

# GML BASED MANIPULATION OF GEODETIC AND MECHANICAL DATA

Marko ŠLJIVARIĆ, Milan REZO, Nikola KRANJČIĆ, Danko MARKOVINOVIĆ

**Abstract:** Ever since the Rulebook on cadaster of infrastructure has been published there has been constantly growing need for efficient data manipulation. However, data manipulation has been an issue for a very long period of time. This paper will provide solution in manipulation of geodetic and mechanical data for City of Zagreb gas plant. All the solutions are given in Geography markup language, GML.

**Keywords:** gas pipeline; geography markup language; geodetic & mechanical data manipulation; Rulebook on Cadaster of infrastructure

## 1 INTRODUCTION

Since the computation power became sufficient to process geospatial data, many geo-data systems have been developed. Some of the early geo-data systems were BUDS [1], GADS [2, 3], NISP [4] or FIRS [5], all designed to retrieve and display facilities using their own query language and internal data structures [6]. [6] proposed approach on manipulation and display of geographic data where all geographic data is treated as relations. [7] proposed model for storage and manipulation of geographic data called GISER whose main advantage was use of procedure-valued attributes. However, all of the authors were developing their own query languages and data structures. Since this was major issue in interoperability, experts needed to find structured data in order to resolve issues in geo-spatial data organization, interoperability, query and processing. First sign of resolving this issue was with the emergence of Geography Markup Language (GML) in year 2000 [8].

[9] created a GML-based interoperable geographical database for conservation of the lunan stone forest landscape to easily manipulate and query different types of data and after the testing the database concluded that it is interoperable with different spatial functions. [10] proposed a framework for feature-level geospatial data sharing, for transportation network data. [10] concluded that proposed framework has a great potential for sharing, accessing, extracting, transporting and displaying distributed geospatial data at the feature level, all in a web browser. [11] proposed a framework for GML storage, processing and interface for users or applications that need geo-spatial data manipulation service. All of the authors' general conclusions were that the GML is useful tool for providing framework to manipulate geospatial data. This paper will present framework to manipulate geodetic and mechanical data, similar to work done in [10] but with upgraded GML schemas and queries.

## 2 CASE STUDY

Over the years, geodetic experts have been collecting data for different infrastructure data, such as gas pipelines, water pipelines, communication cables, etc. Faculty of

Geodesy in Zagreb has been collecting gas pipeline data for the City of Zagreb gas plant in period from 2005 to 2007. Such data will be used in this paper to show how to accelerate manipulation on such data. In order to develop a solution, what type of data geodetic expert should provide to fulfil the task first needs to be defined. According to the City of Zagreb gas plant, geodetic expert should provide them with four different files. Cartographically well-structured file is file which has all the necessary elements to be used in cadaster and to be manipulated in the GML schema. Fig 1. represents cartographically well-structured CAD file with chartered path of gas line and with mechanical data.

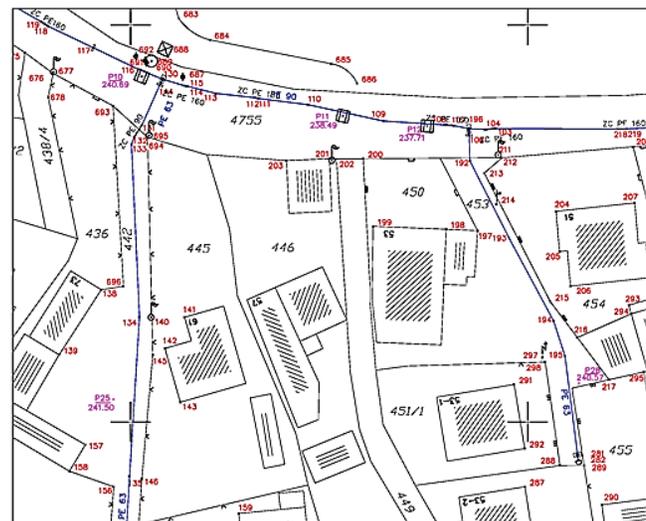


Figure 1 Digital cadaster plan with gas pipeline path

Second file is well defined ASCII file of path or part of the path, as presented in Fig 2. It should contain contractor name, point number, coordinates, altitude, mechanical data code.

Fig 3. shows list of general control network, which is also well defined ASCII file.

The final, fourth file, is list of completed work, with length of gas lines in each of planned path and with information about types of gas lines, as presented in Fig 4.

```

LIN0901063
#COM GPZ 00000006 PE063 02/07/2005 Krvarić do kóbr67
GeoEDT 130 5571404.000 5077843.180 239.3000R 240.450 082 154.30
GeoEDT 144 5571403.950 5077843.079 239.3000R 240.450 079 245.88
GeoEDT 131 5571401.470 5077837.510 239.9000R 241.040 053 237.71
GeoEDT 132 5571400.440 5077835.880 240.1200R 241.220 053 237.71
GeoEDT 133 5571400.100 5077834.970 240.2300R 241.410
GeoEDT 134 5571401.110 5077813.240 241.6200R 242.720
GeoEDT 135 5571399.860 5077793.130 239.8400R 241.000
GeoEDT 136 5571399.250 5077783.040 239.4400R 240.600
GeoEDT 137 5571399.650 5077775.330 238.9900R 240.090
GeoEDT 148 5571400.650 5077768.410 237.9600R 239.060 079 278.29
GeoEDT 149 5571401.760 5077761.370 236.9300R 238.030
GeoEDT 167 5571403.470 5077753.010 235.6500R 236.800
GeoEDT 168 5571406.060 5077742.050 234.0700R 235.170 072 103.27
GeoEDT 169 5571404.480 5077741.590 234.2200R 235.370 053 196.07
GeoEDT 170 5571402.710 5077741.080 234.2800R 235.430 053 196.07
GeoEDT 171 5571400.150 5077739.140 234.0900R 235.190
GeoEDT 172 5571383.150 5077733.760 233.6600R 234.760 079 202.52
GeoEDT 173 5571381.510 5077733.080 233.4800R 234.570 083 199.29
GeoEDT 174 5571381.110 5077732.940 233.4000R 234.490 076 201.22

```

Figure 2 Example of starting LIN file (ASCII file)

```

P1 5571194.420 5077521.550 240.070
P2 5571188.200 5077605.140 237.410
P3 5571172.310 5077667.530 239.710
P25 5571397.840 5077802.840 241.500
P26 5571401.100 5077775.540 240.280
P27 5571403.870 5077735.410 235.020
P28 5571456.480 5077804.890 240.570
GPS2362 5571892.020 5077322.110 178.940

```

Figure 3 List of general control network

```

#COM GPZ 00000001 PE090 28/07/2005 Krvarić
kD=1149,75 [m], hD=1143,04 [m], dH=-60,9 [m]
#COM GPZ 00000002 PE063 18/07/2005 Krvarić do kóbr 4/2
kD=120,85 [m], hD=120,6 [m], dH=2,67 [m]
#COM GPZ 00000003 PE063 06/07/2005 Krvarić do kóbr 18
kD=44,63 [m], hD=44,54 [m], dH=-1,09 [m]
#COM GPZ 00000004 PE063 18/07/2005 Krvarić prema kóbr 40
kD=10,56 [m], hD=10,41 [m], dH=-1,63 [m]
#COM GPZ 00000005 PE063 02/07/2005 Krvarić do kóbr 53/2
kD=44,92 [m], hD=44,74 [m], dH=3,29 [m]
#COM GPZ 00000006 PE063 02/07/2005 Krvarić do kóbr67

```

Figure 4 List of completed work

Having in mind that such structure depends on memory capabilities and PC processing power, next chapters will describe Geography markup language as a better solution to format data and structure. Chapter 4 will provide short overview of new rulebook on cadastre of infrastructure and in Chapter 5 will present proposed GML schema for data management.

### 3 GEOGRAPHY MARKUP LANGUAGE

The OGC (Open Geospatial Consortium) is an international not-for-profit organization committed to making quality open standards for the global geospatial community. These standards are made through a consensus process and are freely available for anyone to use to improve sharing of the world's geospatial data [12]. OGC standards are used in a wide variety of domains including Environment, Defense, Health, Agriculture, Meteorology, Sustainable Development and many more [13]. Among all the other standards for this paper most interesting standard is one on Geography Markup Language, or shorter GML. The Geography Markup Language (GML) is an XML grammar for expressing geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. As with most XML based grammars, there are two

parts to the grammar – the schema that describes the document and the instance document that contains the actual data. A GML document is described using a GML Schema. This allows users and developers to describe generic geographic data sets that contain points, lines and polygons. However, the developers of GML envision communities working to define community-specific application schemas [14] that are specialized extensions of GML. Using application schemas, users can refer to roads, highways, and bridges instead of points, lines and polygons. If everyone in a community agrees to use the same schema they can exchange data easily and be sure that a road is still a road when they view it [15], or as shown in this paper, gas pipelines.

GML represents expanded XML schema, which can define geospatial object classes, features. GML defines content and it does not shuffle content and its presentation [16]. Feature Types are specified in GML application schema. Main application schema *gml.xsd* contains all other GML schemas, but for the majority of usage *feature.xsd* schema is recommended. Current version of GML is 3.2.2.

*Feature.xsd* schema contains sub-schemas: *geometryBasic2d.xsd* schema which defines basic two dimensional geometry type of data and includes *geometryBasic0d1d.xsd* schema for scalar and one dimensional data types as well their connection with values defined in *measures.xsd* in which these values are defined. Next there is a sub-schema *units.xsd* in which units of values are defined. *Dictionary.xsd* schema has all the terminology, and contains *gmlbase.xsd* schema where are all other types of simple, complex or abstract data. Simple data types are defined in *basicTypes.xsd* schema and how this data links between one and another is defined in *xlinks.xsd* schema. *Temporal.xsd* schema is component of *feature.xsd* schema which gives time component to definition of data.

Other GML schemas in *gml.xsd* schema include *dynamicFeature.xsd* schema, *direction.xsd*, *topology.xsd*, *geometryComplexes.xsd*, *geometryAggregates.xsd*, *geometryPrimitives.xsd*, *coverage.xsd*, *valueObjects.xsd*, *grids.xsd*, *coordinateReferenceSystems.xsd*, *datums.xsd*, *coordinateSystems.xsd*, *coordinateOperations.xsd*, *referenceSystems.xsd*, *dataQuality.xsd*, *observation.xsd*, *defaultStyle.xsd* and *temporalReferenceSystems.xsd* and *temporalTopology.xsd* schemas.

Fig 5. shows dependence between GML schemas.

As shown on Fig 5. in whole GML there are only seven schemas not depending on any other schema. They are root schemas of GML:

- *observation.xsd*
- *dynamicFeature.xsd*
- *coverage.xsd*
- *topology.xsd*
- *defaultStyle.xsd*
- *coordinateReferenceSystems.xsd*
- *temporalReferenceSystems.xsd*

Class hierarchy in GML is shown in Fig 6.

On Fig 7 there are presented geometry classes in GML.

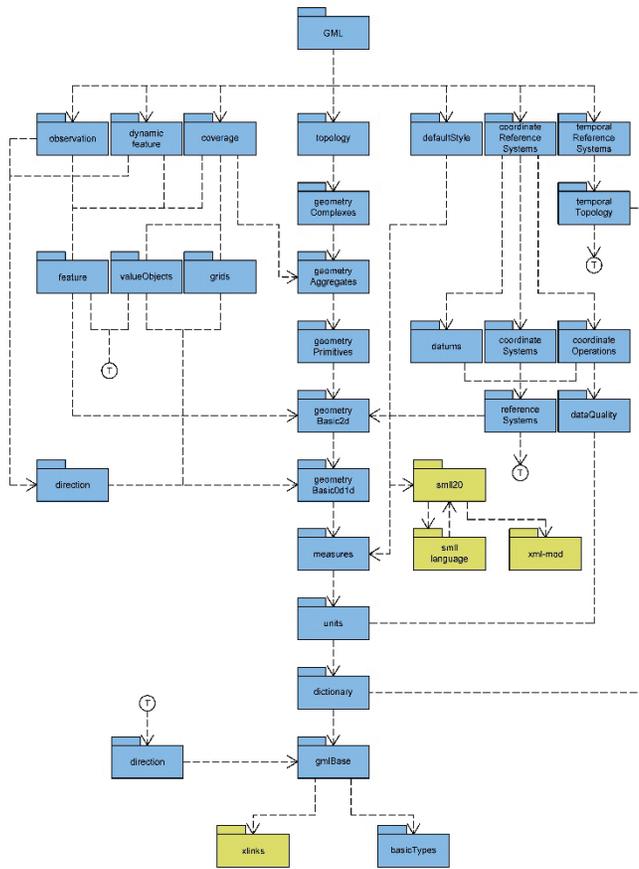


Figure 5 Schema Dependencies [16]

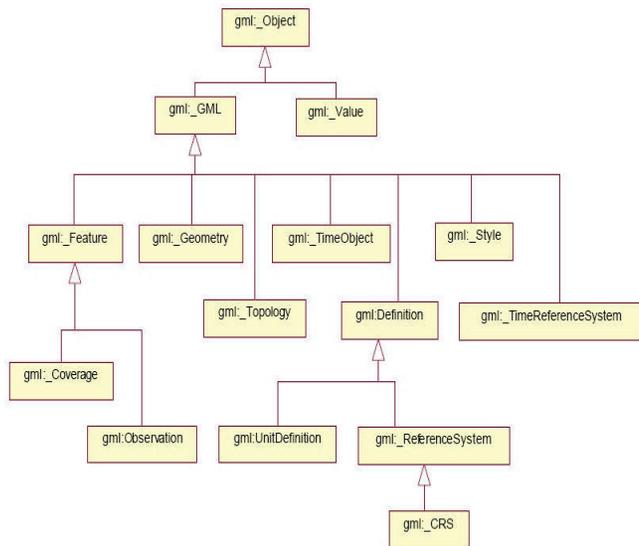


Figure 6 GML Class Hierarchy [16]

#### 4 RULEBOOK ON CADASTER OF INFRASTRUCTURE

On 23<sup>rd</sup> March 2017 the principal of Croatian state geodetic administration brought out Rulebook on cadaster of infrastructure. In the rulebook, how and which data should be collected for cadaster of infrastructure is pointed out. Since in this paper enhanced manipulation of gas pipeline data is shown, further chapter will only present parts of rulebook

regarding gas pipeline. For gas pipeline network it is necessary to obtain data about production, transport, distribution network and connection to such network [17]. Data about basic technical characteristic refers to the type of pressure, material and pipe diameter in millimetres. For every type of infrastructure the data should be collected in Croatian Terrestrial Reference System 1995.55 (HTRS96/TM) and in Croatian Height Reference System 1971.5 (HVRS71). Data about owners, or about administrator of infrastructure that should be collected are name, address, id number and other data necessary for communication such as official mail address, phone number, etc. [17]

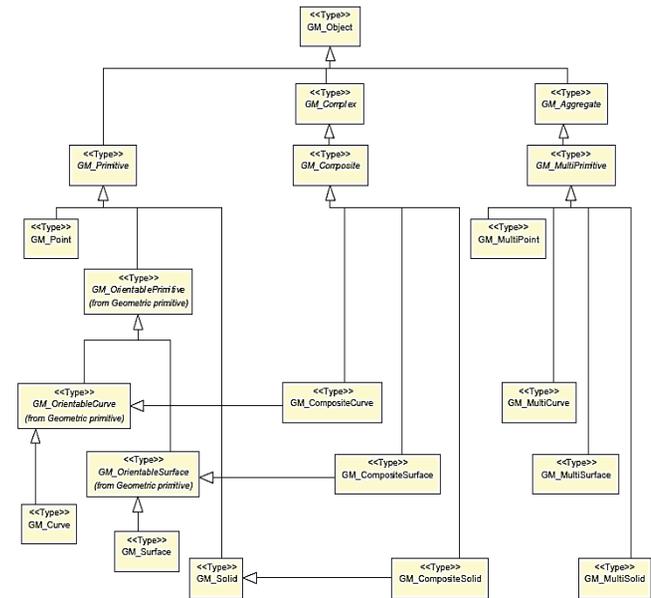


Figure 7 GML Geometry Classes [16]

Cadaster of infrastructure consist of written and numerical part. Written part is:

- list of infrastructure,
- list of owners and infrastructure administrators.

Numerical part is:

- map of infrastructure,
- collection of geodetic elaborate of infrastructure [17]

List of infrastructure contains:

- number of entries within one year and label of confirmed geodetic elaborate of infrastructure of cadaster,
- type of infrastructure and its current use (in usage, abandoned or removed),
- id number of infrastructure and its objects,
- data about owner, or administrator of infrastructure
- names of cadastral municipalities in which infrastructure is settled,
- names of municipalities, settlement and streets in which infrastructure is built,
- house number if the infrastructure is built as home connection,
- length of built infrastructure [17].

List of owners and administrators contains of:

- number of entries of owners or infrastructure administrators,
- data about owners or infrastructure administrators,
- type of infrastructure [17].

Map of infrastructure is graphical plan of infrastructure and related objects of infrastructure. Graphic basic of infrastructural plan is digital orthophoto map in measure 1:2000 or 1:5000 with overlapped digital cadaster map [17].

According to the Rulebook on cadaster of infrastructure geodetic survey contains:

- survey of points of infrastructure in horizontal manner,
- survey of points of infrastructure in vertical manner (for drainage at the bottom of pipe, for other infrastructure at the top of pipes),
- survey of intersections of new infrastructure with existing infrastructure,
- survey of existing infrastructure objects,
- obtaining data about infrastructure owners or infrastructure administrators [17].

Taking the aforementioned into consideration, similar gathering of infrastructural data has been done from 2005 to 2007. Main difference is in geodetic projection since in 2004 was new state horizontal projection datum defined. Since there was a transitional period, the data could be still obtained in old Gauss-Krueger projection.

In the next chapter there is a proposed solution for standardizing process of manipulation of obtained geodetic data for cadaster infrastructure. Main difference in proposed solution is in geodetic projection, but other data is standardized and synced, even in 2005, with new Rulebook of cadaster infrastructure.

## 5 PROPOSED GML SCHEMA

To begin in standardizing manipulation of geodetic data every well-structured GML file should start with schema definition. In example below, the schema used is *feature.xsd*.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns:gpz="C:\0000\ProcGeoInf\Oxygen10\"
  xmlns:gml="http://www.opengis.net/gml"
  targetNamespace="C:\0000\ProcGeoInf\Oxygen10\"
  elementFormDefault="qualified">
  <xs:import namespace="http://www.opengis.net/gml"
    schemaLocation="feature.xsd"/>
```

In order to standardize processes, after defining schema, what type of data is expected and how that data should be structured in order to develop faster model of processing input data should be defined in GML file. Next example shows defining type for point data.

```
<xs:element name="Point" type="gpz:PointType"
  substitutionGroup="gml:_Feature"/>
```

After the whole GML file is completed and data input standardized, the GML should be evaluated. Due to GML basically being XML, it can be evaluated in special software, like XML editor Oxygen 10 or XML Spy. GML document should be well formed and validated. A well-formed document is document that follows the syntax rules specified by the XML 1.0 specification [18]. Valid GML document must be well formed, and it must conform to a document type definition [19]. There are two different document type definition that can be used with GML:

- DTD – The original Document Type Definition
- XML Schema – An XML-based alternative to DTD

In this paper XML Schema *feature.xsd* as document type definition is used. Since the schema is located on official GML pages online, it can be assumed schema does not have errors.

After checking the GML document with SaxonSA and MSXML4.0 modules in Oxygen 10.1 [20], it is clear that GML schema is well formed and validated, while XMLSpy [21] cannot find *SimpleLink* declaration in *gmlBase.xsd* schema which reference to schema *xlinks.xsd*.

## 6 RESULTS

After definition of GML file and filing out GML file with data there are preliminary results. In municipality Mikulići there is 11.5 km of main pipeline and GML data takes about 500 kB of memory space on hard disc. In whole cadaster municipality around 5000 home connections were built. Data for each home connection contains ten points and mechanical data for each connection. Total data for municipality Mikulići is around 25 mB. In the City of Zagreb there are forty cadaster municipalities under the jurisdiction of city gas plant. Using simple math there is at least 10 gB of data. This presents a challenge in data manipulation, data safety, speed of searching, modifying and storing such data. Since the GML files are well structured and well defined, above mentioned issues are easily solved with use of transaction technology which is basic in every relation database, object – relation database, or object-oriented database.

## 7 CONCLUSION

New Rulebook on Cadaster of infrastructure standardize the way data should be gathered. Having in mind that standardizing the process of gathering data results in standardizing the processes of data manipulation. This paper gives an overview of Geography Markup Language, new Rulebook on Cadaster of infrastructure. Proposed GML schema is a direction in which all data manipulation should strive. Main advantage in the data stored using GML is in standardized process of data manipulation, higher data safety and greater speed of searching, modifying and storing such data. Schema and framework proposed in this paper have a better response time and take less memory compared to frameworks proposed in Introduction.

## 8 REFERENCES

- [1] Macri, P. (1973, November). BUDS: The Berkeley Urban Data System. Electronics Research Laboratory, University of California, Berkeley, Memorandum M412.
- [2] Christiani, E. J., et al. (1973, March). An Interactive System for Aiding Evaluation of Local Government Policies. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(2), 141-146. <https://doi.org/10.1109/TSMC.1973.5408495>
- [3] Mantey, P. E., et. al. (1973, June). Information for Problem Solving: The Development of an Interactive Geographic Information System. *IEEE Conference on Communication, vol. II*, Seattle, Washington.
- [4] Parker, J. L. (1971, November). Information Retrieval with Large Scale Geographic Data Bases," *Proc. 1971 ACM-SIGFIDET Workshop on Data Description, Access and Control*, San Diego, Ca. <https://doi.org/10.1145/1734714.1734739>
- [5] Depta, D. J. & Irwin, G. M. (1974, March). *FIRS II - Design Requirements Component*. Weyerhaeuser Company, Woods Product Information Systems, Tacoma, Wash.
- [6] Berman, R. & Stonebraker, M. GEO-OUEL A System for the Manipulation and Display of Geographic Data, University of California, Berkeley.
- [7] Shekhar, S., Coyle, M., Goyal, B., Liu, D.-R., & Sarkar, S. (1997). Data Models in Geographic Information Systems. *Communications of the ACM*, 40(4). <https://doi.org/10.1145/248448.248465>
- [8] <http://www.opengeospatial.org/> (Accessed on 21.05.2019.)
- [9] Zhang, C., Day, M. J., & Peng, Z-R. (2003). GML-Based Interoperable Geographical Databases. *Cartography*, 32(2), 1-16. <https://doi.org/10.1080/00690805.2003.9714249>
- [10] Peng, Z-R. (2005). A proposed framework for feature-level geospatial data sharing: a case study for transportation network data. *International Journal of Geographical Information Science*, 19(4), 459-481. <https://doi.org/10.1080/13658810512331319127>
- [11] Wang, W., Wang, F., Qian, Z., & Zhang, L. (2011). GML Data Management: Framework and Prototype. *16<sup>th</sup> International Conference, DASFAA 2011 International Workshops*, Hong Kong, 101-111. [https://doi.org/10.1007/978-3-642-20244-5\\_10](https://doi.org/10.1007/978-3-642-20244-5_10)
- [12] <http://www.opengeospatial.org/standards/gml>
- [13] Lake, R., Burggraf, D., Trninic, M., & Rae, L. (2004). *Geography Mark-Up Language: Foundation for the Geo-Web*, New York: Wiley.
- [14] <http://old.geotools.org/19708.html> (Accessed on 21.05.2019.)
- [15] <http://map.sdsu.edu/geog583/lecture/unit-6.html> (Accessed on 21.05.2019.)
- [16] Rezo, M. & Šljivarić, M. XML/GML DBMS utemeljeno pohranjivanje i upravljanje geodetskim i strojarskim podacima, Seminarski rad, Geodetski fakultet Zagreb, 2009.
- [17] Državna geodetska uprava: Pravilnik o katastru infrastrukture, 2017, NN 29/2017
- [18] Son, J. H., Kim, J. S., & Kim, M. H. (2005). Extracting the workflow critical path from the extended well-formed workflow schema. *Journal of Computer and System Sciences*, 70, 86-106. <https://doi.org/10.1016/j.jcss.2004.07.001>
- [19] [https://www.w3schools.com/xml/xml\\_validator.asp](https://www.w3schools.com/xml/xml_validator.asp) (Accessed on 21.05.2019.)
- [20] <http://www.oxygenxml.com/> (Accessed on 21.05.2019.)
- [21] <https://www.altova.com/xmlspy-xml-editor> (Accessed on 21.05.2019.)

## Authors' contacts:

**Marko ŠLJIVARIĆ**, doc. dr. sc.  
European University Brčko,  
Technical Faculty,  
Bijeljinska cesta 72-74  
76100 Brčko, Bosnia and Herzegovina  
Tel: +385 91 4444 998  
E-mail: marko.sljivacic@gmail.com

**Milan REZO**, Ph.D. Assistant Professor  
University of Zagreb,  
Faculty of Geotechnical Engineering,  
Hallerova aleja 7,  
42000 Varaždin, Croatia  
Tel. +385 42 408 950  
E-mail: mrezo@gfv.hr

**Nikola KRANJČIĆ**, mag. ing. geod. et geoinf.  
University of Zagreb,  
Faculty of Geotechnical Engineering,  
Hallerova aleja 7,  
42000 Varaždin, Croatia  
Tel. +385 42 408 950  
E-mail: nkranjcic@gfv.hr

**Danko MARKOVINOVIĆ**, Ph.D. Assistant Professor  
University North,  
Jurja Križanića 31b  
42000 Varaždin, Croatia  
E-mail: danko.markovinovic@unin.hr