometres could be correctly calculated. The main indicator in determining whether the journey was correctly entered was the relationship between fields: distance travelled in one stage of the journey and distance travelled by one vehicle in one day (sum of all distances travelled in all journeys of one vehicle in one day). A large number of errors were found. This indicates that the questionnaire should be constructed in the way that it collects a sufficient amount of information, but in the same way does not burden the respondent with complex reasoning.

2.3. Automatic checks of concepts and definitions in data entry program

The data entry software provides three panels for entering data from the questionnaires.

On the first panel, data about vehicle are entered. Data entry application requires an input of region of the sampled vehicle (region to which the address of the owner of the vehicle belongs), ordinal number of the vehicle in the sample for that region, as well as the survey year and survey quarter from the returned questionnaire. The application assigns unique identification that comprises of these data to each vehicle.

On the second panel data on journeys of the specified vehicles are entered.

On the third panel the data on each consignment of goods on the journey are entered. Continuation of the data entry to the panel three depends on the answers given in the field "type of journey" on the second panel. If that answer is 4-empty journey, application disables the user to fill in the data in the third panel.

Checks of concepts and definitions are incorporated in the data entry software. There are auxiliary databases with approved values for certain variables. The program is conceived in that way that it offers a drop down menu for some variables from which the user can choose, such as place of loading/unloading, type of goods, category of dangerous goods etc. For others it reports an error if the value entered is out of range i.e. the entered value for type of transport can be only 1- public transport or 2- transport on own account.

Logical and mathematical relationships were not checked by the data entry software due to difficulties in the organization of the survey and tight deadline. Automatic error localization was performed later, after all data from the two weeks of pilot survey were entered.

2.4. Error localization according to defined logical and mathematical relationships and imputation of missing values.

For the list of variable names and content see Appendix 1.

1. Value of the variable A1.7- type of transport (1-public, 2-on own account), where missing, was imputed according to its value in the previous journey of the same vehicle. If there was no value in the previous journey, values were imputed according to the activity code of the enterprise.

If the activity of the enterprise is transport then the imputed value was 1- public, else imputed value was 2- on own account.

2. Value of the variable A2.10- degree of loading of the vehicle (0- empty journey, 1- not full, 2-full, 9-missing), where blank was imputed in the following way: for journeys of an empty vehicle the imputed value was 0 and for all laden journeys 9.

For future surveys a possibility of imputing the values depending on the carrying capacity, weight of goods and type of goods carried, should be discussed.

3. Loading capacity of the vehicle (variable A1.5) was checked according to the rules: loading capacity of the hauling vehicle should be between 3500kg and 30000 kg. Trailers and semi-trailers should have the loading capacity between 0kg and 25000kg. Maximum loading capacity of joined hauling vehicle with trailer/semi-trailer should be less than 55000 kg.

Detected errors and imputations were corrected by manual insertions of expert assessments of the loading capacity of the vehicle.

4. Relation between the loading capacity (variable A1.5) and weight of goods carried in one consignment of the journey (variable A3.2) was checked by using the following rule:

$$A3.2 < 1.1 * A1.5$$

Detected errors were corrected by manual insertions of expert assessments of the weight of goods carried in one consignment.

5. Relation between variables A2.2 (weight of goods carried in one journey) and A3.2 (weight if goods carried in one consignment) was checked according to the following set of rules:

If the type of journey is A2.1=1¹ then A2.2=A3.2 for a given journey.

If the type of journey is A2.1=2 then $A2.2 = \sum_{i=1}^{\max \text{GoodsN}} A3.2i$ where GoodsN is the ordinal number of the goods transported in one journey.

If the type of journey is A2.1=3 then A2.2=A3.2.

If the type of journey is A2.1=4 then A2.2=0.

Detected errors were corrected by insertions of the values of A2.2 calculated by using the above stated formulas.

For values A2.2>30000 kg where the type of journey is A2.1=1, plausibility checks were performed by examining and correcting these rows manually.

6. Relation between variables A2.5 (distance travelled in one journey) and A3.7 (distance travelled by one consignment in the journey) was checked according to the following set of rules:

If the type of journey is A2.1=1 then A2.5=A3.7.

If the type of journey is A2.1=2 then A2.5=max A3.7i, i= (1, max GoodsN).

If the type of journey is A2.1=3 then A2.5=A3.7.

Detected errors were corrected by insertions of the values of A2.5 calculated by using the above stated formulas.

7. Calculation of tonne-kilometres (variable A2.6) was performed according to these rules:

If the type of journey is A2.1=1 then A2.6= A3.7*A3.2/10 for the given journey.

If the type of journey is A2.1=2 then $A2.6 = \sum_{i=1}^{\max \text{GoodsN}} A3.7i * A3.2i / 10$.

If the type of journey is A2.1=3 then A2.6= A3.7*A3.2/20.

If the type of journey is A2.1=4 then A2.6=0.

Detected errors were corrected by insertions of the values of A2.5 calculated by using the above stated formulas.

8. Calculation of values A1.8.1 - total vehicle kilometres during the survey period-loaded and A1.8.2 - total vehicle kilometres during the survey period-empty was performed according to these rules:

¹ A2_1=1-one stop journeys; A2_1=2-two to four stop journeys; A2_1=3-collection/distribution journeys with more then 5 stops; A2_1=4-empty journeys

$$A1.8.1 = \sum_{i=1}^{\max \text{Journ}N} A2.5 \text{ if the type of journey is } A2.1=1 \text{ or } A2.1=2 \text{ or } A2.1=3.$$

A2.5 - distance travelled in one journey, where max JournN is the number of journeys of the vehicle in the observed period.

$$A1.8.2 = \sum_{i=1}^{\max \text{Journ}N} A2.5 \text{ if the type of journey is } A2.1=4.$$

Detected errors were corrected by insertions of the values A1.8.1 and A1.8.2 calculated by using the above stated formulas.

These logical and mathematical checks showed numerous mistakes, in spite of the manual correction of the questionnaires. It is obvious that a lot of errors can escape human reasoning, whether it is because of insufficient knowledge of the staff, or just random mistakes. Also new errors appeared after the process of data entry.

3. FUTURE IMPROVEMENTS IN DATA EDITING PROCESS

1. Logical and mathematical checks should be adjusted in order to be incorporated in the data entry software. This software should automatically alert the person entering the data from the questionnaire that the variable value entered is either invalid or inconsistent with values of other variables already entered.
2. The data entry software which generates the identification of the vehicle should include the registration number of the vehicle in the vehicle identification. The registration number of the vehicle should be checked in the auxiliary database of vehicles from which the sample was drawn. The inclusion of registration number in the identification of the vehicle is essential for connection between the data collected from the field and the database from which the sample was drawn. This way the quality of the database can be accessed, its quality can be improved by updating it with the information from the field. Also there is a possibility to impute some values which were not filled in the questionnaire from the database.

The information about the vehicle working status i.e. response codes of the vehicles should be entered in the software application, for the purpose of calculation of weights in the phase of estimation. This was overlooked in the creation of the application for the pilot survey.

3. Software should be improved with plausibility checks: checks for things which, whilst unlikely, are nevertheless possible (e.g. national journeys over 750km, a weekly total greater than 7000km). Plausibility checks ought to be kept to an absolute minimum. If, on examining a large number of implausibility checks thrown up by the system, staff find that almost all, whilst unlikely, are in fact correct they then will tend to overlook the few that in fact indicate errors.
4. Derived variables, i.e. those which can be calculated from other variables such as tonne-kilometres or distances travelled in the reference period, should be calculated automatically by a computer, thus avoiding errors.

5. The data entry software should be connected with the ILSE application in order to assign appropriate NUTS3 code to places of origin/destination in the European Union, EFTA countries and candidate countries. The Index of Locations for Statistics in Europe (ILSE) can be considered as an interactive search/conversion tool for postcodes, locality names and railway-station codes, mapping these codes and names to NUTS 3 or NUT 2 regions (NUTS: Nomenclature of territorial unit for statistics). The tool can also search for distances between airport-pairs, as well as for distances between NUTS regions. This application can be used for checking the distances traveled between two places in European Union, EFTA countries and candidate countries.

For national transport, that is, distances traveled in Serbia, checks could be performed by incorporating in the software the link to websites which offer route planning and distance calculations such as ViaMichelin. This way the person entering the data can make logical conclusions whether the answers in the questionnaire about the distances traveled are correct.

4. CONCLUSION

From the errors found in the data from the pilot survey we can conclude that the data entry software for this survey needs to include a lot of checks in order to prevent incorrect data to be entered. The number of errors and severity of errors found after the additional logical and mathematical control show that it is necessary to perform these checks in the process of data entry. However, due to the rather complex logic that needs to be performed, a part of the data editing process should be left to human intervention. The complexity of some problems permits only two possible solutions:

To make some simplifying assumptions in data collection and change questionnaire design so that the respondent is not obligated to determine the type of journey. This would loosen the weight on both respondents and staff but it would also cause the loss of some information.

The other solution is to let the complex problem be checked and solved by staff. This, however, is more time and resources consuming and burdens respondents.

REFERENCES

- [1] Road freight transport methodology, Volume 1, Reference manual for the implementation of Council Regulation N° 1172/98 on statistics on the carriage of goods by road, Eurostat

OBRADA PODATAKA U ISTRAŽIVANJU O PREVOZU ROBE DRUMOM

REZIME

Ovaj rad se bavi procesima obrade podataka koji su korišćeni u pilot istraživanju o prevozu robe drumom, koje je urađeno u Srbiji u toku dve nedelje u 2008. godini. Ovim istraživanjem prikupljeni su podaci o radu teretnih drumskih vozila kako u okviru transportnog sektora nacionalne ekonomije tako i van njega. U radu su navedene greške u podacima iz ovog istraživanja i metode koje su korišćene da bi se one ispravile. Cilj ovog rada je da se uradi analiza efikasnosti ovih metoda i da se razmotre novi metodi za sprečavanje grešaka u unosu i korekciju grešaka u podacima sa terena.

Appendix 1

List of variables:

A1.5 - Loading capacity of the vehicle

A1.8.1 –Total vehicle kilometres during the survey period-loaded

A1.8.2- Total vehicle kilometres during the survey period-empty

A1.7- Type of transport

Possible values: 1-public 2-on own account

A2.1- Type of journey

Possible values: 1-one stop journeys 2-two to four stop journeys 3-collection/distribution journeys with more then 5 stops 4-empty journeys

A2.2 -Weight of goods carried in one journey

A2.5- Distance travelled in one journey

A2.6- Tonne-kilometres performed in one journey

A2.10- Degree of loading of the vehicle

Possible values: 0- empty journey, 1-not full, 2-full, 9-missing

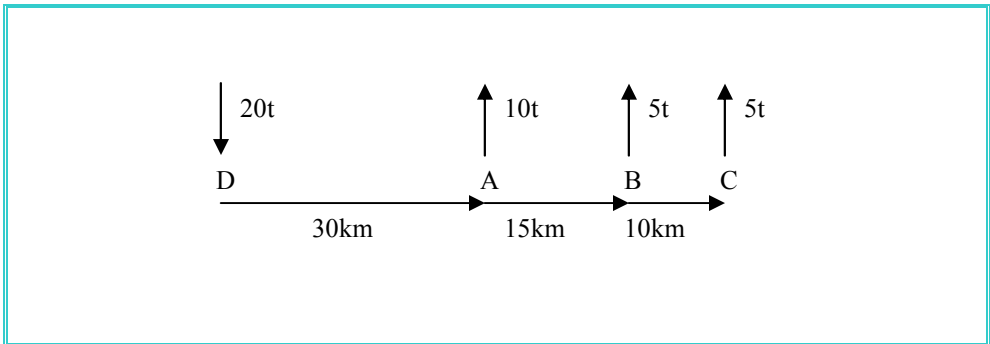
A3.2- Weight of goods carried in one consignment of the journey

A3.7 - Distance travelled by one consignment in the journey

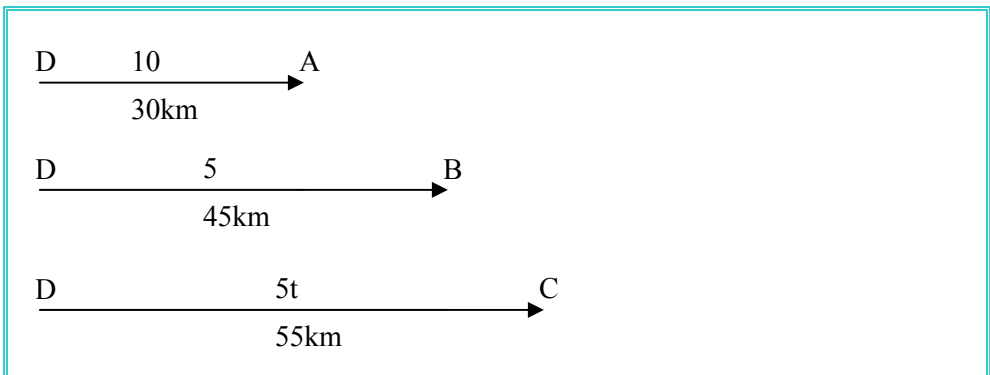
Appendix 2

Restructuring journeys of type 2 (two to four stop journeys) into consignments

A lorry is loaded at the depot (S) with 20 tonnes of goods. It travels 30 kilometres to A where it unloads 10 tonnes, then 15 kilometres to B where it unloads 5 tonnes, then 10 kilometres to C where it unloads the rest (5 tonnes).



The type of journey in this case is A2_1=2 and the journey is restructured by consignments in the following way.



Total tonnes carried on this journey are 20t.

Total distance traveled is 55km.

Total tkm performed are 800 tkm.

$\text{tkm} = 10\text{t} \cdot 30\text{km} + 5\text{t} \cdot 45\text{km} + 5\text{t} \cdot 55\text{km}$.

Jelena Kovijanić
Olga Melovski Trpinac
Željko Lesić

PILOT ISTRAŽIVANJE O PREVOZU ROBE U DRUMSKOM SAOBRAĆAJU U SRBIJI 2008. GODINE

REZIME

Zavod za statistiku Republike Srbije je tokom dve nedelje u 2008. godini (od 23. do 30. marta, i od 31. marta do 6. aprila) sproveo pilot istraživanje o prevozu robe u drumskom saobraćaju. Istraživanjem su prikupljeni podaci o radu teretnih vozila kao što su: težina prevezene robe, pređeni kilometri, realizovani tonski kilometri, mesto polaska/dolaska, vrsta robe itd. Posebno detaljno su prikazani podaci o ukupnom prevozu robe u drumskom saobraćaju sa podelom po načinu organizovanja prevoza na javni drumski saobraćaj i drumski saobraćaj za sopstvene potrebe. Pilot istraživanje je izvršeno u cilju testiranja metodologije koja je rađena prema preporukama Eurostata i koja u potpunosti zadovoljava zahteve za podacima koje je Evropska unija definisala za svoje članice. Takva metodološka usklađenost obezbeđuje uporedivost sa podacima koje Eurostat objavljuje za svoje članice. S obzirom da su podaci dobijeni putem probnog istraživanja, odnosno da je period posmatranja znatno manji od predviđenog metodologijom Eurostata prikazani su samo globalni rezultati na najvišem nivou agregiranja. Poređenje dobijenih rezultata sa ekspertskim procenama dalo je veoma zadovoljavajuće rezultate. To potvrđuje pouzdanost prikazanih podataka na nivou Republike Srbije iz oblasti prevoza robe u drumskom saobraćaju.

UVOD

Drumski prevoz, sa stanovišta dostupnosti korisnicima prevoza, deli se na javni prevoz i prevoz za sopstvene potrebe.

Javni prevoz jeste prevoz koji je pod jednakim uslovima dostupan svim korisnicima prevoznih usluga i obavlja se na osnovu ugovora o prevozu. Prevoz za sopstvene potrebe jeste prevoz koji obavljaju preduzeća, pravna lica, preduzetnici i građani radi zadovoljenja potreba u obavljanju svoje delatnosti.

Pilot istraživanje o prevozu robe u drumskom saobraćaju sprovedeno je u 2008. godini i obuhvata javni i sabračaj za sopstvene potrebe. Drumski saobraćaj za sopstvene potrebe nije bio obuhvaćen redovnim statističkim istraživanjima od 1999. godine, a prema podacima iz ovog pilot istraživanja učestvuje sa 69,2% u ukupnoj količini prevezene robe i sa 36,9% u ukupnim ostvarenim tonskim kilometrima. To ukazuje da se ekonomska kretanja u delatnosti drumskog saobraćaja ne mogu sagledati bez rada saobraćaja za sopstvene potrebe.

U ovom radu, dat je kratak prikaz metodologije pilot istraživanja. Osim toga, prikazani su osnovni rezultati istraživanja: ocene pokazatelja rada u drumskom robnom saobraćaju za 2008. godinu, kao i ocene njihovih grešaka.

METODOLOŠKE OSNOVE

Cilj pilot istraživanja je da se statističko praćenje prevoza robe drumom harmonizuje sa evropskim standardima propisanim u regulativi 1172/98. U istraživanju se prati rad teretnih drumskih vozila, kako u javnom prevozu, tako i u prevozu za sopstvene potrebe. Prikupljaju se sledeći podaci: mesto polaska i mesto dolaska vozila, vrsta prevezene robe, vrsta pakovanja, kategorija opasne robe, stepen iskorišćenja vozila, broj osovina, starost vozila itd.

Jedinice posmatranja su drumska teretna vozila konstruisana za prevoz robe, nosivosti veće od 3,5 tona i tegljači. Nisu uključena poljoprivredna i vojna vozila, i vozila javne administracije, s izuzetkom drumskih vozila u službi javnih servisa kao što su administracija železnice, elektrodistribucija i komunalne službe.

U istraživanju se prati rad vozila u toku 1 nedelje, a istraživanje je sprovedeno u toku 2 nedelje. Izveštajne jedinice su vlasnici vozila, koji daju podatke o radu izabranog vozila.

Okvir za izbor uzorka

Za pripremu okvira za izbor uzorka korišćen je Registar registrovanih drumskih vozila Ministarstva unutrašnjih poslova Republike Srbije (MUP-a). Podaci o registrovanim vozilima su dobijeni posebno za pravna lica i fizička lica.

Baza podataka o vozilima fizičkih lica (preduzetnika), sa stanjem iz marta 2008. sadrži oko 58 hiljada vozila od kojih je oko 23 hiljada nosivosti veće od 3,5 tona ili tegljači. Ova baza sadrži podatke o vozilima u vlasništvu fizičkih lica kao što su: registarska oznaka, vrsta vozila, tip karoserije vozila, nosivost vozila, godina proizvodnje itd. ali ne sadrži podatke o vlasniku vozila. Kako ona ne sadrži podatke o vlasniku, i njegovoj adresi, iz nje nije mogao da se bira uzorak vozila za koji bi se prikupljali podaci. Podaci o broju vozila ove baze korišćeni su u fazi ocenjivanja za korigovanje pondera i izračunavanje ocena parametara za skup svih teretnih vozila nosivosti veće od 3,5 tona ili tegljača.

Baza podataka o vozilima u vlasništvu pravnih lica, koja je dobijena od MUPa je sa stanjem registrovanih vozila na kraju 2006. godine, i sadrži sledeće podatke: matični broj preduzeća, naziv preduzeća, adresa, delatnost preduzeća, registarska oznaka vozila, tip karoserije vozila, nosivost vozila, godina proizvodnje, datum poslednje registracije, broj osovina itd.

Okvir za izbor uzorka čini lista vozila, izdvojenih iz baze podataka vozila u vlasništvu pravnih lica, koja imaju nosivost veću od 3,5 tona ili tegljači. U toku sprovođenja istraživanja, kada su upitnici već bili poslani na teren, utvrđeno je da iz okvira treba isključiti i vozila koja nisu konstruisana za prevoz robe (na osnovu šifre karoserije), kao i poljoprivredna, vojna vozila i vozila javne administracije, s izuzetkom drumskih

vozila u službi javnih servisa kao što su administracija željenice, elektrodistribucija i komunalne službe. Ova vozila su naknadno isključena iz okvira i uzorka.

U cilju dobijanja što pouzdanijih podataka Registar vozila u vlasništvu pravnih lica je povezan sa bazom aktivnih preduzeća, Statističkim poslovnim registrom (SPR), sa stanjem iz marta 2008. Uparivanje je vršeno prema matičnom broju preduzeća kome vozilo pripada. Svi slogovi nisu mogli da se upare zbog razlika u matičnom broju ili zbog toga što je u nekim slogovima u registru MUP-a nedostajao ovaj podatak. Dodatno povezivanje prema imenu preduzeća izvršeno je za velika preduzeća sa velikim brojem registrovanih vozila.

Poređenje šifri delatnosti preduzeća vlasnika vozila iz registra vozila i SPR-a pokazuje da se ovi podaci ne podudaraju uvek. Jedan od razloga može biti taj što delatnost preduzeća iz registra vozila odgovara delatnosti dela preduzeća koji koristi vozilo, dok u SPR-u imamo pretežnu delatnost preduzeća.

U pilot istraživanju korišćena je šifra delatnosti preduzeća iz SPR gde je moguće, a u ostalim slučajevima iz registra vozila.

SPR takođe sadrži podatke o adresi sedišta preduzeća, ali ovi podaci nisu korišćeni direktno već samo kao dodatni kontrolni podaci. Podaci o teritoriji i adresi preuzeti su iz registra vozila, jer je u njemu zabeleženo gde je vozilo registrovano i tu postoji veća šansa da se vozilo nađe.

U tabeli 1. dati su podaci o broju teretnih drumskih vozila u vlasništvu pravnih lica, po klasama nosivosti na kraju 2006. (podaci iz registra MUP-a). Među ovim podacima bilo je 1510 slogova bez informacije o nosivosti vozila. Eliminacijom teretnih vozila koja ne ispunjavaju date kriterijume (nosivost manja od 3,5 tona, vozila koja nemaju upisanu nosivost i vozila sa neodgovarajućom šifrom karoserije ili delatnošću) formiran je okvir za izbor vozila u uzorak. U koloni (3) dat je broj vozila u okviru kod kojih je uspešno izvršeno uparivanje sa SPR-om prema matičnom broju pravnog lica, vlasnika vozila.

Tabela 1. Broj vozila po kategorijama nosivosti u Registru i okviru

	Broj vozila u Registru	Broj vozila u okviru	
		sva vozila u okviru	vozila uparena sa SPR-om
	(1)	(2)	(3)
Ukupno	24993	18861	18405
<3.5 tona	3652		
3.5 – 10 tona	12206	11404	11166
>= 10 tona	5063	4899	4806
Nedostaje nosivost	1510		
Tegljači	2562	2558	2433

Plan uzorka

U pilot istraživanju primenjen je stratifikovan prost slučajni uzorak. Stratumi su definisani međusobnim ukrštanjem klasa nosivosti, delatnosti preduzeća i teritorije.

Prema **nosivosti** vozila su stratifikovana u 3 klase:

- od 3,5 do 10 tona
- 10 i više tona
- tegljači

Nosivost tegljača je jednaka nuli. Ova vozila su namenjena za vuču drugih vozila koja nemaju svoj pogon, prvenstveno prikolica.

Prema **teritoriji** vozila su razvrstana prema tome da li su registrovana u:

- Centralnoj Srbiji bez Beograda
- Beogradu ili
- Vojvodini

Prema **delatnosti** vozila su grupisana u 6 klasa:

- Poljoprivreda i šumarstvo
- Industrija
- Građevinarstvo
- Trgovina i hoteli i restorani
- Saobraćaj
- Ostalo (finasijsko posredovanje, poslovi sa nekretninama, iznajmljivanje, državna uprava i socijalno osiguranje, obrazovanje, zdravstveni i socijalni rad, druge komunalne, društvene i lične usluge).

Podaci o teritoriji preuzeti su iz Registra MUP-a. Podaci o delatnosti preduzeća su prema SPR-u (kod uparenih slogova) ili iz Registra MUP-a (kod neuparenih slogova).

U obe nedelje, u kojima je sprovedeno istraživanje, uzorak od 350 vozila alociran je po stratumima proporcionalno 4. korenu broja vozila u tim stratumima. Razlog za ovakvu alokaciju je namera da se više jedinica alocira u manje stratumima i da se uzorak "raširi" jednako po stratumima. U većini slučajeva, manji stratumima su sa vozilima veće nosivosti. Na ovaj način imaju veću šansu da uđu u uzorak nego proporcionalnom alokacijom. Primenjena alokacija ujedno obebeđuje ocene približno iste tačnosti za domene definisane klasama teritorije, nosivosti i delatnosti.

Izbor uzorka je izvršen korišćenjem permanentnih slučajnih brojeva (PRNUM) u opsegu (0,1] koji su dodati pre izbora svim jedinicama u uzorku. Vozila koja nisu povezana sa SPR-om nisu uzeta u obzir prilikom izbora uzorka. Nakon sortiranja svih jedinica u stratumu prema permanentnim slučajnim brojevima, uzorak željene veličine n_h sastavljen je od prvih n_h jedinica u stratumu.

Permanentni slučajni brojevi omogućavaju jednostavan izbor uzorka i mogućnost da se kontroliše preklapanje uzoraka.

Od 700 izabranih vozila u uzorku je ostalo 678 nakon što je utvrđeno da je iz uzorka treba isključiti vozila koja nisu konstruisana za prevoz robe (na osnovu šifre karoserije), kao i poljoprivredna, vojna vozila i vozila javne administracije, s izuzetkom drumskih vozila u službi javnih servisa kao što su administracija železnice, elektrodistribucija i komunalne službe.

Prikupljanje podataka

Za prikupljanje podataka primenjen je izveštajni metod na bazi uzorka. Izveštajna jedinica je vlasnik vozila koji daje podatke o radu izabranog vozila u toku jedne izabrane nedelje. Izveštajne jedinice su samo pravna lica (preduzeća).

U Tabeli 2. prikazani su podaci koji se odnose na realizaciju uzorka. Broj popunjenih upitnika je oko 47%, najmanje je popunjenih u Beogradu (37%) a najviše u Centralnoj Srbiji (55%). Osnovni razlog nepopunjenih upitnika je neažurnost okvira za izbor uzorka jer je u dosta slučajeva bilo teško uspostaviti kontakt sa preduzećem, vlasnikom (korisnikom) vozila, usled pogrešne adrese. Dodatna pomoćna informacija u spiskovima vozila o adresi preduzeća iz SPR-a (uz adresu iz registra vozila MUP-a) izgleda da nije bila od veće koristi.

Tabela 2. Realizacija uzorka

	Broj vozila				
	planirani uzorak	realizovani uzorak	vozila koja su radila ili nisu radila	otpisana vozila i vozila izvan delokruga istraživanja	neodgovor
	(1)	(2)	(3)	(4)	(5)
Ukupno	678	322	301	21	356
Centralna Srbija	254	143	135	8	111
Beograd	214	80	72	8	134
Vojvodina	210	99	94	5	111

Unos i logička kontrola

Unos podataka je izvršen u Republičkom zavodu za statistiku u programu koji daje izlaz podataka u formatu koji je propisan standardom Eurostata. Struktura i kompleksnost upitnika čiji je unos bio potreban zahteva da program za unos bude takav da obezbedi inteligentan i interaktivan unos sa puno logičke kontrole, što korišćeni program nije obezbedio. Zbog toga je naknadno izvršena sva potrebna logička kontrola i materijal korigovan i pripremljen za dalju obradu. Detaljan opis izvršenih kontrola i korekcija nalazi se u radu Kovijanić (2009).

U narednim istraživanjima program za unos trebalo bi da u većoj meri obezbedi kontrolu podataka. Potrebno je, pre svega, definisati vrste promenljivih i njihov domen (dozvoljene vrednosti), međusobne veze promenljivih (logičke, matematičke i sl.), kao i

statističke ili empirijske metode za identifikovanje podataka koji „iskaču“ iz intervala. Takođe je neophodno da program dozvoljava popunjavanje određenih polja u zavisnosti od prethodno unetih vrednosti, kao i da se sve izvedene vrednosti, kao što su npr. tonski kilometri, računaju automatski (kompjuterski). Dalje je potrebno razviti, u meri u kojoj je to moguće, procedure za automatsku imputaciju ili korekciju podataka. Takođe program bi trebalo da se poveže sa ILSE (Index of Locations for Statistics in Europe) softverom koji će omogućiti šifriranje mesta u EU prema NUTS klasifikaciji i proveru pređenih rastojanja unutar EU.

Ocena parametara

Ocene totala, količnika i proporcija izvedene su po standardnom postupku za stratifikovan prost slučajaj uzorak. Tako se ocene totala izračunavaju sumiranjem ponderisanih vrednosti iz uzorka, pri čemu je ponder za jedinice istog stratuma jednak količniku broja jedinica u stratumu i broja jedinica u uzorku. Ocenjivanje je izvedeno uz sledeće specifičnosti:

1. Nije neophodno da se ponderi izračunavaju za svaki vremenski period (nedelju) istraživanja. Ponderi su izračunati na osnovu podataka za obe nedelje u pilot istraživanju.
2. U ponder se uključuje vremenska komponenta tako što se, u zavisnosti od perioda za koji se traži ocena, ukupan broj vozila u stratumu pomnoži sa brojem relevantnih perioda istraživanja (sa 2 za dve nedelje, sa 13 za kvartal, i sa 52 ako želimo ocenu za celu godinu).
3. Za ocenu rada ukupnog voznog parka (vozila u vlasništvu pravnih i fizičkih lica) iz domena istraživanja, ponderi su korigovani za broj vozila u vlasništvu fizičkih lica (koja nisu mogala da budu birana u uzorak).
4. Implicitno je pretpostavljeno da jedinice za koje nismo dobili odgovor imaju iste karakteristike kao jedinice za koje smo dobili odgovor.

Ponder koji se koristi za podatke o vozilima i putovanjima, za izabrana vozila u svakom stratumu je izračunat na sledeći način:

$$K \cdot \frac{N}{S + S'}$$

Gde je :

N ukupan broj vozila u stratumu.

S+S' je broj vozila u realizovanom uzorku pri čemu je **S** broj upitnika koji se mogu iskoristiti (upitnici koji prijavljuju da je vozilo radilo ili nije radilo u toku posmatrane nedelje), a **S'** je broj sa informacijom da je vozilo otpisano ili izvan delokruga istraživanja (npr. vozilo nosivosti ispod 3,5 tona, specijalno vozilo koje nije konstruisano za prevoz robe, npr. dizalica, iako su prema bazi MUP-a pripadala delokrugu istraživanja).

K je vremenski period za koji su ocene potrebne, K=2 (za 2 nedelje), K=13 (za kvartal) ili K=52 (za godinu).

Ukupan broj vozila **N** (tabela 3) čine sva vozila u vlasništvu pravnih i fizičkih lica koja ulaze u delokrug istraživanja¹, a ne samo vozila iz okvira za izbor uzorka (dela baze MUPa koji je po matičnom broju pravnih lica vlasnika vozila uparen sa SPR-om). Na taj način ocenjuje se ukupan rad svih vozila, uz pretpostavku da je rad vozila koja nisu mogla da budu izabrana u uzorak jednak radu vozila iz okvira za izbor uzorka.

Da bi se broj vozila u vlasništvu fizičkih lica rasporedio po stratumima uvedeno je uprošćavanje da svaki preduzetnik poseduje jedno vozilo. U dobijenoj bazi vozila postojao je podatak o teritoriji i o nosivosti vozila, ali ne i o delatnosti kojom se bavi vlasnik vozila. Unutar svake od tri teritorijalne oblasti (centralna Srbija, Beograd i Vojvodina), vozila svake od tri klase nosivosti (3,5-10 tona, >=10 tona i tegljači) raspoređena su po klasama delatnosti prema strukturi aktivnih preduzetnika iz SPR. Na taj način pretpostavljeno je da se vozila preduzetnika svake kategorije nosivosti raspoređuju po delatnosti kao aktivni preduzetnici u SPR-u, stanje mart 2008.

Tabela 3. Broj vozila u vlasništvu pravnih i fizičkih lica koja ulaze u delokrug istraživanja

	Ukupno	Vozila pravnih lica	Vozila fizičkih lica
	(1)	(2)	(3)
Ukupno	41691	18861	22830
3.5 – 10 tona	28747	11404	17343
>= 10 tona	9297	4899	4398
Tegljači	3647	2558	1089

Razlog uključivanja vozila **S'** je taj što tone i tonski kilometri mogu da budu precenjeni ako se svi upitnici koji sadrže pozitivnu informaciju o statusu vozila ne uključe. Pretpostavlja se da proporcija ovih vozila, nađenih prilikom istraživanja, odgovara proporciji ovih vozila u okviru. U ostalim slučajevima, kada je vozilo prodato ili izdato, preduzeće zatvoreno ili upitnik nije dostavljen, nije moguće koristiti ove upitnike jer ne daju informaciju o tome da li vozilo još uvek postoji, da li se koristi ili ne.

OSNOVNI REZULTATI

Određene su Horvitz-Thompson ocene za totale promenljivih:

- Prevezene tone
- Pređeni kilometri vozila
- Pređeni tonski kilometri

¹ Jedinice posmatranja su drumska teretna vozila konstruisana za prevoz robe, nosivosti veće od 3,5 tona i tegljači. Nisu uključena poljoprivredna i vojna vozila, i vozila javne administracije, s izuzetkom drumskih vozila u službi javnih servisa kao što su administracija železnice, elektrodistribucija i komunalne službe.

i ocena predenog puta 1 tone robe (kao količnik ocene tonskih kilometara i prevezene robe) kako ukupno tako i za javni i saobraćaj za sopstvene potrebe. Osim na nivou cele Srbije ove ocene su određene i za domene definisane klasama teritorije, nosivosti i delatnosti. Dobijeni rezultati su prikazani u tabelama 4, 5 i 6.

Tabela 4. Prevoz robe u drumskom saobraćaju po teritoriji, privrednim delatnostima i nosivosti vozila, 2008.

- Ukupan drumski saobraćaj-

Po teritoriji									
	broj vozila pravnih jedinica i preduzetnika	Ocene parametara				Koefficienti varijacije			
		predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
Ukupno	41691	1537399	108436	9514909	88	11,16	12,23	15,46	16,76
Centralna Srbija	22422	923902	56037	7310460	130	16,10	17,03	19,30	21,96
Beograd	6746	213786	17046	591214	35	22,15	26,65	21,07	26,58
Vojvodina	12523	399711	35353	1613235	46	17,84	22,63	24,63	23,31

Po privrednim delatnostima									
	broj vozila pravnih jedinica i preduzetnika	Ocene parametara				Koefficienti varijacije			
		predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
Poljoprivreda i šumarstvo	1171	53890	2647	232148	88	37,55	22,81	39,23	30,35
Industrija	9792	393941	24141	1226375	51	17,86	31,50	21,33	34,43
Gradevinarstvo	5200	160450	23131	711160	31	15,54	18,09	19,98	20,93
Trgovina	14415	400587	26638	3253947	122	26,91	28,05	33,70	33,73
Saobraćaj	5903	294170	17153	3683589	215	23,89	31,82	25,11	34,97
Ostalo	5210	234361	14727	407689	28	35,47	25,77	24,37	26,24

Po nosivosti vozila									
	broj vozila pravnih jedinica i preduzetnika	Ocene parametara				Koefficienti varijacije			
		predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
3.5-10 tona	28747	1026036	44748	5096912	114	15,81	18,20	25,86	27,43
>10 tona	9297	323436	51708	1823777	35	15,06	19,66	23,11	20,84
Tegljači	3647	187926	11979	2594220	217	14,73	20,45	19,26	22,50

Tabela 5. Prevoz robe u drumskom saobraćaju po teritoriji, privrednim delatnostima i nosivosti vozila, 2008.

- Javni drumski saobraćaj -

Po teritorij								
	Ocene parametara				Koeficijenti varijacije			
	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
Ukupno	459698	33357	6007821	180	22,82	20,05	24,14	22,38
Centralna Srbija	372082	23731	5026406	212	27,44	25,97	27,94	27,36
Beograd	36011	2930	233532	80	36,81	41,39	43,02	25,30
Vojvodina	51605	6696	747883	112	38,93	34,32	46,38	34,61

Po privrednim delatnostima								
	Ocene parametara				Koeficijenti varijacije			
	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
Poljoprivreda i šumarstvo	8352	1238	125548	101	44,49	22,81	48,55	30,44
Industrija	79335	3971	518493	131	35,86	31,5	46,70	54,81
Gradevinarstvo	26344	6539	380410	58	28,28	18,09	33,71	37,21
Trgovina	159577	8245	2007320	243	47,31	28,05	55,73	27,14
Saobraćaj	179112	11741	2887260	246	37,05	31,82	30,38	43,97
Ostalo	6977	1623	88790	55	63,09	25,77	64,55	79,48

Po nosivosti vozila								
	Ocene parametara				Koeficijenti varijacije			
	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
3.5-10 tona	339596	15185	3684552	243	29,87	28,53	36,72	27,60
>10 tona	53584	15368	913509	59	40,94	32,78	42,22	41,50
Tegljači	66517	2804	1409759	503	22,68	27,07	24,90	26,17

Tabela 6. Prevoz robe u drumskom saobraćaju po teritoriji, privrednim delatnostima i nosivosti vozila, 2008.

- Drumski saobraćaj za sopstvene potrebe -

Po teritorij								
	Ocene parametara				Koeficijenti varijacije			
	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
Ukupno	329956	75079	3507088	47	14,14	16,87	19,90	20,45
Centralna Srbija	172711	32307	2284054	71	21,33	27,05	28,27	32,60
Beograd	49148	14116	357682	25	23,53	32,05	22,77	29,83
Vojvodina	108097	28657	865352	30	24,21	27,83	29,06	18,99

Po privrednim delatnostima								
	Ocene parametara				Koeficijenti varijacije			
	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
Poljoprivreda i šumarstvo	10578	1409	106601	76	43,09	36,74	73,92	59,48
Industrija	138171	20170	707883	35	22,94	36,90	21,39	36,38
Gradevinarstvo	29533	16592	330749	20	21,96	25,73	27,47	17,38
Trgovina	78160	18393	1246627	68	30,34	39,83	35,52	35,83
Saobraćaj	38199	5412	796329	147	57,35	79,92	61,88	81,04
Ostalo	35314	13104	318900	24	22,92	28,70	32,34	31,04

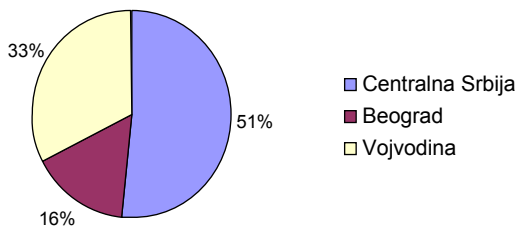
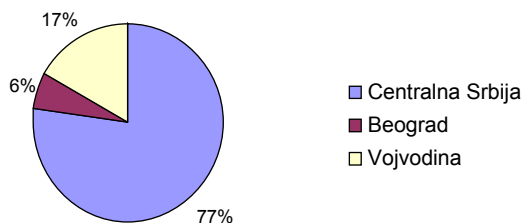
Po nosivosti vozila								
	Ocene parametara				Koeficijenti varijacije			
	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km	predeni km teretnih vozila, hilj.	prevezena roba, hilj. tona	tkm, hilj.	prosečan predeni put jedne tone robe, km
3,5-10 tona	205328	29563	1412359	48	18,95	25,66	30,96	36,14
>10 tona	67785	36340	910268	25	24,49	27,09	31,83	18,32
Tegljači	56843	9175	1184461	129	34,57	26,52	38,85	36,75

Ocene parametara drumskog saobraćaja za 2008. dobijene su samo na osnovu uzorka od 678 vozila, od kojih je u realizovanom uzorku 47%, pri čemu su podaci prikupljeni u toku 2 nedelje. Zbog toga su prikazane ocene samo za najšire klasifikacije teritorije, nosivosti i delatnosti. Ocene koeficijenta varijacije izračunatih ocena nisu manji od 10%. Ipak, osim za ocene javnog i saobraćaja za sopstvene potrebe po klasama delatnosti, koeficijenti varijacije su manji od 50% i podaci se mogu prikazati, uz uzimanje u obzir njihove nevelike preciznosti.

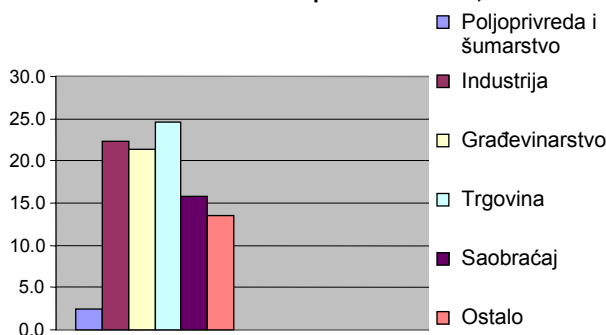
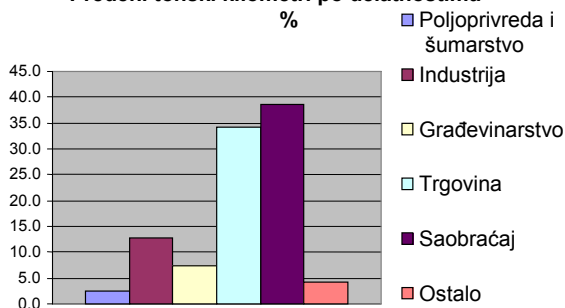
Najveća količina robe prevezena je vozilima registrovanim u centralnoj Srbiji, 51,7%, zatim u Vojvodini 32,6%, dok je učešće Beograda u ukupno prevezenoj količini robe 15,7%.

Što se tiče pređenih kilometara, teretna vozila registrovana u centralnoj Srbiji još više prednjače sa učešćem od 60,1%, dok kilometraža koju su prešla vozila u Beogradu iznosi samo 13,9%.

Ostvareni tonski kilometri vozila registrovanih u centralnoj Srbiji čine 76,8 % od ukupih tonskih kilometara, vozila registrovanih u Vojvodini 17%, a beogradskih vozila 6,2%.

Prevezena količina robe, po teritoriji, %**Prevezeni tonski kilometri, po teritoriji %**

Najveća količina robe prevezena je u delatnostima trgovine i industrije, 24,6% i 22,3% od ukupne količine, dok je najveći deo tonskih kilometara realizovan u delatnosti saobraćaja 38,7%.

Prevezena količina robe po delatnostima, %**Pređeni tonksi kilometri po delatnostima**

Količina prevezene robe u saobraćaju za sopstvene potrebe čini 69,2% od ukupne količine prevezene robe. Kako je prevoz za sopstvene potrebe najviše zastupljen u delatnosti industrije, građevinarstva i trgovine, to obzirom na organizaciju prevoza u tim delatnostima, koji je orijentisan na kraće relacije sa većim brojem obrta, učešće ove vrste prevoza u ukupno realizovanim tonskim kilometrima iznosi 36,9%.

Prosečan pređeni put jedne tone robe iznosi 88 kilometara. Najveći prosečan pređeni put jedne tone robe prešla su vozila koja su registrovana na teritoriji centralne Srbije, 130 kilometara, a najmanji vozila registrovana na teritoriji Beograda, 35 kilometara. Prosečan pređeni put jedne tone robe u javnom saobraćaju je značajno veći od prosečnog puta u saobraćaju za sopstvene potrebe.

ZAKLJUČNE NAPOMENE

Republički zavod za statistiku Srbije planira uvođenje istraživanja o prevozu robe drumom prema metodologiji Evropske unije u jednoj od narednih godina. Ovo istraživanje treba da se sprovodi tokom cele godine, da za svaku nedelju postoji poduzorak vozila za koje se prikupljaju podaci o njihovom jednonedeljnom radu. Rezultati istraživanja bi se objavljivali kvartalno, a kada Srbija bude u obavezi prema EU, prikupljeni podaci bi se u istoj periodici, prema definisanom zahtevu, dostavljali Eurostatu.

S obzirom na složenost istraživanja po sadržaju, načinu sprovođenja i obradi, iskustva sprovedenog Pilot istraživanja o prevozu robe drumom su veoma značajna za pripremu i organizaciju ovog istraživanja.

Jedan od ključnih faktora za uspešno sprovođenje istraživanja je da se za svaku godinu u kojoj se sprovodi istraživanje obezbedi ažurna baza vozila od Ministarstva unutrašnjih poslova. Baza podatke o vozilima koje smo dobili za Pilot istraživanje, koje je sprovedeno u 2008, za vozila u vlasništvu pravnih lica nije bila ažurna, jer je bila sa stanjem kao na kraju 2006. Takođe, za neka vozila nisu postojali svi potrebni podaci (npr. o nosivosti vozila ili matični broj pravnog lica), postojao je i problem prevođenja klasifikacija delatnosti i teritorije jer MUP koristi klasifikacije različite od statističkih. Za vozila u vlasništvu fizičkih lica (preduzetnika) nisu bili na raspolaganju podaci o vlasniku i adresi tako da nisu ni mogli da budu birani u uzorak.

MUP priprema novu bazu vozila koja bi trebalo da bude boljeg kvaliteta. Veoma bi bilo važno da se RZS i MUP sporazumeju o obezbeđivanju potrebnih podataka za istraživanje o prevozu robe drumom. U slučaju da za bazu vozila u vlasništvu fizičkih lica nije moguće dobiti podatke o imenu i adresi vlasnika u celini, dobro bi bilo kada bi potrebni podaci bili raspoloživi bar za vozila izabrana u uzorak.

Na osnovu pilot istraživanja utvrđeno je i:

- da je potrebno izvršiti izmene obrazca za prikupljanje podataka o radu vozila kako bi se obezbedilo jednostavnije i tačnije prikupljanje podatka. Korišćeni obrazac, pripremljen po uzoru na obrazac korišćen u Francuskoj, pokazao se suviše složenim za popunjavanje u našim uslovima;
- da je za vozila za koja se ne prikupe podaci potrebno na terenu utvrditi razlog neobezbeđivanja podataka, i prema definisanim šiframa, upisati ga na spisak vozila izabranih u uzorak. Najviše problema u prikupljanju podatka bilo je u Beogradu s obzirom na mali broj ljudi trenutno angažovanih u Zavodu za informatiku i statistiku Beograd, pa je potrebno razmotriti kako da se prevaziđu mogući problemi u prikupljanju podataka;
- da bi bilo korisno da se podatak o vrsti vozila (kamion/tegljač) prikupi u narednim istraživanjima i na taj način proveriti kvalitet informacije iz baze o vrsti vozila;
- da je u realizovanom uzorku svega 47% od ukupno planiranog broja vozila. Osnovni razlog slabe realizacije je neažurnost baze podataka o vozilima;

- da program za unos podataka treba da sadrži mnogo više logičkih kontrola podataka nego program koji je korišćen u pilot istraživanju;
- da treba razmotriti promenu plana uzorka kako bi se smanjio broj stratuma i na taj način uprostila i ubrzala procedura ocenjivanja traženih parametara o radu vozila, kada istraživanje bude uvedeno kao redovno;
- da je za redovno istraživanje o radu vozila u drumskom saobraćaju potrebno imati veoma dobru organizaciju rada po fazama sa vremenom trajanja svake faze rada kako bi se rezultati obezbedili u predviđenim rokovima.

Uprkos tome što je istraživanje sprovedeno samo za dve nedelje i što je mala realizacija uzorka, bilo je moguće oceniti parametre rada vozila za najšire klasifikacije teritorije, delatnosti i nosivosti vozila i to, za saobraćaj ukupno kao i za javni i sopstveni saobraćaj sa uzoračkim greškama koja se u većini slučajeva kreću od 10%- 50%. Dobijeni rezultati su u skladu sa prethodnim ekspertskim procenama. Osim toga, pokazuju da je učešće sopstvenog saobraćaja značajno i u prevezenoj robi (69,2%) i u tonskim kilometrima (36,9%) i da ga je neophodno obuhvatiti statističkim istraživanjima. Rad saobraćaja za sopstvene potrebe zvanična statistika nije pratila od 1999. godine.

Pilot istraživanje sprovedeno je uz ekspertsku pomoć INSEE, francuske statističke agencije. Zahvaljujemo se kolegama na nizu korisnih saveta i na primedbama koje su doprinele kvalitetu rezultata sprovedenog istraživanja. Istovremeno su nam date i preporuke o metodološkim izmenama koje bi trebalo da se uvedu u cilju poboljšanja rezultata narednih istraživanja.

LITERATURA

- [1] Kovijanić, J. (2009). Data editing in the survey on road freight transport in Serbia, Participants' Contribution Papers, Fourth Balkan Summer School on Survey Methodology, Durres, Albania
- [2] Road freight transport methodology Reference Manual for the implementation of Council Regulation No 1172/98 on statistics on the carriage of goods by road, Eurostat Methodologies and Working papers, 2008.

PILOT SURVEY ON ROAD FREIGHT TRANSPORT IN SERBIA, 2008

ABSTRACT

Statistical Office of the Republic of Serbia, during the two weeks in 2008 (from March 23rd till March 30th, and from March 31st till April 6th), conducted a pilot survey on road freight transport. In this survey data about the work of road freight vehicles were collected, such as: weight of transported goods, distance traveled, tonne kilometers performed, place of departure/arrival, type of goods etc. Data were given for transport of goods in road transport, divided by type of journey (journey in public transport and journey in transport on own account). This pilot survey was performed in order to test the methodology that was made according to the recommendations of Eurostat, and which satisfies fully the requests for data which the European Union had defined for its Member States. Such methodological harmonization provides comparability with the data that Eurostat publishes about its Member States. Considering that data were provided from the pilot survey i.e. that the period of observation is significantly shorter than the one planned in the methodology of Eurostat, only global results on highest levels of aggregation were given. The comparison of these results with expert assessments gave very satisfying results. This confirms the reliability of the given data on road freight transport, on the level of the Republic of Serbia.

PRILOG

Tabela 1. Broj teretnih vozila u osnovnom skupu po teritoriji

	Broj vozila pravnih jedi- nica i pre- duzetnika	Broj vozila pravnih jedinica	Broj vozila u planiranom uzorku	Broj vozila u realizo- vanom uzorku	Broj vozila za koje je dobijen odgovor	Broj vozila izvan domena istraživanja	Neodgovor
Ukupno	41691	18861	678	322	301	21	356
Centralna Srbija	22422	8987	254	143	135	8	111
Beograd	6746	4008	214	80	72	8	134
Vojvodina	12523	5866	210	99	94	5	111

Tabela 2. Broj teretnih vozila u osnovnom skupu po nosivosti

	Broj vozila pravnih jedi- nica i pre- duzetnika	Broj vozila pravnih jedinica	Broj vozila u planiranom uzorku	Broj vozila u realizo- vanom uzorku	Broj vozila za koje je dobijen odgovor	Broj vozila izvan domena istraživanja	Neodgovor
Ukupno	41691	18861	678	322	301	21	357
3,5 – 10,0 tona	28747	11404	281	133	120	13	148
Više od 10,0 tegljači	9297	4899	225	116	109	7	109
	3647	2558	172	73	72	1	99

Tabela 3. Broj teretnih vozila u osnovnom skupu po delatnosti

	Broj vozila pravnih jedi- nica i pre- duzetnika	Broj vozila pravnih jedinica	Broj vozila u planiranom uzorku	Broj vozila u realizo- vanom uzorku	Broj vozila za koje je dobijen odgovor	Broj vozila izvan domena istraživanja	Neodgovor
Ukupno	41691	18861	678	322	300	21	357
Poljoprivreda i šumarstvo	1171	972	82	52	50	2	30
Industrija	9792	6067	143	60	56	4	83
Gradevinarstvo	5200	3382	121	73	72	1	48
Trgovina	14415	3831	123	38	36	2	85
Saobraćaj	5903	3360	125	58	50	8	67
Ostalo	5210	1249	84	41	37	4	43

OSVRT

*Nikola Dobrić
Alpen-Adria Univerzitet, Klauingerfut*

OSVRT

SAVREMENE TENDENCIJE U LINGVOSTATISTICI¹

1. UVOD

Primena kvantitativnih metoda analize u lingvistici sve više je prisutna u raznim oblicima istraživanja. U ovom osvrtu imaćemo u vidu pre svega korpusni pristup lingvistici koji se oslanja na statističke metode i njihovu primenu unutar ovih velikih baza teksta.

Razvoj velikih računarskih baza teksta postala je rasprostranjena pojava u novijoj lingvističkoj praksi. Korpusi, kako se te baze nazivaju, poslednje dve-tri decenije zaokupljaju pažnju mnogih lingvista. Stvaranje korpusa i sakupljanje uzoraka govornog i pisanog jezika praksa je mnogih univerziteta, naučnih instituta, kompanija, ali i privatnih lica. Ponegde su to projekti od posebnog, nacionalnog značaja podržani na visokom državnom nivou.

Prvi računarski korpusi nastali su još šezdesetih godina prošlog veka. Konstrukcija takvih korpusa bila je u to vreme u znatnoj meri tehnički uslovljena, a nagli razvoj koji je posle toga usledio bio je snažno podstican unapređenjem tehnoloških rešenja za obradu i skladištenje teksta. Danas, četrdeset godina posle prvog Brownovog korpusa engleskog jezika, imamo stotine takvih korpusa, najčešće lako dostupnih, ali verovatno i hiljade zatvorenih i privatnih. U pitanju su korpusi raznih jezika i raznih vrsta. Najrasprostranjeniji su korpusi engleskog jezika, ali već se stvaraju i veliki korpusi drugih jezika.

Korpusi opšte vrste i dobre reprezentativnosti čuvaju jedan jezik u njegovoj stvarnoj upotrebi, u raznolikim oblicima i varijetetima u kojim se javlja. Sasvim je jasno koliko je konstrukcija jednog takvog korpusa posebno važna za relativno nerasprostranjen jezik sa malim brojem govornika čiji mnogobrojni dijalekti i sociolekti nestaju gotovo svakodnevno, a čija se semantička i sintaksička struktura nepovratno menja pod globalnim uticajem engleskog jezika.

¹ Autor je u Statističkom društvu Srbije – Klubu statističara Vojvodine 16.04.2009. godine održao predavanje na temu: *Korpusna statistika u lingvistici: Mogućnosti primene u izučavanju jezika.*

2. KORPUS U LINGVISTICI

Ideja korpusa ima svoje važnije početke još davne 1755. godine. Pri konstrukciji jednog od prvih rečnika engleskog jezika (Samuel Johnson, Dictionary of the English Language) sakupljen je korpus zavidnog obima. Sastojao se od zabeleženih razgovora Johnsonovih saradnika, kao i kolekcije književnih dela koje je on smatrao vrednim. Iz ovog korpusa uzimani su relevantni primeri koji su služili za objašnjenje značenja reči.

Iako po današnjim standardima ovakav korpus još ne bi posedovao sve neophodne karakteristike, on ipak predstavlja realizaciju osnovne ideja stvaranja i primene ovakvih skupova jezika. Ukazano je na značaj „hvatanja” jezika u njegovoj stvarnoj upotrebi, pri komunikaciji. (I danas je još uvek, kao i u Johnsonovo doba, jedna od glavnih upotreba velikih korpusa – priređivanje rečnika.)

Korpus u svom osnovnom značenju predstavlja skup teksta, bilo pisanog ili govornog jezika. Dakako, postoji određena distinkcija unutar samog termina. U tradicionalnoj lingvističkoj praksi, korpusom se smatra skup delova jezika koji se odabiraju po precizno određenim lingvističkim kriterijumima. Za razliku od korpusa, *zbirka tekstova* je skup ili arhiva tekstova koji ne moraju biti podvrgnuti selekciji i ne moraju poštovati određene zahteve niti lingvističke kriterijume. U modernoj lingvistici, kao i u daljem tekstu, termin *korpus* prvenstveno se odnosi na *računarski korpus*. On je kodiran i standardizovan, optimizovan za pretragu i analizu, i nalazi se u digitalnom obliku u velikim računarskim bazama. Takav korpus se obično sastoji od više miliona reči iz različitih jezičkih i društvenih izvora. On idealno obuhvata sve moguće pojave jednog jezika, njegovu upotrebu u realnim životnim situacijama, jezik zabeležen u datom vremenskom rasponu i pretočen u digitalni tekstualni oblik.

Unošenje dodatnih lingvističkih podataka u tekstove naziva se kodiranje. Po tome se korpusi razlikuju od ostalih, običnih zbirki tekstova. Ono se u osnovi sastoji iz dve grupe vrlo složenih i mukotrpnih procedura. Prva je *etiketiranje* i odnosi se na komplikovani proces dodavanja dodatnih informacija u korpus. Informacije se mogu odnositi na obeležavanje reči po gramatičkoj kategoriji, rodu, broju, morfološkim i fonološkim karakteristikama, itd. *Parsiranje* je sledeća faza kojom se morfosintaksičke kategorije u tekstu dovode u međusobne sintaksičke odnose višeg stepena. Parsiranjem se određuje sintaksička struktura rečenice. Retki su korpusi koji poseduju ovako poželjan nivo kodiranja. Ovakve procedure zapravo i čine korpus veoma pogodnim u analizi jezika i društva jer dozvoljavaju razne vrste pretraga.

2.1 Vrste korpusa

Postoje korpusi različitih veličina i strukture, kao i različitih namena kada su u pitanju vrste lingvističke analize. Dve su osnovne vrste računarskih korpusa: *opšti* i *specijalni*. Specijalni korpusi po svojoj konstrukciji predstavljaju samo određeni varijetet jezika dok su opšti korpusi mnogo zanimljiviji i podesniji za istraživanje jer odgovaraju široj paleti lingvističkih tema budući da njihov dizajn predstavlja celokupnu formu jednog jezika.

Najvredniji korpusi opšteg tipa su takozvani *monitoring* korpusi. To su korpusi čija se reprezentativnost održava stalnim dodavanjem novih delova jezika i stalnim proširivanjem palete varijeteta u njima. Još jedna vrsta, odnosno podvrsta specijalnih korpusa, su i *paralelni višejezični korpusi*, sastavljeni od uparenih tekstova koji potiču iz dva ili više jezika. Oni se koriste uglavnom u komparativnim i kontrastivnim analizama, kao i u teoriji prevođenja. Takođe se mora imati na umu i *sinhrona* ili *dijahrona* dimenzija jednog korpusa. Dijahroni korpusi proširuju dimenziju analize uključanjem istorijskog razvoja teksta. Dijahroni korpusi, bilo specijalni ili opšti, sačinjeni su od tekstova posebno biranih tako da predstavljaju jedan jezik imajući u vidu njegove promene tokom različitih istorijskih epoha. Neki od njih obuhvataju i veoma duge vremenske raspone, čak i od nekoliko vekova.

2.2 Upotreba korpusa

Jezik je neraskidivo vezan za društvo, društvene aktivnosti i društvene promene. Praćenjem i otkrivanjem promena u jeziku moguće je posredno steći sliku i o društvenim promenama. Skoro da je zbivanje u svakoj sferi društva moguće pored ostalog analizirati i preko jezika. Time se otvaraju raznovrsna pitanja same analize jezika. Tu je važno pitanje reprezentativnosti: gde, kako i na koji način prepoznati, sakupiti i obuhvatiti jezik koji je pravi odraz celog društva, i koji bi samim tim bio pogodan za istraživanje i formulisanje jezičkih zakonitosti. Rešenje je pronađeno u stvaranju velikih računarskih korpusa. Oni omogućuju analizu jezika koja prerasta granice omeđene strogim lingvističkim zahtevima. Korpusi su postali izvor informacija za mnoge naučne discipline, poput leksikologije i drugih lingvističkih disciplina, sociologije, psihologije, ekonomije, politike.

3. OSVRT NA NEKE KORPUSE

Postoje stotine različitih računarskih korpusa pojedinih svetskih jezika, većinom dostupnih preko interneta, bar kada su u pitanju osnovne informacije o njima. Jedan od prvih korpusa novije generacije svakako je *Britanski nacionalni korpus* (British National Corpus) na adresi <http://www.natcorp.ox.ac.uk/>. Možda je on još uvek i najbolji primer monitoring korpusa opšteg tipa, To je veliki korpus koji se sastoji od preko 100 miliona reči iz pisanih i govornih izvora. Izvori za 75% pisanog jezika u njemu uglavnom su informativni tekstovi iz oblasti nauke, religije, ekonomije, filozofije, umetnosti i medija, dok je 25% odvojeno za tekstove preuzete iz raznih književnih dela. Usmeni jezik je zastupljen sa oko 10 miliona reči i sastavljen je od zapisa spontanih i skriptovanih razgovora, javnih govora i usmenog jezika u medijima. Korpus je etiketiran za vrste reči. Većina frekvencijskih pretraga se ne naplaćuje.

Već je pomenuto da danas u svetu postoje na mnogim mestima izgrađeni veliki korpusi jezika. Ovde ćemo, navodeći nešto više detalja o njima, pomenuti sledeće korpusne:

- Korpus savremenog američkog engleskog jezika (Corpus of Contemporary American English): monitoring korpus opšteg tipa sa 360 miliona reči na adresi <http://www.americancorpus.org/>,
- Korpus australijskog engleskog jezika (Australian Corpus of English): monitoring korpus opšteg tipa sa 1 milionom reči na adresi <http://khnt.hit.uib.no/icame>,
- Kembridžov međunarodni korpus (Cambridge International Corpus): višejezični korpus specijalnog tipa sa 275 miliona reči na adresi <http://www.cambridge.org/elt/corpus>,
- Ruski nacionalni korpus (Russian National Corpus): monitoring korpus opšteg tipa sa 150 miliona reči na adresi <http://www.ruscorpora.ru/en/index.html>,
- Nacionalni korpus hrvatskog jezika: monitoring korpus opšteg tipa sa 30 miliona reči na adresi <http://www.hnk.ffzg.hr/>,
- Korpus savremenog srpskog jezika: korpus opšteg tipa sa 24 miliona reči na adresi <http://korpus.matf.bg.ac.yu/prezentacija/korpus.html>,
- Korpus srpskog jezika: korpus opšteg tipa sa 12 miliona reči na adresi <http://www.serbian-corpus.edu.rs/indexns.htm>.

Ovaj popis bi mogao da se proširi i navođenjem drugih poznatih korpusa kao što su: Korpus starogrčkog jezika (Thesaurus Linguae Graecae), Francuska međujezička baza (French Interlanguage Database), Lankasterski korpus mandarinskog kineskog (Lancaster Corpus of Mandarin Chinese), Banka engleskog jezika (Bank of English), Korpus dečjeg jezika (Child Language Data Exchange System), Wolverhempton korpus engleskog poslovnog jezika (Wolverhampton Business English Corpus), Korpus srednjevekovnog engleskog stiha i proze (Corpus of Middle English Prose and Verse).

4. MONITORING KORPUS SRPSKOG JEZIKA

Postoje dva opšta računarska korpusa srpskog jezika o kojima su javno saopšteni osnovni podaci. Jedan od njih je pomenuti *Korpus srpskog jezika*. On sadrži 11 miliona reči. Ovaj korpus poseduje nekoliko odličnih karakteristika. To je pre svega njegova, za jedan korpus poželjna, dijahrona dimenzija, jer se sastoji od tekstova čije vreme nastanka pokriva široku vremensku skalu: od 12. veka pa do modernog doba. Ručno je etiketiran za gramatički status reči, broj grafema, slogova, za fonološku strukturu i za početak i kraj rečenica. Ovakav nivo etiketiranja je veoma redak pa je ovaj korpus pogodan za razne vrste istraživanja jezika. Nedostaci se tiču najvažnije karakteristike korpusa, reprezentativnosti, koju ovaj korpus ne poseduje u dovoljnoj meri. Naime, nepostojanje govornih izvora u korpusu i nedovoljna dimenzija monitoringa, velike su mane. Takođe, Korpus nije dopunjen novim varijetetima jezika u dovoljnom obimu. Ovaj korpus je nemoguće koristiti bez posebne dozvole konstruktora. Važno je napomenuti da se među prvim osnivačima korpusnog pristupa izučavanju jezika u Srbiji i pionirima konstrukcije korpusa srpskog jezika javio dr Đorđe Kostić. On je vizionarski počeo sa konstrukcijom ovog korpusa još 1957. godine.

Korpus savremenog srpskog jezika sada ima oko 24 miliona reči, takođe iz isključivo pisanih izvora. Većina korpusa je neetiketirana a izvori nisu dovoljno reprezentativni. Iako predstavlja zavidno postignuće, još uvek se mora smatrati nedovršenim. To treba da se uzme u obzir kako bi ispravno tumačili rezultati analiza koje se zasnivaju na njegovom korišćenju. Dozvolu za upotrebu ovog korpusa predusretljivo daju njegovi tvorci. Idejni

tvorac ovog korpusa dr Duško Vitas, takođe je jedan od važnih pokretača korpusne lingvistike u Srbiji, sa svojom značajnom popularizatorskom ulogom kada je reč o korpusnom pristupu izučavanjima jezika.

Izgrađenost korpusa srpskog jezika prema sadašnjem stanju jasno ukazuje na potrebu stvaranja jednog relevantnog monitoring korpusa opšteg tipa na nivou državnog projekta, s otvorenim pristupom kada je u pitanju korišćenje, po ugledu na Rusiju ili Hrvatsku. Imajući u vidu njegov potencijalni značaj i nesumnjivu višestruku korist, budući korpus bi morao biti koncipiran kao nacionalni projekat Ministarstva za nauku. Izgradnja takvog korpusa podrazumeva ulaganje znatnog napora kako bi se s uspehom rešili raznovrsni problemi koji prate stvaranje jednog korpusa. Tu bi trebalo brižljivo planirati pojedine faze obuhvata putem velikog broja manjih uzoraka. (Preporučuju se uzorci od po 2000 reči iz što je moguće većeg broja izvora, - pisanih i govornih). Pri tom bi trebalo rešavati i pitanja vezana za dobijanje potrebnih dozvola kada su u pitanju tekstovi zaštićeni pravima njihovih autora i izdavača. Zbog nekomercijalne prirode i otvorenog pristupa istraživačkoj zajednici, a i zbog male veličine pojedinačnih uzoraka, trebalo bi računati i sa eventualnim izdacima za takve namene. Osim toga, to je posao koji konzumira i vreme, a predstavlja dosta administrativnog rada. Kao što je pomenuto, budući korpus srpskog jezika, uz već pomenute poželjne osobine, morao bi da bude monitoring korpus, a to znači kontinuiran rad uz neprekidno dopunjavanje i dograđivanje novim primerima srpskog jezika, prvo sinhronog a potom, svakako, i dijahronog teksta.

Jedna od prvih upotreba ovakvog korpusa mogla bi da bude konstrukcija rečnika srpskog jezika. Primera radi, nova izdanja rečnika engleskog jezika izlaze svakih nekoliko godina kako bi se pratile promene značenja reči i druge promene u jeziku. Teži se izdavanju rečnika koji su što je moguće više prilagođeni korisnicima. Brzina semantičkih promena je znatna pa se očekuje da će većina rečnika engleskog jezika ubuduće izlaziti samo u elektronskom obliku. Poznati su naponi koji se u Srbiji ulažu već skoro pola u stvaranju Rečnika srpskog jezika. Taj trud je do sada doveo samo do polovine ukupne konstrukcije. Dugo vreme utrošeno u konstrukciji jednog rečnika nepovoljno se odražava na kvalitet konačnog rezultata: može da se izgubi kontakt sa modernim jezikom jer značenja definisana pre nekoliko godina mogu prirodnim evolutivnim procesima u jeziku biti potpuno promenjena. Nada za spašavanje već uloženog truda i izgledni završetak ovog jezičkog projekta veka a i svih sličnih budućih leksikografskih poduhvata u Srbiji leži skoro sigurno u konstrukciji pogodnog korpusa srpskog jezika.

5. KORPUSNA ANALIZA

Osnovni parametar zakonitosti unutar jednog korpusa, a samim tim i zakonitosti unutar jezika i društva, jeste frekventnost referentnih jezičkih jedinica. Referentne jezičke jedinice, iz perspektive korpusne lingvistike, mogu biti raznolike, od fonema, morfoloških nastavaka, gramatičkih jedinica, leksičkih skupova, pojedinačnih reči do složenijih značenjskih jedinica.

Iz definicije statističkog skupa („Skup svih pojedinačnih slučajeva na kojima se izvesna pojava statistički posmatra zove se statistički skup ili statistička masa ili populacija.” Njegić, R. i Žižić, M. (1987) *Osnovi statističke analize*, Savremena administracija, Beograd, str. 14.) sledi da je i korpus, bez obzira na vrstu kojoj pripada, istovremeno i jedan poseban javni oblik statističkog skupa. Sama korpusna analiza podrazumeva nekoliko faza, pri čemu se u svakoj od njih koriste odgovarajuće statističke procedure. Nekoliko osnovnih koraka korpusnog istraživanja isključivo se zasnivaju na primeni statističkih metoda. (Koriste se standardne statističke metode, kao i posebni modeli specijalno razvijeni za potrebe lingvističkih istraživanja: *Markovljev* model predvidljivost

jezičkih sekvenci, Šenonov model jezičke ekonomije engleskog jezika i, naročito popularna među istraživačima, Zipfova hipoteza o odnosu distribucije frekvencije prema rangu.). Rezultati dobijeni statističkom analizom uključuju se u dalju lingvističku analizu i jezičku teoriju što, pri uspješnom istraživanju, rezultira konačnim odgovorima na ranije postavljena otvorena pitanja.

Pri statističkoj obradi podataka u jezički baziranim istraživanjima koriste se posebno dizajnirani računarski statistički paketi među kojima su najviše upotrebljavani SPSS (Statistical Package for Social Sciences) i R statistički program. SPSS je nadaleko poznati statistički softver. On se prvenstveno se koristi od strane univerzitetskih ili istraživačkih organizacija budući da je neophodno prethodno pribavljanje licenci. R program dolazi potpuno besplatan i upravo je namenjen privatnim korisnicima. Iako poseduje solidne analitičke mogućnosti kao mana se može istaknuti činjenica da operiše putem programskog jezika što relativno usporava tok analize, i opterećuje istraživača.

* * *

Savremena lingvostatistika u velikoj meri oslanja se na korišćenje velikih računarskih jezičkih baza – korpusa. Uz opšte, razvijaju se i posebne statističke metode koje se koriste specijalno u lingvistici. Pojavljuju se i specijalizovani softverski paketi prilagođeni lingvostatističkoj analizi. Kada je reč o korpusnoj analizi srpskog jezika sasvim je sigurno da će primena statističkih metoda analize biti ubuduće od višestrukog značaja. Ukoliko bi se našlo dovoljno društvene podrške za izgradnju računarskog korpusa srpskog jezika, s poželjnim karakteristikama o kojima je bilo reči u ovom osvrtu, sigurno je da bi u tom poslu značajnu ulogu dobila i statistika. To, pored ostalog, ukazuje i na potrebu uključivanja statistike u nastavne planove osnovnih studija, master programa i doktorskih usavršavanja na filološkim fakultetima u Srbiji, kao i na značaj korpusne analize. A to potvrđuje važnost izučavanja lingvostatistike i praćenje savremenih tendencija u ovoj atraktivnoj naučnoj oblasti.

LITERATURA

- [1] Baayen, R.H. (2008) *Analyzing Linguistic Data. A Practical introduction to Statistics Using R*. Cambridge University Press, New York.
- [2] Biber, D., Conrad, S., Reppen, R. (2000) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge, Cambridge University Press.
- [3] Bod, R., Haz, J., Jannedy, S. (2003) *Probabilistic Linguistics*. The MIT Press, Cambridge, Massachusetts, London.
- [4] Johnson, K. (2008) *Quantitative Methods in Linguistics*. Blackwell Publishing, Oxford.
- [5] Lončar, S. (2009) *Pregled softvera za statističku analizu lingvističkih podataka*. Predavanje u Statističkom društvu Srbije, Klub statističara Vojvodine, (umnoženi materijal).
- [6] Oakes, M. (1998) *Statistics for Corpus Linguistics*. Edinburgh Textbooks in Empirical Linguistics, Edinburgh University Press, Edinburgh.

PRIKAZ NAUČNOG DOGAĐAJA

Dr Luka Filipović
Beogradska poslovna škola, Beograd

XXXV SIMPOZIJUM O OPERACIONIM ISTRAŽIVANJIMA

U periodu od 14 – 17. septembra 2008. godine održan je u Soko Banji Simpozijum o operacionim istraživanjima, već dugo poznat kao SYM-OP-IS 2008. u organizaciji Saobraćajnog fakulteta. Suorganizatori simpozijuma bili su fakulteti iz Beograda: Mašinski, Matematički, Rudarsko-geološki, Ekonomski, Organizacionih nauka, instituti iz Beograda: Matematički, Ekonomski, “Mihajlo Pupin”, zatim Vojska Srbije i Društvo operacionih istraživača. Pokrovitelj Ministarstvo za nauku i tehnološki razvoj Republike Srbije, kao i donatori “SAP WEST BALKANS”, “Saobraćajni institut CIP”, “BAS”, “PEKABETA”, “DONCAFE GROUP” svi iz Beograda i “ELEKTROSRBIJA” iz Kraljeva doprineli su da se organizacija i realizacija simpozijuma uspešno pripremi i završi. Programskim odborom rukovodio je prof. dr Dušan Teodorović, a organizacionim odborom doc. dr Branka Dimitrijević.

Ukupno je podneto 170 radova od kojih je 161 rad odobren i štampan u Zborniku, a 9 radova je dato kao saopštenje uz rezime i ključne reči. Interesantno je dati i strukturu broja autora po radovima. Tako je po 1 autor bio u 56 radova, po 2 autora u 67 radova, po 3 u 32 rada, po 4 u 11 radova i sa 5 autora bilo je 4 rada. Ukupno je bilo 298 autora koji su radeći sami ili kao koautori, ove godine pripremili 170 radova. Autori, pretežno iz Srbije bili su i iz Kanade, Grčke, Crne Gore, BiH, Slovenije, Turske, Rumunije i Makedonije. Iz činjenice da je većina fakulteta i univerziteta u Srbiji koncentrisana u Beogradu i ove godine bilo je najviše autora iz Beograda, zatim iz Novog Sada, Subotice, Čačka, Kragujevca i drugih.

Radovi koji su štampani u Zborniku prošli su proceduru recenzije.

I na XXXV SYM-OP-IS-u bilo je predavanja po pozivu. To su prof. Markos Papageorgiou iz Grčke i prof. Pierre Hansen iz Kanade.

Rad po sekcijama

Na XXXV SYM-OP-IS-u rad i aktivnosti realizovani su u 25 sekcija. U daljem dat je naziv sekcije i broj radova:

1. Ekonomija (6)
2. Ekonomski modeli i ekonometrija (7)
3. Ekspertski sistemi (2)
4. Elektronsko poslovanje (7)
5. Energetika (1)
6. Finansije (9)
7. Geoinformacioni sistemi (5)
8. Građevinarstvo (14)
9. Informacioni sistemi (9)
10. Istraživanje i razvoj (4)
11. Istraživanje podataka (DATA MINING) (4)
12. Logistika (10)
13. Matematičko programiranje (6)
14. Meko računarstvo (7)
15. Menadžment (13)
16. Poljoprivreda i šumarstvo (4)
17. Primena OI u odbrani (24)
18. Rudarstvo i geologija (1)
19. Saobraćaj, transport i komunikacije (13)
20. Simulacija (5)
21. Statistički modeli (6)
22. Stohastički modeli i vremenske serije (3)
23. Upravljanje proizvodnjom i zalihama (2)
24. Upravljanje rizikom (5)
25. Višekriterijumska analiza optimizacija (3)

Iz činjenice da je bilo više bliskih sekcija sa manjim brojem radova, iste su spajane pa se u istoj sali u vremenu predviđenom za određeni broj referata odvijalo rad. Tako se od 25 sekcija radilo u 18 delova na kojima su predsedavali rukovodioci i organizatori sekcija.

Rad XXXV SYM-OP-IS-a počeo je 15. septembra u 11^h plenarnom sednicom na kojoj je predsedavao prof. dr Dušan Teodosijević. Na istoj sednici su saopštili svoje radove gosti po pozivu i to: prof. Markos Papageorgiou: "A rolling-Horizon quadratic-programing approach to the signal control problem in large-scale congested urban road networks" i prof. Pierre Hansen: "Edge realizability of graphs".

Najviše radova bilo je u sekciji “Primena OI u odbrani (24)”, zatim “Grsđevinarstvo (14)”, “Menadžment (13)”, “Saobraćaj, transport i komunikacije (13)”, “Logistika (10)”, itd.

Pregledom sekcija, kao i radova po sekcijama i generalno, može se bez preterivanja konstatovati brojnost iz koje se vidi da SYM-OP-IS-2008 nastavlja nivo, zahteve, strogost, originalnost i već utvrđene kriterijume kao i kodeks očuvanja i poboljšanja tradicionalnih postavki Simpozijuma. Brojni autori, od onih koji su prisutni od početka rada SYM-OP-IS-a, su postavljali i organizovali sekcija, uvodili mlađe istraživače, davali mnoštvo originalnih radova, pa sve do mlađih istraživača, koji počinju na najbolji način svoje stručne i naučne karijere, po ko zna koji put SYM-OP-IS dokazuje u punoj meri svoje postojanje.

Zbornik radova je priča za sebe. Mnogi istraživači različitih struka naći će uvek neku novinu, neku ideju za sebe, pa će i oni od sledećih godina biti učesnici Simpozijuma.

Sekcije “Statistički modeli” i “Stohastički modeli i vremenske serije” sa ukupno 9 radova

U sekciji “Statistički modeli” prihvaćeni su i realizovani sledeći radovi sa autorima:

1. Zakon velikih brojeva kao podrška odlukama u osiguranju /Ivana Simeunović, Slavica Dabetić/;
2. Moguće interpretacije koeficijenta asimetrije i spljoštenosti /Jelena Stanojević/;
3. Analiza glavnih komponenta na makroekonomskim indikatorima /Julija Cerović, Vladimir Vasić/;
4. Redukovani BOOTSTRAP-T intervali poverenja za aritmetičku sredinu sa primenom u istraživanju tržišta /Vesna rajić, Dragan Lončar/;
5. Stope slobodnih radnih mesta u Beogradu i Banatu /Olgica Bošković/;
6. Statističko modeliranje rizičnih pojava u osiguranju /Dajana Vindžanović/;

U sekciji “Stohastički modeli i vremenske serije”:

1. Detekcija haosa: Algoritmi isoftverska implementacija /Vladica Stojanović, Milan Božinović/;
2. Ekstremi u realizacijama stohastičkog procesa sa procenom efekata /Tomislav Zečević, Luka Filipović/;
3. Autokovarijansa i autokorelacija u analizi vremenskih serija /Nikola Janković, Marija Kerkez/;

Svaki od radova iz ove dve sekcije sadrži dobru analizu problematike, razradu tematike sa određenim poboljšanjima, korekcijama i proširenjima, sa brojnim primerima primene na procese i pojave.

Dovoljno je da pažljivi i istrajni čitalac posveti mnogo vremena da pročita sve radove iz svih oblasti koje us bile u sastavu sekcija. Rekli bismo da je to teško ostvarivo. Međutim, ocene rukovodioca sekcija, kao i učesnika po sekcijama bile us više nego vredne i značajne za dalji razvoj Operacionih istraživanja, kako u teorijskom pogledu tako i u punoj primeni, koja se širi na sve nove procese i nove discipline pojedinih oblasti.

Organizacija Simpozijuma

Čast da organizuje XXXV Simpozijum o operacionim istraživanjima pripala je Saobraćajnom fakultetu, jer je to godina neke vrste jubileja, a više od toga Simpozijum je skoro i zaboravio dečje bolesti, mladalačku hrabrost i odvažnost i ulazi u period pune ozbiljnosti, vrhunskih radova, ustaljenog ciklusa organizacije, redakcije radova, izbora sekcija i dr. No, ređamo godine života SYM-OP-IS-a , ali priliv mlađe generacije ne pomera prosek starosti učesnika. Po ovome Simpozijum će biti većito srednjih godina sa divnom mešavinom iskustva priznatih naučnih radnika i mladosti koja dolazi.

Organizator je sa puno umešnosti i rutine rešavao tekuće poslove, bez obzira na određene teškoće, koje redovno prate pripremu Simpozijuma. Soko Banja je u kišnim uslovima pokazala blagonaklonost i domaćinsko ophođenje. Znači sigurno da će Simpozijum kad tad doći u Soko Banju i u sunčanim danima.

Druženje, zajednička večera i stručne rasprave po temama, sve noviji radovi, mlađi učesnici, postaje zalog za dug život Simpozijuma.

IN MEMORIAM

*Prof. dr Branislav Boričić
Ekonomski fakultet, Beograd*

In memoriam

PROF. DR VLADISLAV MILOŠEVIĆ (1934-2009)

Rođen je 1934. godine u Kragujevcu. Osnovnu školu i gimnaziju je završio u Beogradu. Diplomirao je na Prirodno-matematičkom fakultetu u Beogradu 1961. godine. Magistrirao je 1967. godine, a 1976. godine odbranio doktorsku disertaciju na Ekonomskom fakultetu u Beogradu.

Godine 1963. zaposlio se kao profesor gimnazije u Francuskoj. Od 1964. godine radio je u Zavodu za statistiku SR Srbije kao naučni saradnik. Kao asistent započinje svoju akademsku karijeru na Ekonomskom fakultetu 1969. godine. U zvanje docenta je izabran 1977. godine, u zvanje vanrednog profesora 1985, a u zvanje redovnog profesora 1991. godine.

Boravio je 1977. godine na Statističkom institutu univerziteta u Parizu, i radio u seminaru za Teorijsku statistiku.

U nastavi na osnovnim studijama na našem fakultetu predavao je Teorijsku statistiku i Statistiku II. Izvodio nastavu na odeljenju fakulteta u Šapcu i na Vojnoj akademiji u Beogradu. Bio je angažovan na poslediplomskim studijama na našem fakultetu i van njega. Na magistarskom kursu za Statističku analizu Ekonomskog fakulteta u Beogradu predavao je više predmeta: Matematičku statistiku, Teoriju statističkog zaključivanja, Analizu varijanse i planiranje eksperimenata i Diskriminacionu analizu. Bio je angažovan na diplomskim studijama na Vojnoj akademiji u Beogradu, gde je držao predavanja iz Teorije verovatnoće i statističkih metoda, na Ekonomskom fakultetu u Kragujevcu držao je predavanja iz Teorije verovatnoće, na Jugoslovenskom Institutu za ekonomska istraživanja u Beogradu držao je predavanja iz Teorije distribucija. Rukovodilac je magistarskog kursa Statistička analiza od 1992. godine.

Član je Matematičkog instituta u Beogradu, kao i Statističkog društva Srbije i njegove sekcije za Teorijsku statistiku.

Penzionisan je 2001. godine.

Pored ovih formalnih i egzaktnih, ali važnih podataka o profesoru Miloševiću, mi koji smo sa njim radili i imali tu sreću da sa njim saradujemo, bili sa njim «na ti» i obraćali mu se sa «Vlado», moramo se setiti i njega iz onog drugog ugla koji po pravilu nije sadržan u biografijama.

Govoriti o Vladi kao svestranoj ličnosti ne bi bilo lako. Trebalo bi dobro poznavati sport, jer Vlada je u svojoj mladosti bio atletičar i nikada ga nije napustio zdravi takmičarski duh olimpizma, trebalo bi biti umetnik, jer Vlada je pronicao u najtananije finese likovne i muzičke umetnosti, književnosti i poezije, trebalo bi biti filolog, jer Vlada je bio poliglota, trebalo bi biti gastronom, jer Vlada je poznavao tajne pripreme i konzumiranja hrane, treba biti matematičar i statističar, pa će samo o toj dimenziji njegove ličnosti ovde biti reči.

Pre svega bih govorio o Vladi kao čoveku, kolegi, profesoru, naučniku, o svakodnevnoj komunikaciji sa njim.

Vlada je bio posvećeni profesor kakvih je danas malo na univerzitetu. Apsolutno predat svojoj svetoj misiji učitelja, svojim knjigama, beleškama, idejama, a nikada raspoložen da pravi kompromise sa neznanjem, površnošću i neprofesionalnošću. Upravo takav, obrazovao je desetine generacija statističara Srbije i bivše Jugoslavije, pružajući im bazična znanja, ali i jasno definišući najvažnije principe profesije.

Strpljivošću i sistematičnošću Vuka Karadžića, Vlada je sakupljao interesantne, neobične i intrigirajuće zadatke iz trigonometrije, algebre, analize, geometrije, teorije verovatnoća, statistike, raspravljao o njima i njihovim rešenjima, tražio im odgovarajuće mesto u svojim zbirkama i udžbenicima.

Ne mali broj puta je započinjao svoju borbu odbrane profesije usamljen, a završavao je pobednički. Pobede je odnosio snagom svojih čistih i snažnih argumenata, jer za njega drugačiji od ovakvih argumenti nikada nisu ni postojali. Stoga bih rekao da je Vlada imao lep i srećan život, jer je sreću spoznao u učenju i obrazovanju sebe i drugih, u vladanju u skladu sa svetim moralnim i profesionalnim načelima. I to je, reklo bi se, bio jednostavan recept životne sreće, koji svi tobože znamo i tobože primenjujemo. Ali za stvarnu primenu takvog recepta su potrebni znanje, kompetentnost, hrabrost i profesionalno poštenje, čime je Vlada raspolagao u izobilju.

U Vladi smo, kao starijem kolegi, uvek nalazili pouzdanu podršku, jer njegovi saveti i komentari su uvek bili otvoreni, iskreni, neposredni i, kao takvi, neprocenjivo korisni i vredni. Vladu je krasila osobina istinoljubivosti i to osobina istinoljubivosti po svaku cenu. On nije bio sagovornik koji vam saopštava ono što želite da čujete, već je saopštavao ono što je zaista mislio i želeo. Zato je komunikacija s njim bila dragocenija od ostalih.

Već njegovim odlaskom u penziju i čestim zdravstvenim incidentima, na fakultetu ostaje praznina kakvu ostavljaju za sobom samo posvećeni naučnici i predani pedagozi.

Prerano je Vlada ostavio i svoju suprugu Biljanu, kćerku Mionu i sina Miloša, na koje je bio ponosan i zbog kojih je bio srećan suprug i otac, praveći im nepravedan gubitak koji više nije moguće nadoknaditi.

*Prof. dr Ljiljana Petrović
Ekonomski fakultet, Beograd*

In memoriam

PROF. DR VLADISLAV MILOŠEVIĆ (1934-2009)

Profesor Vladislav Milošević je predavao predmet Teorijska statistika na redovnim studijama Ekonomskog fakulteta u Beogradu od 1977. godine do odlaska u penziju 2001. godine. Pored toga predavao je i predmete Matematička statistika, Teorija statističkog zaključivanja i Diskriminaciona analiza na posleddiplomskim studijama. Držao je nastavu na odeljenju fakulteta u Šapcu, na Ekonomskom fakultetu u Kragujevcu i na Vojnoj akademiji u Beogradu. Bio je mentor i član mnogih komisija za odbranu magistarskih radova i doktorskih disertacija.

Na stručnom usavršavanju bio je na Statističkom institutu Univerziteta u Parizu. Objavio je preko 50 naučnih i stručnih radova iz oblasti teorijske i primenjene statistike u domaćim i stranim časopisima i izlagao više puta na sastancima statističara Evrope i Međunarodnog statističkog instituta (1987. u Solunu, 1989 u Parizu). Radovi profesora Miloševića su bili zapaženi u stručnim krugovima. Bio je član Statističkog društva Srbije Matematičkog Instituta u Beogradu.

Objavio je 12 udžbenika (univerzitetskih i srednjoškolskih) i 27 stručnih iz oblasti statistike i matematike. Bio je i recenzent velikog broja udžbenika. Udžbenike profesora Miloševića pored studenata Ekonomskog fakulteta koriste i studenti drugih fakulteta jer su napisani na visokom stručnom nivou uz praćenje savremenih dostignuća i najnovije literature u oblasti statističkih nauka. Treba pomenuti i veliku zaslugu profesora Miloševića u stvaranju bibliotečkog fonda Ekonomskog fakulteta u oblasti statistike.

U naučnim i stručnim radovima profesor Milošević je proučavao različite oblasti verovatnoće i statistike. Magistarski rad pod naslovom "Karakteristične funkcije aleatornih promenljivih" i doktorska disertacija pod naslovom "Transformacije aleatornih promenljivih" (odbranjeni na Ekonomskom fakultetu u Beogradu) obuhvatali su slučajne promenljive, posebno njihove karakteristične funkcije, transformacije slučajnih promenljivih i slučajnih vektora. U više naučnih i stručnih radova su razmatrane granične teoreme karakterističnih funkcija slučajnih promenljivih i izračunavanja momenata i centralnih momenata diskretnih slučajnih promenljivih preko rekurentnih obrazaca. U svojim radovima je ukazivao i na primenu Teorije verovatnoće u fizici i drugim naučnim oblastima.

Na kongresima matematičara, fizičara i astronoma Jugoslavije bili su zapaženi njegovi radovi iz primene uopštenih funkcija u matematičkoj statistici i o teoriji odlučivanja kao teorijskoj osnovi statističkog zaključivanja.

Proučavao je i mešavine funkcija raspodela slučajnih promenljivih predstavljenih Lebeg-Stijltesovim integralom i rezultate izlagao na XVII sastanku statističara Evrope u Solunu 1987. godine.

Više radova je posvećeno problemu slučajnog izbora iz fizičkog i statističkog osnovnog skupa. Proučavao je i problem robustnosti u Bajesovoj analizi, definisao način merenja robustnosti Bajesove funkcije odluke u odnosu na apriornu raspodelu i definisao aposteriornu robustnost i rizik robustnosti.

Brojni studenti, magistranti, doktoranti, sa poštovanjem će se sećati svog dragog profesora. Kolege će ga pamtiti kao vrsnog stručnjaka i dragog kolegu.

*Prof. dr Dubravka Pavličić
Ekonomski fakultet, Beograd*

In memoriam

PROF. DR MILEVA ŽIŽIĆ (1938-2009)

U trenutku kada se kao kolektiv opraštamo od naše koleginice, profesorice Mileve Žižić, nije nimalo lako u nekoliko reči sažeti njen životni i profesionalni put, kao i doprinos oblasti kojoj je posvetila pune četiri decenije svoga života.

Ceo svoj radni vek profesorica Žižić je provela sa studentima. Sve vreme se nesebično trudila da im prenese svoje znanje i ljubav prema statistici, zbog čega je bila ne samo veoma cenjena nego i omiljena kao predavač. Studenti su poštovali i njenu objektivnost na ispitima, kojih je, kao što znamo, bilo mnogo – bez preterivanja se može reći i na desetine hiljada. Ipak, treba reći da njena objektivnost nikada nije prelazila u rigidnost, jer je profesorica Žižić imala sluha za studente sa privatnim problemima i nalazila načina da im pomogne, a da pri tom ne odstupi od svojih ustaljenih kriterijuma. Kao višegodišnji rukovodilac na fakultetu, u svojstvu prodekana za nastavu ili dekana, tesno je saradivala sa predstavnicima studenata, uvek spremna da udovolji svim prihvatljivim zahtevima. Možemo sa sigurnošću reći da su hiljade diplomiranih ekonomista bili i njeni đaci, od kojih su mnogi svoj diplomski rad napisali pod njenim mentorstvom. O profesorici Žižić kao pedagogu mnogo govori i to što mi se više puta dešavalo u privatnim kontaktima da su moji sagovornici, kada bi saznali da sam profesor statistike, evocirali uspomene na studentske dane i bez izuzetka sa osmehom i poštovanjem pomenuli njeno ime, ponekad i sa neskrivenim ponosom što su bili njeni studenti.

Profesorica Žižić je predavala i na posle diplomskim studijama i bila mentor ili član komisije većem broju magistranata i doktoranata. Nekima od njih, koji su izabrali univerzitetsku karijeru, kasnije je recenzirala udžbenike, nastavljajući da prati njihov pedagoški rad.

Kada je reč o istraživačkom opusu profesorice Žižić, on obuhvata veći broj naučnih i stručnih radova napisanih samostalno ili u koautorstvu, objavljenih u časopisima ili zbornicima sa brojnih savetovanja u zemlji i inostranstvu, kao i nekoliko udžbenika i zbirki zadataka iz statistike. Pokušala sam da ih hronološki i tematski grupišem u nekoliko celina.

U magistarskom radu i doktorskoj disertaciji profesorica Žižić primenjuje statističke metode u analizi razvoja poljoprivredne proizvodnje, a zatim u analizi problema formiranja, raspodele i upotrebe dohotka individualnih domaćinstava, ukazujući na značaj utvrđivanja realnog dohotka i posebno analizirajući probleme njegovog utvrđivanja.

Sledećih nekoliko godina profesorica Žižić je svoje istraživanje usmerila na oblast analize vremenskih serija. U monografiji i radovima koje je publikovala tih godina, izučavala je metode koje obezbeđuju efikasne prognoze kretanja ekonomskih veličina. Reč je o relevantom istraživanju imajući u vidu da predviđanje budućeg kretanja vremenske serije predstavlja jedan od osnovnih zadataka savremene analize vremenskih serija. Ovim problemom profesorica Žižić se bavila sa tri aspekta. Prvo, razmatrala je različite metode izravnjanja, čija se popularnost zasniva na jednostavnosti primene. Drugo, kako najveći broj makroekonomskih vremenskih serija pokazuje tendenciju stohastičkog kretanja, posebno je izučavala složene modele kojima se obuhvata dinamika značajnog stohastičkog trenda. Treće, budući da sezonske varijacije često predstavljaju značajan izvor fluktuacija u kretanju ekonomskih veličina, bavila se problemom modeliranja različitih vrsta sezonskih varijacija u cilju dobijanja pouzdanih prognoza.

Profesorica Žižić je objavila i nekoliko radova u kojima se razmatraju problemi primene statističke metodologije u praksi ali i mogućnosti da se na bazi dobijenih rezultata donose različite poslovne odluke. Izdvojila bih dva rada, napisana u koautorstvu. U prvom se analizira primena modela višestruke regresije u oblasti poljoprivrede, ukazuje na činjenicu da su pretpostavke za primenu modela često narušene, analizira način pravilnog izbora i kompozicije modela, i ističe značaj obazrive interpretacije rezultata. U drugom radu se razmatra upotreba tabele i faktora kontingencije za određivanje optimalne marketing strategije u preduzeću.

Poslednjih godina svog radnog veka profesorica Žižić je svoje interesovanje usmerila na istraživanja u oblastima finansijskih tržišta i osiguranja. U okviru finansijskih tržišta bavila se uticajem varijacija prinosa hartija od vrednosti i stepena njihovog slaganja na ukupnu rizičnost portfolija i problemima diverzifikacije portfolija kako bi se uz dati rizik maksimizirao očekivani prinos ili za dati očekivani prinos minimizirao rizik. Tu su i radovi u kojima se razmatra primena statističkih metoda u neživotnom osiguranju i problemi upravljanja rizikom kamatne stope.

Svakako treba pomenuti i radove koji sadrže istorijske preglede razvoja statistike i naučno-nastavne publicistike u ovoj oblasti u Jugoslaviji počev od II svetskog rata. U njima se afirmišu autori koji su postavili temelje razvoju statistike kod nas, i oni koji su kroz udžbeničku literaturu doprneli njenoj popularizaciji i masovnijoj primeni. Ovoj grupi autora danas nesumnjivo pripada i profesorica Žižić zbog doprinosa koji je dala u svojim publikovanim radovima i, pre svega, brojem i kvalitetom udžbenika i zbirki koje je u koautorstvu napisala tokom četiri decenije duge i plodne karijere vrsnog pedagoga.

Profesorica Žižić je koautor četiri udžbenika i zbirke zadataka iz osnova statističke analize namenjenih studentima prve godine ekonomskog fakulteta, udžbenika poslovne statistike koja se izučavala na drugoj godini, i udžbenika statistike za srednje škole ekonomskog usmerenja. O visokom stručnom i pedagoškom kvalitetu ovih učila govore ne samo pozitivne recenzije i prikazi u stručnim časopisima i pohvalna mišljenja studenata iskazana u anketama, već i činjenica da su pojedini udžbenici imali brojna izdanja i da su godinama korišćeni i na ekonomskim fakultetima pojedinih univerziteta u bivšoj Jugoslaviji. Njen udžbenik za srednju školu, preveden i na mađarski jezik, decenijama se koristi u nastavi, što je dokaz nesumnjivih pedagoških sposobnosti autora da jezikom i stilom materiju učine pristupačnom i interesantnom i za mlađi uzrast.

Mileva Žižić je bila istaknuti član našeg kolektiva, cenjena i omiljena među kolegama i zaposlenima u vannastavi. Podatak da je birana za šefa katedre za statistiku i matematiku, da je u dva mandata izabrana za prodekana i da je bila dekan fakulteta, jasno pokazuje da je kolektiv u njoj prepoznao osobu velikih organizacionih sposobnosti, lojalnu i energičnu, koja je uvek na prvo mesto stavljala interese fakulteta i umela da se za njih izbori. Profesorica Žižić je često govorila da je fakultet njena druga kuća, i to u njenom slučaju nije bila puka fraza. Poznato je da je u svojstvu dekana svakoga dana među prvima dolazila na fakultet i to vreme posvećivala vannastavi, a zatim se usredsređivala na pitanja vezana za normalno odvijanje nastave, štiteći uvek renome fakulteta i interese studenata. Ako bi se dogodilo da u nastavi na nekom predmetu iskresne problem, nije izbegavala neprijatnu dužnost da razgovara sa odgovornim kolegama. Nije razumela niti htela da prihvati da bilo ko od kolega može ličnine interese ili međusobne nesuglasice da pretpostavi interesima fakulteta.

Nalazila je vremena i da prati profesionalne rezultate kolega, posebno mladih, a oni najuspešniji su sticali njenu bezrezervnu podršku. Posebno je vodila računa o blagovremenim izborima u viša zvanja, smatrajući to dužnošću fakulteta prema vrednim i uspešnim članovima kolektiva.

Donošenje odluka na fakultetu, velikoj i po strukturi složenoj ustanovi, često nije lako, a neretko biva praćeno žučnim diskusijama na katedrama ili Nastavno-naučnom veću. Umela je profesorica Žižić u njima da učestvuje, energično brani svoje stavove i ponekad emotivno reaguje. Ali, nezavisno od toga da li biste se sa njom slagali ili ne, činjenica je da su njene kritike bile upućene isključivo stavovima i argumentima druge strane, a nikada osobama koje su ih zastupale. Zato bi se već pri sledećem susretu, sa njoj svojstvenim srdačnim osmehom i čvrstim stiskom ruke, pozdravila sa jučerašnjim neistomišljenicima, pokazujući tako zavidan stepen akademске tolerancije. Ponekad je u verbalnim duelima bivala i lično povređena. Ali, i u ovim slučajevima imala je Mileva ljudske širine da oprost i zaboravi.

Sa profesoricom Žižić sam neposredno saradivala na istom predmetu 25 godina. Veoma je volela svoj predmet, a veliki broj studenata na prvoj godini nije doživljavala kao teret, već kao šansu da populariše statistiku i doprinese omasovljenju njene primene.

Njen odnos prema starijoj kolegini, profesorici Njegić je redak primer odanosti i poštovanja koje je Mileva izražavala do kraja svog života, nalazeći vremena da i pored velike zauzetosti posećuje profesorku u penziji ili je barem pozove telefonom.

Prema mlađim kolegama je bila zahtevna kada je reč o nastavnim obavezama i zaštiti interesa predmeta. Nikoga nije favorizovala, insistirala je na dobrim kolegijalnim odnosima, dogovoru po svim pitanjima u vezi sa predmetom i jedinstvenom kriterijumu na ispitima, štiteći uvek ugled predmeta ali i interese studenata. Ako bi imala primedbi, uputila bi ih direkto, bez okolišavanja, i na taj način je uz međusobno puno uvažavanje uspevala da brzo otkloni većinu nesporazuma. Nije nas smatrala ličnim saradnicima, već saradnicima na predmetu, zbog čega nije zloupotrebljavala autoritet koleginice sa višim nastavnim zvanjem. Posebno želim da naglasim da je, čak i u periodima kada je bila na odgovornim rukovodećim mestima - na fakultetu, Beogradskom univerzitetu ili Republičkom zavodu za statistiku, Mileva Žižić svoje nastavne obaveze, predavanja i ispite, samostalno obavljala. Podsticala je našu kreativnost u nastavi, rado prihvatila predložene inovacije, i bila ponosna kada bi se studenti pohvalno o nama izražavali. U radu na udžbeniku je bila veoma kooperativna, otvorena za nove ideje i spremna da bez trunke sujete sasluša i prihvati sugestije mlađih kolega.

Ipak, nije uvek bilo lako pratiti njenu ogromnu radnu energiju, na kojoj sam joj zavidela. Često sam se pitala koliko sati ima njen radni dan i kako je moguće da sve stiže.

Bogata radna biografija profesorice Žižić navodi na pomisao da je ova vredna i sposobna žena imala uz to i idealne životne uslove koji su joj omogućavali da se u potpunosti posveti svom radu, a to svakako nije bio njen slučaj. Ona je naime imala i porodične obaveze, uvek bila požrtvovana supruga i majka. A ni dugogodišnji život u skućenom stambenom prostoru nije uspeo niti da naruši porodični sklad, niti da je spreči da profesionalno napreduje. Na sinove Branka i Zorana prenela je svoju profesionalnu ambiciju i ljubav prema porodičnom životu, pa je nakon njihovih diploma i prvih poslovnih uspeha doživela i radost da postane baka.

Kada je posle 40 godina napornog rada otišla u penziju, planirala je da sa suprugom Radulom uživa u dužim boravcima na Durmitoru i u Herceg Novom, družeći se sa rodbinom i prijateljima i prateći odrastanje svojih unuka Irene, Maje i Nikole. Nažalost, planove je prekinula Radulova teška bolest, nega bolesnika i konačno gubitak supruga koji je Mileva veoma teško podnela. Ubrzo zatim razbolela se i sama. O svojoj bolesti je govorila otvoreno i pribrano, ne tražeći od sagovornika sažaljenje. Izgubila je bitku trudeći se do poslednjeg trenutka da ne optereti porodicu. U našem poslednjem telefonskom razgovoru mi je rekla: Moram da se borim, nema mi druge. To je, može se reći, bio životni moto profesorice Žižić, hrabre i istrajne žene, velikog radnika i čestitog i dobrog čoveka.

Neka je slava i hvala profesorici Milevi Žižić.

*Prof. dr Branislav Boričić
Ekonomski fakultet, Beograd*

In memoriam

PROF. DR MILEVA ŽIŽIĆ (1938-2009)

Teško je oproštati se od dugogodišnjeg prijatelja, saradnika, roditelja, poznanika,... Mi se danas opraštamo od naše drage koleginice koja je ceo svoj radni vek provela u ovoj kući, zajedno sa nama, deleći, kako one lepe, tako i one teške dane, a bilo je i jednih i drugih. U profesoricu Milevu Žižić uvek smo mogli naći čvrst oslonac, dobrog savetnika i iskrenog prijatelja.

Mileva Žižić je rođena 1938. godine u Čantaviru u Bačkoj. Osnovnu školu i gimnaziju završila u Herceg Novom. Zamišljam devojčicu tokom oskudnih posleratnih godina na potezu od Zelenike do Igala kako sanjari o budućnosti. Sećam se nebrojenih razgovora sa profesoricom Žižić u kojima se ponovo i ponovo vraćala toj obali, jadranskom biseru.

Diplomirala je na Ekonomskom fakultetu u Beogradu 1960, magistrirala 1967. i doktorirala 1973. godine. Dolazi na Ekonomski fakultet u Beogradu kao asistent 1960; u zvanje docenta birana 1973, a u zvanje vanrednog profesora 1979. godine. Redovni profesor postaje 1985. godine. Predavala na predmetu Osnovi statističke analize. Nastavu držala na Vojno-tehničkoj akademiji. U okviru poslediplomske nastave, izvodila je nastavu iz više predmeta na kursovima Monetarni, bankarski i finansijski menadžment, Fiskalni menadžment, Statistička analiza, Upravljanje i informacioni sistemi i Aktuarstvo. Rukovodila poslediplomskim kursom Demografija. Na usavršavanju boravila u Parizu (1985), Londonu (1979. i 1981.) i Istočnom Berlinu (1985). Bila rukovodilac Odeljenja Ekonomskog fakulteta u Šapcu (1979-1981), prodekan (1989-94), dekan Ekonomskog fakulteta (1994-96) i prorektor Univerziteta u Beogradu (1995-97). Član Ekonomskog saveta Vlade SR Jugoslavije (1995). Šef Katedre za statistiku i matematiku. Član Naučnog društva ekonomista (i član Predsedništva) i Statističkog društva Jugoslavije. Od 1997. član Filozofsko-ekonomskog društva Lomonosov u Moskvi. Obavljala dužnost direktora Republičkog zavoda za statistiku. U penziji je od 2001. godine.

I pored ovako bogatog i posvećenog profesionalnog života, profesorica Žižić je uspela paralelno, kao malo ko drugi, a to je posebno teško na ovom podneblju kada su u pitanju žene, da živi i jedan primeran porodični život. U porodičnom gnezdu koje je svila sa svojim Radulem, podigla je Zorana i Branka, doživela da od njih vidi unuke i otišla svom Radulu ponovo u zagrljaj.

Na kraju, nakon ove čiste biografske faktografije, ne možemo a da se ne setimo da je profesorica Mileva Žižić, ostavljajući za sobom dubok trag uglednog profesora, održavši hiljade časova i ispita, rukovodioca, doprinevši unapređenju uslova rada i kvaliteta nastave na Fakultetu i Univerzitetu, politički angažovane ličnosti, delujući u skladu sa svojim stavovima i principima.

Opraštamo se, govoreći „večnaja pamjat“, što bi značilo, a u slučaju profesorice Mileve Žižić to je nesumnjivo, „Zauvek u najlepšem sećanju generacija đaka, studenata, kolega,...“

Slava joj!

CIP – Каталогизacija u publikaciji
Narodna biblioteka Srbije, Beograd

311

STATISTIČKA revija / glavni i odgovorni
urednik Miladin Kovačević. - God. 1, br. 1
(1951) - . - Beograd (Milana Rakića 5) :
Statističko društvo Srbije, 1951 - (Beograd
: Republički zavod za statistiku Srbije). -
25 cm

Tromesečno
ISSN 0039-0534 = Statistička revija
COBISS.SR-ID 46175239

STATISTICAL REVIEW

YEAR LVIII

Number 1-4

FOR 2009

CONTENS

ORIGINAL PAPER

An it definition of a statistical survey..... *Branko Jiriček* 3

EXPERT ARTICLES

Data editing and imputation in the survey on road freight transport in
 Serbia *Jelena Kovijanić* 15
 Pilot Survey on road freight transport in Serbia, 2008.....
 *Jelena Kovijanić, Olga Meloski Trpinac, Željko Lesić* 24

REVIEW

Contemporary trends in lingvo-statistics..... *Nikola Dobrić* 43

REVIEW OF SCIENTIFIC EVENT

XXXV Symposium on Operational Researches..... *Luka Filipović, PhD.* 51

IN MEMORIAM

Vladislav Milošević, *PhD.* *Branislav Boričić, PhD.* 57
 Vladislav Milošević, *PhD.* *Ljiljana Petrović, PhD.* 59
 Mileva Žižić, *PhD.* *Dubravka Pavličić, PhD.* 61
 Mileva Žižić, *PhD.* *Branislav Boričić, PhD.* 65