

# Geographia Technica



Technical Geography  
an International Journal for the Progress of Scientific Geography

**Volume 14, Geographia Technica No. 1/2019**

[www.technicalgeography.org](http://www.technicalgeography.org)

**Cluj University Press**

## Editorial Board

Okke **Batelaan**, Flinders University Adelaide, Australia  
Yazidhi **Bamutaze**, Makerere University, Kampala, Uganda  
Valerio **Baiocchi**, Sapienza University of Rome, Italy  
Gabriela **Biali**, "Gh. Asachi" University of Iasi, Romania  
Habib **Ben Boubaker**, University of Manouba, Tunisia  
Gino **Dardanelli**, University of Palermo, Italy  
Ioan **Donisa**, "Al.I.Cuza" University of Iasi, Romania  
Qingyun **Du**, Wuhan University, China  
Massimiliano **Fazzini**, University of Ferrara, Italy  
Oleg **Horjan**, Agrarian State University, Republic of Moldova  
Edward **Jackiewicz**, California State University, Northridge CA, USA  
Shadrack **Kithiia**, University of Nairobi, Kenya  
Jaromir **Kolejka**, Masaryk University Brno, Czech Republic  
Muh Aris **Marfai**, Universitas Gadjah Mada, Yogyakarta, Indonesia  
Béla **Márkus**, University of West Hungary Szekesfehervar, Hungary  
Jean-Luc **Mercier**, Université de Strasbourg, France  
Yuri Sandoval **Montes**, Universidad Mayor de San Andrés, La Paz, Bolivia  
Maria **Nedealcov**, Inst. of Ecology-Geography, Republic of Moldova  
Dušan **Petrovič**, University of Ljubljana, Slovenia  
Hervé **Quénot**, Université de Rennes 2 et CNRS, France  
Marieta **Staneva**, Pennsylvania State University, USA  
Wayan **Suparta** Pembangunan Jaya University, Indonesia  
Gábor **Timár**, Eötvös University Budapest, Hungary  
Eugen **Ursu**, Université de Bordeaux, France  
Changshan **Wu**, University of Wisconsin-Milwaukee, USA  
Chong-yu **Xu**, University of Oslo, Norway

## Editor-in-chief

Ionel **Haidu**, University of Lorraine, France

## Editorial Secretary

Marcel Mateescu, Airbus Group Toulouse, France  
George Costea, Yardi Systemes, Cluj-Napoca, Romania

## Online Publishing

Magyari-Sáska Zsolt, "Babes-Bolyai" University of Cluj-Napoca, Romania

# **Geographia Technica**



**Technical Geography**

**an International Journal for the Progress of Scientific Geography**

**2019 – No. 1**

**Cluj University Press**

**ISSN: 1842 - 5135 (Printed version)**

**ISSN: 2065 - 4421 (Online version)**

© 2019. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording or any information storage and retrieval system, without permission from the editor.

Babeş-Bolyai University  
Cluj University Press  
Director: Codruța Săcelean  
Str. Hașdeu nr. 51  
400371 Cluj-Napoca, România  
Tel./fax: (+40)-264-597.401  
E-mail: [editura@editura.ubbcluj.ro](mailto:editura@editura.ubbcluj.ro)  
<http://www.editura.ubbcluj.ro/>

Asociatia Geographia Technica  
2, Prunilor Street  
400334 Cluj-Napoca, România  
Tel. +40 744 238093  
[editorial-secretary@technicalgeography.org](mailto:editorial-secretary@technicalgeography.org)  
<http://technicalgeography.org/>

Cluj University Press and Asociatia Geographia Technica  
assume no responsibility for material, manuscript, photographs or artwork.

# Contents

## *Geographia Technica*

Volume 14, Issue 1, spring 2019

*An International Journal of Technical Geography*

ISSN 2065-4421 (Online); ISSN 1842-5135 (printed)

### **MODELING THE RELATIONSHIP BETWEEN WEATHER PARAMETERS AND CHOLERA IN THE CITY OF MAROUA, FAR NORTH REGION, CAMEROON**

Mouhaman ARABI, Moise Chi NGWA  
(Maroua, Cameroon & Baltimore, USA)..... 1  
DOI: 10.21163/GT\_2019.141.01

### **EVALUATION OF SOIL EROSION PROCESS AND CONSERVATION PRACTICES IN THE PARAGOMINAS-PA MUNICIPALITY (BRAZIL)**

Denis Conrado da CRUZ, José María Rey BENAYAS, Gracialda Costa FERREIRA, André Luis MONTEIRO, Gustavo SCHWARTZ  
(Alcalá de Henares, Spain & Belém, Brazil) ..... 14  
DOI: 10.21163/GT\_2019.141.02

### **VARIABILITY ANALYSIS OF TEMPORAL AND SPATIAL ANNUAL RAINFALL IN THE MASSIF OF AURES (EAST OF ALGERIA)**

Adel KHENTOUICHE, Hadda DRIDI  
(Batna, Algeria) ..... 36  
DOI: 10.21163/GT\_2019. 141.03

### **THE MULTIPLE DATA AND GEOGRAPHIC KNOWLEDGE APPROACH TO A LIQUID TOXIC ROAD ACCIDENT MITIGATION – THE TWO BLOCKS GIS DATA PROCESSING FOR AN OPERATIVE INTERVENTION**

Jaromír KOLEJKA, Petr RAPANT, Jana ZAPLETALOVÁ  
(Brno & Ostrava-Poruba, Czech Republic) ..... 49  
DOI: 10.21163/GT\_2019. 141.04

### **ASSESSMENT OF METAL CONTAMINATION IN COASTAL SEDIMENTS OF THE LÉVRIER BAY AREA, ATLANTIC OCEAN, MAURITANIA**

Mohamed El Houssein LEGRAA, Abdelkader MOHAMED SALECK, Ali YAHYA DARTIGE, Mohamed Lemine CHEIKH ZAMEL, Mohamed Mahmoud Ould ABIDINE, Hassan ERRAOUI, Zeinebou SIDOUMOU  
(Nouakchott, Mauritania & Tétouan, Morocco) ..... 65  
DOI: 10.21163/GT\_2019. 141.05

**ROAD NETWORK BASED COMMUNITY DETECTION. CASE STUDY FOR AN EASTERN REGION OF AUSTRO-HUNGARIAN MONARCHY**

Zsolt MAGYARI-SÁSKA (Gheorgheni, Romania) ..... 82  
DOI: 10.21163/GT\_2019. 141.06

**REGIONAL DISPARITIES IN HUNGARIAN URBAN ENERGY CONSUMPTION – A LINK BETWEEN SMART CITIES AND SUCCESSFUL CITIES**

Zoltán NAGY, Tekla SEBESTYÉN SZÉP, Dóra SZENDI  
(Miskolc, Hungary) ..... 92  
DOI: 10.21163/GT\_2019. 141.07

**GIS-BASED APPROACH TO IDENTIFY THE SUITABLE LOCATIONS FOR SOIL SAMPLING IN SINGAPORE**

Mărgărit-Mircea NISTOR, Harianto RAHARDJO, Alfrendo SATYANAGA,  
Eng-Choon LEONG, Koh Zhe HAO, Aaron Wai Lun SHAM, Hongjun WU  
(Singapore) ..... 103  
DOI: 10.21163/GT\_2019. 141.08

**DROUGHT EVALUATION WITH NDVI-BASED STANDARDIZED VEGETATION INDEX IN LOWER NORTHEASTERN REGION OF THAILAND**

Tanutdech ROTJANAKUSOL, Teerawong LAOSUWAN  
(Maha Sarakham, Thailand) ..... 118  
DOI: 10.21163/GT\_2019. 141.09

**DECISION SUPPORT FOR CYCLING INFRASTRUCTURE PLANNING: A CASE STUDY IN PŘEROV, CZECHIA**

Aleš RUDA (Brno, Czech Republic)..... 131  
DOI: 10.21163/GT\_2019. 141.10

**CHARACTERIZATION OF HEAT WAVES: A CASE STUDY FOR PENINSULAR MALAYSIA**

Wayan SUPARTA, Ahmad Norazhar Mohd YATIM  
(South Tangerang, Indonesia & Sabah, Malaysia) ..... 146  
DOI: 10.21163/GT\_2019. 141.11

**MOUNTAIN TRACKS DEVELOPMENT METHODOLOGY FOR ADVENTURE RECREATION ACTIVITIES IN GURGHU MOUNTAINS, ROMANIA**

Adrian TORPAN, Mihai VODA  
(Jönköping, Sweden & Targu Mures, Romania) ..... 156  
DOI: 10.21163/GT\_2019. 141.12

## MODELING THE RELATIONSHIP BETWEEN WEATHER PARAMETERS AND CHOLERA IN THE CITY OF MAROUA, FAR NORTH REGION, CAMEROON

Mouhaman ARABI<sup>1</sup>, Moise Chi NGWA<sup>2</sup>

DOI: 10.21163/GT\_2019.141.01

### ABSTRACT:

Locally measured variables such as temperature and rainfall have been positively associated with increased cholera incidence in multiple studies. The study examined the effect of rainfall, temperature and relative humidity on cholera incidence in the City of Maroua, located in the Far North of Cameroon. The relative individual contribution of rainfall, temperature and humidity and then their combined contribution on the occurrence of cholera were examined. Using monthly time series of cholera epidemiological data, average monthly rainfall (mm), average monthly air temperature (degree Celsius), and average monthly relative humidity (percent) data from 1996 to 2011. We implemented Generalized Additive Modeling (GAM) procedures to measure the contribution of each weather parameter to the incidence of cholera and identified the most influential parameter on cholera incidence. We found that taken individually, rainfall, temperature and humidity are correlated with cholera incidence but temperature seems more determinant since it has a higher deviance explained (26%). Furthermore, the association between temperature and rainfall in the multivariate model with interaction has the highest deviance explained is 64.4%, the lowest AIC is 1911.963, and the highest R<sup>2</sup> (0.6). These results indicate that statistical time series models in general and the Generalized Additive models in particular should lead to a better understanding of the disease mechanism that can assist in the planning of public health interventions. These results contribute also to the growing debate on climate and cholera.

*Key-words:* Cholera, Maroua, Cameroon, GAM model.

### 1. INTRODUCTION

Cholera is an acute diarrheal infection caused by ingestion of food or water contaminated with the bacterium *Vibrio cholerae*. Cholera remains a global threat to public health and an indicator of inequity and lack of social development. Researchers have estimated that every year, there are roughly 1.3 to 4.0 million cases, and 21 000 to 143 000 deaths worldwide due to cholera (Ali et al., 2015). The disease which has a short incubation period of two hours to five days and this enhances the potentially explosive pattern of cholera outbreaks (WHO, 2012) has been scrutinized since the birth of epidemiology, and it is still a subject of intense interest for modern-day epidemiologists (Codeço & Coelho, 2006). During the 19<sup>th</sup> century, cholera spread across the world from its original reservoir in the Ganges delta in India. Seven cholera pandemics have killed millions of people across all continents with the seventh pandemic that began in 1961 in the Celebes, Indonesia,

---

<sup>1</sup> University of Maroua, Advanced School of Engineering, Department of Environmental Sciences, P.O. Box: 46 Maroua, Cameroon; mouhamanarabi@hotmail.com

<sup>2</sup> Johns Hopkins Bloomberg School of Public Health, Department of International Health, 21205, Baltimore, Maryland, USA; mngwa2@jhu.edu

currently ongoing. Recently, its re-emergence has been noted in parallel with the ever-increasing size of vulnerable populations living in unsanitary conditions, such as peri-urban slums, as well as camps for internally displaced people or refugees, where minimum requirements of clean water, sanitation, and hygiene are not met (WHO, 2012).

Cholera has a strong relationship with weather events (Lipp et al., 2002) but displays a complex relationship to extremes of weather conditions (McMichael, 2015). Locally measured variables such as temperature and rainfall have been positively associated with increased cholera incidence in multiple studies (Reyburn et al., 2011). The relationship between climate and cholera has been a subject of study for a very long time. However, during the past 2 decades, development of modern tools and technologies has led to fascinating observations sparking new interest in the role of weather and climate in infectious disease dynamics (Lipp et al., 2002). Climate constrains the range of infectious diseases, while weather affects the timing and intensity of outbreaks (Dobson & Carper, 1993). A link was observed between rainfall and the occurrence of cholera in the city of Douala, Cameroon (Guévert et al., 2010) but no statistical measure of the relationship was provided.

In the Equatorial Monsoon climate zone of Cameroon, which includes the city of Douala, a Spatial Autoregressive Poisson regression model was used (Ngwa et al., 2016-a) to identify statistically significant associations between the risk of transmission and several environmental factors, including the presence of major water body, highway, as well as the average daily maximum temperature and precipitation levels over the preceding two weeks. The synergistic effect of sunshine hours and temperature in cholera outbreaks was demonstrated using monthly means time series data for cholera in Matlab in Bangladesh (Islam et al., 2009). A study of the influence of rainfall and temperature on the appearance and increase in the number of cases of cholera outbreak in Lusaka Zambia using an explicative model and time series analysis found that all epidemics showed a seasonal trend coinciding with the rainy season (November to March) and attributed 4.9% of the risks to temperature and 2.4% to rainfall (Fernandez et al., 2009). Another study in Kolkata India (WHO, 2011) examined the relationship between temperature (°C), rainfall (mm), and relative humidity (%) and occurrence of diarrhoea and cholera using 10 years (1999–2008) retrospective data. Using univariate descriptive analysis of the collected data followed by univariate and bivariate time series analysis and ARIMA modeling, they found that some relationships between the outcome variables and the predictor variables might exist but could neither determine nor explain the exact nature of the relationships due to some important limitations of this retrospective study.

A study on the impact of climate variability on health in the Far North of Cameroon (Kometa et al., 2013) revealed a strong positive correlation between changes in rainfall, temperature, humidity and the incidence of vector-borne diseases particularly cholera and meningitis but did not provide an appropriate statistical correlation analysis. Still, in the Sudano-Sahelian climate subzone of Cameroon statistically significant relationships were found between average daily maximum temperature and rainfall levels and cholera transmission with a two weeks lag (Ngwa et al., 2016-b). However, the relative contribution of each of the weather parameters (Rainfall and temperature) to the occurrence of cholera in the City of Maroua is poorly understood. Likewise, we found no study that has quantified the relation between relative humidity and the incidence of cholera in the City of Maroua. As such, gaps in knowledge exist as to the relative contribution of rainfall, temperature, and relative humidity in the City of Maroua. This study addresses this gap and examined the relationship between rainfall, temperature and relative humidity and cholera



incidence in the city of Maroua, Far North Cameroon. In other words, the study intended to answer the following questions: How do rainfall, temperature and relative humidity and cholera incidence correlate in the city of Maroua and what is the individual contribution of each weather parameter to the incidence of cholera? Which of rainfall, temperature and relative humidity has the strongest influence on cholera incidence in the city of Maroua? Answers to these questions could help better understand the disease mechanisms in this sahelian city with implications on the operational planning public health interventions strategies. The paper has therefore two main objectives: 1) examine the individual contribution of rainfall, temperature and relative humidity to the occurrence of cholera; 2) examine the combined contribution of all three parameters and the effect of interaction between them.

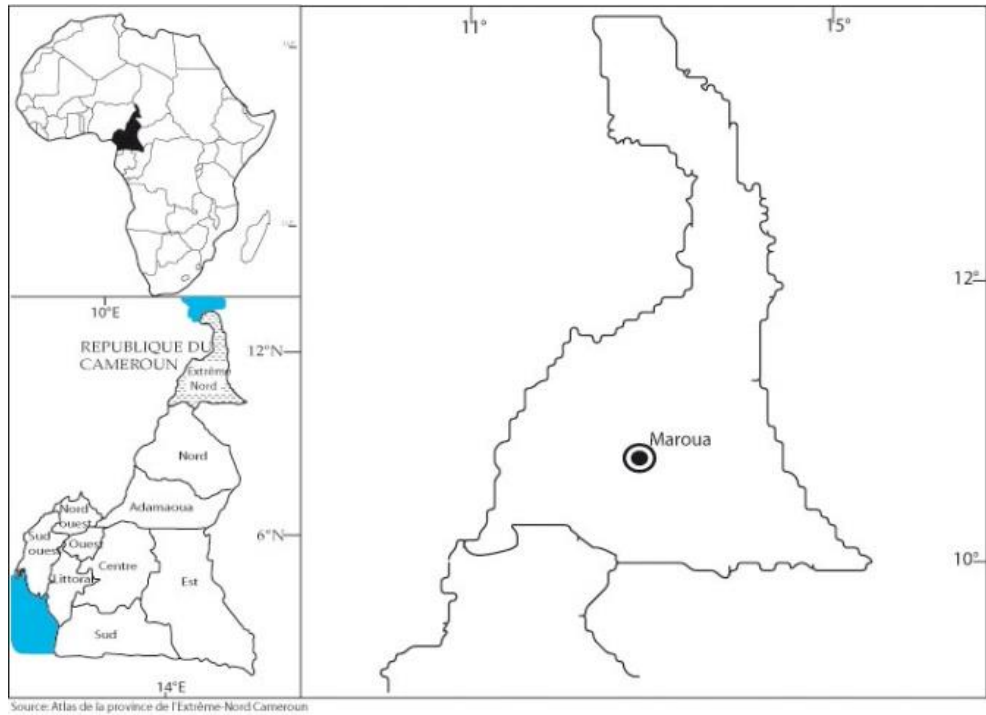
To explore the relationships between the response and explanatory variables, Generalized Additive Models (GAM) have provided a flexible framework of modelling (Osei et al., 2012). The study of the geographical distribution of disease incidence and its relationship to potential risk factors has provided, and continues to provide, rich ground for the application and development of statistical methods and models (Bailey, 2001). As emphasized by Haidu (2016), the use of statistical techniques in this paper contributes to the development of research in Technical Geography.

## **2. THE STUDY SETTING AND DATA**

The study is conducted in the city of Maroua, the capital city of the Far North Region of Cameroon (**Fig. 1**). It lies between latitude  $10^{\circ}35'50''N$  and  $10^{\circ}35'50.0''N$  and longitude  $14^{\circ}18'57''E$  and  $14^{\circ}18'57.0''E$ . The average altitude in the city is 423 m with a surface area of 466500 ha (4665 km<sup>2</sup>). The total population was estimated in 2010 at about 215 000 inhabitants (BUCREP, 2010). The climate is Sudano-Sahelian and the average annual temperature is 28°C (temperatures reach their highest levels between January and May), with actual temperatures fluctuating throughout four distinct seasons. There is one dry season and one wet season. These are further broken down based on average temperatures, yielding four distinct periods in the area: dry and relatively cool from November to January as the region experiences a shade of winter as one moves further north, dry and hot from January to April, torrentially rainy from April to June, and relatively cool and sporadically wet from June to November (IRD-MINRESI-INC, 2000).

The weekly cholera epidemiological data from 1996 to 2011 were obtained from the Regional Delegation of Public Health for the Far North. The data for all years were collected through health facility based on passive public surveillance systems, which turns active during outbreaks. In brief, during outbreaks, cholera data are collected at health facilities and sent to health districts where they are compiled and sent to the regional monitoring service (Regional Public Health Delegation) in paper format (Ngwa et al., 2016-b). The cholera clinical case data used in this study were obtained at the Regional Public Health Delegation for the Far North. The weekly epidemiological incidence data for Maroua Urban Health District was extracted from the regional database and aggregated to a monthly time series data. Average monthly rainfall (mm), average monthly air temperature (degree Celsius), and average monthly relative humidity (percent) data from 1996 to 2011 were obtained from the meteorological services (1996 to 2005) and the Institute of Agronomic Research in Maroua (2007-2011). These two dataset were merged to form a

single weather time series data. Missing values were completed by computing the average between the previous and the following month's data.



**Fig. 1.** Location of the study site.

### 3. METHODOLOGY

#### 3.1. Statistical analyses

Simple descriptive statistics are used to capture the incidence of cholera in the city of Maroua. Total number reported cases, periods of occurrence, and various waves are discussed, the mean, standard deviation, median, minimum, maximum, range, skewness, kurtosis test and standard error of various time series were explored.

#### 3.2. Modeling

The Generalized Additive Modeling (GAM) procedures were implemented to measure the contribution of each weather parameter as well as the combined contribution of all three parameters to the incidence of cholera. The GAM was used because of its ability to deal with highly non-linear and non-monotonic relationships between the response and the set of explanatory variables (Guisan et al., 2002) as it is the case in the context of this study. Furthermore, (GAM) models provide a powerful class of models for modeling nonlinear effects of continuous covariates in regression models with non-Gaussian responses (Osei et al., 2012). A GAM takes into account an additive relation between the response and predictors of a model and rules out any linearity assumption (Hastie & Tibshirani, 1990). A series of stepwise regression models were performed between rainfall, temperature and relative humidity, and the incidence of cholera (Chatfield, 1975). The GAM model is used

in this study to estimate the independent contribution of each rainfall, temperature and relative humidity and determined the effect of the interaction between these variable to cholera incidence in the city of Maroua.

**3.2.1. Model fitting**

Univariate GAM models were fitted on a stepwise procedure using thin plate regression spline to measure and determine separately the individual effect of humidity, temperature and rainfall on cholera incidence. Two multivariate GAMs were also fitted on a stepwise procedure to measure and determine the contribution of associated and interacting weather parameters. The first regression model was fitted to measure and determine the contribution of the association of “humidity against temperature”, “humidity against rainfall”, “temperature against rainfall” and “humidity against temperature vs rainfall, respectively. The second model repeated the same association but added the interaction term between the three weather parameters to the association.

**3.2.2. Evaluating the contribution of each weather parameter**

Step-wise procedures were used to evaluate the contribution of each weather parameter and identify which one influences cholera incidence the most, that is, the GAM that best describes cholera incidence. Single variable GAMs were first run separately for each of the three variables and the result was ranked in order of deviance explained. Separate univariate GAMs were therefore produced first to measure the individual relationship between cholera incidence and rainfall, cholera incidence and temperature, and cholera incidence and humidity. All the analysis and modeling process were done using the R studio computing software version097.551 (RStudio Team, 2012). The following functions were applied: “summary, ggplot2, mgcv, vis.gam”. We chose R because it is freely available open source software.

**3.2.3. Measuring the combined contribution of three weather parameters**

At the second stage, GAMs were produced to measure the combined effect of the three weather parameters on cholera incidence considering the interaction between them. Non-parametric relationships between response and predictor variables were expressed in terms of smooth functions developed using thin plate regression splines. In this study, we used an iterative process combining the penalized thin plate regression splines because they tend to give the best Minimum Square Error (MSE) performance (Wood, 2006). In addition, the tensor product smooths often perform better than isotropic smooths when the covariates of a smooth are not naturally on the same scale. As such, their relative scaling is arbitrary (Wood, 2003). Different combinations of bivariate GAMs were later sequentially run for all three variables first with simple association (Y 1) and then with association and interaction term (Y 2) following:

$$Y1 = GAM (Y \sim S (X1) + S(X2)) \tag{1}$$

$$Y 2 = GAM (Y \sim S(X1)+S(X2)+S(X3)+TE (X1, X2, X3) \tag{2}$$

Where “Y” represents the number of cholera cases reported; “S” represents the smoother which exists purely to help set up the model using spline based smooths; “X1”, “X2”, and “X3” represent rainfall, relative humidity and temperature, respectively; and “TE” the tensor smoother, which exists purely to help set up the model using tensor product based smooths.

The result was ranked in order of deviance explained. Then each of the remaining variables was added in turn to the two variables model and the deviance explained recalculated. Each variable was checked with a p-value at 0.1% significance level. The sensitivity of results to changes in parameters in the smoothing process was also checked. The residuals were checked to confirm that their distribution was approximately normally distributed with zero mean and that they exhibited no clear relationship with the predictor variables or fitted values. The Durbin-Watson test statistic was used to test the independence of GAM residuals.

### 3.2.4. Model selection

The selection of the model is important, as under-fitting a model may not capture the true nature of the variability in the outcome variable, while an over-fitted model loses generality (Snipes & Taylor, 2014). In this study the model selection was based on the Akaike Information Criterion (AIC), (Akaike, 1973) as a way to compare different models on a given outcome. The chosen model is the one that minimizes the Kullback - Leibler distance between the model and the truth defined as:

$$AIC = -2 (\ln (\text{likelihood})) + 2 K \quad (3)$$

Where likelihood is the probability of the data given a model and K is the number of free parameters in the model. AIC scores are shown as  $\Delta AIC$  scores, or difference between the best model (smallest AIC) and each model (so the best model has a  $\Delta AIC$  of zero) (Burnham & Anderson, 2002). To assess statistical significance of the derivative (the rate of change), we examined the asymptotic normality and the point wise 95% confidence interval.

## 4. RESULTS AND DISCUSSION

### 4.1. Descriptive statistics

From 1996 to 2011, the city of Maroua was mainly hit by two waves of cholera outbreaks during this period, a total of 2259 cases of cholera were reported. The first wave lasted from July 1996 to October 1998 (**Fig. 2**) and the second occurred from August 2010 to November 2011 (**Fig. 3**). The outbreaks occurred mainly between July and November and the peaks of cases were generally recorded in August and (**Fig. 2 & 3**). The characteristics of the data used are shown in **Table 1**.

**Table 1.**  
**Characteristics of the weather time series variables used in the city of Maroua.**

	mean	sd	median	trimmed	min	max	range	skew	kurtosis	se
Cholera case	11,77	52,48	0	0,05	0	542	542	6,85	57	3,79
temperature	29,61	3,98	29	29,39	20,33	42,8	42,8	0,55	0,36	0,29
rainfal	67,53	88,41	22,5	51,28	0	325,5	325,5	1,22	0,31	6,38
humidity	48,02	22,81	43,83	47,74	11,75	86,1	86,1	0,12	-1,55	1,65

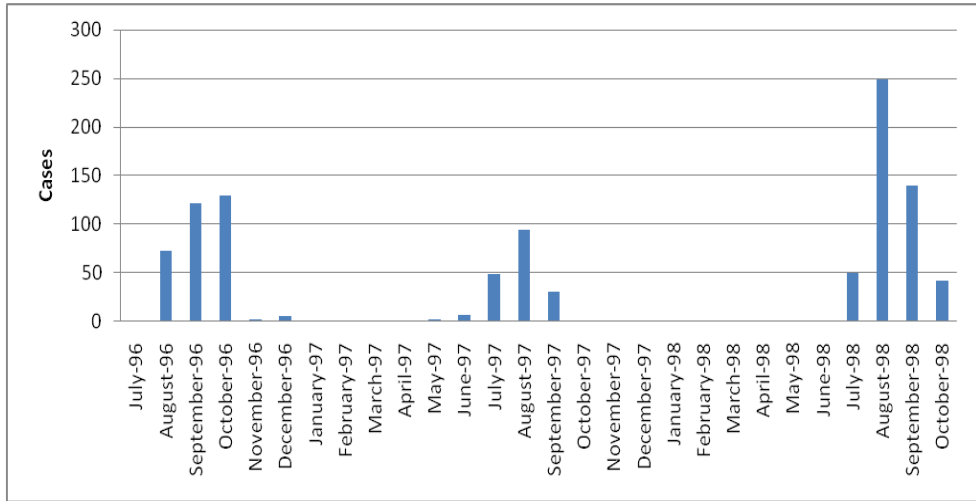


Fig. 2. Monthly number of cases in Maroua from March 1996 to May 1998.

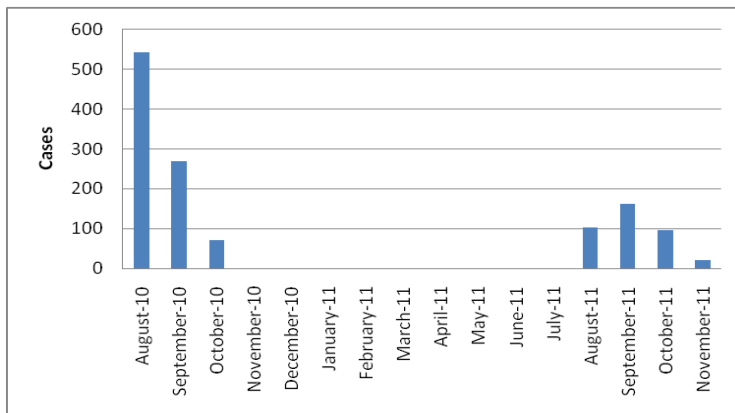


Fig. 3. Monthly number of cases in Maroua from August 2010 to November 2011.

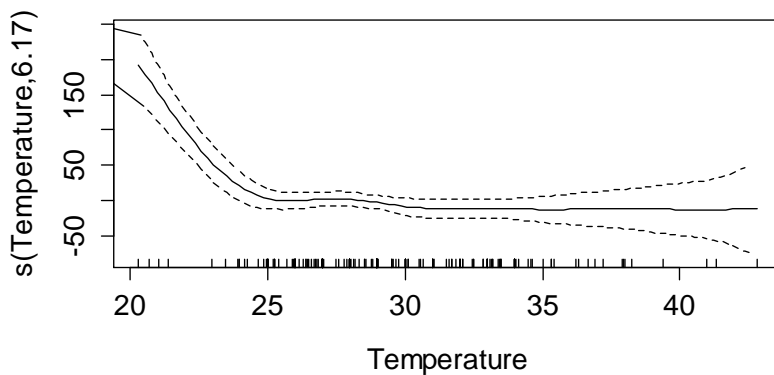
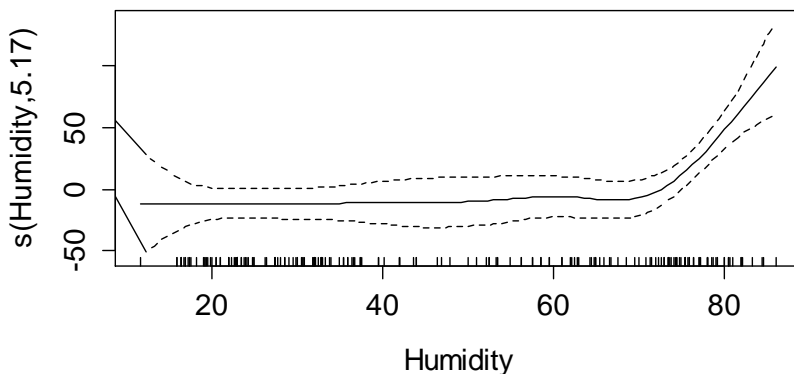
#### 4.2 The effect of individual weather parameters

As shown in **Table 2** below that summarizes the results from each of the models, the  $R^2$  (coefficient of determination) and the deviance explained of temperature are the highest (0.243) and (26.8%) respectively; AIC (2021.273) of temperature is the lowest. We note from the Table 2 above that all the three parameters are correlated with cholera incidence with temperature being the most determinant parameter. It has the highest deviance explained (26.8%). In the results of the individual GAM models shown as plots (**Fig. 4, 5, 6**) for the effect of the weather parameters on cholera, the broken lines represent point-wise 95-percent confidence envelopes around the fit. We observed that cholera cases mainly occurred when temperatures are between 21° and 30 ° (**Fig. 4**) and relative at humidity

between 79 and 82 (**Fig. 5**). The curve in Fig. 5 below shows multiple inflections for rainfall, which would suggest the effects of temperature that interacts strongly over rainfall. When associating cholera incidence with weather variables separately, we noted that the correlation is positive for all the three parameters. However, both the correlation (0.243) and the deviance explained (26.8%) by temperature are higher than the correlation with humidity (0.176) and the deviance explained (19.9%) by humidity alone which are very weak and the correlation with rainfall (0.0644) and the deviance explained (13.6%) by rainfall alone which are the least important. Yet in Zambia, the attributable risks were 4.9% for temperature and 2.4% for rainfall (Fernandez et al., 2009).

**Table 2.****Performance of the weather parameters models.**

	Std error	R 2	Deviance explained	p-value	AIC
Humidity	3.438	0.176	19.9 %	1.09e-06	2036.569
temperature	3.295	0.243	26.8 %	2.03e-09	2021.273
Rainfall	3.593	0.0644	13.6 %	0.00265	2055.812

**Fig. 4.** The effect of temperature on cholera incidence.**Fig. 5.** The effect of humidity on cholera incidence.

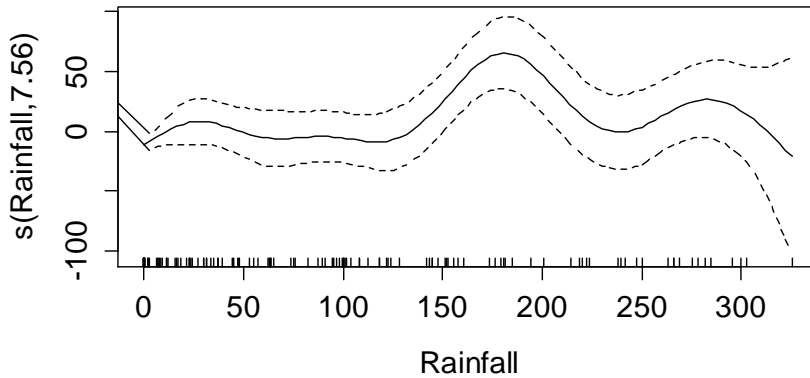


Fig. 6. The effect of rainfall on cholera incidence.

### 4.3. The effect of associated weather parameter

#### *Model 1: Using tensor smooth without interaction*

Table 3 below summarizes the results of the first set of GAM models that considered only the association of the weather parameters. From table 3 below we note that among the association of two weather parameters, “temperature against rainfall” has the highest deviance explained, the lowest generalized cross-validation (GCV) and AIC but the highest  $R^2$ . This means a very strong rainfall-temperature interaction. The association of all the three parameters yields the highest deviance explained and the highest  $R^2$ , but the lowest AIC and GCV.

Table 3.

Estimates of the GAM without interaction.

	std error	R 2	DevExp	F	AIC
Humidity + Temperature	3,036	0,358	43,6	5,238	2000,22
Humidity + Rainfall	3,048	0,353	40,8	28,41	2000,771
Temperature + Rainfall	2,504	0,563	60,4	12,92	1926,857
Humidity + Temperature + Rainfall	1,264	0,889	94,1	816029	1704,733

#### *Model 2: Using tensor smooth with interaction*

The results presented in table 4 below are from the second set of GAM models that takes into account both the association of weather parameters and the interaction between them. All variables selected in the final model were statistically significant. The couple “temperature vs. rainfall” still gives evidence of strong interaction according to  $R^2$  and Deviance explained. The results presented in table 4 below are from the second set of GAM models that takes into account both the association of weather parameters and the interaction between them. All variables selected in the final model were statistically significant. One can observed that the  $R^2$  and the deviance explained are higher while the

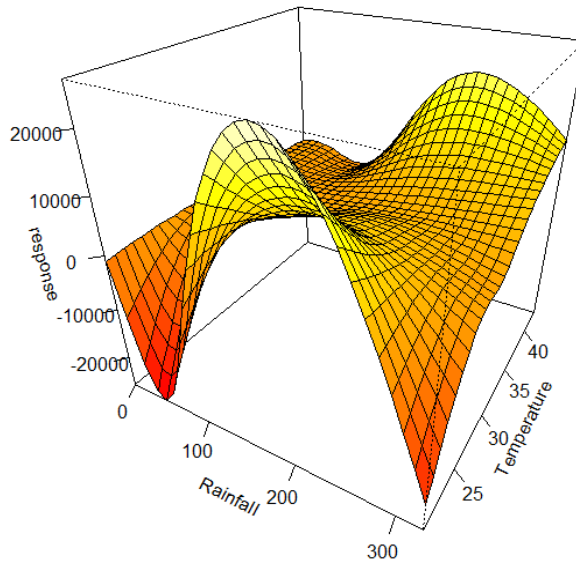
standard error, the GVC and the AIC are lower than the first model. The couple “temperature vs. rainfall” still gives evidence of strong interaction.

**Table 4.**

**Estimates of the final GAM with association and interaction.**

	std error	R 2	DevExp	AIC
Humidity + Temperature	2,987	0,378	45,5	1999,331
Humidity + Rainfall	2,815	0,448	52,6	1979,349
Temperature + Rainfall	2,395	0,6	64,4	1911,963
Humidity + Temperature + Rainfall	1,128	0,911	95,4	1661,502

Interaction effects between the parameters are shown as perspective plots (**Fig. 7**) to see how the linear predictor or expected response varies with two predictors, if all the others were held fixed at some value. The y-axis reflects the relative importance of each parameter of the model, and for the interaction effects, this is presented on the z-axis. The perspective plot views of the GAM show the results of the best-fitting smooths for the variables included in the model. The interaction between the two variables is presented as a surface; the z-axis shows the response and the relative importance of each variable is presented in the x- and y- axis. We observe from the perspective plots that the response remain at 0 producing no effect for “humidity against rainfall”; a little positive response is observed for “humidity and temperature” and a higher response is observed for the association between rainfall and temperature (**Fig. 7**).



**Fig. 7.** The interaction between rainfall and temperature.



When we consider the association between two weather parameters, the association and interaction between rainfall and temperature has a higher correlation  $R^2$  (0.6) and deviance explained (64.4%) than the association and interaction between the other parameters. The importance of rainfall as a driver of the seasonal cycle of cholera is implied by its waterborne transmission, the dose-dependent nature of infection, and the decline of cases during the monsoon season (Pascual et al., 2002). However, warm temperatures associated with the presence of water bodies favor the growth of *Vibrio cholerae* (Lipp et al., 2002) and that is due to the fact that higher temperature favors the growth of phytoplankton (Patz et al., 2005; Steffen, 2005). The correlation coefficient  $R^2$  is higher (0.911) when all three parameters are simultaneously associated with 95.4% of deviance explained when considering interaction. Similar results were found in Sinazongwe District of Southern Zambia (Phiri et al., 2015) where they noted that increased precipitation was associated with the occurrence of cholera outbreaks with a Spearman rank correlation  $r = .86$  between the number of cholera cases and the amount of precipitation (Spearman  $r=0.86$ ;  $P<.01$ ). On the other hand, *V. cholera* infection was associated with higher Relative Humidity ( $>80\%$ ) at  $29^\circ\text{C}$  temperature with intermittent average (10 cm) rainfall (Rajendran et al., 2011). Heavy rainfall indirectly influenced the *V. cholera* infection, whereas no correlation was found with high temperature.

#### 4. CONCLUSION

In this study we found that in Maroua city temperature has the highest coefficient of determination and that cholera is most likely to occur at temperature between  $21^\circ\text{C}$  and  $35^\circ\text{C}$  when the average monthly rainfall is above 100 mm. The effect of humidity appears to be very negligible. The bivariate association and interaction between rainfall and temperature has a higher coefficient of determination and deviance explained. The multivariate association with interaction of all three parameters with cholera has higher values for both the deviance explained and the coefficient of determination. These results indicate that statistical time series models should lead to a better understanding of the disease mechanism in this sahelian city, thus contributing to the growing debate on the relationship between weather parameter and the occurrence of cholera.

#### ACKNOWLEDGEMENTS:

Mouhman Arabi was supported in part by the University of Maroua and Moise Chi Ngwa was supported in part by the DOVE-Project at Johns Hopkins Bloomberg School of Public Health.

#### REFERENCES

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B.N. Petrov., and F. Csaki (Eds.), *Second International Symposium on Information Theory*. Budapest: Symposium, Academia Kiado.
- Ali, M, Nelson, A.R., Lopez, A.L., Sack, D. (2015). PLoS Neglected Tropical Diseases 9(6): e0003832. doi:10.1371/journal.pntd.0003832.
- Bailey, T.C. (2001). Spatial statistical methods in health. *Cadernos de Saúde Pública*. 17(5), 1083-1098. <https://dx.doi.org/10.1590/S0102-311X2001000500011>.

- BUCREP (2010) La population du Cameroun en 2010.  
[http://www.bucrep.cm/index.php/en/?option=com\\_phocadownload&view=file&id=18:population-du-cameroun-en-2010&Itemid=39&lang=fr](http://www.bucrep.cm/index.php/en/?option=com_phocadownload&view=file&id=18:population-du-cameroun-en-2010&Itemid=39&lang=fr).
- Burnham K. P., & Anderson D. R. (2002) *Model Selection and Multi-model Inference A Practical Information-Theoretic Approach*, Second Edition, Springer-Verlag, New York.
- Chatfield C. (1975) *The analysis of time series: theory and practice*. Chapman & Hall; New York, London.
- Codeço CT, Coelho FC (2006) Trends in cholera epidemiology. *PLoS Medicine* 3(1): e42  
10.1371/journal.pmed.0030042.
- Dobson, A. Carper, R., (1993) Biodiversity, *Lancet*, 342, 1096–1099.
- Fernandez, L., Bauernfeind, A., Jimenez, J.D., Gill, C.L, El Omeiri, N and Guilbert, D.H. (2009) Influence of temperature and rainfall on the evolution of cholera epidemics in Lusaka, Zambia, 2003-2006: analysis of a time series. *Trans R Soc Trop Med Hyg.*, 103(2), 137-43,  
DOI: 10.1016/j.trstmh.2008.07.017.
- Guévart E, Noeske J, Gérémié SG, Fouda AB, Mouangue A, Manga B. (2010) Weather and cholera: epidemic in Douala, Cameroon in 2004. *Médecine tropicale: revue du Corps de santé colonial* 70(4),407-8
- Guisan, A., Edwards, T.C-Jr., & T. Hastie. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, 157(2-3), 89-100.
- Haidu, I. (2016). What Is Technical Geography. *Geographia Technica*, 11(1), 1-5. DOI: 10.21163/GT\_2016.111.01
- Hastie, T. & Tibshirani, R. (1990). Generalized Additive Models. *Monographs on Statistics & Applied Probability*, Chapman & Hall.
- IRD-MINRESI/INC (2000) Atlas de la province extrême-nord Cameroun ; IRD Éditions.
- Islam, M.S., Sharker, M.A., Rheman S., Hossain S., Mahmud Z.H., Islam M.S., Uddin A.M., Yunus M., Osman M.S., Ernst R., Rector I., Larson C.P., Luby S.P., Endtz H.P., Cravioto A. (2009) Effects of local climate variability on transmission dynamics of cholera in Matlab, Bangladesh. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 103(11) 1165-70.  
10.1016/j.trstmh.2009.04.016.
- Kometa, S. S., Mathias, A. T. E., Humphrey, N. N. and Amawa, S. G. (2013) Human health vulnerability to climate variability: The cases of Cholera and Meningitis in some urban areas of the far north region of Cameroon. *Journal of Geography and Geology*, 5(1),116-125.
- McMichael, A.J., (2015) Extreme weather events and infectious disease outbreaks, *Virulence*, 6:6, 543-547, DOI: 10.4161/21505594.2014.975022.
- Lipp, E., Huq, A. & Colwell, R.R. (2002) Effects of Global Climate on Infectious Disease: the Cholera Model. *American Society for Microbiology*. Vol 15 No 4.
- Ngwa, M.C., Liang, S., Kracalik, I.T., Morris, L., Blackburn, J.K., Mbam, L.M., Ba Pouth, S.F.B., Teboh, A., Yang, Y., Arabi, M., Sugimoto, J.D. & Morris, J.G., (2016-a) Cholera in Cameroon, 2000-2012: Spatial and Temporal Analysis at the Operational Health District and Sub Climate Levels. *PLoS Neglected Tropical Diseases*. DOI:10.1371/journal.pntd.0005105.
- Ngwa, M. C., Liang, S., Mbam, L. M., Mouhaman, A., Teboh, A., Brekmo, K., Mevoula, O. & Morris, J. G., (2016-b) Cholera public health surveillance in the Republic of Cameroon - opportunities and challenges. *Pan African Medical Journal*. 24(222).
- Osei, F.B., Dulker A.A. & Stein A. (2012) Bayesian structured additive regression modelling of epidemic data: application to cholera. *BMC Medical Research Methodology*. 12 (118), <http://www.biomedcentral.com/1471-2288/12/118>.

- Pascual M, Bouma MJ, Dobson AP (2002) Cholera and climate: revisiting the quantitative evidence. *Microbes and Infection* 4:237–245.
- Patz, J.A., Campell-Lendrum, D., Holloway, T., & Foley, J.A., (2005), Impact of regional climate change on human health. *Nature*, 438(17), 310-317.
- Phiri, P., Nzala, S.H, Baboo, K. S. (2015) Factors associated with the recurring cholera outbreaks in Sinazongwe, district of southern Zambia. *Medical Journal of Zambia*, 42(4), 184-192.
- Rajendran, K., Sumi, A., Bhattachariya, M.K., Manna, B., Sur, D., Kobayashi, N. & Ramamurthy, T., (2011) Influence of Relative Humidity in *Vibrio Cholerae* Infection: A Time Series Model. *Indian Journal of Medical Research*, 133: 138-45.
- Reyburn, R., Kim, D. R., Emch, M., Khatib, A., von Seidlein, L., & Ali, M., (2011) Climate Variability and the Outbreaks of Cholera in Zanzibar East Africa: A Time Series Analysis. *American Journal of Tropical Medicine and Hygiene*, 84(6), 862-869.
- RStudio Team (2012). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.  
URL <http://www.rstudio.com/>
- Snipes, M., & Taylor, D. (2014) Model Selection and Akaike Information Criteria: An Example from Wine Ratings and Prices. *Wine Economics and Policy*. 3(1), 3-9, doi.org/10.1016/j.wep.2014.03.001.
- Steffen, W. (2006) *Stronger evidence but new challenges: climate change science 2001-2005*. 28 p. Canberra, A.C.T. : Australian Greenhouse Office.
- WHO (2011) *Relationship Between Climate Variability And Occurrence Of Diarrhoea And Cholera A Pilot Study Using Retrospective Data From Kolkata, India*; Technical Report National Institute Of Cholera & Enteric Diseases (Niced) Kolkata, India.
- WHO (2012) *Cholera Fact sheet N°107* available online from:  
<http://www.who.int/mediacentre/factsheets/fs107/en/index.html>.
- Wood S.N. (2006) *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC Press.
- Wood, S.N. (2003) Thin plate regression splines. *Journal Royal Statistics Society B*, 65(1), 95-114.

## **EVALUATION OF SOIL EROSION PROCESS AND CONSERVATION PRACTICES IN THE PARAGOMINAS-PA MUNICIPALITY (BRAZIL)**

*Denis Conrado da CRUZ<sup>1</sup>, José María Rey BENAYAS<sup>1</sup>, Gracialda Costa FERREIRA<sup>2</sup>,  
André Luis MONTEIRO<sup>3</sup>, Gustavo SCHWARTZ<sup>4</sup>*

DOI: 10.21163/GT\_2019.141.02

### **ABSTRACT:**

Over the last decades, the natural environment has been degraded at a much greater speed than its own resilience. Lack of knowledge about soil natural limitations and mismanagement can increase their degradation and nutrient losses by erosion. The objective of this study was to estimate and map soil vulnerability to erosion through the Universal Equation of Revised Soil Loss (RUSLE) and based on the ecodynamic concept of physical and biotic environment analysis, and finally to evaluate conservation practices in the municipality of Paragominas with the economic database of IBGE / SIDRA. In the two analyzed methods the percentage of area with low and high potential and erosivity estimation were similar. The estimation of low and low-moderate loss and vulnerability represents about 77% (15,064 km<sup>2</sup>) of the territory by RUSLE and 60% (11,485 km<sup>2</sup>), by ecodynamic concept. The high to very high soil loss zones represent only 3% (642 km<sup>2</sup>) and 2.7% (584 km<sup>2</sup>), in the RUSLE and ecodynamic concept, respectively. Most of the variables analyzed in both methods presented low estimation values of loss and erosivity potential. The soil and slope attributes, for example, obtained exactly 79% (15,377 km<sup>2</sup> - RUSLE) and 80% (15,572 km<sup>2</sup> - ecodynamic concept), except for the climate and factor R attributes, in both methods the vulnerability potential and erosion, were only 1.5% (292 km<sup>2</sup>) and 1.3% (253 km<sup>2</sup>), based on the ecodynamic concept and RUSLE respectively. The geospatial analysis of the use practices correlated with the economic data showed an intense use of agricultural activities, logging and mining, which caused severe environmental damages, considering that 45% (8,773.3 km<sup>2</sup>) of the municipality have already been deforested and converted into other uses. The municipality still has 47% (9,182 km<sup>2</sup>) of its territory covered by altered primary vegetation and 23% (4,441 km<sup>2</sup>) by secondary vegetation, important information to subsidize decision-making processes related to ecological-economic strategies for the management of natural resources in the study area.

*Key-words: Agriculture, Environmental variables, Potential for soil loss, Amazon*

## **1. INTRODUCTION**

Erosion is a process by which soil and rocks wear out, then transported by natural or anthropogenic agents and deposited elsewhere (Verheijen *et al.*, 2009), it can pollute water, cause land degradation, reduce soil fertility and increase the loss of organic matter (Cerdan *et al.*, 2010). Changes in the management and use of land resulted from human activities can increase soil erosion causing irreversible damage to the environment (Fiorio *et al.*,

---

<sup>1</sup> University of Alcalá, Life Sciences Department, 28871, Alcalá de Henares, Spain, [conrado\\_denis@hotmail.com](mailto:conrado_denis@hotmail.com); [josem.rey@uah.es](mailto:josem.rey@uah.es).

<sup>2</sup> Amazon Rural Federal University, 66.077.830, Belém Brazil, [gracialdaf@yahoo.com.br](mailto:gracialdaf@yahoo.com.br).

<sup>3</sup> Geoforest Consultancy, 66063-040, Belém-Pa, Brasil, [monteiroandre7@gmail.com](mailto:monteiroandre7@gmail.com).

<sup>4</sup> Forest Ecology and Management Embrapa Eastern Amazon, 70.770.901, Belém, Brazil, [gustavo.schwartz@embrapa.br](mailto:gustavo.schwartz@embrapa.br).

2016; Da Silva *et al.*, 2016). The vegetation withdrawal due to the insertion of plantations can increase up to 600% soil erosion (Chaplot *et al.*, 2005). Soil loss brings concern all over the planet, as it is a source for food production (Abel *et al.*, 2005).

Production of food and other goods for human needs, combined with population growth and agricultural intensification, has resulted in severe land degradation and, particularly, soil erosion. Between 1961 and 2010 the world population duplicated from 3 billion to 6 billion inhabitants, while the crop production increased four times in almost the same area (FAO, 2017). It is estimated that since the beginning of stable agriculture, ca. 430 million hectares was damaged due to soil erosion (Lal, R, 2017).

Agricultural production in Brazil was fostered through a National Development Plan in the 1970s aiming to improve the country economy (Becker, 2005). Such Plan was developed to attract entrepreneurs to the Brazilian Amazon, with the government releasing taxes on rural credits to incentivize investment (Kohlhepp, 2002). That period was marked by huge agricultural, industrial, and silvicultural projects besides infrastructure actions (Hall, 1989; Serra, 1998). The intense advance of the agricultural sector and the territory occupation was evident in many Amazon municipalities, leading to a drastic change in the landscape. For example, Paragominas has already 45% (8,773.3 km<sup>2</sup>) of its land deforested (INPE/Prodes, 2018), causing loss of soil organic matter and nutrients and soil erosion (Angima, 2003). Lack of knowledge about soil limitations and inadequate soil management practices can trigger or speed up its erosion in agricultural areas (Navas *et al.*, 2005). Beyond land use activities, soil loss and degradation are affected by a combination of different environmental factors including geology (rock types), geomorphology (landform), vegetation (e.g. its cover and structure), pedology (soil types), and climate (e.g. rainfall) (Arnesen, 2009). The awareness about risks to environment, economy and livelihoods have resulted in research, technology and production practices to minimize soil erosion (Bakker *et al.*, 2008).

Understanding and quantifying erosion processes are important steps in the decision making process regarding the best management to be adopted. Methods with empirical models are used, for example, the Universal Soil Loss Equation (USLE), which emerged in the late 1970s (Wischmeier & Smith, 1978) in the United States, but it is widely used in Brazil. This method presents great accuracy, as considers soil type, soil morphology, rainfall, cultivation practices and management, allows quantification and regionalization of the area with higher risk of erosion (Wischmeier & Smith, 1978; Bertoni, 2005; 2012). It was later revised and adapted by several authors “Revised Universal Soil Loss Equation” (RUSLE). Another technique widely used to analyze the natural vulnerability to soil loss was based on the ecodynamic of Tricart (1977), modified by Crepani *et al.*, (2001), which uses the morphogenesis/pedogenesis relation integrated to satellite images, assigning values of vulnerability/stability to each thematic class.

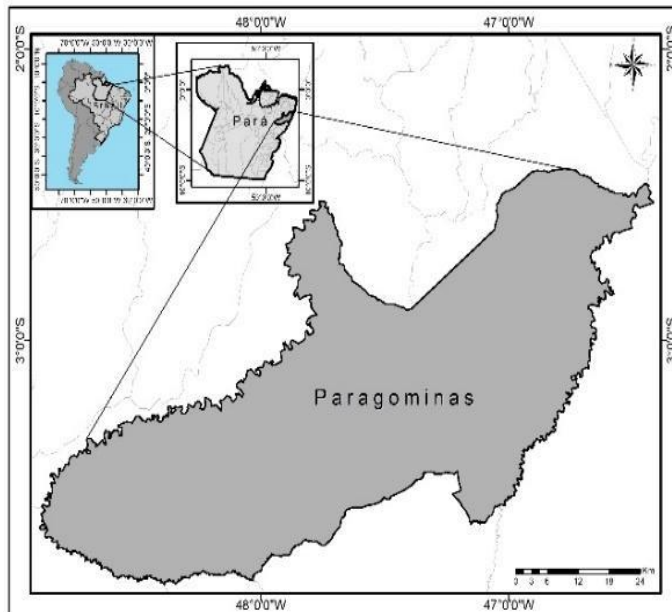
The use of remote sensing and geographic information systems contribute significantly to monitoring, mapping, and managing landscapes (Ferreira, 2008), particularly in large and remote areas. For example, in Brazil where, besides having huge territorial extension, the access to some areas is also a constraint, what emphasizes the need of a constant input in this field in order to monitor effectively (Assad & Sane, 1998; Câmara *et al.*, 2001). This study aims to estimate and map vulnerability to soil erosion through geospatial analysis, by the model of the Universal Equation of Soil Loss Revised (RUSLE), ecodynamic concept of analysis of physical and biotic environment and conservation practices evaluation in Paragominas municipality in the Brazilian Amazon. We analyzed the local attributes of geology, geomorphology, vegetation, pedology, and climate, by means of satellite imagery

analysis, thematic maps, and rainfall data, then a map of land was obtained which represents a range of levels of vulnerability to soil erosion. The analysis presented here can contribute to support decision makers regarding ecological-economic strategies for natural resource management in the study area and other parts of Brazilian Amazon.

## 2. MATERIAL AND METHODS

### 2.1. Study site

Paragominas municipality has an area of 19,465 km<sup>2</sup> and it is placed in Pará state, Brazil. Its original vegetation was mainly formed by Dense Ombrophilous Forest (Watrin, 1992). The predominant soil type is yellow latosol, rich in clay and has low fertility. The climate is warm and humid, with annual average temperature of 26.3°C (**Fig. 1**).



**Fig.1.** Paragominas municipality (Source: authors).

### 2.2. Estimation and mapping of soil erosion

Soil erosion estimations were carried on the concept proposed by Tricart (1977), adapted by Crepani (2001) and by the Revised Universal Soil Loss Equation - RUSLE (Wischmeier and Smith, 1978). Both methods evaluate soil erosion with variables such as: rainfall; soil type; land-use class; geology and geomorphology.

### 2.3. Soil erosion based on the ecodynamic concept

We used the method based on the ecodynamic concept (Tricart, 1977), adapted by Crepani *et al.*, (2001) to Brazil. This concept considers the balance between soil formation processes (pedogenesis) and erosion processes (morphogenesis). It applies a range of soil erosion vulnerability values to the target land areas that are analyzed (stability or instability). Areas where pedogenesis and morphogenesis predominate have a value around 1 and 3, respectively, and areas where both processes are balanced have values around 2 (**Table 1**).

**Table 1.**

Classes of soil erosion vulnerability.	
Class	Vulnerability Value
Low Vulnerability	1 > - 1.6
Low-moderate vulnerability	1.6 - 1.9
Moderate vulnerability	1.9 - 2.1
High-moderate vulnerability	2.1 - 2.4
High vulnerability	2.4 - 3

Source: Elaborated by the author, adapted from Crepani *et al.*, (2001), by ecodynamic concept (Tricart, 1977).

Vulnerability values of soil erosion (between 1 and 3) were assigned based on attributes related to geology, geomorphology, pedology, vegetation types, and climate variables. They resulted in five thematic maps obtained by means of the Map Algebra of ArcGis 10.1 (Esri, 2012). For a given land area, the overall vulnerability (V) was:

$$V = G + SL + S + Vg + C / 5 \tag{1}$$

Where:

G = vulnerability due to Geology, SL = vulnerability due to Slope, S = vulnerability due to Soil, Vg = vulnerability due to Vegetation, and C= vulnerability due to Climate.

### 2.3.1. Geology attribute

Vulnerability due to Geology depends upon rock type. We used the geological database of the Brazilian Geological Service (CPRM, <http://www.cprm.gov.br/>), at a 1:100.000 spatial scale. The CPRM provides data in georeferenced vector format of aerogeophysical projects available in Geobank (<http://geobank.cprm.gov.br/>). Values of soil erosion vulnerability related to rock types in the study area are reported in **Table 2**.

**Table 2.**

Values of soil erosion vulnerability related to rock types.	
Rock types	Value
Quartzites or metaquartzites	1.0
Rhyolite, Granite, Dacite	1.1
Granodiorite, Quartz Diorite, Granulites	1.2
Migmatite, Gneiss	1.3
Phonolite, Nepheline syenite, Trachyte, Syenite	1.4
Andesite, Diorite, Basalt	1.5
Anorthosite, Gabbro, Peridotite	1.6
Mylonitos, Muscovite Quartz, Biotite, Shale chlorites	1.7
Pyroxene, Kimberlite amphibolite, Dunite	1.8
Hornblende, Tremolite, Shale actinolite	1.9
Shale Staurolite, Granatiferous shale	2.0
Phyllite, Metassiltite	2.1
Slate, Metargilite	2.2
Marbles	2.3
Quartz Sandstone or orthoquartzites	2.4
Conglomerates, Subgraywacke	2.5
Greywackes, Arkose	2.6
Siltstones, Mudstones	2.7
Husk	2.8
Calcareous, Dolomites, Marls, Evaporites	2.9
Soft-bottom sediments: Alluvium, Colluvium, Sands, etc.	3.0

Source: Elaborated by the author, adapted from Crepani *et al.*, (2001), by ecodynamic concept.

### 2.3.2. Slope attribute

The source of soil erosion vulnerability related to slope, was adapted by Crepani *et al.*, (2001) study (**Table 3**). These data derive from SRTM (Shuttle Radar Topography Mission) available on EMBRAPA (Brazilian Corporation of Agricultural Research, (<https://www.embrapa.br/territorial/>), calculated using the ArcGis 10.1 Solpe tool.

**Table 3.**

**Values of soil erosion vulnerability related to slope.**

Degree	%	Value/ Vuln.	Degree	%	Value/ Vuln.	Degree	%	Value/ Vuln.
<2	<3.5	1.0	9.9-11.2	17.4-19.8	1.7	19.1-20.4	34.6 - 37.2	2.4
2-3.3	3.5-5.8	1.1	11.2-12.5	19.8-22.2	1.8	20.4-21.7	37.2 - 39.8	2.5
3.3-4.6	5.8-8.2	1.2	12.5-13.8	22.2-24.5	1.9	21.7-23.0	39.8 - 42.4	2.6
4.6-5.9	8.2-10.3	1.3	13.8-15.2	24.5-27.2	2.0	23.0-24.4	42.4 - 45.3	2.7
5.9-7.3	10.3-12.9	1.4	15.2-16.5	27.2-29.6	2.1	24.4-25.7	45.3 - 48.1	2.8
7.3-8.6	12.9-15.1	1.5	16.5-17.8	29.6-32.1	2.2	25.7-27	48.1 - 50	2.9
8.6-9.9	15.1-17.4	1.6	17.8-19.1	32.1-34.6	2.3	>27	>50	3.0

Source: Elaborated by the author, adapted from Crepani *et al.*, (2001), by ecodynamic concept.

### 2.3.3. Pedology attribute

The vulnerability related to the pedology attribute (**Table 4**) refers to mapping of soil units according to Crepani *et al.*, (2001), updated with Prado (2001) nomenclature. This attribute database was obtained from IBGE (Brazilian Institute of Geography and Statistics), based on the new Brazilian System of Soil Classification (EMBRAPA, 1999).

**Table 4.**

**Soil erosion vulnerability values related to soil types.**

Soil Classes	Vuln.	Soil classes	Vuln.
Yellow Latosol	1	Spodosol	2
Red-Yellow Latosol	1	Neosol Litólicos	3
Red Latosol	1	Neossolos Flúvicos	3
Latosol Brunos	1	Neossolos Regolíticos	3
Latosol (...) Humic	1	Neossolos Quartzarênicos	3
Latosol Bruno (...) Humic	1	Vertisol	3
Acrisol	2	Organosols	3
Acrisol Luvisol Alisol Nitosol	2	Gleysol	3
Acrisol Nitosol	2	Gleysols Plinthosol	3
Luvisol	2	Plinthosol	3
Chernozem	2	Rocky Outcrop	3
Planosol	2	-	-

Source: Elaborated by the author, adapted from Crepani *et al.*, (2001), with new nomenclature from Prado (2001).

### 2.3.4. Vegetation and use of class attribute

The definition of values on vulnerability to erosion regarding vegetation attribute was identified under the forest canopy density using satellite imagery determined 21 class units, following the Crepani *et al.*, (2001) definition (**Table 5**).

It was used satellite imagery from Landsat 8 TM, orbits/points 222/62, 222/63, 223/62, and 223/63 (2015 data), with the lowest cloud coverage of the period, available in the website Glovis from NASA. From these satellite imageries were produced a classification using the supervised Maximum Likelihood method. This method uses mean and variance of the data set for the classification decision rule and for this reason a considerable number of pixels is required for each region to be classified. First the use class was defined, then



samples (ROI's) of the areas to be classified in the image were generated (**Table 6**). With the selected ROI's for use class the image was classified using the software ENVI 4.7, afterward vulnerability values of classes were set as follows in **Table 7**.

**Table 5.**

<b>Soil erosion vulnerability values related to vegetation types.</b>	
<b>Vegetation Class</b>	<b>Vuln.</b>
Dense Ombrophilous Forest	1.0
Open Ombrophilous Forest	1.0
Mixed Ombrophilous Forest	1.0
Semideciduous Seasonal Forest	1.4 - 1.7
Campinaranas formation	1.4 - 1.7
Forested Savanna and Steppe Savanna	1.4 - 1.7
Dense tree steppe, with or without palms	1.4 - 1.7
Wooded Savanna and Steppe Savanna da e Wooded Steppe Savanna	2.0
Deciduous Seasonal Forest	2.0
Wooded Campinarana	2.0
Wooded Steppe	2.0
Buritizal with fluvial and/ or lake influence	2.0
Wooded Campinarana with or without palms	2.4 - 2.6
Savanna Park, Wooded Savanna Park	2.4 - 2.6
Campinarana and Steppe with shrub size	2.4 - 2.6
Vegetation under marine influence (Sandbanks)	2.4 - 2.6
Vegetation under fluvial and/or lake influence	2.4 - 2.6
Montane Refuge and High Montane Refuge	2.4 - 2.6
Woody- grassy Savanna, Woody-grassy steppe savanna, and Woody-grassy steppe	3.0
Woody-grassy campinarana	3.0
Vegetation under herbaceous marine influence	3.0
Montano and High-Montano refuges	3.0
Cloud/ Shadow/ exposed soil/ sparse vegetation	3.0

Source: Elaborated by the author, adapted from Crepani *et al.*, (2001), by ecodynamic concept.

### **2.3.5. Climate attribute**


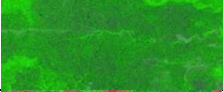
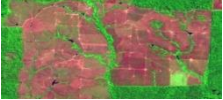




About the attribute climate (pluviometry), 21 classes were classified on soil erosion (**Table 8**). The areas with lower annual pluviometric rate and higher extension of rainy season were classified with values around stability (1.0) to intermediary values of vulnerability/stability (2.0). On the other hand, areas with higher annual pluviosity rate and shorter rainy season vulnerability presented values around 3.0. In order to assign the vulnerability values to the climate, the precipitation data of the last 18 years of the municipality station, was used in the system of the national meteorological institute (INMET-<http://www.inmet.gov.br/portal/>). The average precipitation was calculated between the months of January 2000 and November 2018 (**Fig. 2**).

### **2.4. Revised Universal Soil Loss Equation (RUSLE) method.**

Through the RUSLE equation, the main spatial distribution factors responsible for soil erosion were performed in the GIS environment. Information plans were embedded into the database and manipulated through the geoprocessing tools in ArcGis (**Fig. 3**).

Table 6.

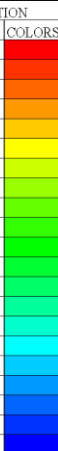
## Land use patterns and Vegetal cover used in the satellite image classification.

Typological classification	Pattern identified in satellite image	Description
Primary Forest		Forests that have passed through interventions in the past, but have their primary structure conserved.
Secondary Forest		Forests that have undergone deforestation or degradation processes and are currently in regeneration.
Pasture		Livestock pasture areas with low biomass, characterized areas with healthy pastures and degraded pastures.
Agriculture		Agricultural plantations mechanized, which presupposes high technological level and family agriculture areas.
Deforestation		Areas in which all vegetation cover was removed, leaving the soil exposed.
Urban area		Areas of urban agglomeration, with industrial estates, streets, buildings and highways.
Hydrography		Drainage with rivers and springs.

Source: Elaborated by the author.

Table 7.

## Soil Vulnerability regarding vegetation.

LANDSCAPE UNITES	AVERAGE	DEGREE OF VULNERABILITY	DEGREE OF SATURATION			
			RED	GREEN	BLUE	COLORS
U1	3.0	VULNERABLE	255	0	0	
U2	2.9		255	51	0	
U3	2.8		255	102	0	
U4	2.7		255	153	0	
U5	2.6		255	204	0	
U6	2.5	MODERATE VULNERABILITY	255	255	0	
U7	2.4		204	255	0	
U8	2.3		153	255	0	
U9	2.2		102	255	0	
U10	2.1		51	255	0	
U11	2.0	MEDIUM STABLE VULNERABILITY	0	255	0	
U12	1.9		0	255	51	
U13	1.8		0	255	102	
U14	1.7		0	255	153	
U15	1.6		0	255	204	
U16	1.5	MODERATE STABLE	0	255	255	
U17	1.4		0	204	255	
U18	1.3		0	153	255	
U19	1.2		0	102	255	
U20	1.1		STABLE	0	51	
U21	1.0	0		0	255	

Source: Elaborated by the author adapted from Crepani *et al.*, (2001), by ecodynamic concept.

Table 8.

**Soil vulnerability regarding pluviometric rate.**

Pluviometric Intensity mm/month	Vuln.	Pluviometric Intensity mm/month	Vuln.	Pluviometric Intensity mm/month	Vuln.
< 0 50	1.0	200 - 225	1.7	375 - 400	2.4
50 - 75	1.1	225 - 250	1.8	400 - 425	2.5
75 - 100	1.2	250 - 275	1.9	425 - 450	2.6
100 - 125	1.3	275 - 300	2.0	450 - 475	2.7
125 - 150	1.4	300 - 325	2.1	475 - 500	2.8
150 - 175	1.5	325 - 350	2.2	500 - 525	2.9
175 - 200	1.6	350 - 375	2.3	> 525	3.0

Source: Elaborated by the author adapted from Crepani *et al.*, (2001), by ecodynamic concept.

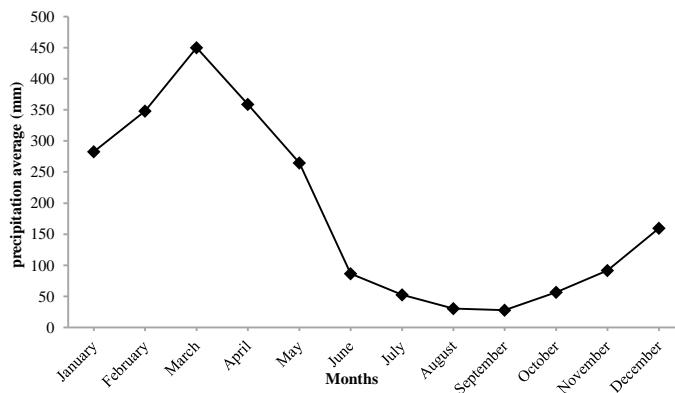


Fig. 2. Precipitation average from 2000 to 2018, INMET data. (source: authors).

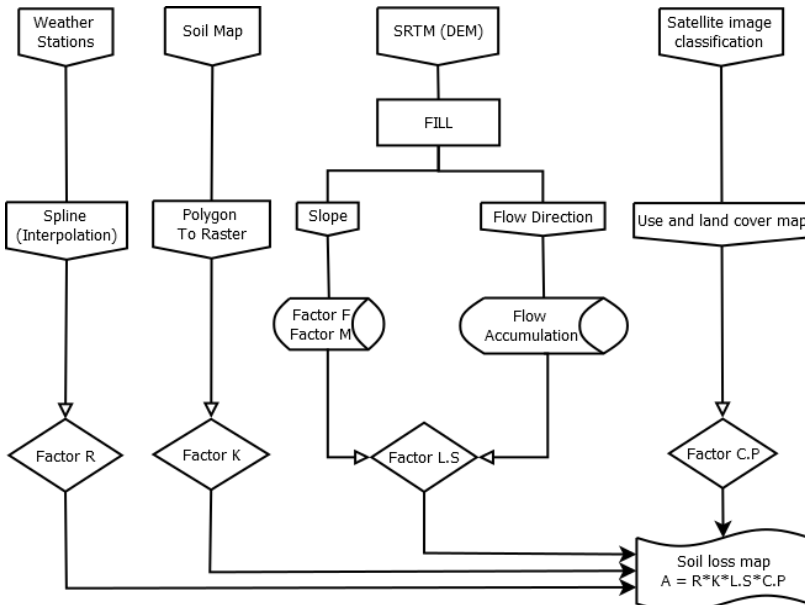


Fig. 3. Flowchart of the technique used in the generation of the soil loss map (source: authors).

(2)

$$A = R \times K \times L \times S \times C \times P$$

Being: A = Soil loss calculated per unit area, (Mg.ha<sup>-1</sup>.ano<sup>-1</sup>); R = Rainfall factor: rainfall erosion index, (Mg.ha<sup>1</sup>.ano<sup>1</sup>); K = soil erodibility factor (MJ/ha.mm/h); L = slope length factor, (m); S = Slope degree factor, (%); C = Use and management factor (dimensionless); P = Conservationist practice factor (dimensionless).

#### 2.4.1. Factor R (Erosivity by rains)

In order to determine the annual rainfall erosivity, annual average precipitation data of the 28 rainfall stations contained in Paragominas were used for the period from 2000 to 2018. **Fig. 2** shows the monthly average rainfall erosivity using the formula used in Amazon:

(3)

$$EI_{monthly} = 42,307 (Pm^2/Pa) + 42,77$$

Where: Pm = monthly precipitation and; Pa = annual precipitation average.

The average annual rainfall erosivity is obtained by the sum of the monthly erosivity average of each season. The map was generated in the ArcGis environment, by inserting the table (dbase format) with spatial distribution of the stations (UTM coordinates of the stations) and their respective values of calculated erosivity. The interpolation of the values representing the spatial variation of erosivity was done through the ArcGis Spline tool.

#### 2.4.2. Factor K (Soil Erodibility)

The evaluation of soil erodibility was obtained from IBGE Geoscience Center (Brazilian Institute of Geography and Statistics) in scale 1:5000.000, based on the new Brazilian System of Soil Classification (EMBRAPA, 1999). The soil classes were grouped, generating a map of soil types, where the value of K was associated for each type.

#### 2.4.3. Factor LS (Topographic factor)

In this factor the length of slope L represents the distance between the point which originates the surface flow to the point where the slope decreases enough for sediment deposition to occur. The slope gradient (S) refers to slope variation in slope intervals, these two parameters (LS) are represented as a single topographic factor, defined as the rate of soil loss per unit area of a standard plot of 22.13 m in length and 9% of slope (Wischmeier and Smith, 1978). Calculated by means of the following steps:

$$F = \frac{\sin\beta/0.0896}{3(\sin\beta)^{0.8}+0.56} \quad m = \frac{F}{(1+F)} \quad L \left( \frac{\lambda}{22.13} \right)^m$$

$$S_{(i,j)} = \begin{cases} 10,8 \sin \beta (i, j) + 0,03 & \tan \beta (i, j) < 0,09 \\ 16,8 \sin \beta (i, j) + 0,5 & \tan \beta (i, j) > 0,09 \end{cases}$$

In factor L the  $\lambda$  is the slope length, m the slope length exponent and  $\beta$  the slope angle. The slope length is defined as the horizontal distance from which originates the surface flow to the point where the deposition begins or where the flow flows into a channel, and at factor S the angle  $\beta$  is taken as the mean angle of all sub -redes in the steepest direction (**Fig. 4**).

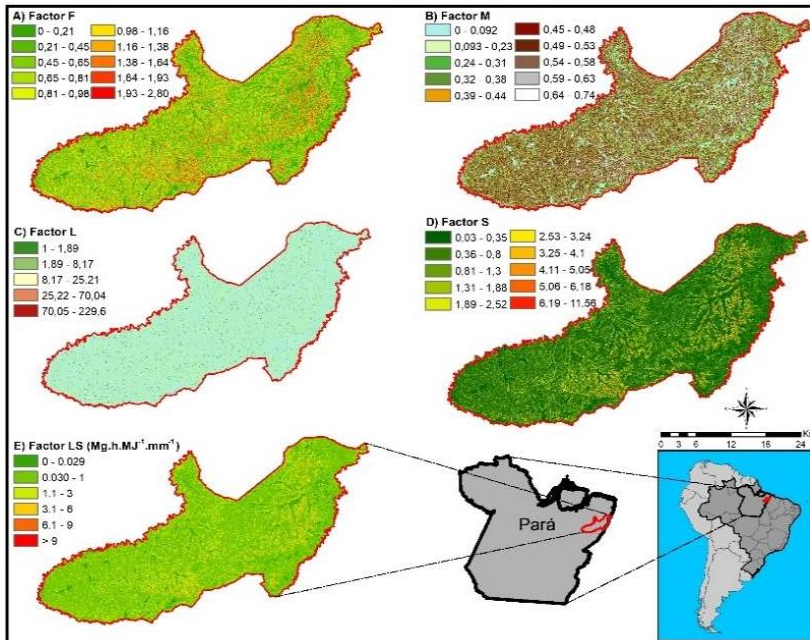


Fig. 4. Erosivity factors for the L.S calculation (source: authors).

#### 2.4.4. Factor C (Use/soil management) and P (Conservationists practices)

Factor C is the expected relationship between the soil loss of a cultivated land under given conditions and the corresponding losses of a land kept continuously uncovered and cultivated (Bertoni, 2005). While the factor P is the relationship between losses in soils with a given conservationist practice and those where the crop is planted in the slope direction. A factor P of 0.01 was considered for areas with primary, secondary and urban vegetation, 0.09 for agriculture, 0.45 for pasture, 1 for deforestation and 0 for water. To elaborate the map of factor C, it was necessary to classify the satellite image for soil use and to assign C and P values for each type of use (Donzelli *et al.*, 1992), then convert to a raster format, where  $CP = C.P$  and CP factor specializations were obtained from the numerical reclassification of the vegetation cover and land use maps (Table 1) for the year 2017.

#### 2.5. Conservation management and Practices

In order to evaluate the activities developed in Paragominas, a mapping was carried out on the land use, a classification carried out by the TerraClass Project, then developed and executed by the Regional Center of the Amazon (CRA). The data base of this project are the mapped deforested areas published by the PRODES Project - Monitoring of the Amazon Forest by Satellite. TerraClass analyzes the possible causes of tree cutting considering the following classes: annual agriculture; unobserved area; urban area; mining; occupation mosaic; pasture with soil exposed; clean pasture; dirty grass; regeneration with grass; reforestation; secondary vegetation; forest and not forest. This mapping counts a series of 10 years analysis of use and coverage (2004, 2008, 2010, 2012 and 2014), and to evaluate the current situation an unsupervised classification was performed in an image from 2017, described in Table 6.

In order to analyze the use and coverage of the soil obtained by the classifications carried out by satellite images, we also used the data base of the IBGE Automatic Recovery System - SIDRA. This system has an economic historical series of data since 1974, in this component we analyzed only the years of 2004, 2008, 2010, 2012, 2014 and 2017. Another important practice developed in Paragominas is mining and to evaluate this component, it was used data from the National Department of Mineral Production (DNPM), available through the Geographic Information System of Mining – SIGMINE.

### 3. RESULTS AND DISCUSSIONS

After analysis, for the entire area of Paragominas municipality, a map with values of vulnerability to soil erosion was obtained for each following category: geology, slope, pedology, vegetation and climate. Then, a general map of vulnerability to soil erosion and estimates of soil loss was created, according to the method of ecodynamics proposed by Tricart (1977) and adapted by Crepani *et al.*, (2001), and by RUSLE.

#### 3.1. Analysis of the vulnerability based on the ecodynamic concept

##### 3.1.1. Geology attribute

The basic information of geology is the cohesion degree of integrated rocks from the ecodynamics (Tricart, 1977). It means that in the most cohesive rocks the processes of weathering and pedogenic formation prevailed, while the less cohesive rocks are more susceptible to erosive processes. It is considered that in rocks with little cohesion erosive processes can prevail, while in very cohesive rocks the processes of weathering and soil formation must prevail (Crepani *et al.*, 2001). For this attribute, it was identified that only 1% (212 km<sup>2</sup>) of the municipality area has low vulnerability to erosion, with rocks of granite, granodiorite, gneiss and schist, with value between 1.1 to 1.3 these rocks are more weather resistant (Gomes, 2000), since the igneous rocks are more resistant to temperature rises.

Paragominas municipality presents ca. 61% (11,874 km<sup>2</sup>) of its territory in the moderate-high attribute class (value 2.5) with sedimentary rocks of the sandstone, argillite, and silt types. This material can be composed from angular grains to sub-rounded enveloped by clay matrix of infiltration (Frostick, 1984), which can lead to greater erosive process. In 38% (7,397 km<sup>2</sup>) of the municipality area there is a high vulnerability to soil erosion (value 3). These sediments have smaller interfluves (of higher-intensity dissection), for this reason receive vulnerability value higher for the geology attribute (Crepani, *et al.*, 2001) (**Fig. 5 A**).

##### 3.1.2. Declivity attribute

Declivity is the relief slope regarding the horizon that has a direct relation with transformation speed from potential to kinetic energy. Thus, the higher the slope the faster the potential energy of rainwater becomes kinetic energy and higher the water masses velocity and their transport capacity, responsible for relief erosion (Watson & Lafflen, 1986). With respect to this attribute, it was verified that 46% (8,954 km<sup>2</sup>) of the municipality has declivity lower than 2% (value 1). In 34% (6,618 km<sup>2</sup>) of the municipality the slope is between 2% and 6%, with low-moderate vulnerability (value 1.5). The regions presenting a slope between 6% and 20% represent 13% (2,530 km<sup>2</sup>) of Paragominas, moderate vulnerability value. Lang *et al.*, (1984) observed that zones with 9% of slope presented greater erosion between furrows of a topsoil compared to an area of 3% slope.

Paragominas is located in the morphostructural domain of plateaus in non-folded sedimentary sequences (IBGE, 1996), characterized by flattened structural surfaces, with average altitude around 200 m. For this reason, only 5% (973 km<sup>2</sup>) of the municipality has a slope between 20% and 50% (moderate-high vulnerability), so regions having a slope higher than 50% (high vulnerability) correspond to only 2% (289 km<sup>2</sup>) of the total area (**Fig. 5 B**).

### **3.1.3. Soil attribute**

Approximately 79% (15,337 km<sup>2</sup>) of Paragominas soil area are composed of latosols, which are mineral soils, deep, well-drained, with B horizon latosolic, usually cohesive, quite hard when dry, mainly in the AB and BA horizons (Rodrigues *et al.*, 1991; Embrapa, 1999). Therefore, there are more stable soils and resistant soils to erosive processes with value 1 of vulnerability. Only 2% (389 km<sup>2</sup>) of its territory has argisol soils (vulnerability value 2), since their soils are moderately stable in relation to erosive processes. This type of soil has a different textural gradient between A and B horizons, this can lead to soil loss by difficulty in infiltration, Schaefer *et al.*, (2002), in their study identified nutrients losses by erosion on simulated rainfall conditions, with different surface coverages on argisol soils.

The vulnerable soil erosion (vulnerability value 3) of the municipality corresponded to the regions with Gleysol and Plintosol representing 2% and 17% (3,698 km<sup>2</sup>), respectively (**Fig. 5 C**). The plintosols, for example, are mineral soils formed through water percolation restriction. These soils are poorly drained with excessive plinization (Rodrigues, 2003), while gleysol is characterized by high gley status of soils, which results in a reduced moisture regime whereby soil waterlogging for a long period during the year (Embrapa, 1999).

### **3.1.4. Vegetation attribute (Soil use)**

The composition and structure of the vegetation are important components in soil losses analyzes (Gomes, 2000). Deforestation is one of the anthropogenic actions that changes the forest structure and speeds up processes of soil erosion. According to the PRODES project, up to 2017 the accumulated deforestation in Paragominas reached ca. 45% (8,744 km<sup>2</sup>) of its territory. These areas have been converted to others land uses, and were intended for livestock raising and crop production, mostly grain cultivation, with practices characterized by monoculture, intense mechanization and agrochemical inputs (Alves, *et al.*, 2014). After processing the satellite imagery, 9% (1,752 km<sup>2</sup>) of the municipality was identified in the high vulnerability category presenting sparse vegetation or exposed soil (**Fig. 5 D**)

The most stable Paragominas areas correspond to about 76% (14,793 km<sup>2</sup>) of the municipality, which is occupied by ombrophilous forest with altered primary and secondary vegetation under advanced successional stage, thus, low potential for vulnerability to soil erosion. In this advanced stage of succession, there is a greater production of litter (Pezatto & Wisniewski, 2006; Barbosa & Faria, 2006) that may favour soil protection, since it is the main form of nutrient return to the soil and moisture retention (Espig *et al.*, 2009).

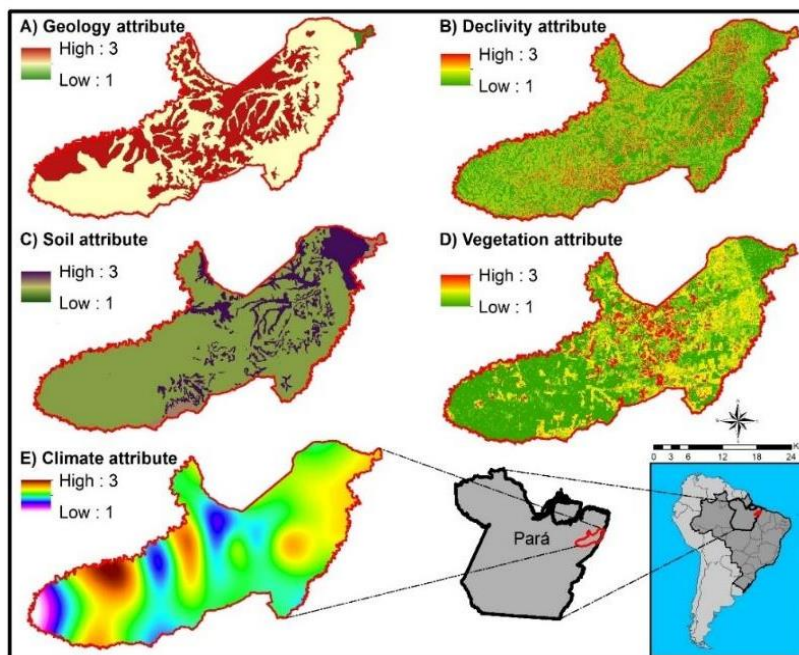
### **3.1.5. Climate attribute**

The potential capacity of rainfall to accelerate the process of soil erosion is related to the precipitation intensity in a region (Bertoni & Lombardi Neto, 2008; Guerra, Silva & Botelho, 2009). The climate in humid equatorial Amazon is very favourable to the vegetal production benefiting the forest protection, acting as a huge thermostat avoiding extremes

of temperature (Schmidt, 1947), and by means of rainfall and temperature the climate controls the weathering of a region (Mota, *et al.*, 2009). The erosion process caused by rainfall (Santos, *et al.*, 2010) is the result of soil particles surface disaggregation, due to the energy of the drop's impact and by the surface runoff force. The detachment and transport of sediments promotes soil losses by floods (Bertol *et al.*, 2007; Bertoni & Lombardi Neto, 2012).

Rainfall directly influences erosion processes, and in Paragominas the pluviometric intensity average (mm/moth) is smaller than 1 mm during the dry season (June to November) and more than 1,000 mm in the rainy season (December to May). As a consequence, the entire municipality has value 2 - 2.4 moderate (52%; 10,044 km<sup>2</sup>), to high-moderate 2.5 - 2.7 value (39%; 7,591 km<sup>2</sup>) vulnerability to soil erosion for the climate attribute (**Fig. 5 E**).

Only 1.5% (311 km<sup>2</sup>) of the municipality has a value of 1.7-1.9 vulnerability considered medium-low and another 8% (1,557 km<sup>2</sup>) considered to be a high vulnerability potential to soil erosion (2.8 - 3 vulnerability value). As the dry and rainy period in the municipality are well defined they do not cause a significant impact on soils covered by vegetation. The rainfall of a given region is considered a risk factor due to the production of sediments per unit of drainage area which raises with the increase of the drainage area, the larger the river basins, the greater the possibility of possible erosions (Oakes *et al.*, 2012).



**Fig. 5.** Map of attributes responsible for vulnerability to soil erosion (*source: authors*).

## 3.2. Estimation of soil loss according to (RUSLE)

### 3.2.1. Factor R (Erosivity of rains)

Soil erosivity of this factor involves the disintegration of soil particles, transported and deposited by rainfall and surface runoff of water on the soil (Crepani, 2004). The values of



this factor for the 28 rainfall stations vary from 5,621 to 17,540 MJ.mm/ha.h, with an average of 13,500 MJ.mm/ha.h, standard deviation of 1,630 MJ.mm/ha.h. Paragominas has a very heterogeneous R-factor, however, the highest and lowest values of erosivity estimation indices are found in the southeast of the municipality, which are well-vegetated zones (**Fig. 6 A**). According to Da Silva (2004), the highest values for the R-factor were found in the Amazon region, in that study the strong class of erosivity are between 7,000 to 9,800 MJ.mm/ha.he very strong are higher than 9800. In Paragominas only 1.7 % (325 km<sup>2</sup>) in the strong category and 97.1% (18,892 km<sup>2</sup>) in the very strong class.

### **3.2.2. Factor K (Soil erodibility)**

This factor is related to the soil and understanding its characteristics and properties is primordial, since its composition can affect the velocity of infiltration, water storage capacity, permeability, transport by rain, runoff, splash, dispersion and abrasion (De Lima, 2010). The soils most likely to undergo laminar erosion are in a small part of the northeast of the municipality and in the administrative limits (rivers boundaries) present in the soils Argisols (2%, 389 km<sup>2</sup>) and Gleysol (2%, 389 km<sup>2</sup>) a small part of the municipality, most of the municipality is Latosol type (79%, 15,377 km<sup>2</sup>) and has a low factor K (**Fig. 6 B**).

### **3.2.3. Factor LS (Topographic factor)**

Considered as one of the factors of high relevance, the factor has a strong influence (Wischmeier & Smith, 1978), because the volume of the floods is directly related to the degree of slope of the terrain, however in the municipality of Paragominas there is not a high degree of slope. **Fig. 6 C** shows the distribution of the factor LS, the lowest LS value was between 0 - 0.029 t h Mh<sup>-1</sup> mm<sup>-1</sup> corresponding to 8% (1,677 km<sup>2</sup>) of the municipality, while the higher value that was above 9 t h Mh<sup>-1</sup> mm<sup>-1</sup>, represents only 0.03% (7 km<sup>2</sup>) of Paragominas. The majority of Paragominas (80%, 15,539 km<sup>2</sup>) is between 0 to 1 t h Mh<sup>-1</sup> mm<sup>-1</sup>, with an average of 0.2 t h Mh<sup>-1</sup> mm<sup>-1</sup> and a standard deviation of 0.9 t h Mh<sup>-1</sup> mm<sup>-1</sup>.

### **3.2.4. Factor C (use/soil management) and P (conservationist practices)**

Brazil loses per year tons of soil from the surface layers, which are dragged into the streams, rivers, lakes and lowlands, resulting in an increase in bed volume and a decrease in the soil covered by vegetation (Dlamini *et al.*, 2011; Podwojewski *et al.*, 2011). When spatializing the factor CP values with soil types, it was observed that the lowest values are found in the latosols, Dystrophic Argilubic Pintossol, and Gleysol soils, while the Petroferric Eutrophic have a higher CP value. Values range from 0 to 1, with an average of 0.12 and a deviation of 0.21. Paragominas has an area of 71% (13,749 km<sup>2</sup>) with a CP factor between 0 - 0.086, where the primary and secondary vegetation is located, between 0.087 - 0.047 only 6% (1,076 km<sup>2</sup>) is found in cattle ranching regions and in 24% (4,640 km<sup>2</sup>) of Paragominas found a CP factor greater than 0.048 zones under agricultural cultivation (**Fig. 6 D**).

In the classification of land use and cover, the expressive classes were primary vegetation 47% (9,228 km<sup>2</sup>), secondary vegetation 22,9% (4,463 km<sup>2</sup>) and pasture 21,8% (4,252 km<sup>2</sup>) respectively, however the latter (pasture) has a higher CP value, followed by deforestation (1.7%, 337 km<sup>2</sup>) and agriculture (5.6% 1.086 km<sup>2</sup>). The data of intense agricultural production (Schlesinger, 2010) indicate that areas under crop production are increasingly overused around the world, exhausting soils capacity and making them less resilient and more vulnerable to erosion (Mazzali, 2000).

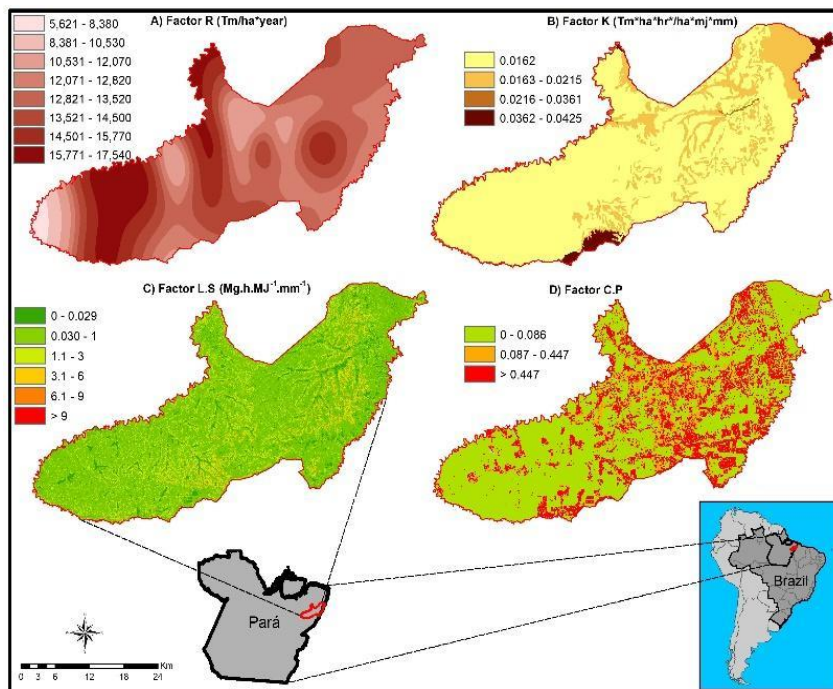


Fig. 6. Erosion estimation calculated by the RUSLE method (source: authors).

### 3.3. Soil loss estimation and vulnerability (Ecodynamic concept and RUSLE method)

These methods are widely applied in Brazil, allowing quantification and regionalization of the area with the highest risk of soil erosion. Paragominas presents ca. 77% (15,064 km<sup>2</sup>) of its territory with low-moderate (t ha<sup>-1</sup> year<sup>-1</sup>) degree of soil loss calculated by the RUSLE method, not so different from the method based on the ecodynamic concept, which was 60% (11,485 km<sup>2</sup>) with low and low-moderate degree of erosion of vulnerability to soil erosion (1 - 1.8 value). Similar situation was found for areas with high soil loss, with high soil loss, which represented only 3% (642 km<sup>2</sup>) degree of erosion according to RUSLE and 2.7% (584 km<sup>2</sup>) according to the ecodynamic concept, (Table 8, Fig. 7). This erosion can cause flow of superficial layers carrying organic matter, nutrients and seeds, resulting in high production costs (Parker *et al.*, 1995).

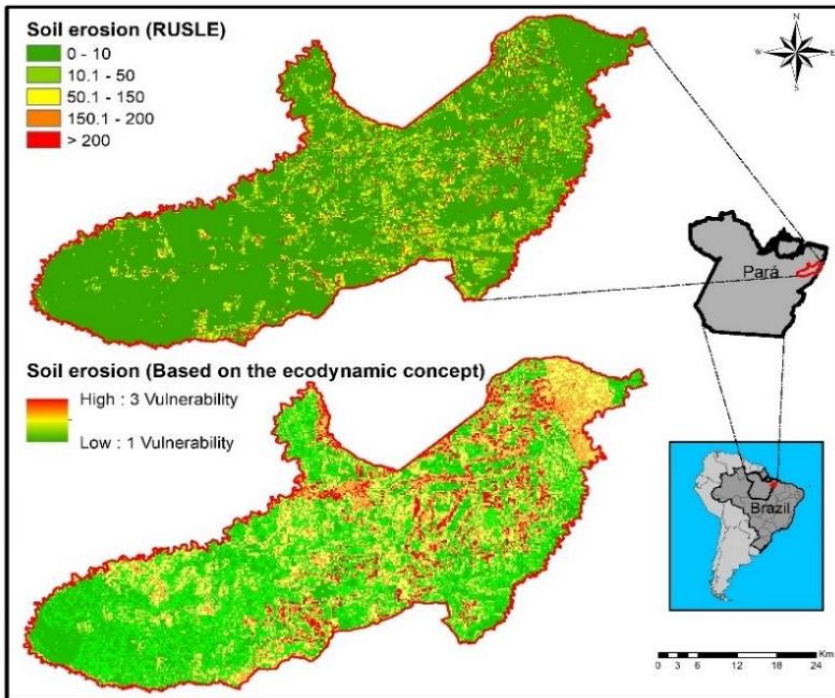
In almost all variables analyzed in the two methods with respect to their potential of soil erosivity the values are similar. Soil and slope attribute, for example, obtained 79% (15,377 km<sup>2</sup>) and 80% (15,572 km<sup>2</sup>) respectively of the municipality with low vulnerability value and erosivity estimation. Similar situation was noted for the low vulnerability degree for the vegetation attribute (soil use and cover), the result of erosivity value was 76% (14,793 km<sup>2</sup>) for the ecodynamic concept and 71% (13,820 km<sup>2</sup>) for calculation of RUSLE (CP). The only exception was in relation to the climate and factor R attribute, in both methods the rainfall data are used to generate the index vulnerability and erosion estimation, only 1.5% (292 km<sup>2</sup>) and 1.3 (253 km<sup>2</sup>) of Paragominas is considered as low vulnerability grade, according to the ecodynamic concept and RUSLE respectively. Consequently, for this factor, vulnerability values and erosion estimation were more significant throughout the municipality.

**Table 8.**

**Estimates of loss and vulnerability to soil erosion in Paragominas.**

RUSLE method				Ecodynamic model method			
Loss of soil (t ha <sup>-1</sup> year <sup>-1</sup> )	Degree of erosion	Municipal area		vulnerability values	Degree of erosion	Municipal area	
		Km <sup>2</sup>	%			Km <sup>2</sup>	%
0-10	low-moderate	15,064	77.4	1-1,5	low	4,672	24.3
10.1-50	moderate	2,093	10.8	1,6-1,8	low-moderate	6,813	35.4
50.1-150	high-moderate	1,665	8.6	1,9-2,1	moderate	5,256	26.9
151.1-200	high	218	1.1	2,2-2,4	high-moderate	2,141	10.7
>200	very high	424	2.2	2,5-3	high	584	2.7
-	-	19,465	100	-	-	19,465	100

Source: Elaborated by the author.



**Fig. 7.** Soil erosion soil calculated by the method based on the ecodynamic concept and RUSLE (source: authors).

### 3.4. Management and Soil conservation

The conservation practices aim to control soil and water losses in areas with agricultural activities, for example, without altering the productive capacity of the soil. Thus, it is important to adapt the soil conservation to the occupation of the area according to its capacity of use, so that the management practices can favor the erosion control, improving the water infiltration capacity in the soil, reducing the surface runoff that leads to the formation of aggregates and minimizing then the impact of rain drops.

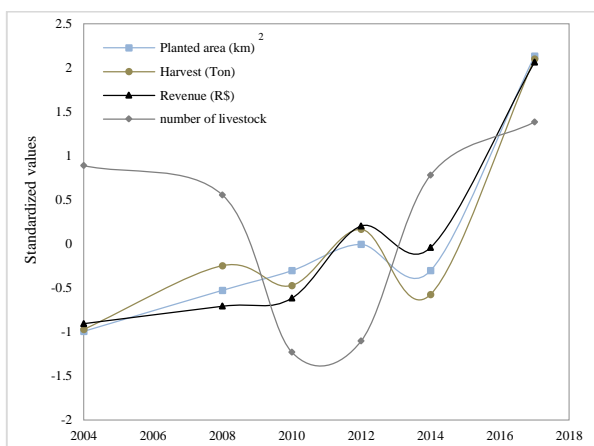
The conservation processes can be mechanical, edaphic and vegetative, depending on the cropping system. However, for a good result it is necessary to apply them

simultaneously, since each one develops a function and solves a part of the problem. The mechanical practices use artificial structures for the conduction or interception of surface runoff, while edaphic practices are related to the cropping system, controlling erosion and contributing to better soil fertility. On the other hand, the vegetative activities combat erosion based on the protection of the soil against the action of precipitation using the vegetation. For this reason, the maintenance of the adequate vegetation coverage in the soil is one of the basic principles for conservation, a process that is hampered by deforestation.

Paragominas was born in the troubled development process of the Brazilian Amazon, which was encouraged by the government in order to develop the region economically (Mahar, 1979). With around 50 years it owns already ca. 45% (8,773.3 km<sup>2</sup>) from its territory deforested and converted into other uses (INPE/Prodes, 2018). Over the years these deforested areas have been converted into other uses and the municipality is currently characterized mainly by agriculture, logging and mining extraction (especially bauxite).

In the classification analyzed in this study for the period from 2004 to 2017 we noticed that between 65% and 70% of Paragominas is covered by primary vegetation altered and secondary vegetation. In the last year (2017), for example, about 47% (9,110 km<sup>2</sup>) of the Paragominas territory was covered by altered primary vegetation, 23% (4,406 km<sup>2</sup>) by secondary vegetation and 5.5% (1,074 km<sup>2</sup>) correspond to agricultural activity (**Table 9**).

The values in mapped area were not so different from the agricultural areas available in the SIDRA system. However, when we analyzed the standardized agricultural production data of planted area, harvested tone and revenue, we noticed that in 2008 there was an increase in the cultivated area and in the quantity of harvested product, despite the decrease in revenue. The Worst scenario was identified in the year 2010, when crop area and revenue decreased while acreage increased. In 2012, it was the year of the best agricultural performance, according to SIDRA data, since there was a small increase in the area produced, with an increase in the area collected and a better performance of the income, during the same period the production of head of cattle was one of the worst. Another atypical event was in 2014, in that year, production and harvest decreased, but revenue grew (**Fig. 8**). The trend of inverse proportionality between agricultural and livestock production was also noticed, that is, when there is a decrease in livestock production, agricultural production rises, evidenced from 2008.



**Fig. 8.** Planted area, harvest and revenue of the agriculture data and number of livestock (source: authors).

Table 9.

**Land use classification (TerraClass of 2004 to 2014 and sentinel-2 of 2017)**

Class	TerraClass (km <sup>2</sup> )					Unsupervised Classification (km <sup>2</sup> )
	2004	2008	2010	2012	2014	2017
Annual Agriculture	172	449	683	835	1,019	1,074
Unobserved area	1275	2,378	264	1,702	859	-
Urban area	15	23	29	31	38	39
Deforestation	1,029	55	64	16	10	333
Forest	11,180	10,757	10,645	10,602	10,490	9,110
Hydrography	26	50	50	50	26	45
Mining	-	7	18	31	-	69
Occupation mosaic	27	23	12	7	44	24
Non-Forest	7	7	7	7	7	-
Others	16	5	34	8	4	166
Pasture with exposed soil	-	-	0	0	-	-
Clean pasture	2,871	2,775	2,678	2,569	3,148	-
Dirty pasture	540	712	418	334	441	4,199
Reforestation	-	-	134	252	208	-
Regeneration with pasture	707	224	1,005	281	405	-
Secondary vegetation	1,601	2,002	3,423	2,741	2,767	4,406
Total	19,465	19,465	19,465	19,465	19,465	19,465

Source: Elaborated by the authors.

Paragominas was one of the Amazon municipalities with higher rate of illegal logging for many years. According to Imazon (Institute of Man and the Environment of the Amazon), during the period from 2007 to 2012, ca. 74% (960 km<sup>2</sup>) of the municipality's logging was carried out without authorization, in all that years of monitoring the logging without permission was higher than the authorized one (Fig. 9). When analyzing the official data of wood production in m<sup>3</sup> and revenue of SIDRA/PEVS - Production of Plant Extraction and Silviculture, we noticed that there is a significant difference between the production of wood in m<sup>3</sup> in relation to the generated revenue. In the period from 2008 to 2012 the revenue increased, while the harvest decreased, that difference is quite evident when comparing the area of logging monitored by Imazon with the data of harvest and revenue in 2008, that year the logging in the municipality was 80% (601 km<sup>2</sup>) of the total (749 km<sup>2</sup>), and the data of the PEVS also show an overestimation of the revenue in relation to production (Fig. 10), this can be explained by the illegality in the sector.

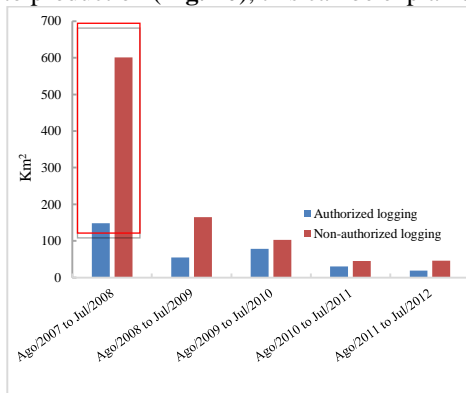


Fig. 9. Logging monitoring (Imazon).

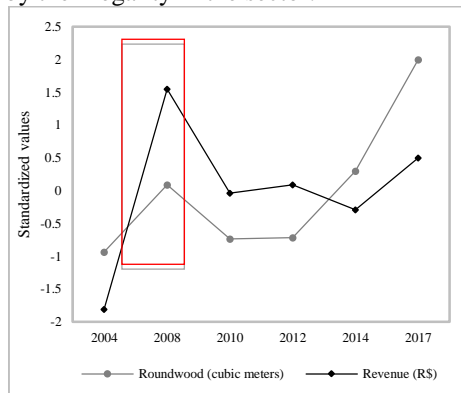


Fig. 10. Logging data (PEVS).

Another quite significant activity mapped out in this study was mining. This activity has existed for a long time, but in the past, it did not require sophisticated technological processes. Before the areas were mined with a semi-mechanized extraction system, today they present a complex set of large equipment. One of the ores with a lot of potential in the Amazon is bauxite, which produces aluminum, which is used in various segments, such as: packaging, transportation, civil construction, electricity, consumer goods, machinery and equipment and others. It is important to point out that mining activity significantly improves the economy of a region, but also causes great degradation, since it unbalances the environment in huge extension of land and changes the soil components (Reis, 1999).

Paragominas has one of the greater enterprises of mineral activity of the state. According to DNPM (2018), the municipality presents 43% (8,308 km<sup>2</sup>) of its territory under mining process, with large majority (39%, 3,246 km<sup>2</sup>) having a research permit, 27% (2,318 km<sup>2</sup>) with a request of mining and 25% (2,070 km<sup>2</sup>) in concession (Fig. 11 A). Regarding the ore class, bauxite represents ca. 72% (5,998 km<sup>2</sup>) of the total in mining processes classified by DNPM (Fig. 11 B). As for the use of the extracted minerals (53%, 4,425 km<sup>2</sup>) there is no information about the use, for the metallurgy activities are destined 25% (2,035 km<sup>2</sup>) and about 1,726 km<sup>2</sup> (21%) of area is destined for extraction of ore for industry (Fig. 11 C).

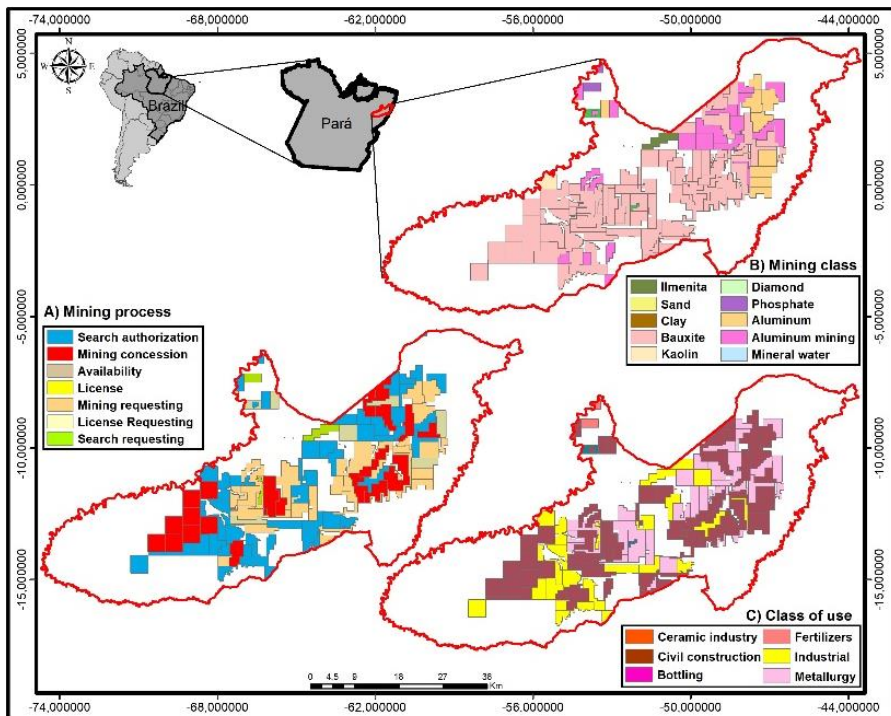


Fig. 11. DNPM mining data (source: authors).

#### 4. CONCLUSION

The attributes analyzed in the two methods presented in their majority low and intermediate vulnerability estimation and potential of soils loss, it requires attention to avoid such regions to become areas with high potential for erosion vulnerability in the future. For the two analyzed methods (Ecodynamic concept and RUSLE) only 3% of Paragominas had a high potential of erosivity. The identification of regions with vulnerability to soil erosion potential is efficient to assist decision making and territorial management, in order to answer important environmental questions, determine low operating costs, plan activities in management practices and environmental conservation of the municipality.

Monitoring erosion over large area extensions is a costly process, so geoprocessing becomes a useful tool for estimating soil loss. The multicriteria analysis using GIS tools (Geographic Information System) were extremely important in this study, because they mapped and estimated the vulnerability potential to soil erosion for the Paragominas municipality, however there are not many studies on the subject in the Amazon region, therefore it is necessary further researches with field analysis in the region to corroborate the results. Soil erosivity is a natural process, however, inadequate human actions regarding soil use generate irreversible environmental degradation. The intense erosion process, for example, leads to soil impoverishment and pollution of water networks, causing economic, social and environmental problems on a scale from local to the global. Thus, the results of this study reveal the need for greater attention in the areas of greater environmental risk.

Good data collection and analysis of rainfall, topographic information, land cover and management system lead to significant results to be obtained from areas susceptible to degrading erosive processes. The studies associated with soil erosion are fundamental, both for agricultural conservation practices, as to subsidize the planning environmental, in which economic practices must be calculated under conservationist principles.

#### REFERENCES

- Abel, W., Gilson Filho, M., José, R.S., José, P. And José, L.S., 2005. Determinação de Fatores da Equação Universal de Perda de Solo em Sumé, Pb. *Revista Brasileira De Engenharia Agrícola e Ambiental*, 9(2).
- Alves, L., Carvalho, E. And Silva, L., 2014. Diagnóstico Agrícola Do Município de Paragominas, Pa. *Embrapa Amazônia Oriental-Boletim de Pesquisa E Desenvolvimento (Infoteca-E)*.
- Angima, S., Stott, D., O'neill, M., Ong, C. and Weesies, G., 2003. Soil Erosion Prediction Using Rusle For Central Kenyan Highland Conditions. *Agriculture, Ecosystems & Environment*, 97(1), Pp. 295-308.
- Arnesen, A.S., 2009. Análise Da Vulnerabilidade Natural à Erosão na Bacia Hidrográfica do Rio Tijucas/Sc Através de Técnicas de Geoprocessamento: Um Subsídio à Governança Territorial.
- Assad, E.D. And Sano, E.E., 1998. Sistema de Informações Geográficas: Aplicações Na Agricultura.
- Bakker, M.M., Govers, G., Van Doorn, A., Quetier, F., Chouvardas, D. And Rounsevell, M., 2008. The Response of Soil Erosion and Sediment Export to Land-use Change in Four Areas of Europe: The Importance of Landscape Pattern. *Geomorphology*, 98(3), Pp. 213-226.
- Barbosa, J.H.C. And De Faria, S.M., 2006. Aporte de Serrapilheira ao Solo em Estágios Sucessionais Florestais na Reserva Biológica de Poço das Antas, Rio de Janeiro, Brasil. *Rodriguésia*, 461-476.
- Becker, B.K., 2005. Geopolítica da Amazônia. *Estudos Avançados*, 19(53), Pp. 71-86.
- Bertoni, J., & Lombardi Neto, F. (2005). Conservação do solo. ícone. São Paulo.
- Bertoni, J., & Lombardi Neto, F. (2008). Conservação do solo, 7ª Edição, Editora ícone. São Paulo.
- Bertoni, J.; Lombardi Neto, F. Conservação do solo. 8.ed. São Paulo: Ícone, 2012. 360p.

- Bertol, O.J., Rizzi, N.E., Bertol, I. And Roloff, G., 2007. Perdas de Solo e Água e Qualidade do Escoamento Superficial Associadas à Erosão entre Sulcos em área Cultivada sob Semeadura direta e Submetida às Adubações Mineral e Orgânica. *Revista Brasileira de Ciência do Solo*, 31(4).
- Câmara, G., Moreira, F.R., Barbosa, C., Almeida Filho, R. And Bönisch, S., 2001. Técnicas de Inferência Geográfica. *Câmara, G.; Davis, C.; Monteiro, Amv; Paiva, Ja.*
- Cerdan, O., Govers, G., Le Bissonnais, Y., Van Oost, K., Poesen, J., Saby, N., ... & Klik, A. (2010). Rates and spatial variations of soil erosion in Europe: a study based on erosion plot data. *Geomorphology*, 122(1-2), 167-177.
- Crepani, E., Medeiros, J.D., Hernandez Filho, P., Florenzano, T.G., Duarte, V. And Barbosa, C.C.F., 2001. *Sensoriamento Remoto e Geoprocessamento Aplicados ao Zoneamento Ecológico-Econômico e ao Ordenamento Territorial*. Inpe São José dos Campos.
- Crepani, E., Medeiros, J. D., & Palmeira, A. F. (2004). *Intensidade pluviométrica: uma maneira de tratar dados pluviométricos para análise da vulnerabilidade de paisagens à perda de solo*. São José dos Campos: INPE.
- Chaplot, V., Giboire, G., Marchand, P., & Valentin, C. (2005). Dynamic modelling for linear erosion initiation and development under climate and land-use changes in northern Laos. *Catena*, 63(2-3), 318-328.
- Da Silva, A. M. (2004). Rainfall erosivity map for Brazil. *Catena*, 57(3), 251-259.
- Da Silva Barros, P. P., Fiorio, P. R., Martins, J. A., & Demattê, J. A. M. (2016). Comparação entre o uso e ocupação e perda de solo, nos anos de 1995 e 2010, na microbacia hidrográfica do Ceveiro. *Revista Ambiência*, v. 12, n. 2, p. 513-523.
- De Lima, E. R. V. (2010). Erosão do solo: fatores condicionantes e modelagem matemática. *CADERNOS DO LOGEPA*, 3(1).
- Dlamini, P., Orchard, C., Jewitt, G., Lorentz, S., Titshall, L., & Chaplot, V. (2011). Controlling factors of sheet erosion under degraded grasslands in the sloping lands of KwaZulu-Natal, South Africa. *Agricultural Water Management*, 98(11), 1711-1718.
- Embrapa, 1999. Centro Nacional De Pesquisas de Solos. Sistema Brasileiro de Classificação de Solos. *Rio de Janeiro: Embrapa/Solos*, 412p.
- Espig, S.A., Freire, F.J., Marangon, L.C., Ferreira, R.L.C., Freire, Maria Betânia Galvão dos Santos and Espig, D.B., 2009. *Sazonalidade, Composição e Aporte de Nutrientes da Serapilheira em Fragmento de Mata Atlântica*. São José dos Campos: Scielo Brasil.
- ESRI, A. (2012). ArcGIS 10.1. Environmental Systems Research Institute, Redlands, CA.
- Fao - Food and Agriculture Organization of the United Nations. Statistical Databases. Agriculture2017.
- Ferreira, L.G., Ferreira, N.C. and Ferreira, M.E., 2008. Sensoriamento Remoto da Vegetação: Evolução E Estado-Da-Arte. *Acta Scientiarum.Biological Sciences*, 30(4).
- Fiorio, P. R., da Silva Barros, P. P., de Oliveira, J. S., & Nanni, M. R. (2016). Estimativas de perda de solo em ambiente SIG utilizando diferentes fontes de dados topográficos Estimates of soil loss in a GIS environment using different sources of topographic data. *AMBIÊNCIA*, 12(1), 203-216.
- Frostick, L. E., Lucas, P. M., & Reid, I. (1984). The infiltration of fine matrices into coarse-grained alluvial sediments and its implications for stratigraphical interpretation. *Journal of the Geological Society*, 141(6), 955-965.
- Gomes, A.R., 2000. Avaliação Da Vulnerabilidade À Perda De Solo Em Região Semiárida Utilizando Sensoriamento Remoto E Geoprocessamento-Área Piloto De Parnamirim-Pe, Inpe - Instituto Nacional De Pesquisas Espaciais.
- Guerra, A. T., da Silva, A. S., & Botelho, R. G. M. (2009). Erosão e conservação dos solos: conceitos, temas e aplicações. Bertrand Brasil.
- Hall, A.L., 1989. *Developing Amazonia: Deforestation and Social Conflict in Brazil's Carajás Programme*. Manchester University Press.
- Ibge - Instituto Brasileiro De Geografia E Estatística, 1996. *Mapa Geomorfológico das Folhas Sa. 23v-A E Sa 23 V-C*.



- Inpe – Instituto Nacional de Pesquisas Espaciais. *Projeto Prodes - Monitoramento da Floresta Amazônica Brasileira por Satélite*. Disponível em: <<http://www.obt.inpe.br/prodes/index.html>>. Acessado em: 30/09/2018.
- Kohlhepp, G., 2002. Conflitos De Interesse no Ordenamento Territorial da Amazônia Brasileira. *Estudos Avançados*, 16(45), Pp. 37-61.
- Lang, K.J., Prunty, L., Schroeder, S. And Disrud, L., 1984. Interrill Erosion as an Index of Mined Land Soil Erodibility. *Transactions of The Asae*, 27(1), Pp. 99-0104.
- Lal, R. (2017). Soil erosion by wind and water: problems and prospects. In *Soil erosion research methods* (pp. 1-10). Routledge.
- Lepsch, I. F. (2016). *Formação e conservação dos solos*. Oficina de textos.
- Mahar, D.J. 1979. *Frontier development policy in Brazil: a study of Amazonia*. Praeger, Nova York
- Mazzali, L., 2000. *O Processo Recente De Reorganização Agroindustrial: Do Complexo À Organização" Em Rede"*. Unesp.
- Mota, G. V., Vitorino, M. I., Cunha, A. C. D., Kuhn, P. A. F., Ferreira, D. B. D. S., Santos, D. M., ... & Souza, J. R. S. D. (2009). Precipitação sazonal sobre a Amazônia oriental no período chuvoso: observações e simulações regionais com o RegCM3.
- Navas, A., Machín, J. And Soto, J., 2005. Assessing Soil Erosion in a Pyrenean Mountain Catchment Using Gis And Fallout 137 Cs. *Agriculture, Ecosystems & Environment*, 105(3), Pp. 493-506.
- Oakes, E. G. M., Hughes, J. C., Jewitt, G. P. W., Lorentz, S. A., & Chaplot, V. (2012). Controls on a scale explicit analysis of sheet erosion. *Earth Surface Processes and Landforms*, 37(8), 847-854.
- Parker, D.B., Michel, T.G. And Smith, J.L., 1995. Compaction and Water Velocity Effects on Soil Erosion in Shallow Flow. *Journal of Irrigation and Drainage Engineering*, 121(2), Pp. 170-178.
- Pezzatto, A.W. And Wisniewski, C., 2006. Produção de Serapilheira Em Diferentes Seres Sucessionais da Floresta Estacional Semidecidual No Oeste Do Paraná. *Floresta*, 36(1).
- Podwojewski, P., Janeau, J. L., Grellier, S., Valentin, C., Lorentz, S., & Chaplot, V. (2011). Influence of grass soil cover on water runoff and soil detachment under rainfall simulation in a sub-humid South African degraded rangeland. *Earth Surface Processes and Landforms*, 36(7), 911-922.
- Prado, H. Solos do Brasil. 2.ed. Piracicaba: H. Prado, 2001. 220p.
- Rodrigues, T.E., Junior, Oliveira. R. C De, Silva, J. M. L Da, Valente, M.A. and Capeche, C.L., 1991. *Caracterização Físico-Hídrica Dos Principais Solos Da Amazônia Legal*.
- Rodrigues, T. E., Silva, R. D. C., Da Silva, J. M. L., De Oliveira Junior, R. C., Gama, J. D. F., & Valente, M. A. (2003). Caracterização e classificação dos solos do município de Paragominas, Estado do Pará. *Embrapa Amazônia Oriental-Documentos* (INFOTECA-E).
- Santos, G. G., Griebeler, N. P., & Oliveira, L. F. C. D. (2010). Chuvas intensas relacionadas à erosão hídrica.
- Schaefer, C.E.R., Silva, D.D., Paiva, K.W.N., Pruski, F.F., Albuquerque Filho, M.R. and Albuquerque, M.A., 2002. Perdas de Solo, Nutrientes, Matéria Orgânica e Efeitos Microestruturais Em Argissolo Vermelho-Amarelo Sob Chuva Simulada. *Pesquisa Agropecuária Brasileira*, 37(5), Pp. 669-678.
- Schlesinger, S. (2010). Onde pastar?: o gado bovino no Brasil. Rio de Janeiro: Fase.
- Schmidt, J. C. J. (1947). O clima da Amazônia. Instituto brasileiro de geografia e estatística.
- Serra, M.A., 1998. *The Social Impacts of Regional Development Policies In Eastern Amazonia: A Case Study Of Parauapebas*, London School of Economics and Political Science.
- Tricart, J., 1977. Ecodinâmica. *Série Recursos Naturais E Meio Ambiente*. Supren/Ibge.
- Verheijen, F. G., Jones, R. J., Rickson, R. J., & Smith, C. J. (2009). Tolerable versus actual soil erosion rates in Europe. *Earth-Science Reviews*, 94(1-4), 23-38.
- Watrin, O.D.S. And Da Rocha, A., 1992. Levantamento Da Vegetação Natural E Do Uso da Terra No Município De Paragominas (Pa) Utilizando Imagens Tm/Landsat. *Embrapa Amazônia Oriental-Séries Anteriores* (Infoteca-E).
- Watson, D. And Lafen, J., 1986. Soil Strength, Slope, And Rainfall Intensity Effects on Interrill Erosion. *Transactions of The Asae*, 29(1), Pp. 98-0102.
- Wischmeier, W. H., & Smith, D. D. (1978). Predicting rainfall erosion losses-a guide to conservation planning. *Predicting rainfall erosion losses-a guide to conservation planning*. USDA, 62 pp.

## **VARIABILITY ANALYSIS OF TEMPORAL AND SPATIAL ANNUAL RAINFALL IN THE MASSIF OF AURES (EAST OF ALGERIA)**

*Adel KHENTOUCHE<sup>1</sup>, Hadda DRIDI<sup>1</sup>*

DOI: 10.21163/GT\_2019. 141.03

### **ABSTRACT:**

The spatial and temporal variations of precipitation in the Aures Massif of Algeria from 1974 to 2009 were investigated using a geostatistical approach. This diachronic approach did not allow a real cyclical differentiation but reveals two phases obvious distinct wet and dry, very marked. Annual rainfall variability is associated with disruption stationary in the series. The use of the rainfall index (PCI) clearly explains the state of variability and the temporal behavior of the annual rains. The climate of the region was in the semi-arid upper floor, and had considerable rainfall. However, this situation has been reversed since 1991. It is predicted that rainfall will decrease from 50 to 105 mm/year on altitudes and the foothills, from 13 to 50 on Saharan areas. Statistical tests reveal breaks around 1991-1994. We notice a severe drought which began in 1991. The results indicated that the spatial pattern of precipitation was primarily the local climate effect significant type, and inside the massif, altitude and latitude are the two factors that control this variability that installs and starts to take the form of climate change

*Key-words: Variability, Drought, Breaking, Rain phase, Standardized index of rain.*

### **1. INTRODUCTION**

Precipitation is one of the most important climate factors affecting human economies and terrestrial ecosystems (Xoplaki et al., 2004; Santos et al., 2007; Pauling et al., 2006). Increasing evidence indicates that temporal and spatial variations in precipitation have been taking place at the global scale (Semenov & Bengtsson, 2002; Haidu, 2003; Labat, 2005), regional scale (Cannarozzo et al., 2006; Khon et al., 2007; Pauling, et al., 2006) and local scale over the past few centuries (Le Quesne et al., 2006).

The variability of the time series is treated with the help of various methods based mainly on specialized programs to detect the year of rupture and the way of segmentation or dividing this time series (Haidu, 2004; Haidu & Magyari-Sàska, 2009). In this study we have based on some statistical laws including the Hubert segmentation method (Hubert, 2000) to determine the year announcing the beginning of change in the behavior of the time series.

Algeria is a country of the subtropical zone of Northern Africa. Its climate is very different between its regions (North-South, East-West). It is of Mediterranean type on the entire North fringe that includes the coastline and the Tellian Atlas (warm and dry summers, humid and cool winters), semi-arid on highlands in the centre of the country. The precipitation is characterized by a very significant spatial and temporal variability while annual rainfall tranche decreases as one move towards southern latitudes. They fall less than 100 mm in the South of the Saharan Atlas. In desert regions, the spatio-temporal

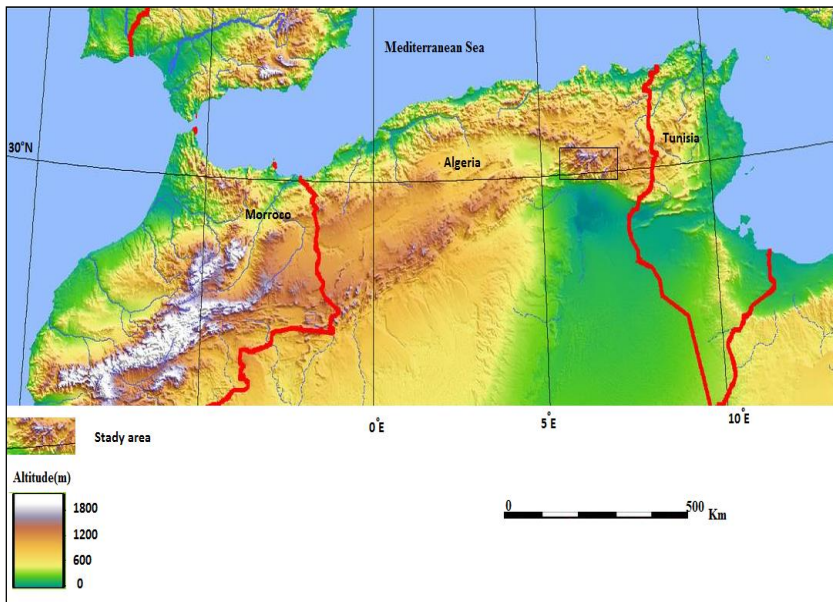
---

<sup>1</sup> *Laboratory of natural risks and area development, Earth sciences Institute, Batna 2 University, Algeria.05000, adel.khentouche@yahoo.com, hadda.dridi@gmail.com*

pattern of precipitation has an especially strong influence on eco-hydrological processes. In addition to the rains decrement from the North to the South, there is also a decrease from the East to the West. Mountain area is a special land space unit, with remote geographical location, large resource gradient, disaster proneness, ecological vulnerability, and other characteristics (Jansky et al., 2002). The Aures massif, which is a separate and important physical unit by its size extent, does not escape from these influences concerning the spatial distribution of precipitation nor from the global changes influences which are animated by global temperatures increase. This study aims to characterize the pluviometry and drought dynamics in the Aures massif through searching change-points and trends in time series.

## 2. STUDY AREA

It extends between the longitudes 5°24'to 7° E and the latitudes 34°45' to 36°N. The studied zone is located in the North-East of the Algerian territory (**Fig. 1**) and it covers 1203 Km<sup>2</sup>. It belongs to the Highlands of Constantine. The studied zone is a mountainous region belongs to the semi-arid floor and it is facing very serious problems of water lack. Precipitation is characterized by a spatial and temporal distribution very irregular from one station to another.

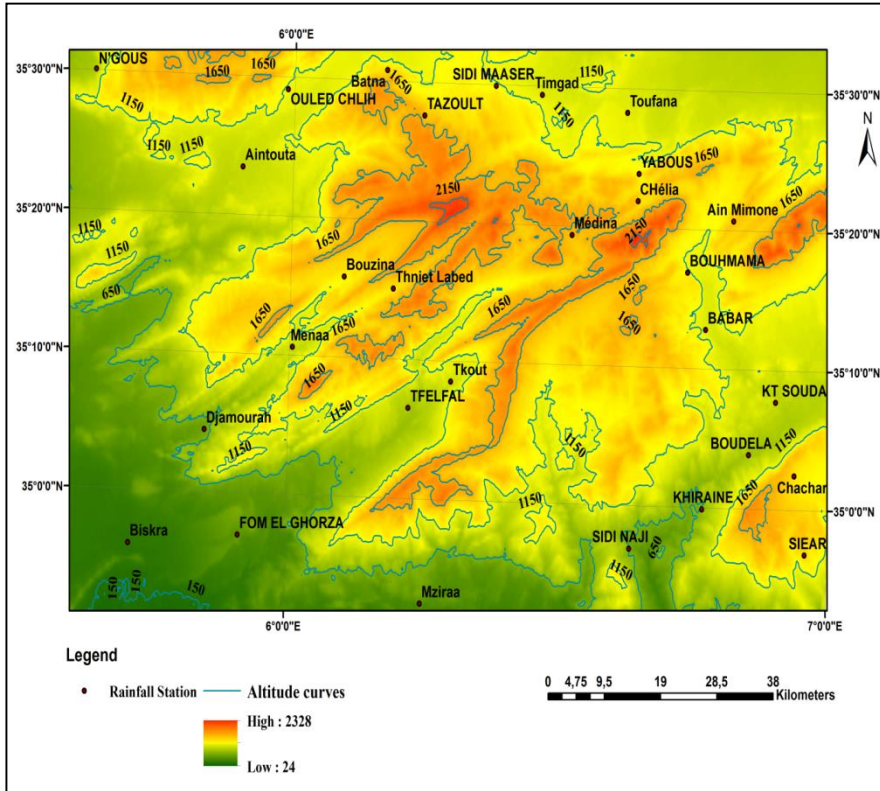


**Fig.1.** Situation of the studied zone.

## 3. DATA AND METHOD

Data come from two agencies responsible for the pluviometric network in Algeria: The National Agency of Hydraulic Resources and the National Office of Meteorology. To realize this study, many pluviometric stations have been retained in order to form the most complete annual data base and the most representative as possible of the studied zone. So, 32 stations have been selected with sufficient time series well distributed in the studied

zone and strictly criticized by statistical methods (homogeneity and bridging test) to study the pluviometric regime trend. We have used many statistical tests of the change-point in the stationarity as well as the test on sequential trends in order to account the temporal and spatial evolution of the pluviometric regime. The standardized index which is widely used in the rainfall variations studies will give the distinction between the phases and the frequency of dry and humid years. The observed stations locations were shown in **Fig. 2**.



**Fig. 2.** Map with the pluviometric stations location used in the study.

### 3.1. Statistical Analysis

This study was carried out by the application of the statistical tests of rupture (point of change of time series of precipitation) detection on annual time scale. The choice of this method is based on the robustness of their bases. The tests were carried out with allow to characterize, as well as possible, the evolution of climate parameters; and identify the pivotal years of climate change. KhronoStat software (Boyer, 2002) is adapted to all variables (climatic, hydrological, and meteorological). However, it requires complete series with no gaps. Its choice in this study is justified by the robustness of its tests and also by its success through several similar studies. It can evolve on an annual, monthly or daily scale depending on the needs expressed.

The second test category concerns the homogeneous character of the series (Pettitt test, Buishand test, Hubert test, Bayesian methods or Lee & Heghinian test): they relate to the

detection of breaks in a time series. KhronoStat is a statistical model developed by IRD (Research Institute for Development) at the House of Water Sciences (MSE) of Montpellier. It was developed as part of a study on climate variability in West and Central Africa and is oriented on the analysis of hydroclimatic series.

### 3.2. Statistical tests

**Pettit’s test.** Pettit test is a rank-based test for detecting significant changes in the mean of time series data when the exact time of change is unknown. The test is considered robust to changes in the distributional form of time series and relatively powerful compared to Wilcoxon-Mann-Whitney test, cumulative sum and cumulative deviations. Furthermore, Pettitt test has been widely adopted to detect changes in climatic and hydrological time series data.

The null and alternative hypotheses will be reformulated as follows.

Ho: the T variables follow one or more distributions that have the same location parameter.

Two-tailed test: Ha - there is a time  $t$  when there is a change of location parameter in the variables.

Left-tailed test: Ha - there is a time  $t$  when the location parameter in the variables is reduced by D.

Right-tailed test: Ha - there is a time  $t$  when the location parameter in the variable is augmented by D.

The statistic used for the Pettit’s test is computed as follows:

$$U_{t, n} = \sum_{i=1}^t \sum_{j=t+1}^n D_{ij} \tag{1}$$

$$D_{ij} = -1 \text{ if } (x_i - x_j) > 0, D_{ij} = 0 \text{ if } (x_i - x_j) = 0, D_{ij} = 1 \text{ if } (x_i - x_j) < 0 \tag{2}$$

**Buishand’s test.** Buishand’s test is suitable for variables following any form of distribution whose properties have been mainly studied for the normal case. For this study, Buishand focuses on the case of the two-tailed test and the U statistic. For U statistic, the null and alternative hypotheses are given by;

Ho - the T variables follow one or more distributions that have the same mean.

Two-tailed test: Ha - there exists a time  $t$  when variables change in mean.

U of Buishand is defined by:

$$U = \frac{\sum_{k=1}^{n-1} S_k / D_x}{n(n+1)} \tag{3}$$

$$S_k = \sum_{t=1}^k ([X_i] - \bar{X}) \tag{4}$$

where the terms  $S_k$  and  $D_x$  are respectively partial sum and standard deviation given respectively by equations (3) and (4).

In case of rejection of the null hypothesis, no estimate of the break date is proposed by this test. In addition to this procedure, the construction of a control ellipse makes it possible to analyze the homogeneity of the series of  $(x_t)$ . The variable  $S_k$ , defined above, follows a normal distribution of zero mean and variance

$[k(N - k)\sigma^2] / N$ ,  $k = 0 \dots N$  under the null hypothesis

**Hubert Segmentation.** The principle of this procedure is to "split" the series into  $m$  segments ( $m > 1$ ) so that the average calculated on any segment is significantly different from the average of the segment ( $s$ ) neighbours. Such a method is suitable for looking for multiple changes of mean.

Segmentation is defined as follows.

Any series  $x_i$ ,  $i = i_1, i_2$  with  $i_1 \geq 1$  and  $i_2 \leq N$  where ( $i_1 < i_2$ ) constitutes a segment of the initial series of  $(x_i)$ ,  $i = 1 \dots N$ .

Any partition of the initial series in  $m$  segments is a segmentation of order  $m$  of this series. From a particular segmentation of order  $m$  practiced on the initial series, we define:

$i_k$ ,  $k = 1, 2, \dots, m$

$$* N_k = I_k - I_{k-1} \quad (5)$$

$$X_K = \frac{\sum_{i=i_{(k-1)+1}}^{i=i_k} X_i}{N_K} \quad (6)$$

$$Dm = \sum_{K=1}^K d_K \quad (7)$$

$$d_k = \sum_{i=i_{(k-1)+1}}^{i=i_k} ([X_i - \bar{X}_k])^2 \quad (8)$$

The segmentation retained must be such that for a given order  $m$  of segmentation, the quadratic difference  $Dm$  is minimum. This condition is necessary but not sufficient for the determination of the optimal segmentation. It must be added to the constraint that the *Averages* of two contiguous segments must be significantly different. This constraint is satisfied by application of the **Scheffé test**. For a given segmentation order, the algorithm determine the optimal segmentation of a series that is such that the deviation  $Dm$  is minimal. This procedure can also be interpreted as a stationary test, the null hypothesis being the studied series is non-stationary. If the procedure doesn't produce acceptable segmentations of order bigger or equal to two, the null hypothesis is accepted

**Bayesian method of Lee and Heghinian.** The Bayesian method of Lee and Heghinian aims at confirming or invalidating the hypothesis of a change of mean in the series. It is a parametric approach whose application on a series requires a normal distribution of the values of this one. The absence of rupture in the series constitutes the null hypothesis. The procedure is based on the following model:

$$X_i = \begin{cases} \mu + \varepsilon_i & I = 1, \dots, \tau \\ \mu + \sigma + \varepsilon_i & I = \tau + 1, \dots, N \end{cases} \quad (9)$$

where the  $\varepsilon_i$  are independent and normally distributed, of zero mean and variance  $\sigma^2$ . The variables  $\tau$ ,  $\mu$ ,  $\delta$  and  $\sigma$  are unknown parameters;  $\tau$  and  $\delta$  represent respectively the position of the break in time and the amplitude of the change on the average. The possible change (the position and the amplitude) corresponds to the mode of the posterior distributions of  $\tau$  and  $\delta$ .

The method thus provides the probability that the rupture occurs at the moment  $\tau$  in a series where it is assumed a priori that there is indeed a change at an indeterminate time.

**Table 1. Detection break results applied to pluviometric series**

Station	Period	Segmentation (Hubert)				Buichand	Pettit	Lee - Heghinian
		Start	End	Mean	Std			
Chélia	1974-2009	1974	1992	536.27	95.83	Rejected	1992	1992
		1993	2009	380.96	116.33			
Yabous	1974-2009	1974	1992	502.28	92.47	Rejected	1992	1992
		1993	2009	346.53	100.35			
Toufana	1974-2009	1974	1991	500.38	94.78	Rejected	1991	1991
		1992	2009	342.66	101.27			
Bouhmama	1974-2009	1974	1992	530.82	72.21	Rejected	1992	1992
		1993	2009	389.92	90.45			
Medina	1974-2009	1974	1992	474.86	72.96	Rejected	1992	1992
		1993	2009	355.10	76.31			
Timgade	1974-2009	1974	1992	399.84	105.61	Rejected	1992	1992
		1993	2009	269.77	71.28			
Babar	1974-2009	1974	1991	475.97	80.61	Rejected	1991	1991
		1992	2009	264.63	90.30			
A. Mimoun	1974-2009	1974	1991	471.65	79.57	Rejected	1991	1991
		1992	2009	277.76	71.65			
S. Mansser	1974-2009	1974	1991	364.03	64.32	Rejected	1991	1991
		1992	2009	199.62	77.86			
Kheiran	1974-2009	1974	1992	310.28	98.70	Rejected	1992	1992
		1993	2009	180.87	102.49			
Tkout	1974-2009	1974	1991	324.87	79.68	Rejected	1991	1991
		1992	2009	185.28	55.53			
F. Lgherza	1974-2009	1974	1991	91.70	25.60	Rejected	1991	1991
		1992	2009	44.58	14.32			
A. Touta	1974-2009	1974	1991	351.53	78.99	Rejected	1991	1991
		1992	2009	228.25	70.95			
Biskra	1974-2009	1974	1991	143.41	27.38	Rejected	1991	1991
		1992	2009	81.67	36.29			
Bouzina	1974-2009	1974	1991	303.66	72.52	Rejected	1991	1991
		1992	2009	177.55	52.81			
Menaar	1974-2009	1974	1991	323.13	74.22	Rejected	1991	1991
		1992	2009	215.56	55.98			
Tifelfel	1974-2009	1974	1991	250.49	63.39	Rejected	1991	1991
		1992	2009	167.86	68.92			
Batna	1974-2009	1974	1991	455.47	64.59	Rejected	1991	1991
		1992	2009	260.42	98.13			
Chechar	1974-2009	1974	1991	304.55	89.80	Rejected	1991	1991
		1992	2009	164.68	54.34			
Djemoura	1974-2009	1974	1991	159.66	44,036	Rejected	1991	1991
		1992	2009	98.57	21.63			
Mziraa	1974-2009	1974	1993	60.24	12.88	Rejected	1992	1993
		1994	2009	33.71	14.18			
Kh.S.Nadji	1974-2009	1974	1991	67.54	14.74	Rejected	1991	1991
		1992	2009	29.21	10.38			
TH. Abed	1974-2009	1974	1991	341.77	73.22	Rejected	1992	1991
		1992	2009	195.33	55.56			
Doucen	1974-2009	1974	1991	79.24	35.37	Rejected	1991	1991
		1992	2009	45.14	21.70			
Merouana	1974-2009	1974	1991	374.47	63.28	Rejected	1991	1991
		1992	2009	205.51	46.77			
Tazoult	1974-2009	1974	1991	390.37	71.79	Rejected	1991	1991
		1992	2009	231.41	92.24			
Siar	1974-2009	1974	1991	111.00	30.81	Rejected	1991	1991
		1992	2009	47.66	18.17			
K.Souda	1974-2009	1974	1991	184.16	35.92	Rejected	1991	1991
		1992	2009	110.84	7.49			
Boudela	1974-2009	1974	1991	249.06	33.80	Rejected	1991	1991
		1992	2009	156.20	30.27			
O.Chlih	1974-2009	1974	1991	377.50	66.96	Rejected	1991	1991
		1992	2009	220.27	79.34			

## 4. RESULTS AND DISCUSSION

### 4.1. Rain phases and breaks

The null hypothesis  $H_0$  was rejected in the case of rank correlation test (confidence interval 90%-95%) for all the rainfall stations. The results obtained showed a trend effect between the successive values of certain time series. It has been concluded that these series, which we are going to analyse, are devoid of randomness. The used methods to highlight this rupture were: Buishand's U-statistic, Pettitt's test, Lee and Heghinian's Bayesian method, and Hubert's segmentation method. Break detection results are reported in **Table 1**.

In all studied zone and for all of the tests, the change-point (rupture) occurred during the years 1991 and 1994 for most pluviometric stations. These rains, in turn, have experienced considerable decrease of more than 40% for the majority of the stations. The pluviometric deficits during this period are considerable and important and the rainfall values of each station are expressed by a significant and varied standard deviation. The breaks observed were included two separate phases: the first one is dominated by humid influences, while the second presents dry trends.

### 4.2. Rainfall index

To better endow study of arguments, the defined rainfall index was calculated as a reduced centered variable (Lamb, 1982).

$$SPI = \frac{(P_a - P_m)}{\sigma_p} \quad (10)$$

where:  $SPI$  is the standardized index of rain of the year  $a$ ,  $P_a$  is the pluviometry of the year  $a$ ,  $P_m$  is the average annual pluviometry on the reference period 1974-2009 and  $\sigma_p$  is the standard deviation of the pluviometry on the same reference period.

This index reflects a pluviometric excess or deficit for the considered year relatively to the reference period. It also highlights the intensity of the pluviometric deficit or excess. Pluviometric deficit corresponds to the difference between the rain of a given year and the normal over a long period (36 years in the present study). The tables and the drought index graph were established to confirm both the applied statistical tests for change-points detection and the calculated pluviometric deficit (or excess) they allowed to visualize the dry or humid period after the change-points.

The observation of the available data allowed generating some remarks. The general behaviour of the rainfall regime is centered on the "near normal" category. The frequency of years in this class ranged from 18 to 29 years, or about 51 to 82% for the entire period from 1974 to 2009, which explains the dominance of this category **Table 2**. It is interesting to mention that the stations of the arid and semi-arid zone were recorded very low frequency values but with a single rhythm. This observation indicated a fair compensation for wet and dry effects for the whole period. It is noted that the number of years of the two following classes (moderately humid, moderately dry) remains relatively balanced.

The deviation between the humid and dry years seems clearly significant and inclines to the humid influences for all the stations. The most representative example is the case of O. Chlih station. (**Table 3**). For all the stations, the  $SPI$  values were concentrated in the humid categories, but dry years were very limited.



**Table 2. Results of pluviometric index calculation for the period (1974-2009).**

STATION	< 2- Extremely dry	-1,5 to -99 Very dry	-1,0 to -1,49 Moderately dry	-0,99 - 0,99 Near to Normal	1,0-1,49 Moderat. humid	1,5-1,99 Very humid	2,0 >Extremely humid
S. NAJI		1	5	25	4		1
Babar			6	23	6		1
Tkout	1	1	5	22	5	1	1
Tfelfal			10	18	6	1	1
F. HORZA		1	4	23	5	1	2
Medina	1	1	5	23	5		1
Chelia		1	6	21	5	3	
Timgad	1	1	2	25	4	1	2
Tazoult		3	1	24	5	3	
S. Maanser		2	5	23	4	2	
SIEAR			4	26		6	
O.Chlih		2	3	21	9	1	
N'Gaous			5	23	6	1	1
Merouan			6	23	4	3	
A.Mimoun	1	2	5	21	4	3	
K. souda			3	27		5	1
Boudela		2	4	23	5	2	
Kheiran			6	24	3	1	2
Bouhmam	1	1	3	24	6	1	
Chechar			6	24	4		2
Th.Abed		3	3	23	4	2	1
Batna		3	4	22	7		
Doucen		1	1	29	1	2	2
Bouzina		1	5	26	1	1	2
Mena		2	5	22	2	2	3
Djamourah			4	25	5		2
Biskra			8	22	3	3	
Yabous		1	5	24	3	3	
A.Touta		2	5	22	4	3	
A.Baida		2	7	23	2	2	
Mziraa	1	2	3	23	5	1	1
Toufana		2	5	23	5	1	

**Table 3. Results of pluviometric index calculation for the period (1992-2009).**

Station	> to 2- Extremel y dry	-1,5to -1,99 Very dry	-1,0 to -1,49 Moderately dry	-0,99 to 0,99 Near to normal	1,0 to 1,49 Moderately humid	1,5 to 1,99 Very humide	>To 2 Extreme ly humide
S. Naji		1	5	12			
Babar		3	3	12			
Tkout		1	5	12			
Tfelfal			9	7	2		
F.Ghorza		1	4	13			
Médina	1	2	3	11			1
Chélia		2	4	10			
Timgad	1	1	1	14			1
Tazoult		3	1	13		1	
S. Manser	2		7	9			
SIEAR			4	14			
O. chlih		3	2	12	1		
N'gaous			2	8	3	3	2
Merouana		3	3	12			
A. Minoune	1		7	10			
K.Souda			3	15			
Boudela		2	4	12			
Kheiran			6	11			1
Bouhma ma	1	1	3	12	1		
Chahar			6	12			
Th.Abed		2	4	12			
Batna		3	5	10			
Doucen			1	16			
Bouzina		1	5	12			
Menaa		2	4	12			
Djamoura		1	3	14			
Biskra		3	5	9			
Yabous		3	4	10			
A.Touta		3	5	11			
A.Baida		2	6	9			
Mziraa	1	2	3	11			
Toufana		2	5	11			

More than 50% of the years are deficient in the second phase for the whole of the representative stations in the massif. Drought years have been started in the end of the year 1991 to date, (Table 3). However, there were sometimes isolated humid years whose annual pluviometry were more than the arithmetic average in the drought period. The figure below (Fig. 3) shows the annual variations of the rainfall index (SPI).

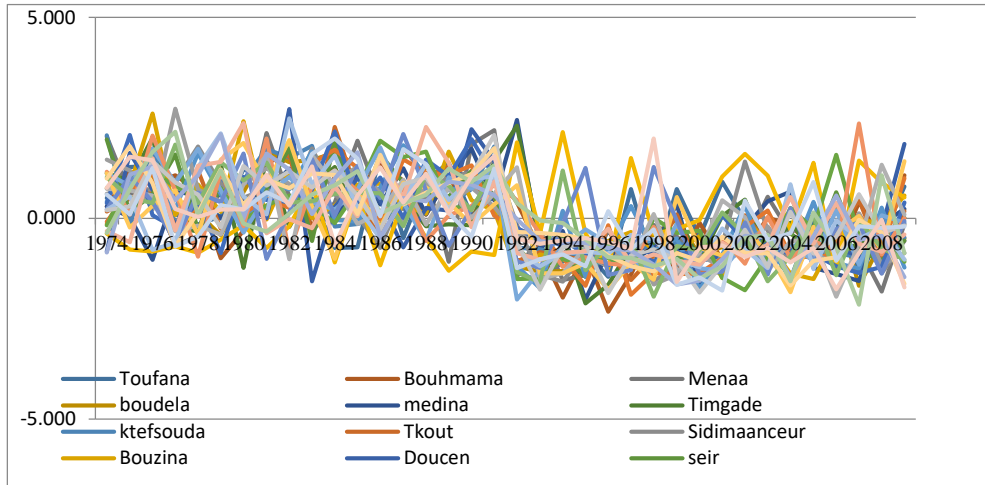


Fig. 3. Pluviometric index values (1974-2009).

It could be noticed that more than 50% of the years were deficient in the second phase for the whole of the representative stations in the massif (Table 3, Fig. 3). These deficits vary from one year to another and from one station to another according to different proportions. For some stations, the number of the deficient years may reach 70% especially at the station of Timgad. Then, he graphic representation of SPI allowed to highlight the sequence of the humid years and dry years period (Fig. 3). The SPI values were inclined to humidity in the first phase and steered for drought in the second phase.

### 4.3. Spatial distribution of annual rains

The representation of the rain’s distribution was realized starting from a geostatistical approach (simple kriging), the cartographic representation takes into account Hubert’s segmentation (annual averages before and after the change-point).

The examination of the Fig. 4, a and b, allows generating the some remarks. The stations located in the North and the North-West mark high values relatively to the Southern stations. The deficient years are gathered under the latitude 35° (Fig. 4-a). The concave shape clearly indicates the drought extension which seems very important in South northward (rise of the Saharan effects).

Wet trends in some northern stations and in the mid-range recorded a remarkable decline in spatial extension that would be occupied by the extension of southern dry effects: migration of dry effects to the center and the north (the spread of the red colour), (Fig. 4-b).

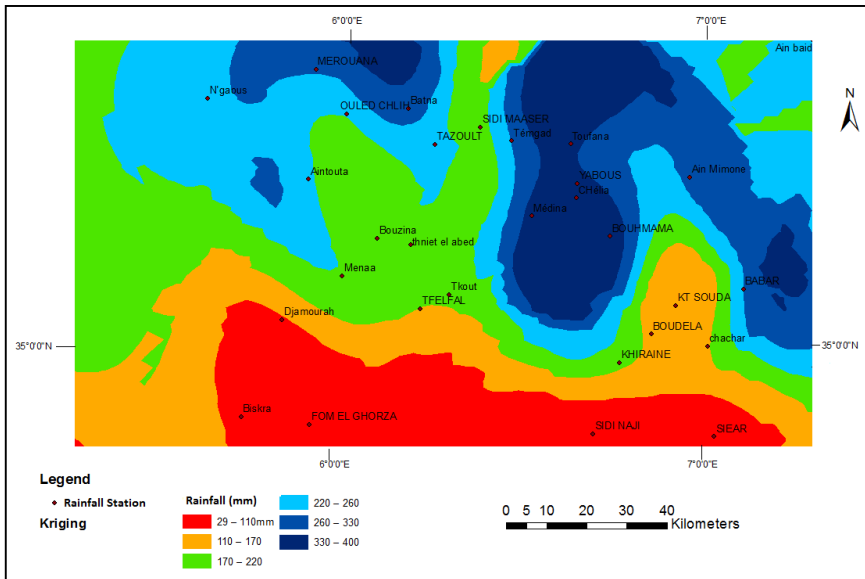


Fig.4- a. Isohyets of the 1974-1991.

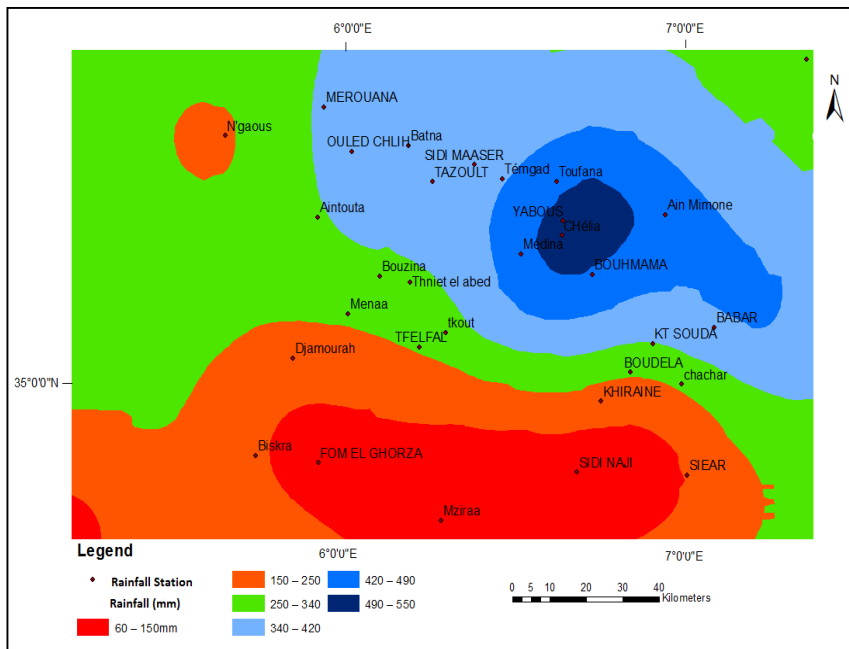


Fig. 4- b. Isohyets of the 1992-2009.

## 5. CONCLUSION

The study of the rupture allowed to locate a change in the pluviometric regime during the decade 1991-1994 for most studied pluviometric stations. The observed change-points can include two distinct phases; the first one is dominated by humid influences, the second presents dry trends. Precipitation is spread unevenly, both spatially and temporally, in the Mediterranean Sea region (Funatsu et al. 2007; Nastos et al. 2013). Its areal distribution is controlled by both small-scale and large-scale processes. At the large scale, the Mediterranean region is affected both by middle attitude cyclones and subtropical highs (Lionello et al., 2006).

The spatial distribution of precipitation has been established according to the Hubert segmentation algorithm which considers the year of rupture as a separation between two wet and dry phases. This distribution showed a significant decrease in precipitation for the second phase (1992-2009), particularly on the northern part of the massif. The rainfall regime is controlled by the geographic factors effect (Wang et al., 2016.) and the studied zone showed a high variability and accelerated drought.

Obtained results lead us to validate the hypothesis of climatic change and more specifically that a rainfall deficit is being installed at our studied zone. Studying the spatial and temporal variability of precipitation controlled by environmental factors may be helpful for an evaluation of the effect of soil and water conservation.

## ACKNOWLEDGEMENT

Special thanks are extended to Risk and Land Planning Laboratory of the University of Batna 2

## REFERENCES

- Boyer, F. (2002) Logiciel KhronoStat d'analyse statistique de séries chronologique, IRD UR2, Programme 21FRIEND AOC, Equipe Hydrologie UMRGBE, Université de Montpellier II, Ecole des Mines de Paris. <http://www.hydrosciences.org/mytech/khronostat.html>
- Buishand, T.A. (1982) Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology*, **58** (1-2), 11-27.
- Cannarozzo, M, Noto, L.V. &Viola, F. (2006) Spatial distribution of rainfall trends in Sicily (1921–2000). *Physics and Chemistry of the Earth*, **31**(18), 1201–1211.
- Funatsu, B. M., Claud, C. & Chaboureau, J.-P. (2007) Potential of advanced microwave sounding unit to identify precipitating systems and associated upper-level features in the Mediterranean region: Case studies, *J. Geophys. Res.*, **112**, D17113, doi:10.1029/2006JD008297.
- Haidu, I. (2003) *Significativité des tendances pluviométriques pluriannuelles*. Publication de l'Association Internationale de Climatologie, Université des Sciences et Technologies de Lille, **15**, 260-268.
- Haidu, I. (2004) *Modèle de variabilité climatique indépendant de l'échelle temporelle*. In. vol. Climat, mémoire du temps, Actes de XVII-ème Colloque de l'Association Internationale de Climatologie. Université de Caen Basse-Normandie, 43-46.
- Haidu, I. & Magyari-Sàska, Z. (2009) Animated Sequential Trend Signal Detection in Finite Samples. *Proceedings of the ITI 2009 31st Int. Conf. on Information Technology Interfaces* (Ed. LuzarStiffler, V; Jarec, I; Bekic, Z), June 22-25, 2009, Cavtat, Croatia.
- Hubert, P. (2000) The segmentation procedure as a tool for discrete modelling of hydrometeorological regimes, *Stoch. Environ. Res. Risk. Assess.*, **14** (4-5), 297-304.
- Jansky, L., Ives, J.D., Furuyashiki, L. & Watanabe, L. (2002) Global mountain research for sustainable development. *Global Environmental Change*, **12**, 231–239.

- Khon, V.C, Mokhov I.I., Roeckner, E. & Semenov, V.A. (2007) Regional changes of precipitation characteristics in northern Eurasia from simulations with global climate model. *Global and Planetary Change*, **57**(1-2), 118–123.
- Labat, D. (2005) Recent advances in wavelet analyses. Part I: A review of concepts. *Journal of Hydrology* **314** (1-4), 275–288.
- Lamb, P. J. (1982) Persistence of sub-Saharan drought. *Nature*, **299**, 46-47.
- Lionello, P., Malanotte-Rizzoli, P. & Boscolo, R. – Editors, (2006) Mediterranean Climate Variability. Vol. 4, Elsevier, ISBN: 9780444521705, 438 pp.
- Nastos, P., Kapsomenakis, J. & Douvis, K. (2013) Analysis of precipitation extremes based on satellite and high-resolution gridded data set over Mediterranean basin, *Atmos. Res.*, 131, 46–59, doi:10.1016/j.atmosres. 04.009.
- Pauling, A., Luterbacher, J., Casty, C. & Wanner, H. (2006) Five hundred years of gridded high-resolution precipitation reconstructions over Europe and the connection to large-scale circulation. *Climate Dynamics* **26**(4), 387–405.
- Le Quesne, C., Stahle, D.W., Cleaveland, M.K., Therrell, M.D., Aravena, J.C. & Barichivich, J. (2006) Ancient *Austrocedrus* tree-ring chronologies used to reconstruct central Chile precipitation variability from a.d. 1200 to 2000. *Journal of Climate* **19**, 5731–5744.
- Santos, J.A., Corte-Real, J., Ulbrich, U., & Palutikof, J. (2007) European winter precipitation extremes and large-scale circulation: a coupled model and its scenarios. *Theoretical and Applied Climatology*, **87**(1-4), 85–102.
- Semenov, V. & Bengtsson, L. (2002) Secular trends in daily precipitation characteristics: greenhouse gas simulation with a coupled AOGCM. *Climate Dynamics*, **19**(1-4), 123–140.
- Wang, J.F., Zhang, T.-L. & Fu, B.-J. (2016) A measure of spatial stratified heterogeneity. *Ecological Indicators*, **67**, 250–256.
- Xoplaki, E., Rouco, J.F.G., Luterbacher, J., & Wanner, H. (2004) Wet season Mediterranean precipitation variability: Influence of large-scale dynamics and trends. *Climate Dynamics* **23**(1), 63–78.

# **THE MULTIPLE DATA AND GEOGRAPHIC KNOWLEDGE APPROACH TO A LIQUID TOXIC ROAD ACCIDENT MITIGATION – TWO BLOCK GIS DATA PROCESSING FOR AN OPERATIVE INTERVENTION**

*Jaromír KOLEJKA<sup>1</sup>, Petr RAPANT<sup>2</sup>, Jana ZAPLETALOVÁ<sup>1</sup>*

DOI: 10.21163/GT\_2019.141.04

## **ABSTRACT :**

One of the current tasks of disaster management is to effectively counter toxic accidents on traffic communications. The paper demonstrates the procedure of the use of geographic data and knowledge with GIS technology for the operational mitigation of accident impacts on the traffic communication with leakage of toxic substance. A simulated leakage of toxic liquid substance on a highway in the Czech Republic was chosen as an example.

The process is divided into two units. In the first preparatory block, data on soils and the geological environment are analysed and purpose oriented pre-processed. The data layer generally describes the expected movement of pollutants, e.g. predominant surface runoff, or predominant infiltration and/or a balanced combination of both of them. In the second operational unit, a location of the accident is precisely identified and the estimation of possible routes of pollutant runoff is performed with respect to the current status of the territory. Key points on these routes are identified with the aim to select mitigation measures and optimum access routes modelled for intervention techniques to reach key points in order to prevent contamination of water bodies.

*Keywords: Data pre-processing, Operational data processing, GIS, Routes of pollutant run-off, critical points, best access routes, risk management.*

## **1. INTRODUCTION**

The leakage of toxic or other harmful substances occurs on roads and rails very often. On roads rather more numerous leakages can be expected, due to the nature of road freight transport (the capacity of each lorry) being predominant in smaller volumes than on railways. On the other hand, the roads usually go much closer to places where there is a permanent or occasional concentration of population potentially threatened by accidents related to the leakage of harmful substances.

Risk categorisation of roads and highways of the Czech Republic using statistically registered accidents was carried out under the European programme of assessing the safety of roads EuroRAP (European Road Assessment Programme) (EuroRAP is an international non-profit organisation founded in 2003 in Belgium, whose members are motoring associations, national and regional administrators of roads and investors, invited experts and commercial organisations). According to these statistics, motorways and expressways seem to be the safest roads. Media information, however, gives the impression that the most

---

<sup>1</sup> *Institute of Geonics, Czech Academy of Science, Ostrava, Department of Environmental geography, Drobného 28, CZ-60200 Brno, Czech Republic, jkolejka@centrum.cz, jana.zapletalova@ugn.cas.cz*

<sup>2</sup> *VSB-Technical University of Ostrava, Faculty of Mining and Geology, Department of Geoinformatics, 17. Listopadu 2172/15, CZ-70800 Ostrava-Poruba, Czech Republic, petr.rapant@vsb.cz*

serious accidents occur mainly on motorways, where these events are greatly mediated. Strictly speaking, it is obvious that such accidents can happen anywhere where hazardous substances are transported by road or by railway. So, such events are of greater amounts there and the number of potential sites of their creation is endless. Research teams are trying (with the assistance of modern information technology) to construct standardised procedures to streamline remediation post emergency interventions in the field. Yet, of course, it is impossible to prepare an exhaustive number of accident scenarios that, despite using scientific knowledge and available geospatial technologies, would cover all possible locations. The standardised scenario must therefore be formulated in such a way that it could be started when the disaster management staff receives the information about such an incident. This paper aims to demonstrate the procedure for completing such a scenario using GIS technology on the example of a simulated accident on a highway, and to demonstrate its applicability under specific long-term and short-term territory conditions.

## **2. THEORETICAL AND METHODOLOGICAL BACKGROUND**

The ADR Decree (European Agreement Concerning the International Carriage of Dangerous Goods by Road) sets out the general conditions for the transportation of dangerous substances on the roads. Its revised version came into force in the EU member states in January 1, 2015. The methodology and proposed procedures dealing with the risks of possible consequences of accidents in the transport of dangerous substances by road are based on this directive. Generally, transportation of hazardous and toxic substances is regulated by international legislation, which is based on EU Regulations (RID, ADR, ICAO, ADN, IMDG Code). General principles for transporting dangerous materials are common to all transport modes of such substances. Their dangerousness is given by their physical-chemical properties, toxicity or ecotoxicity. A number of different methods and studies dealing with the transportation of hazardous substances show that the transport of hazardous substances is a very relevant topic for today. As shown by the statistics, 39% of all toxic accidents occur during transportation and another 8% during loading and unloading (Bernartík, 2006). For example, in Italy, the calculation of the risk of accidents during transport of these substances was done by Fabiano, Currò, Palazzi, & Pastorino (2002).

Light fuel oils (LPG), gasoline, diesel, chlorine, ammonia and other industrial gases are transported in the Czech Republic on roads. Among accidents of vehicles carrying toxic substances generally prevail the accidents of liquid substances (Chudová & Blažková, 2007). The calculation of the risk of accidents when carrying dangerous substances was done in many countries, including the Czech Republic (Krejčí & Bambušek, 2012). The authors use data from the National traffic census and statistical data. They state that 1,669 million tkm of dangerous goods were transported on roads in the Czech Republic in 2010. 101 traffic accidents of vehicles carrying hazardous materials were reported for the same year (Krejčí & Bambušek, 2012).

The key work dealing with the research into toxic accidents on roads is represented by the Guidelines for quantitative risk assessment (Purple Book) (De Haag & Ale, 2005) which, in its second part, deals with risk evaluation of hazardous substances transportation. A quantitative risk assessment approach of the transport of dangerous substances is also referred to in a complex view by Nicolet-Monnier & Gheorghe (1996). They formulate their ideas applying existing knowledge to meet the needs of risk management. Høj & Kröger (2002) conducted a risk analysis of transport (in general, not just hazardous



substances) on roads and rails. They noted that the risk assessment was mostly devoted to special sites on the transport network, such as tunnels or bridges. The issue of risk in tunnels is also addressed by Diamantidis, Zuccarelli & Westhäuser (2000). They sorted out the obtained and reviewed documents from specific events in the form of guidelines for decision making processes and they also elaborated thematic terminology suitable for communication among experts of risk management. Some works indicate the need to standardise procedures for disaster management (Fabiano et al, 2002).

Goerlandt & Montewka (2015) note that the number of applications designed for risk management does not take into account the theoretical issues related to crisis management (e.g., set of definitions and terminology or perspectives of risk). This can cause a number of problems. The behaviour of people in disastrous situations has an individual character, but it demonstrates many common features (Burns & Slovic 2012; Fujiki & Renard, 2018). The vulnerability of densely settled urban areas to technological risks is especially very complex (Ștefănescu, Botezan & Crăciun, 2018). The behaviour of certain persons can even be predicted in specific situations (De Dominics et al, 2015). In view of this, it is necessary to create and search for appropriate forms of warning in the event of a crisis situation, in order to prevent additional damage. From a practical point of view (with regard to the unpredictable part of human behaviour), the construction of mitigation scenarios for road toxic accidents must be formulated so as to minimise the extent of the threat at the outset through the use of equipment and professional rescue teams.

Considerable attention is devoted to risk assessment of accidents with leakage of hazardous substances and the use of GIS tools in the professional community. Using GIS tools, it is possible to describe, visualise and model the past and ongoing emergencies, including accidents on the roads (Zhang, Hodgson & Erkut, 2000). There is a wide spectrum of GIS studies dealing with the issue of toxic accidents on roads ranging from the optimisation of transport routes of hazardous substances in order to minimise the risk of accidents and minimisation of negative impacts on the environment, through the studies modelling course of events, and designing effective intervention after the mitigation and elimination of the impact of the accident. Kolejka (2010) looked into the localisation of key areas for intervention units in case of leakage of liquid toxic substances in urban environments. Bubbico, Di Cave & Mazzarotta (2004) used data on the population, the local environment and the current meteorological situation for evaluating the risk of potential toxic accidents and the need for effective intervention. A GIS-based study dealing with the risk of hazardous material transportation on roads was completed by Huang (2004) as an example of the type of frequented use of GIS technologies in the risk assessment of transport on roads. Questions of the organisation of evacuations of endangered persons by means of decision-making support systems (DSS) based on GIS were dealt with by de Silva & Eglese (2000). The studies dealing with intentional toxic accidents on roads as terrorist acts are also included (Maschio et al, 2009). A part of the evaluation of the risk of transport of hazardous substances may be represented by visualisation of such risks. Van Raemdonck et al, (2013) highlight this on the example of Flanders. Their work focuses on two components: a) assessment of the probability of an accident, and b) the evaluation of consequences of an accident (if it occurs). A similar method was also used by Verma & Verter (2007). Risk prevention and effective elimination of the consequences of accidents can be crucial especially in the transport of hazardous (toxic) substances. Cadar, Boitor & Dumitrescu (2017) confirmed that an increase in the Annual Average Daily Traffic

(AADT) generates an increase in the total number of road accidents. Ivan & Haidu (2012) documented the time concentrations of road accidents in rush hours.

Many programmes for risk modelling associated with toxic accidents were developed. These include, e.g., the programmes PHAST, RMP COMP, ALOHA, CHARM, EFFECTS, ROZEX and others. The CAMEO programme and the subsequent mapping software MARPLOT (Mapping Application for Response, Planning and Local Operational Tasks) was developed for dispersion modelling of hazardous substances. It was developed by the United States Environmental Protection Agency (EPA) and the National Oceanic and Atmospheric Administration (NOAA). In the Czech Republic, the software ROZEX alarm has been developed (<http://www.tlp-emergency.com/rozex.html>). It is an application enabling model leakage of dangerous chemical substances, quickly generating the necessary information for the intervening bodies of the Integrated Rescue System. Besides the characteristics of hazardous chemicals, the programme works also with GIS systems and enables dangerous zones in the map layer to be depicted.

The frequency of transport of dangerous goods is permanently increasing. Its growth is mainly due to the globalisation of industrial production. Especially the transfer of freight from rail to roads is very strongly manifested. It is generally estimated that about 20% of vehicles transport hazardous or toxic substances. Such an accident, depending on the nature of the accident site and the transported hazardous or toxic substances, can directly threaten the inhabitants in settlements, as well as waterways, soils and underground waters. The consequences of accidents of trucks transporting fuel are especially dangerous. Usually the point is that several tens of tons of fuel may escape into the surrounding environment. The worst accidents are in the settlements, where the spilled fuel may leak into the sewage system. Its evaporation causes high concentrations of flammable gases that spread in sewers, uncontrollably explode and cause subsequent fires (e.g. Sydney 1987 - see Tuma, 2000).

The essential condition for quick implementation of efficient scenarios for optimising intervention after the occurrence of the event is represented by the access to publicly available data, GIS technology, quality and reliable functioning communication between the disaster management staff and field intervention unit. The below presented procedure is a concise and revised version of the detailed user manual published in book form in 2015 (Kolejka, Rapant et al, 2015). The preliminary research results were demonstrated to the scientific community at a conference in Brno, Czech Republic in 2015 (Kolejka, Rapant, Zapletalová, 2016).

### **3. THE OPTIMISING SCENARIO FOR FIELD INTERVENTION AFTER A TOXIC DISASTER ON THE ROAD**

#### **3.1 Data sources**

The accident on a road connected with leakage of harmful (toxic) substances (in this case liquid) should be viewed from two standpoints:

1. The event occurred in a particular area, whose features have an effect both on the spread of liquid pollutants into the area and on the accessibility of suitable intervention sites available for the technique.
2. The event occurred at a specific moment of time, which corresponds to the time-varying features of the territory.

Taking these facts into account, it is necessary to divide the mitigation procedure into two working units:

A - Preparatory block, during which time conventionally constant data about the territory will be collected, evaluated and processed for future operational application. The purpose of the process is to create data layers representing potential conditions for the movement of liquid pollutants in the area in compliance with realistic, time-varying area features at the moment of accident.

B - Operational block, whose activity is started after the emergency staff receive information about the disaster. Now, the prearranged data from the preparatory block are used in the context of current conditions of the area, both positional and instantaneous. Data sources concerning the stable features of the territory and having an influence on the behaviour of liquid pollutants are listed in **Table 1**.

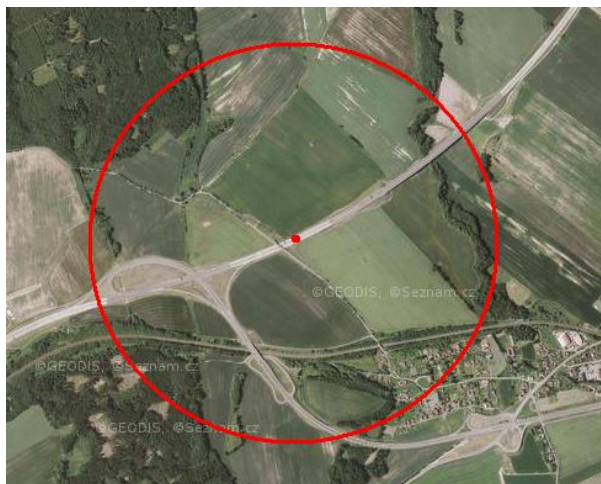
**Table 1**

**Data sources for use in preparatory and operational blocks for supporting crisis management in response to a toxic disaster on the road**

No.	branch of knowledge	title of geodata	administrator	data contents and method of application
Preparatory block				
1	geology	Geological map of ČR 50	CGS (Czech Geological Survey)	The digital map is available at a resolution corresponding to the scale of 1:50 000, even though it was apparently constructed (before generalisation) with more detailed documentation in the scale of 1:25 000 and locally at 1:10 000. Sometimes it is necessary to correct mapped areas according to the valley network and contour lines (Source: <a href="http://www.geology.cz">http://www.geology.cz</a> )
2	soil science	Soil water retention capacity and hydrological soil groups CR 50	Research Institute for Soil and Water Conservation (VÚMOP)	Digital soil maps differentiate areas of classes of soil water retention capacity and filtration coefficient. Seamless maps are available at a resolution corresponding to the scale of 1:50 000 though they were apparently constructed (before generalisation) with more detailed documentation in a scale of 1:5 000. Woodlands are omitted. (source: <a href="http://www.vumop.cz/">http://www.vumop.cz/</a> )
Operational block				
3	Inland transport roads	Road map	Directorate of Roads and Motorways (RSD)	Digital Road map presents roads in layers of different road classes.
4	topography	ZABAGED	State Administration of Land Surveying and Cadastre (ČÚZK)	Geodatabase ZABAGED in both in raster and vector formats represents a basic topographic maps of the Czech Republic in the scale 1:10 000

No.	branch of knowledge	title of geodata	administrator	data contents and method of application
Operational block				
5	orientation	coloured orthophotomap	Cenia	Continuously updated colour aerial orthophotomap shows the current situation of the territory. Data can be retrieved from the map server. It represents a resolution of about 1 m.
6	geomorphology	Digital model of relief of 4th generation	ČÚZK	The digital terrain model was completed using lidar technology for the total territory of the Czech Republic. Its vertical resolution is a few dms, horizontal less than 5 m. Source: . <a href="http://www.cuzk.cz/">http://www.cuzk.cz/</a>
7	land use	ZABAGED – forests, meadows, built-up areas, communications and other	ČÚZK	Individual classes of land use are registered in the geodatabase ZABAGED in separated data layers. Layers of forest, meadows, arable land and built-up areas are relevant for scenarios. The data resolution corresponds to the map scale 1:10 000 Source <a href="http://www.cuzk.cz/">http://www.cuzk.cz/</a>
8	hydrology	the water management map 50	T. G. Masaryk Water Research Institute (VÚV)	The map shows the river network, other water objects and watersheds in the resolution of 1:50 000. Source: <a href="http://www.vuv.cz">http://www.vuv.cz</a>
9	Hydrometeorology	Saturation indicator	Czech Hydrometeorological Institute (CHMU)	The saturation indicator represents an estimation of the current saturation of the territory with water, usually at 8:00 a.m. local time. It can also be generated during the day to different hours repeatedly. It is derived using a simple model of the balance of rainfall, runoff and evapotranspiration. Its main task is the detection of potential risk of flash floods development and occurrence. Source: ČHMÚ- <a href="http://hydro.chmi.cz/hpps/main_rain.php?mt=ffg">http://hydro.chmi.cz/hpps/main_rain.php?mt=ffg</a>
10	Administrative division	CS0_US_communes	Czech Statistical Office (CSO)	Digital map layer presents: administrative limits of communities on the territory of the Czech Republic. Source: database of Fire Rescue Service of the Czech Republic (HZS ČR)

Both procedure units are demonstrated on the example of a simulated accident on highway D1 (**Fig. 1**).



**Fig. 1.** Visualisation of the segment of the seamless colour orthophotomap surrounding the simulated accident on the D1 motorway at the 312<sup>th</sup> kilometre from Prague. It shows the location of the event in the centre from a circle with a radius of 1000 m

(Source: <http://mapy.cz/letecka?x=17.7855914&y=49.5958452&z=16>)

### 3.2 Data processing

A – the preparatory block of geo-data processing is focused on the classification of territorial units throughout the Czech Republic from the viewpoint of the expected behaviour of a liquid pollutant (with undifferentiated viscosity) and on the evaluation of the territory penetration level for the intervention technique. In terms of the movement of pollutants in the terrain and the possibility of stopping this movement, it is important whether the liquid substance will predominantly soak into the soil and geological environments, or whether it will dominantly flow down the surface or process infiltration into the ground and the downslope surface run-off will be more or less equally significant.

**The tendency of soil cover to support liquids soaking from the surface, or conversely its resistance to liquid penetration** may generally be derived from soil properties recorded on the map of soil water retention capacity of the Czech Republic at a scale of 1:50 000 and on a map of hydrological soil groups in the Czech Republic 1:50 000. The data are distributed by the data administrator – namely the Research Institute for Soil and Water Conservation (VÚMOP) - to customers in the ESRI shape file format at a resolution corresponding to a scale of 1:50 000, mainly by the layout of topographic map sheets of basic topographic maps of the Czech Republic at a scale of 1:50 000 (not required). The map of water retention capacity highlights the tiered possibility of soaking, or rather of runoff liquids according to the soil water regime (and, of course, the nature of the leaking substance). The interpreted map of water retention capacity of soils (according to **Table 2**) represents an intermediate stage for further data processing.

**The evaluation of hydrological soil groups in terms of the risks of support, or seepage of liquid pollutant**, can be done in analogy with the previous case, through expert assessment. It is convenient to implement it off-line for the entire territory of the Czech Republic already in the preparatory stage of the procedure for a possible case of similar incidents elsewhere in the country.

Table 2

**Targeted soil classification of the water retention capacity according to soil influence on the runoff of liquid pollutant (R - dominated by surface runoff, I - prevails infiltration, N - balanced ratio of infiltration and surface runoff)**

Designation of soil environment	Group soils	Water retention capacity	Water capacity (l/m <sup>3</sup> )	Goal-directed characteristics
R	Group 1	low	100–160	strongly supports surface runoff and causes poor terrain penetrability for intervention technique
R	Group 2	lower middle	100–160	supports surface runoff, which is not sufficient for good penetrability for intervention technique
N	Group 3	middle	100–220	supports surface runoff quite a lot, which means conditional patency
I	Group 4	upper middle	220–320	supports the infiltration of liquid, which is sufficient for good patency
I	Group 5	high	above 320	strongly supports the infiltration of liquid, which allows a good terrain penetrability for intervention technique

However, the derived map itself tentatively indicates the behaviour of pollutants with regard to the long term soil moisture content and the soil's ability to absorb or 'repel' flowing harmful substances. In other words, 'the final water retention capacity reflects the average depth of soil profile and water content. They characterise the actual amount of water which the soil is capable of retaining'. The completed evaluation of soil water retention capacity in terms of risk of support, or prevention soaking of liquid pollutant, is a matter of expert assessment and can be done off-line.

**The Map of hydrological soil groups based on the size of the infiltration coefficient of soil** is based on grain composition of soils and their mechanical effect on the possibility of infiltration, or rather of liquid runoff. The content of the map 'Hydrological soil groups' (according to VÚMOP: Map and data products - hydrological characteristics) is specifically interpreted in the form of a map describing the anticipated character of the liquid pollutant runoff according to the hydrological soil groups. Conversion of the original map content is purposely oriented into three groups (**Table 3**) and the relevant maps in the GIS.

**The integration of partial evaluations of the soil and geological environments forming a general overview of the expected behaviour of liquid pollutants in the potentially affected area** runs through an overlay of partially interpreted maps in the raster format and completes combinations of letters. If in the three-digit combination is mostly 'I' - the result should be marked 'I' (the predominant process is the infiltration of liquid), if 'R' prevails – the surface runoff of liquid prevails) if 'N' dominates - it results finally in 'N' (neutral behaviour of liquids with a tendency of surface runoff), if all three letters are represented in the three-digit combination - 'N' is the result.

**Table 3**

**Targeted soil classification of the stipulated hydrological groups according to their influence on the liquid pollutant runoff (the table of abbreviations - see Table 2).**

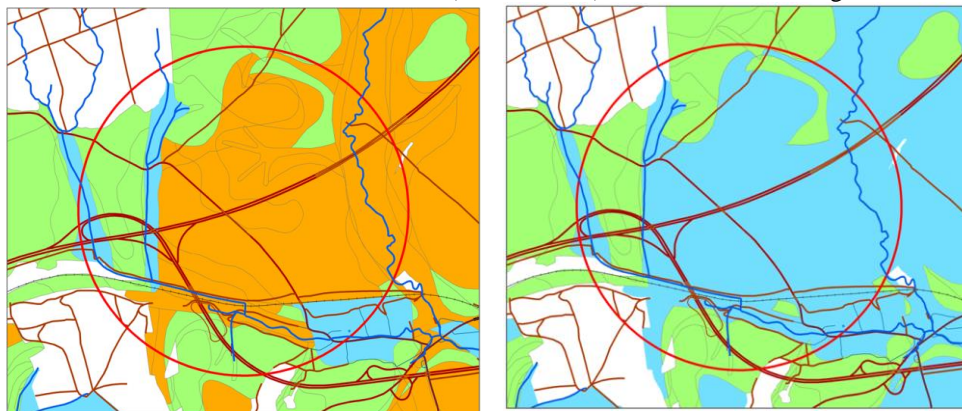
Designation of soil environment	Group soils	Infiltration level	Speed of infiltration (mm/min)	Specific characteristics
I	Group A	high	> 0.20	it supports very well the infiltration of rainwater / pollutant and causes a good penetrability of the terrain by the intervention technique
I	Group B	middle	0.10 – 0.20	It supports well the infiltration of rainwater / pollutant and also causes a good penetrability of terrain by the intervention technique.
N	Group C	low	0.05 – 0.10	It supports the surface runoff of water / pollutant quite a lot, which is sufficient for conditioned terrain penetrability
R	Group D	Very low	< 0.05	it represents an almost impervious environment strongly supporting surface runoff and causes poor penetrability for intervention techniques

**Table 4**

**Targeted classification of rocks according to their influence on runoff of liquid pollutants (the test area around Bělořín example).**

Label of geological environment	Purpose characteristic	Classification of rocks (With numbers used in the picture/table of the geological map)
I	Rocks and soils insufficiently supporting surface runoff and strongly supports soaking	deluvial sandy-loamy and clayic-loamy sediments (6) deluvial loamy-stony sediments (7) glacifluvial sandy gravels (10), fluvial sandy gravel terraces (13), loamy stony eluvium (15), Tertiary sands and sandy gravels (16), calcareous sands (18) alluvial fans (42)
N	Rock and soil support in average surface runoff as well as soaking	loess loam (8), picrites basalts, tuffs (29), conglomerates, sandstones, mudstones (30), shales, siltstones, offals (33), offal (34), assorted conglomerates (36) limestones (41) , landslides (43)
R	Rocks and soils intensely supporting surface runoff and not much soaking	peat (2), fluvial sandy loam sediments (4) deluvial sandy loam sediments (5) glacialustrineclays (11), clay-loamy eluvium (14), Tertiary calcareous clays (17), sandstones and claystones of Ždánice-Hustopeče formation (19), claystones and silicites of menilite formation (20), claystones of sub-menilite formation (21), sandstones and conglomerates of Stráže type (22), gray calcareous claystones and sandstones of Frydek strata (23), calcareous claystones of Dub formation (24) , claystone and sandstones of Neměčice formation (26), calcareous claystones, sandstones and conglomerates of Těšín-Hradiště formation(28).

The integrated output constitutes a derived character of the resulting pollutant behaviour. The procedure is performed in two alternatives: for a dry territory situation (**Fig. 2 - left**) and for the wet territory condition (**Fig. 2 - right**), depending on the nature of the topic water saturation of the area. This information is provided in cartographic form at a resolution of 1 km<sup>2</sup> by CHMI on its website for the entire territory of the Czech Republic. In the case of water saturation of the area (wet situation), the value 'N' is changed to 'R'.

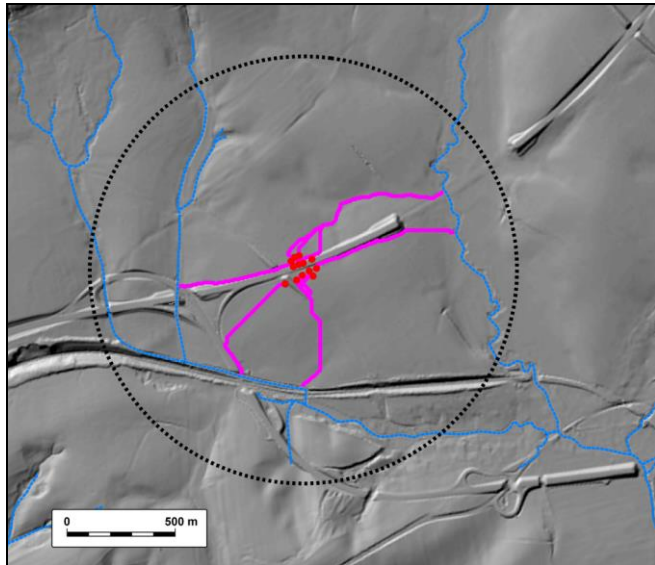


**Fig. 2.** Visualisation of areas with different behaviour of liquid pollutant under dry - left and under wet - right territory conditions (light blue – surface run-off; light green – soaking; orange – surface run-off; white - forests), the circle radius is 1000 m (Source: CGS, VÚMOP, ČÚZK).

The character of the movement of liquid pollutants in the environment is also affected by the slope gradient and surface roughness. The surface roughness varies during the seasons of the year and depends on the state of the vegetation cover. The agricultural plots show the most significant changes; it is not possible to ensure reliable information about the state of the surface in advance for any place in the state. The situation is different in the case of the slope. A currently available 4<sup>th</sup> generation digital terrain model constructed on the base of aerial laser scanning has a horizontal resolution of app. 5 m for the entire territory of the Czech Republic. Its application is expected in the operating block of the procedure.

Operational block B is started by adopting the report of the accident stating its position. In the narrow space around the accident site in a digital terrain model, an estimated number of points are entered. Using the hydrological modelling in GIS, the immediately possible routes of runoff toward the nearest surface water receptors are derived (**Fig. 3**). The identification of several possible routes may cause scattering of attention disaster staff, which sends an intervention unit (or better more units) to the site of the incident. Meanwhile, the incident commander already ascertains on site which of the modelled pollutant runoff routes is true. This will be then the focus of the intervention unit with the aim to eliminate the impacts of the accident. Emergency units are primarily tasked with stopping the spread of liquid toxic pollutants. They cannot do it in any place, but only in places accessible to the technique. The motion of the technique is limited not only by the natural parameters of the terrain (high slope, watercourses, long-soaked sites or currently waterlogged places with limited infiltration of liquids) but also by anthropogenic objects and by the different forms of land use.



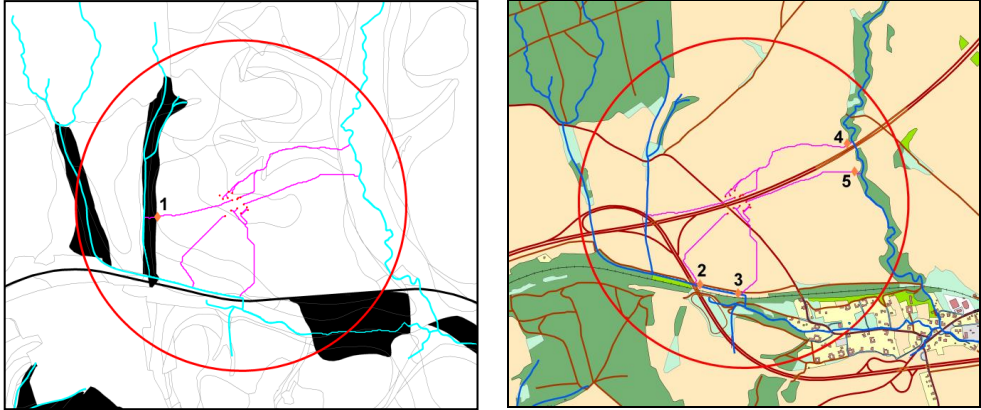


**Fig. 3.** Example of an operative search of possible outflow routes of liquid pollutants (in purple) downslope from the accident site (red dots located around the accident site) within a circle with a radius of 1000 m around the crash site showing watercourses (in blue) as well (Source: ČÚZK).

The main task is to prevent direct contamination of watercourses by moving liquid pollutants. "Key points" must be identified (localities where the route of pollutant changes one type of environment to another - e.g. sites with domination infiltration or surface runoff).

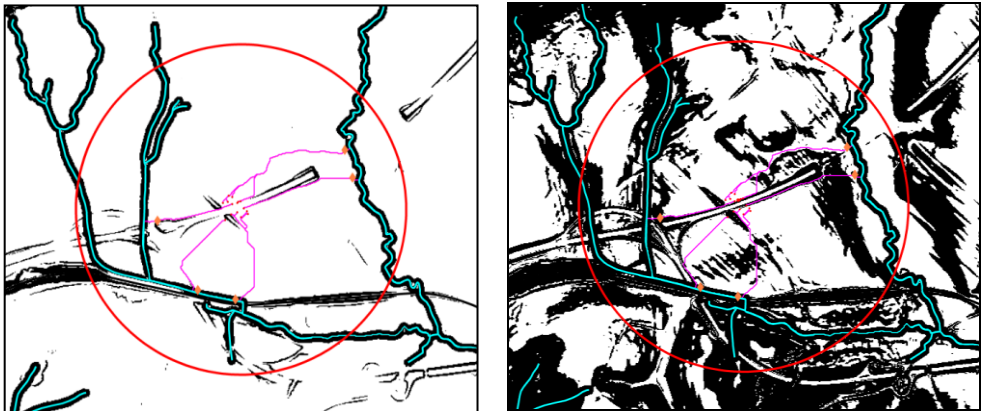
Depending on the environment in which the movement of pollutants takes place, it is necessary to select a form of intervention (e.g. damming of routes in areas where the surface runoff dominates and not allowing its entry into the environment with a predominance of soaking, which would result in the contamination of ground waters or the draining fluid if it has already penetrated to the infiltration area). If the water courses are surrounded by forest, which is impervious to the technique, it is necessary to prevent the entry of liquid pollutants into the forest. Key points are thus determined either by the data analysis in GIS such as the intersections of potential routes of the pollutant movement together with the boundaries of the different types of the environments (depending on the nature of the soil and geological structure) and the forest limit. It is much easier to create a new point shape file within an on-screen operation of the eligible disaster staff personnel. The intervention of the field emergency unit must then be directed over these points into a higher altitude. In the case of this simulated accident 5 key points to 5 routes of the anticipated movement of liquid pollutant were identified. Point no. 1 is at the place where section with a predominant infiltration passes to a section of a dominant surface runoff (**Fig. 4 - left**). The points no. 2, 3, 4 and 5 are on the edge of the forest, through which a watercourse passes affected by contamination (**Fig. 4 - right**).

It is necessary to find out the applicable access routes for the intervention technique to these five key points operationally. It is also necessary to take into account all the supposed barriers to technique movement and to use safe routes of access. The slope gradient plays an important role.



**Fig. 4.** The distribution of key points on the routes of anticipated flow of liquid pollutant towards the water receptors (left - a key point no. 1 on the physically different types of the natural environment for the movement of pollutants, right - key points 2, 3 on the direct contact with the potential water receptor, 4 and 5 on the edge of the forest area with the potential water receptor inside).

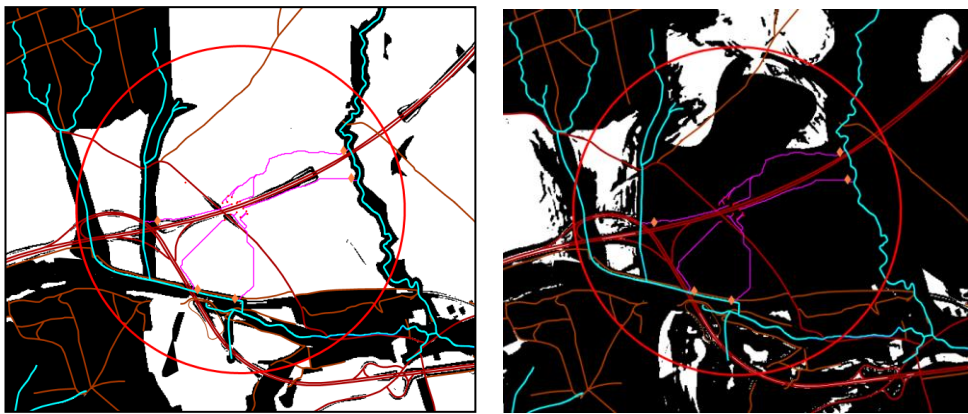
There is no uniform standard for reliable passage of terrain by different types of intervention techniques and success usually depends on the skills of the driver and the current state of the technique. At least in general, the areas in which intervention technique should be avoided can be designated. These can be, in the case of a slope, represented by a gradient above  $15^\circ$  under dry area conditions (**Fig. 5 - left**) and  $7^\circ$  under wet conditions (**Fig. 5 - right**). Areas of such critical gradients can be derived from the digital terrain model.



**Fig. 5.** Slope gradient as a barrier for the access of intervention techniques (left - inclination over  $15^\circ$  under dry conditions; right - slope inclination above  $7^\circ$  under wet conditions).

Another obstacle to the movement of intervention techniques arises from the nature of the soil and the geological environment. This environment becomes an obstacle when rainwater infiltration is not allowed (because of long-term or momentary saturation with water). Therefore the areas where different features affecting the movement of liquid

pollutants were identified in the preliminary block can be applied. The areas of the dominant surface runoff of the liquid pollutant in dry and wet conditions represent the barrier. Another type of barrier is rivers and some forms of land use (buildings, walls, fenced gardens, railway, forest, wetlands, etc.) depending on the particular area (see e.g. Hasnat, Islam & Hadiuzzaman, 2018). Through the summation of data layers of individual barriers it is possible to gain an integral layer of barriers to the access of the emergency units. Such a layer may be prepared already in the initial preparatory block for the entire country. Due to the demanding character of the countrywide processing of huge data files, and the relative simplicity of available data interpretation - both pre-elaborated in a preparatory block (character movement of liquid pollutants) and taken from available public sources (DTM, river network, land use), this data can be processed only for the area around the site of the accident after reporting its position. It is important to know the course of existing roads passable for the intervention technique in the terrain (**Fig. 6**).



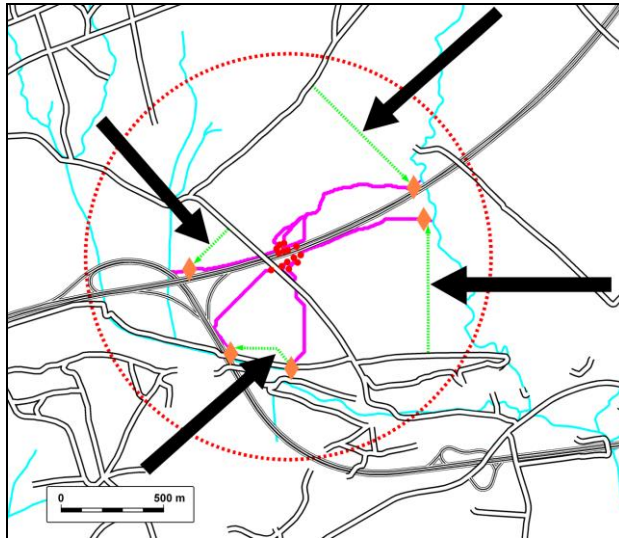
**Fig. 6.** Territorial distribution of all identified types of barriers for assessment of intervention teams in the territory towards the key points on the routes of liquid pollutant flow (left – under dry conditions, right - under wet condition) with presentation of the existing road network.

Searching for the optimal route access by the intervention technique to key points is done by standard tools in ArcGIS technology (version 10.2, Cost Distance and the Cost Path in Spatial Analyst). The input data layers to this procedure are:

1. The layer of barriers to the intervention technique access (with alternatives for dry or wet) conditions converted into binary form (masks),
2. The layer of the road network (on which the technique can come as close as possible to the critical points),
3. The layer of critical points which are the localities of the last chance for adequate intervention for prevention of the liquid pollutant spread (**Fig. 7**).

It is evident from **Fig. 6-right** that under wet conditions (with long-standing and simultaneously instantaneous water saturation of soil and geological environment after the previous precipitation period) the accessibility for the intervention technique is at high risk, and the emergency unit has to be aware of this. The problem can then be solved by deploying special techniques (belt, air), or pedestrian access for the intervention team. For an alternative under dry conditions, finding optimal routes for the intervention technique is possible using ArcGIS tools. The incident commander at the site then directs the operating

units on the routes along which is the liquid pollutant moves (**Fig. 7**). According to the cadastral maps and the actual flow of liquid pollutant, the incident commander determines to which municipalities a warning against the threat of direct contamination by the flowing harmful substances, or indirectly by contaminated water in the water recipient (if it fails to stop the dripping) will be sent.



**Fig. 7.** Optimal access routes from existing road network (in green) avoiding all types of barriers to key points on lines of expected flow of liquid pollutant for intervention teams searched operationally using ArcGIS tools.

#### 4. CONCLUSIONS

The demonstrated procedure of selected steps of a disaster team is based solely on publicly available geodata in the Czech Republic, their interpretation and processing. The division is into two working units: the preparatory block and the operational block, and it is necessary to be able to timely and effectively counter the threat of contamination of waterways by toxic liquid that leaked into the environment after an accident on the road, at any point on the road network. This data should be permanently available for the disaster staff - its workplace GIS (both national and regional), or unrestricted access should be organised. According to current Czech legislation it is possible. The preparatory block is the most time-consuming - it deals with the qualified estimation of the movement character of liquid pollutants in the area and also the penetrability of the terrain for intervention techniques with regard to the soil and geological environment for two alternatives: under dry and wet conditions. Such data layer should be compiled and available for the entire country. Due to the nature of the data, however, it is applicable only in case of an accident outside of large forest units (necessary soil data are missing here).

The operational block uses pre-processed data from the preparatory block and operatively applies them to the nearby site of the accident with regard to the current state of saturation of the territory with water, which affects both the character of the movement of the liquid pollutant and the area penetrability for the intervention technology. Integrated

data processing in order to optimise the intervention (choice of methods of intervention and their location) takes a matter of a few minutes. The intervention unit will be sent to the pre-selected sites and then operatively directed after being found on the sitethrough which the toxic liquid moves. Like this, the units can choose the method of action (draining, damming drainage), and the place of intervention with regard to the type and accessibility for the technique.

## ACKNOWLEDGEMENT

This work was supported by the project “Disaster management support scenarios using geoinformation technologies” [No VG20132015106], programme Safety Research promoted by Ministry of Interior, Czech Republic; by VSB-Technical University of Ostrava, Faculty of Mining and Geology, Department of Geoinformatics, 17. Listopadu 2172/15, CZ-70800 Ostrava-Poruba, Czech Republic.

## REFERENCES

- Bernatík, A., 2006. Prevention of major accidents II. Ostrava, Association of Fire and Security Engineering with headquarters at VŠB (Prevence závažných havárií II. Ostrava, Sdružení požárního a bezpečnostního inženýrství se sídlem VŠB) – Technická univerzita Ostrava.
- Bubbico, R., Di Cave, S. & Mazzarotta, B., 2004. Risk analysis for road and rail transport of hazardous materials: a GIS approach. *Journal of Loss Prevention in the Process Industries*, 17 (6), 483–488.
- Burns, W. J. and Slovic, P., 2012. Risk perception and behaviours: anticipating and responding to crises. *Risk Analysis*, 32 (3), 579–82.
- Cadar, R. D., Boitor, M. R. & Dumitrescu, M., 2017. Effects of traffic volumes on accidents: The case of Romania’s national roads. *Geographia Technica*, 12 (2), 20-29.
- De Dominics, S., *et al.*, 2015. We are at risk, and so what? Place attachment, environmental risk perceptions and preventive coping behaviours. *Journal of Environmental Psychology*, 43 (1), 66–78.
- De Haag, P. U. & Ale, B.J.M., 2005. Guidelines for quantitative risk assessment (Purple book). Publication Series on Dangerous Substances (PGS 3). Guidelines for quantitative risk assessment. Ministerie van Binnenlandse Zaken en Koninkrijkrelaties/Ministerie van Verkeer en Waterstaat, The Hague.
- De Silva, F.N. & Eglese, R W., 2000. Integrating simulation modelling and GIS: spatial decision support systems for evacuation planning. *Journal of the Operational Research Society*, 51 (4), 423–430.
- Diamantidis, D., Zuccarelli, F. & Westhäuser, A., 2000. Safety of long railway tunnels. *Reliability Engineering & System Safety*, 67 (2), 135–145.
- Fabiano, B., *et al.*, 2002. A framework for risk assessment and decision making strategies in dangerous food transportation. *Journal of Hazardous Materials*, 93 (1), 1–15.
- Fujiki, K. & Renard, F., 2018. A geographic analysis of post-disaster social impacts on a municipal scale – A case study of a potential major flood in the Paris region (France). *Geographia Technica*, 13, (2), 31-51.
- Goerlandt, F. & Montewka, J., 2015. Maritime transportation risk analysis: Review and analysis in light of some foundational issues. *Reliability Engineering & System Safety*, 138 (6), 115–134.
- Hasnat, M. M., Islam, M. R. & Hadiuzzaman, M., 2018. Emergency response during disastrous situation in densely populated urban areas: A GIS based approach. *Geographia Technica*, 13 (2), 74-88.

- Høj, N. P. & Kröger, W., 2002. Risk analyses of transportation on road and railway from a European Perspective. *Safety Science*, 40 (1–4), 337–357.
- Huang, B., 2004. A GIS based route planning for hazardous material transportation. *Journal of Environmental Informatics*, 8 (1), 49–57.
- Chudová, D. & Blažková, K., 2007. Transport of hazardous substances from the point of view of emergency planning of the area Ostrava. (Přeprava nebezpečných látek z pohledu havarijního plánování území. Ostrava, Sborník vědeckých prací Vysoké školy báňské) – Technické univerzity Ostrava. LIII(1): řada bezpečnostní inženýrství, pp. 9–14.
- Ivan, K. & Haidu, I., 2012. The spatio-temporal distribution of road accidents in Cluj-Napoca. *Geographia Technica*, 7(2), 32–38.
- Kolejka, J., 2010. Digital geographic data utilization in simulating toxic road accident. *Riscuri si catastrofe*, 8 (1), 11–24.
- Kolejka, J., Rapant, P., et al., 2015. Scenarios to support crisis management with geoinformation technologies (Scénáře podpory krizového řízení geoinformačními technologiemi), 1st ed., Brno, Soliton.
- Kolejka, J., Rapant, P. & Zapletalová, J. 2016. The GIS support to measures on the ground in case of leakage of liquid pollutant on the road. In: *Proceedings of the 23<sup>rd</sup> Central European Conference. Central Europe Area in View of Current Geography*. Masarykova univerzita, Brno, pp. 179–190.
- Krejčí, L. & Bambušek, M., 2012. Risky transport of dangerous goods by road in the Czech Republic (Rizikovost přepravy nebezpečných věcí silniční dopravou v ČR). Available at: [http://pernerscontacts.upce.cz/27\\_2012/Krejci.pdf](http://pernerscontacts.upce.cz/27_2012/Krejci.pdf).
- Maschio, G., et al., 2009. A GIS tool for the emergency management of terrorist actions in the transport of hazardous substance. *Journal of Loss Prevention in the Process Industries*, 22 (5), 625–633.
- Nicolet-Monnier, M. & Gheorghe, A., 1996. Quantitative Risk Assessment of Hazardous Materials Transport Systems: Rail, Road, Pipelines and Ship. Springer Science & Business Media.
- Raemdonck, K., Macharis, C. & Mairesse, O., 2013. Risk analysis system for the transport of hazardous materials. *Journal of Safety Research*, 45 (1), 55–63.
- Sobotka, P., 2015. Accident information on the roads of the Czech Republic per year 2014 (Informace o nehodovosti na pozemních komunikacích České republiky za rok 2014). 19 p., Available at: [www.policie.cz/soubor/2014-12-informace-pdf.aspx](http://www.policie.cz/soubor/2014-12-informace-pdf.aspx).
- Ștefănescu, L., Botezan, C. & Crăciun, I., 2018. Vulnerability analysis for two accident scenarios at an upper-tier Seveso establishment in Romania. *Geographia Technica*, 13 (1), 109–118.
- Tůma, J., 2000. Catastrophic Techniques of 20th Century Fears (Katastrofy techniky děsící 20. Století). Praha, Academia.
- Verma, M. & Verter, V., 2007. Railroad transportation of dangerous goods: Population exposure to airborne toxins. *Computers & Operations Research*, 34 (5), 1287–1303.
- Vonásek, V. & Lukeš, P. (eds.), 2015. Statistical Yearbook 2014 Czech Republic. Fire protection, Integrated Rescue System, Fire & Rescue Service of the Czech Republic. Praha, Ministerstvo vnitra, Generální ředitelství Hasičského záchranného sboru.
- Zhang, J., Hodgson, J. & Erkut, E., 2000. Using GIS to assess the risks of hazardous materials transport in networks. *European Journal of Operational Research*, 121 (2), 316–329.

## ASSESSMENT OF METAL CONTAMINATION IN COASTAL SEDIMENTS OF THE LÉVRIER BAY AREA, ATLANTIC OCEAN, MAURITANIA

*Mohamed El Houssein LEGRAA<sup>1</sup>, Abdelkader MOHAMED SALECK<sup>2</sup>,  
Ali YAHYA DARTIGE<sup>3</sup>, Mohamed Lemine CHEIKH ZAMEL<sup>3</sup>, Mohamed Mahmoud  
Ould ABIDINE<sup>1</sup>, Hassan ERRAOUI<sup>4</sup> and Zeinebou SIDOUMOU<sup>1</sup>*

DOI: 10.21163/GT\_2019.141.05

### ABSTRACT:

This work presents an assessment of marine pollution due to heavy metals accumulated in surface sediments collected from the Lévrier bay coastline in NW Atlantic, Mauritania. The samples were subjected to HCl and HNO<sub>3</sub> digestion, and an atomic absorption spectrophotometer was used to determine the concentrations of all metals (Cd, Pb, Cu and Zn) with the exception of Hg that was analysed by DMA-80 Direct Mercury Analyzer. The ranking of the metals at all studied sites was Al>Zn > Cu >Pb> Cd>Hg, and all the sediments samples displayed lower concentrations than the calculated worldwide mean of unpolluted sediments. The data were classified according to seasons. . Eight common pollution indices (i.e. EF, Igeo, QoC(%), CF, SPI, PLI, C<sub>d</sub> and PERI) were applied to ascertain the sediment quality. The enrichment (EF) factor resulting from Pb, Cu and Zn were <1 in surface sediments at all sampling sites and for Cd at Cap Blanc, indicating no enrichment . EF values in COMECA and Baie de l'Etoile for Cd were 1 < EF < 3, indicating a low enrichment. EF values for Cd at IMROP were 3 < EF < 5 indicated a moderate enrichment. The value of the geoaccumulation (Igeo) factor from all metals analysed was less than 0 in all sampling sites, which classified the sediments as uncontaminated. The value of contamination (CF) factor for all elements analyzed was lower than 1 across all sampling sites, implying a low contamination. The values of Sediment Pollution (SPI) Index were 0<SPI<2, indicating a natural status of sediments in COMECA, Baie de l'Etoile and Cap Blanc. The SPI values at IMROP were 2<SPI<5, which showed a low pollution of sediments on this site. The value of the pollution load (PLI) index was <1, which indicated that no metal pollution existed in sediments in all studied sites. The values of the contamination degree (C<sub>d</sub>) were <6 in all sampling sites, indicating a low degree of contamination. The values of the Potential Ecological Risk (PERI) index were 21.5<PERI <28.4 in all sampling sites, indicating a low ecologic risk. These results provide an important reference for future comparisons in the West African sub-region on the state of contamination of marine and coastal areas.

**Key-words:** Heavy metals, Pollution indexes, Sediments, Lévrier bay, Mauritania.

---

<sup>1</sup>Department of Biology, Faculty of Science and Technology, UR / EBIOME, University of Nouakchott AL Aasriya, Mauritania [legraahoussein@yahoo.fr](mailto:legraahoussein@yahoo.fr), [mintsidoumou@yahoo.fr](mailto:mintsidoumou@yahoo.fr), [hmd108@yahoo.fr](mailto:hmd108@yahoo.fr)

<sup>2</sup>Department of Geology, Faculty of Science and Technology, University of Nouakchott AL Aasriya, Mauritania [saleckakm@yahoo.fr](mailto:saleckakm@yahoo.fr)

<sup>3</sup>National Office for Sanitary Inspection of Fisheries and Aquaculture Products (ONISPA), Mauritania [alydartige@yahoo.fr](mailto:alydartige@yahoo.fr), [cheikhzamel@gmail.com](mailto:cheikhzamel@gmail.com)

<sup>4</sup>Department of Earth Sciences, Faculty of Science and Technology, Laboratory of Environmental Geosciences, Abdelmalek Essaidi University, Tétouan, Morocco [h.erraoui@fstt.ac.ma](mailto:h.erraoui@fstt.ac.ma)

## 1. INTRODUCTION

The Mauritanian coastline is under increasing anthropogenic pressure from the land. Urbanization, fishing, port facilities, oil exploration and exploitation all present potential threats to the richness of the fishery resources. All these activities are likely to generate pollution, chronic or accidental.

The Lévrier bay is home to important fishing and industrial activities. It is considered the area where the country's fishing and industrial activities are most intense. This situation is at the origin of disturbances of the marine environment; mainly due to industrial discharges loaded with organic and inorganic pollutants such as trace metals, some of which may be toxic to marine fauna and flora (L. Chouba et al., 2007). Coastal areas can be seriously affected by chemical contamination due to the long hydrologic residence time and consequent long-term retention of pollutants and atmospheric deposits that originate from (Alharbi & El-Sorogy, 2017; Liu et al., 2018; Zhang et al., 2017).

Introduction of heavy metals into the marine environment induces their accumulation in sediments. Many studies have dealt with heavy metal pollution worldwide in coastal areas discharge, oceanic dumping and aeolian processes along littoral zones. Most of these studies have used heavy metal analysis in sediments (Ghasemi, Moghaddam, Rahimi, Damalas, & Naji, 2018; Liu et al., 2018; Zhang et al., 2017). There are several methods for evaluating metal pollution in sediment and for risks caused by trace metals (Bastami et al., 2015; Izah, Bassey, & Ohimain, 2017; Sharifinia, Taherizadeh, Namin, & Kamrani, 2018; Vu et al., 2017; Yu et al., 2017).

The purpose was to provide preliminary information on environmental conditions and risks from metal contamination, and to evaluate the spatial and temporal variability of data covering 32 months during the cold and hot season. We also verified whether metal concentrations did were not significantly higher in sediment of the Lévrier bay, a coastal water located on the north shoreline of Mauritania. The main objectives of the present study were to (a) evaluate the spatial and temporal distribution of trace metals in surface sediments impacted by anthropogenic activities, (b) compare the level of pollution in the Lévrier bay coastal area with neighboring coasts and coasts elsewhere in the world, (c) compare heavy metals concentrations with the sediment quality guideline values and estimate their possible effects on the aquatic life of the study area (d) determine the ecological risk of Zn, Pb, Cu, and Cd in surface sediments using pollution indices and (e) make recommendations in terms of management to decision-makers.

## 2. MATERIALS AND METHODS

### 2.1. Study area

The Mauritanian coastline covers nearly 720 km. The coastal waters are the scene, among others, of two major phenomena. The first one is the occurrence of the thermal front resulting of the meeting of the two main currents of the zone namely the Canary current and the northern equatorial counter-current commonly called the Guinea current. The second one is the coastal upwelling that is one of the most important oceanographic factors in shaping the hydrological regime and structures of water bodies in northwestern Africa. It is at the origin of a high abundance of phytoplankton in the system which is the basis of the complex food web of the marine environment. This coastal Upwelling explains in part the very great wealth in species of the Mauritanian waters which have a reputation as being one of the richest water in fish in the world. The Lévrier Bay, located in the extreme north-west



of Mauritania, is the only large natural bay on the Mauritanian coast and one of the largest on the west coast of Africa (Wagne, BRAHIM, DARTIGE, & SÉFRIOUI, 2011). The bay is home to important fishing and industrial activities. It is considered the area where the fishing and industrial activities of the country are most intense. The bay provides shelter to four different ports including that of the most important industrial and mining society of Mauritania (IMSM). The city of Nouadhibou, the economic capital of the country, has more than forty fishing companies specialised in fish processing. This situation is at the origin of disturbances of the marine environment; mainly due to industrial discharges loaded with organic and inorganic pollutants such as trace metals (cadmium, lead, mercury, etc.) which are often toxic to marine fauna and flora (Lassaad Chouba & Mzougui-Aguir, 2006). Studies on the hydrology of Mauritanian marine waters have identified four major seasons, the most important of which are the cold (January to May) and hot (August to October) seasons interspersed with two inter-seasons, the cold-hot season (June to July) and the hot-cold season (November to December) (Wagne, 2013; Wagne et al., 2011).

We conducted a seasonal survey of six heavy metals (Cd, Pb, Hg, Cu, Zn and Al) accumulated in sediments during 32 months, from January 2013 to October 2016 during both the cold and the hot season at four stations located in the Lévrier bay coast (**Fig. 1**).

Sites were positioned along transects using a Garmin 585 global positioning system (GPS) with a resolution of about 3.5 m:

- Baie de l'Etoile ( $21^{\circ} 02' 19.2''$  N.  $17^{\circ} 01' 36.8''$  W); is located in northwestern Africa. It is the only African bay to the north to be characterized by a significant area of salt marshes spartines;

- Cap Blanc ( $20^{\circ} 46' 47.58''$  N.  $17^{\circ} 03' 30.24''$  W); Nature Reserve, home to one of the world's largest threatened colonies of monk seals;

- COMECA's sampling point is located outside Cansado Bay and more specifically on the southwest side of Cansado Point not far from the oil port and near the metallurgical company "COMECA" ( $20^{\circ} 50' 24.7''$  N.  $17^{\circ} 02' 03.6''$  W);

- IMROP's sampling point is located in the southeast of Cansado bay ( $20^{\circ} 51' 26.2''$  N.  $17^{\circ} 01' 52.0''$  W). It is the site of all the urban discharges of the bay, in particular the Cansado effluents containing domestic sewage.



**Fig. 1.** Location of the study area and sampling sites

## 2.2. Trace metal-based indices

To assess the quality of the sediments collected in this work, we used the Enrichment factor (EF), Quantification of contamination QoC (%), Contamination factor (CF), Geoaccumulation index (Igeo), Sediment Pollution Index (SPI), contamination degree (C<sub>d</sub>), potential ecological risk index (PERI) and pollution load index (PLI). These indices are the most common quantitative methods to assess pollution compared to the concentrations of the elements in the background samples taking into account the bioavailability of the various trace elements (Li et al., 2013). The abbreviations, terminology, algorithms and descriptions of metal pollution indices used in this study are listed in **Table 1**.

**Table 1.**

**Pollution indices, abbreviations, algorithms, descriptions and terminology**

Algorithm	Description	Terminology
$EF = \left( \frac{\left( \frac{C_n}{C_{ref}} \right)_{Sample}}{\left( \frac{B_n}{B_{ref}} \right)} \right) \quad (1)$	<p>C<sub>n</sub>: the content of the metal in the examined environment</p> <p>C<sub>ref</sub>: the content of the examined element in the reference environment</p> <p>B<sub>n</sub>: the concentration of the metal in the background</p> <p>B<sub>ref</sub>: the content of the reference element in the reference environment</p>	<p>EF &lt; 2 depletion to minimal enrichment</p> <p>EF = 2–5 moderate enrichment</p> <p>EF = 5–20 significant enrichment</p> <p>EF = 20–40 very high enrichment</p> <p>EF &gt; 40 extremely high enrichment</p>
$QoC(\%) = \left[ \frac{(C_n - B_n)}{C_n \times 100} \right] \quad (2)$	<p>QoC: The index of contamination</p> <p>C<sub>n</sub>: the concentration of the metal in the sample</p> <p>B<sub>n</sub>: the concentration of the metal in the background</p>	<p>If QoC values of heavy metal where positive, indicating that heavy metal contamination is anthropogenic in nature</p>
$I_{geo} = \log_2 \left[ \frac{C_n}{(1.5 \times B_n)} \right] \quad (3)$	<p><i>C<sub>n</sub></i>: the measured concentration of metal (n) in the sample</p> <p>B<sub>n</sub>: the geochemical background concentration of the metal(n)</p> <p>The factor 1.5 used to minimize the effect of possible variation in the background values</p>	<p>I<sub>geo</sub> &lt; 0 uncontaminated</p> <p>0 &lt; I<sub>geo</sub> &lt; 1 uncontaminated to moderately contaminated</p> <p>1 &lt; I<sub>geo</sub> &lt; 2 moderately contaminated</p> <p>2 &lt; I<sub>geo</sub> &lt; 3 moderately to strongly contaminated</p> <p>3 &lt; I<sub>geo</sub> &lt; 4 strongly contaminated</p> <p>4 &lt; I<sub>geo</sub> &lt; 5 strongly to extremely contaminated</p> <p>I<sub>geo</sub> &gt; 5 extremely contaminated</p>

$CF = \frac{Ch}{C_f^i} \quad (4)$	<p>n: number of heavy metals                  CF: contaminant factor                  Ch: heavy metal concentration                  Cs: heavy metal background reference values (Zn=64. Cu=26. Pb=10. Cd=0.16)</p>	<p>CF&lt;1 low contamination                  1&lt;CF&lt;3 moderate contamination                  3&lt;CF&lt;6 considerable contamination                  CF&gt;6 very high contamination</p>
$PLI = \sqrt[n]{(CF_1 \times CF_2 \times CF_3 \dots \times CF_n)} \quad (5)$	<p>n: number of heavy metals                  CF: contaminant factor</p>	<p>PLI&gt;1 metal pollution is present                  PLI&lt;1 There is no metal pollution</p>
$SPI = \frac{\sum (FE_m \times W_m)}{\sum W_m} \quad (6)$	<p>EF: ratio between sediment total content in a given sample and average shale concentration of a metal <i>m</i>;                  W: toxicity weight of metal <i>m</i> (Zn=1. Cu=2. Pb=5. Cd=300)</p>	<p>0&lt;SPI&lt;2 natural sediments                  2&lt;SPI&lt;5 low polluted sediments                  5&lt;SPI&lt;10 moderately polluted sediments                  10&lt;SPI&lt;20 highly polluted sediments                  SPI&gt;20 dangerous sediments</p>
$C_d = \sum_{i=1}^n CF_i \quad (7)$	<p>C<sub>d</sub>: sum of the contaminant factor (CF)</p>	<p>C<sub>d</sub>&lt;6 low degree of contamination                  6&lt;C<sub>d</sub>&lt;12 moderate degree of contamination                  12&lt;C<sub>d</sub>&lt;24 considerable degree of contamination                  C<sub>d</sub>&gt;24 high degree of contamination</p>
$PERI = \sum E_r^i \quad (8)$ $E_r^i = T_r^i \times C_f^i \quad (9)$	<p>T<sub>r</sub><sup>i</sup>: toxicity coefficient (Zn=1. Cu=Pb=5. Cd=30. Hg=40)                  CF: contaminant factor</p>	<p>PERI&lt;95 low potential ecologic risk                  95&lt;PERI&lt;190 moderate ecologic risk                  190&lt;PERI&lt;380 considerable ecologic risk                  PERI&gt;380 very high ecologic risk</p>

### 2. 3. Sampling and analytical procedure

Sediment samples were taken from the shoreline at low tide using PVC tubes with an internal diameter of 58 mm (quantity 132 ml). An effort was made not to disturb the edges of the sample and to transfer all of the core into glass bottles (maroon colour) for heavy metal analyzes (including Al as a marker) or polyethylene bags for granulometry using a stainless steel spatula. The samples were then kept whole without sieving. Then kept refrigerated at 4-10 ° C for 6-48 hours and frozen at -50 ° C after arrival at the laboratory. Samples were mineralized in closed Teflon bombs. Weigh precisely about 0.2 g of dry weight to be analyzed; 10 ml of concentrated HNO<sub>3</sub> nitric acid (for analysis) was added;

the reactors were hermetically closed and left at room temperature for at least 1 hour. They are then heated by microwave according to a well-defined program.

The analysis of Cd, Pb, Cu, Zn and Al was done by the graphite furnace method and the Hg analyzes were made by a DMA-80 Direct Mercury Analyzer. At the end of the neutralization, the mineralizer was poured into a 50 mL Teflon tube which was supplemented with ultrapure water to the required volume and filtered. This procedure was validated by performing the same extraction on a standard HISS-1 certified sediment (National Research Council of Canada, certified reference materials, marine sediments) and the results obtained were compared to the certified values in **Table 2**. All samples were run in triplicate and all elements were subsequently expressed in mg kg<sup>-1</sup> dry weight.

**Table 2.**  
**Measured and certified values of trace metals and detection limits (mgkg<sup>-1</sup>DW) in the sediments standard reference material BCR16-ERM278K.**

	<b>Cd</b>	<b>Pb</b>	<b>Hg</b>	<b>Cu</b>	<b>Zn</b>	<b>AL</b>
Measured	17±1	590±30	8.4±1	800±40	2983±100	7.3±0.5
Certified	18 ± 0.5	609 ± 14	8.6 ± 0,4	838 ± 16	3060±60	---
Detection limits	0.02	0.26	0.006	0.25	0.29	---

## 2.4. Statistical analysis

All data were statistically processed in STATISTICA ver. 6, SPSS for Windows version 20.0 (2008) and the graphics were drawn in ArcGis 10.3 and OriginPro 8. Heavy metal concentrations in sediments were compared to sediment quality standards, such as, ISQGS and USEPA, to illustrate the state of pollution in the study area. To assess the ecological health risks of trace metals in the study area, chemical pollution indices were also calculated.

## 3. RESULTS AND DISCUSSION

### 3.1. Trace metals concentration in superficial sediments at Lévrier bay

#### 3.1.1. Environmental gradients

Samples taken from the various sites in the Lévrier bay consist for up to 43% of sand (63µm-2mm) and for 16% of silt (2-63 µm). Samples taken at Cap Blanc and COMECA were mainly made up of sand (55 and 45%, respectively) the grain size of which varied between 63µm and 2mm. These two sites, therefore, consist of rather homogeneous sand but with a marked discontinuity composed of silt at Cap Blanc (35%) and COMECA (22%).

Samples taken at the Baie de l'Etoile and IMROP contained the lowest levels of silt (2 and 2.7%, respectively) the grain size of which one varied between 2  $\mu\text{m}$  and 63  $\mu\text{m}$ . The intense mining activity for about sixty years would be at the origin of this enrichment in fine materials of this place (Maigret, J, 1972). As reported in many studies, decreased concentrations of heavy metal in sediments are strongly and positively correlated to grain size. Thus, a high level of concentration of heavy metals is generally associated with the fine-grained material (<63  $\mu\text{m}$ ) compared with size grains > 63  $\mu\text{m}$ , since the size of their surfaces has an influence on the collection and transport of inorganic constituents (Krauskopf, 1955; Romankevich, 1984).

### 3.1.2. Variation of heavy metal concentrations in sediments

Mean concentrations of the heavy metals (mean  $\pm$  SEM) in sediments of the Lévrier bay by individual sites are presented in **table 3**.

The highest Zn concentration of  $1.900 \pm 0.620$  mgkg<sup>-1</sup> (DW) was found at Cap Blanc and the lowest values were observed at IMROP ( $1.219 \pm 0.424$  mgkg<sup>-1</sup> DW). The highest and lowest Cu concentrations were observed at Baie de l'Etoile and COMECA ( $0.590 \pm 0.007$  and  $0.084 \pm 0.003$  mgkg<sup>-1</sup> DW, respectively). The highest concentrations of Pb were found at COMECA ( $1.671 \pm 0.534$  mgkg<sup>-1</sup> DW) and the lowest concentrations were found Cap Blanc ( $0.251 \pm 0.111$  mgkg<sup>-1</sup> DW). For Cd, Baie de l'Etoile exhibited the highest concentration ( $0.130 \pm 0.070$  mgkg<sup>-1</sup> DW) and the lowest concentrations were observed at Cap Blanc ( $0.076 \pm 0.021$  mgkg<sup>-1</sup> DW). However, the Hg values were the lowest among the metals analyzed regardless of the sampling site. The ranking of the heavy metals in the sediments of the Lévrier bay was Al > Zn > Cu > Pb > Cd > Hg (**table. 3**).

**Table 3.**

**Concentraions of heavy metal (mean $\pm$ Standard deviation: mg/kg DW) in sediments for the studied sites**

Sites	Cd	Pb	Hg	Cu	Zn	Al
CO	$0.090 \pm 0.018$	$1.671 \pm 0.534$	$0.016 \pm 0.002$	$0.084 \pm 0.003$	$1.349 \pm 0.011$	$3.038 \pm 0.988$
BE	$0.130 \pm 0.070$	$1.000 \pm 0.400$	$0.017 \pm 0.002$	$0.590 \pm 0.007$	$1.283 \pm 0.045$	$3.113 \pm 1.023$
CB	$0.076 \pm 0.021$	$0.251 \pm 0.111$	$0.020 \pm 0.002$	$0.573 \pm 0.083$	$1.900 \pm 0.620$	$3.635 \pm 0.240$
IM	$0.082 \pm 0.012$	$0.805 \pm 0.028$	$0.019 \pm 0.002$	$0.352 \pm 0.008$	$1.219 \pm 0.424$	$0.886 \pm 0.009$

Heavy metal concentrations in superficial sediments varied between cold and hot seasons (**Fig.2**), with increased Cd and Hg concentrations samples collected during the cold season at all sampling sites and for Al in the samples taken from COMECA and Baie de l'Etoile. During the hot season, the concentrations of Pb, Cu and Zn were higher at all sampling sites, and also for Al taken at Cap Blanc and IMROP. The discharge of domestic and industrial wastewater at sea containing metals such as Cd, Cu and Zn, may provide a dominant and constant input into the sediments of Lévrier bay, thus masking the natural seasonal patterns.

The absence of significant variations in the concentration of metals in surface sediments between sampling dates can be explained by deposits (of natural and / or anthropogenic elements) and post-depositional remobilization processes, with modification of the diagenetic activity, likely to influence the concentrations. (with or without a seasonal cycle) (Ridgway & N.B. PRICE, 1987). However, many factors related to the natural variability during the seasons in the physical and chemical characteristics of sediments (especially between cold and hot seasons) can also strongly affect the retention capacity of heavy metals and thus the concentration of metals in the surface sediments.

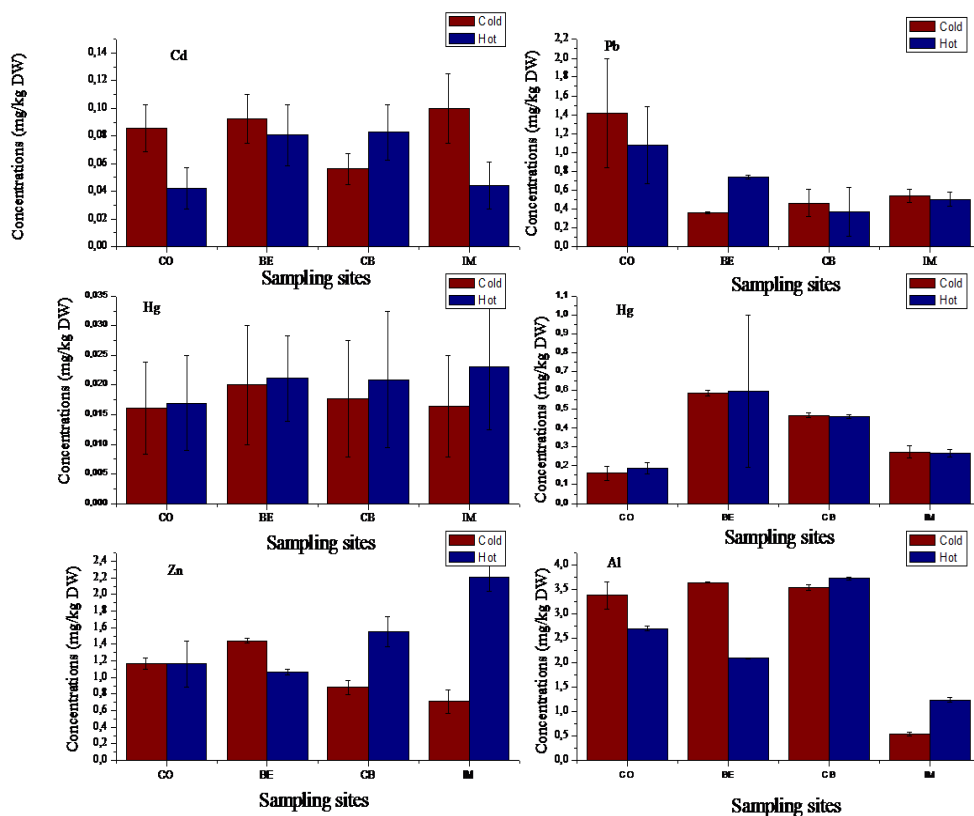


Fig. 2. Heavy metal concentrations in sediments for the studied sites and seasons.

The average concentrations of metals analysed in this study and previous work on the northwest Atlantic coasts are summarized in **Table 4**. The average concentrations of all analyzed heavy metals in this study are significantly lower than values for other areas of the sub-region.

Table 4.

**Comparison of trace metal concentration in sediments at the Lévrier bay and other areas along the North West Atlantic coast**

Sampling sites	Cd	Pb	Hg	Cu	Zn	Al	Reference
Mauritania							
Comeca	0.090±0.018	1.671±0.534	0.016±0.002	0.084±0.003	1.349±0.011	3.038±0.988	Present work
BaieEtoile	0.130±0.070	1.000±0.400	0.017±0.002	0.590±0.007	1.283±0.045	3.113±1.023	
Cap blanc	0.076±0.021	0.251±0.111	0.020±0.002	0.573±0.083	1.900±0.620	3.635±0.240	
Imrop	0.082±0.012	0.805±0.028	0.019±0.002	0.352±0.008	1.219±0.424	0.886±0.009	
Banc d'Arguin	0.16	5.4	---	4	26	2.90	(Nolting, Ramkema, & Everaarts, 1999)
Marocco Atlantic coast							
Estuaries							
Loukous	0,15-1,60	31,81-49,66	---	1,63-32,33	64,81-115,27	12564-48567	(Cheggour et al., 2001)
Sebou	1,15	---	---	20	217	---	
Bouregreg	1,64	---	---	18,3	178	---	
Oum Rabia	1,45			16,2	222		
Lagoon	3.6	18-92	---	juil-65	25-221	---	
Sénégal							
Dakar coast	0.65	---	---	9.98	15.5	8.5	(Diop et al., 2012)
Spain							
Southern Atlantic Coast	0.38	16.09	0.20	30.09	173.45	---	(Usero, Morillo, & Gracia, 2005)

### 3.2. Toxicity assessment

Sediment contamination and contaminant bioavailability were assessed using the sediment quality guideline values recommended by the United States Protection Agency (USEPA, 2014) and the Standard Canadian Interim Marine Sediment Quality (ISQG, 1995). The concentrations of the heavy metals analyzed in the sediments were compared with the thresholds levels (ERL, ERM, PEL, LAL and HAL) for adverse effects of these metals on the biota in the study area.

For an assessment of the environmental aspects of the sediments analyzed and their possible effects on aquatic life in the coastal fringe of the Lévrier bay, it was necessary to use the reference values or standards in relation to the obtained results.

**Table 5.**

**EPA guidelines for sediments, comparison of standard concentrations ( $\text{mg}^{-1}\text{Kg}$ ) for threshold impact of pollutants with sediments of the sampling sites**

USEPA guidelines			Studied sites (Lévrier bay)				
	No polluted	Moderately polluted	Heavily polluted	COMECA	Baie de l'Etoile	Cap Blanc	IMROP
Cd	-	-	>6	0,090±0,018	0,130±0,070	0,076±0,021	0,082±0,012
Pb	<40	40-60	>60	1,671±0,534	1,000±0,400	0,251±0,111	0,805±0,028
Hg	<1,0	-	>1,0	0,016±0,002	0,017±0,002	0,020±0,002	0,019±0,002
Cu	<25	25-50	>50	0,084±0,003	0,590±0,007	0,573±0,083	0,352±0,008
Zn	<90	90-200	>200	1,349±0,011	1,283±0,045	1,900±0,620	1,219±0,424

The concentration of the elements studied (Cd, Pb, Hg, Cu and Zn) were evaluated by comparison with the United States Protection Agency (USEPA, 2014), sediment quality guideline values (**Table 5**), which shows that sediments from all sites sampled in this study were considered unpolluted. The comparison of metal concentrations analyzed in this study with the ERL, ERM, ISQG and PEL, LAL and HAL guideline values are presented in Table 6.

The overall analysis (**Table 6**) shows that the Cu, Zn, Cd and Pb concentrations are lower than all the guide values given above. Therefore, these elements cannot cause adverse effects on the biota of the study area.

Based on guideline values mentioned, a sample may cause a toxic effect once the concentration of one of the metals analysed reaches or exceeds ERL and ISQG limit values (Bakan & Özkoç, 2007). Thus, the guide values of four metals (Cd, Cu, Pb and Zn), presented in the **Table 6** mentioned above, were used as a reference to illustrate the quality of the samples studied.

For the Hg and Al analyzed and in the absence of guide values, their possible effects on aquatic life have not been verified. In addition, these values have been applied by many authors in environmental studies. As an example, in the work done by Caeiro et al, in 2005 on the analysis of elements Cd, Cu, Pb, Cr, Hg, Al, Zn and As in the samples of sediments harvested from the estuary Sado at Portugal, 3% of the samples analyzed could have significant effects on marine organisms while 47% could have moderate effects.

Bakan et al, in 2007 applied the guideline values to evaluate the impact of heavy metals (Cu, Zn, M. Cr. Cd. Pb. and Ni) in surface sediments from the Turkish coast of the Black Sea on biota. These authors confirmed that the values obtained are generally low to moderate and therefore posed no threat to biota. Pedro et al (2008), carried out a study concerning the effects of certain metals (Cd, Co, Cr, Cu, Hg, Ni, Pb and Zn) on the Tagus estuary in Portugal. Their study showed that the metals analyzed had only a weak effect.



**Table 6.**

**Comparison of standard concentrations for threshold impact of pollutants with sediments of the studied sites (mg/kg DW).**

Standard		Zn	Cu	Pb	Cd
NOAA sediment quality	ERM	410	270	218	9.6
(Long, Macdonald, Smith, & Calder, 1995)	ERL	150	34	46.7	1.2
Canada environment (CCME, 1999)	PEL	271	108	112	4.2
	ISOGS	124	18.7	30.2	0.7
(USEPA, 2002)	LAL	5	2	2	0.07
	HAL	410	270	218	9.6
Sediment of the study area	COMECA	1.349±0.011	0.084±0.003	1.671±0.534	0.090±0.018
	Baie de l'Etoile	1.283±0.045	0.590±0.007	1.000±0.400	0.130±0.070
	Cap Blanc	1.900±0.620	0.573±0.083	0.251±0.111	0.076±0.021
	IMROP	1.219±0.424	0.352±0.008	0.805±0.028	0.082±0.012

NOAA: National Oceanic and Atmospheric Administration (USA), ERL: concentration below which adverse effect would be rarely observed, ERM: concentration where biological effects may occur, ISQG: quality guideline value, PEL: concentration above which adverse effects are expected to occur frequently.

### 3.3. Quality assessment with pollution indices

#### 3.3.1. Spatial variation in trace metal-based indices

The use of the many indices and approaches is recommended for a better assessment of the quality of sediments and their development. They are fast and relatively simple to apply (Kwon and Lee., 1998). The use of these tools gives more performance and confidence in decisions about human health and marine and coastal ecosystems (Caeiro et al., 2005). For the calculation of certain indices, we have also used the average values of the metallic elements in the Earth's crust recommended by Taylor (1964) as background values. These average values are recent and calculated from several previous studies. In addition, these values are similar to those proposed by the authors (Turekian & Wedepohl, 1961; Wedepohl, 1995). Indices for elements in the sediments were calculated using Al as a reference element (Turekian & Wedepohl, 1961). Al is a reliable indicator of the contribution of terrestrial and crust-derived materials (Ackermann., 1980; Chester., 2000; Lide., 2003) to estimate the anthropogenic impact on sediments and to differentiate between the anthropogenic and natural sources of metals (Sinex & Helz, 1981; Sutherland,

2000). Results of pollution indices (i.e. EF, Igeo, QoC(%), CF, SPI, PLI, CD and PERI) are presented in **Tables 7, 8**.

The Enrichment Factor (EF) was initially developed to study the origin of the elements in the different compartments of the sea environment (Duce, Hoffman, & Zoller, 1975; Zoller, W. H, Hoffmann, G. L, & Duce, R. A, 1974), but it was gradually developed and expanded to other areas such as soils, lake sediments, tailings, etc. The enrichment factor (EF) ranged from  $0.88 \pm 0.04$  to  $3.44 \pm 0.12$  for Cd, from  $0.07 \pm 0.02$  to  $0.91 \pm 0.03$  for Pb, from  $0.02 \pm 0.01$  to  $0.06$  for Cu and from  $0.05$  to  $0.19 \pm 0.02$  for Zn. The EF for Pb, Cu and Zn were  $< 1$  in surface sediments of all sampling sites and for Cd at Cap Blanc, which denotes “no enrichment” by these trace metals (**Table 7**). The EF values at COMECA and Baie de l'Etoile for Cd were  $1 < EF < 3$  which indicated a “low enrichment”. The EF values for Cd in sediments at IMROP were  $3 < EF < 5$  which indicated a “moderate enrichment”. The geoaccumulation index (Igeo) is used to evaluate heavy metal pollution in sediments (Muller., 1979, Banat et al., 2005). The results of the geoaccumulation factor (Igeo) of all metals analyzed are shown in **Table 7**. Igeo values of trace metals studied in sediments were less than 0 in all sampling sites, so they are classified as ‘uncontaminated’ (**Table 7**). The contamination factor (CF), which describes the contamination by a given toxic substance in a basin (Hakanson, 1980), is also used to assess contamination levels in relation to average concentrations of the respective trace metals in the environment i.e. sediment to the measured background values from previous studies with similar geological origin or uncontaminated sediments (Sutherland, 2000; Tijani, Jinno, & Yoshinari Hiroshiro, 2004; Uriah & Shehu, 2014). The contaminant factor (CF) ranged from  $0.380 \pm 0.030$  to  $0.650 \pm 0.050$  for Cd, from  $0.031 \pm 0.040$  to  $0.209 \pm 0.030$  for Pb, from  $0.002$  to  $0.011$  for Cu, from  $0.018$  to  $0.027 \pm 0.007$  for Zn and from  $0.011$  to  $0.044 \pm 0.003$  for Al. The CF values for all elements analysed were  $< 1$  in surface sediments of all sampling sites, which indicates a “low contamination” by these trace metals (**Table 8**). Negative Igeo values and positive CF values suggest that the metal was imported by human activities, but that the metal in question has not yet reached the threshold of pollution due to the dilution in coarse sediments (Bhutiani, Kulkarni, Khanna, & Gautam, 2017; Oliveira, Palma, & Valença, 2011) or caused by relatively low levels of contamination of some metals in some cores and the background variability factor (1.5) in the Igeo equation (Abraham & Parker, 2007; Müller, G., 1979). Quantification of contamination (QoC (%)) represents the lithogenic metal (Asaah, Victor A., Akinlolu, F, & Suh, Cheo E, 2006). The QoC values for heavy metals Cd, Pb, Hg, Cu, Zn and Al were negative for all sampling sites, indicating that anthropogenic activities are not the causes of contamination at the majority sites in the study area (**Table 7**). The pollution load index (PLI) provides more complete information on the toxicity of the metals studied in the respective samples analyzed (Bhutiani et al., 2017; Ghaleno, 2015; Tomlinson, Wilson, Harris, & Jeffrey, 1980; Yang, Xu, Liu, He, & Long, 2011). The pollution load index (PLI) ranged from  $0.124 \pm 0.005$  to  $0.166 \pm 0.013$  (**Table 8**). The PLI values in sediments of all studied sites were  $< 1$  which indicates that there was no metal pollution (**Table 8**). The PLI followed the order of the Baie de l'Etoile > Cap Blanc > COMECA > IMROP. The Sediment Pollution Index (SPI) is defined as the linear sum of the metal enrichment factors weighted by their respective toxicity weights. These metal toxicity weights are based on their different relative toxicity and are inversely proportional to the lithogenic limits of the average shale. This is based on the fact that metal concentrations in unpolluted sediments should not exceed average shale values (Singh et al., 2002). The Sediment Pollution Index (SPI) ranged from  $0.856$  to  $3.363$  in surface sediments of the study area (**Table 8**). The SPI values at COMECA, Baie de l'Etoile and

Cap Blanc were  $0 < \text{SPI} < 2$  which showed “natural sediments”. The SPI value at IMROP was  $2 < \text{SPI} < 5$  indicated “low polluted sediments”. The contamination degree ( $C_d$ ) is the sum of all contamination factors, which provides information about all contamination in a particular sampling location (Singovszka, Balintova, Demcak, & Pavlikova, 2017). The contamination degree ( $C_d$ ) ranged from  $0.688 \pm 0.052$  to  $0.817 \pm 0.143$  (Table 8). The  $C_d$  values in sediments of all sampling sites were  $< 6$  which indicated a “low degree of contamination”. The Potential Ecological Risk Index (PERI) is commonly used as a tool for the diagnosis of heavy metal pollution in sediments (Hakanson, 1980), as it takes into account the levels and toxicities of heavy metals (Ntakirutimana, Du, Guo, Gao, & Huang, 2013). However, risk levels may be overstated because of the use of total content rather than species-related data. The geo-accumulation index shows the level of enrichment of heavy metals without taking into account the biological effect of heavy metals, while the potential ecological risk index (PERI) includes the toxicity of heavy metals to mitigate this deficiency. Therefore, combining these two evaluation methods can make the assessment results of the heavy metal pollution more comprehensive and rational (Yu et al., 2017). The potential ecological risk index (PERI) varied between 21.50 and 28.14 in the study area (Table 8). PERI values were  $< 90$  at all sampling sites, then according to Hakanson (1980) these sites were classified as being at “low ecologic risk”. The PERI ranking order was Baie de l’Etoile  $>$  COMECA  $>$  IMROP  $>$  Cap Blanc (Table 8). Such indices offer useful information, provided that their limitations are recognized. So, the overall results of this study showed the lowest levels of heavy metal contamination in the sediments at the Lévrier bay. The systematic monitoring and evaluation of these marine and coastal ecosystems is a priority for safeguarding ecological security. The data in this study provide a basis for future investigations of trace metal pollution and its possible impacts on the sensitive ecosystems.

Table 7.

Mean values ( $\pm$  SEM) of heavy for metals enrichment factor (EF), geoaccumulation index (Igeo) and quantification of contamination (QoC(%)) in surface sediment at each sampling site along the Lévrier bay.

Site	Cd			Pb			Hg	Cu			Zn			Al	
	EF	Igeo	QoC(%)	EF	Igeo	QoC(%)	QoC(%)	EF	Igeo	QoC(%)	EF	Igeo	QoC(%)	Igeo	QoC(%)
CO	3.44 $\pm 0.12$	- 2.33 $\pm 0.35$	- 211 .69	0.91 $\pm 0.03$	- 3.36 $\pm 0.31$	- 541 .42	- 383 .75	0.06	- 9.19 $\pm 0.48$	- 315 12.91	0.19 $\pm 0.02$	- 6.50 $\pm 0.16$	- 591 3.77	2.67 $\pm 1.54$	- 260 9.47
BE	1.24 $\pm 0.19$	- 1.97 $\pm 0.44$	- 165 .19	0.57 $\pm 0.06$	- 5.34 $\pm 0.98$	- 136 0.63	- 288 .90	0.00	- 7.11 $\pm 0.02$	- 919 8.51	0.05	- 6.19 $\pm 0.10$	- 547 7.65	2.66 $\pm 1.54$	- 277 1.35
CB	1.75 $\pm 0.05$	- 2.15 $\pm 0.20$	- 188 .29	0.32 $\pm 0.10$	- 5.06 $\pm 0.42$	- 182 6.79	- 314 .24	0.03	- 7.52 $\pm 0.21$	- 117 43.88	0.05 $\pm 0.01$	- 6.75 $\pm 0.56$	- 567 7.30	2.54 $\pm 1.47$	- 216 4.10
IM	0.88 $\pm 0.04$	- 2.31 $\pm 0.49$	- 177 .46	0.07 $\pm 0.02$	- 4.82 $\pm 0.58$	- 143 5.77	- 305 .47	0.02 $\pm 0.01$	- 8.02 $\pm 0.12$	- 202 14.23	0.06 $\pm 0.02$	- 6.16 $\pm 0.32$	- 469 2.23	3.53 $\pm 2.04$	- 919 0.69

Table 8.

Mean values ( $\pm$  SEM) of heavy metals for Contaminant factor (CF), Sediment Pollution Index (SPI), pollution load index (PLI), contamination degree ( $C_a$ ) and potential ecological risk index (PERI) in surface sediments at all sampling sites from the Lévrier bay.

Site	CF					SPI	PLI	$C_a$	PER I
	Cd	Pb	Cu	Zn	Al				
COMEC A	0.450 $\pm$ 0.0 50	0.209 $\pm$ 0.0 30	0.002	0.019	0.037 $\pm$ 0.0 02	3.36 $\pm$ 0.0 12	0.133 $\pm$ 0.0 01	0.726 $\pm$ 0.1 08	21.6 5
Baie de l'Etoile	0.650 $\pm$ 0.0 50	0.125 $\pm$ 0.0 50	0.011	0.018	0.038 $\pm$ 0.0 04	1.22 $\pm$ 0.0 19	0.166 $\pm$ 0.0 13	0.817 $\pm$ 0.1 43	28.1 4
Cap Blanc	0.381 $\pm$ 0.0 60	0.031 $\pm$ 0.0 40	0.010 $\pm$ 0.0 02	0.027 $\pm$ 0.0 07	0.044 $\pm$ 0.0 03	1.71 $\pm$ 0.0 05	0.139 $\pm$ 0.0 08	0.688 $\pm$ 0.0 52	21.5 0
IMROP	0.380 $\pm$ 0.0 30	0.101 $\pm$ 0.0 01	0.006	0.021 $\pm$ 0.0 03	0.011	0.86 $\pm$ 0.0 04	0.124 $\pm$ 0.0 05	0.692 $\pm$ 0.1 06	21.9 1

#### 4. CONCLUSION

The physicochemical parameters collected from superficial sediment in the Lévrier bay vary little from one site to another and do not show an apparent accumulation of heavy metals in the samples analysed. The heavy metal concentrations in superficial sediments varied between sampling sites, with IMROP showing the highest levels of Zn recorded and the lowest values of Hg compared to all other sampling sites. In contrast, no significant seasonal variation between the four sampling sites was found. It can be seen that the mean concentrations of all studied heavy metals in this study were much lower than those in other areas along the NW African coast. All sampling sites that we studied were considered unpolluted by Cd, Pb, Hg, Cu and Zn, when the chemical contamination in the sediments was compared with international standards. Our findings showed that in the overall analysis of the sediments concentrations of Cu, Zn, Cd and Pb were below the corresponding guides values of the ISOGS, ERM, PEL, LAL, HAL, indicating no harmful effect on the biota in the study area. The ecological health risk assessment of trace metals using chemical pollution indices revealed none to moderate enrichment for heavy metals analysed in surface sediments (Enrichment factor), uncontaminated sampling sites (Geoaccumulation index), low contamination (Contamination factor), and natural to low polluted sediments (Sediment Pollution Index). The pollution load index (PLI) at all sampling sites indicates no pollution, and the contamination degree of pollution ( $C_a$ ) values in sediments of all sampling showed a low degree of contamination. According to mean values of the potential ecological risk index (PERI), all sampling sites were classified as being at low ecological risk. The quantification factor of contamination (QoC) indicated that anthropogenic activities were not the main cause of contamination at the majority sites in the study area. Our findings should raise the awareness of the contamination status at the Lévrier bay's coast, and provide a valuable benchmark for future comparisons in the area. They are of great importance for the design of long-term management and good governance policies of marine and coastal areas of Mauritania. Finally, it is important that local authorities act responsibly and enforce existing laws and regulations to protect marine and coastal biodiversity. Without a doubt, this study will contribute to the limited knowledge of sediment contamination by heavy metals in the Lévrier bay and the sub-region.

#### ACKNOWLEDGEMENTS

This study results from a collaboration between the University of Nouakchott Al Aasriya (Mauritania), Abdel Malek Essaadi University (Morocco), Mauritanian Institute of Oceanographic and Fisheries Research (IMROP) and the National Office of Sanitary Inspection of Fishery Products

and Aquaculture (ONISPA). The authors would like to devote a special thanks to these institutions for courtesy, kindness and help. We also would like to thank Rouane Omar Hassane from the Laboratory of Environmental Monitoring Network (LRSE) of Ahmed Ben Bella University 1 (Algérie) for his support. The reason why a large number of authors (7) contributed to the realization of this manuscript takes place in their specialization as well as their respective roles in the different parts of this document.

## REFERENCES

- Abraham, G. M. S., & Parker, R. J. (2007). Assessment of heavy metal enrichment factors and the degree of contamination in marine sediments from Tamaki Estuary, Auckland, New Zealand. *Environmental Monitoring and Assessment*, 136(1-3), 227-238. <https://doi.org/10.1007/s10661-007-9678-2>
- Alharbi, T., & El-Sorogy, A. (2017). Assessment of metal contamination in coastal sediments of Al-Khobar area, Arabian Gulf, Saudi Arabia. *Journal of African Earth Sciences*, 129, 458-468. <https://doi.org/10.1016/j.jafrearsci.2017.02.007>
- Asaah, Victor A., Akinlolu, F., & Suh, Cheo E. (2006). Heavy metal concentrations and distribution in surface soils of the Bassa Industrial Zone 1, Douala, Cameroon. *Arabian Journal for Science and Engineering*, 31, 147-158.
- Bakan, G., & Özkoç, H. B. (2007). An ecological risk assessment of the impact of heavy metals in surface sediments on biota from the mid-Black Sea coast of Turkey. *International Journal of Environmental Studies*, 64, 45-57.
- Bastami, K. D., Afkhami, M., Mohammadzadeh, M., Ehsanpour, M., Chambari, S., Aghaei, S., ... Baniamam, M. (2015). Bioaccumulation and ecological risk assessment of heavy metals in the sediments and mullet *Liza klunzingeri* in the northern part of the Persian Gulf. *Marine Pollution Bulletin*, 94(1-2), 329-334. <https://doi.org/10.1016/j.marpolbul.2015.01.019>
- Bhutiani, R., Kulkarni, D. B., Khanna, D. R., & Gautam, A. (2017). Geochemical distribution and environmental risk assessment of heavy metals in groundwater of an industrial area and its surroundings, Haridwar, India. *Energy, Ecology and Environment*, 2(2), 155-167. <https://doi.org/10.1007/s40974-016-0019-6>
- Caeiro, S., Costa, M. H., Ramos, T. B., Fernandes, F., Silveira, N., Coimbra, A., ... Painho, M. (2005). Assessing heavy metal contamination in Sado Estuary sediment: An index analysis approach. *Ecological Indicators*, 5(2), 151-169. <https://doi.org/10.1016/j.ecolind.2005.02.001>
- CCME. (1999). Canadian environmental quality guidelines. Available online at <http://ceqg-rcqe.ccme.ca/>.
- Cheggour, M., Chafik, A., Langston, W. J., Burt, G. R., Benbrahim, S., & Texier, H. (2001). Metals in sediments and the edible cockle *Cerastoderma edule* from two Moroccan Atlantic lagoons: Moulay Bou Selham and Sidi Moussa. *Environmental Pollution*, 115(2), 149-160.
- Chouba, L., Kraiem, M., Njimi, W., Tissaoui, C. H., Thompson, J. R., & Flower, R. J. (2007). Seasonal variation of heavy metals (Cd, Pb and Hg) in sediments and in mullet, *Mugil cephalus* (Mugilidae), from the Ghar El Melh Lagoon (Tunisia). *Transitional Waters Bulletin*, 1(4), 45-52.
- Chouba, Lassaad, & Mzougui-Aguir, N. (2006). Les métaux traces (Cd, Pb, Hg) et les hydrocarbures totaux dans les sédiments superficiels de la frange côtière du golfe de Gabès.
- Diop, C., Dewaele, D., Toure, A., Cabral, M., Cazier, F., Fall, M., ... Diouf, A. (2012). Étude de la contamination par les éléments traces métalliques des sédiments cotiers au niveau des points d'évacuation des eaux usées à Dakar (Sénégal). *Revue des sciences de l'eau*, 25(3), 277. <https://doi.org/10.7202/1013107ar>
- Duce, R. A., Hoffman, G. L., & Zoller, W. H. (1975). Atmospheric trace metals at remote northern and southern hemisphere sites: pollution or natural? *Science*, 187(4171), 59-61.

- Ghaleno, O. R. (2015). Potential ecological risk assessment of heavy metals in sediments of water reservoir case study: Chah Nimeh of Sistan. *Proceedings of the International Academy of Ecology and Environmental Sciences*, 5(3), 89.
- Ghasemi, S., Moghaddam, S. S., Rahimi, A., Damalas, C. A., & Naji, A. (2018). Ecological risk assessment of coastal ecosystems: The case of mangrove forests in Hormozgan Province, Iran. *Chemosphere*, 191, 417-426. <https://doi.org/10.1016/j.chemosphere.2017.10.047>
- Hakanson, L. (1980). An ecological risk index for aquatic pollution control. a sedimentological approach. *Water Research*, 14, 975-1001.
- ISQG. (1995). Interim Sediment Quality Guidelines. Environment Canada, Ottawa, p. 9.
- Izah, S. C., Basse, S. E., & Ohimain, E. I. (2017). Assessment of Pollution Load Indices of Heavy Metals in Cassava Mill Effluents Contaminated Soil: a Case Study of Small-scale Processors in a Rural Community in the Niger Delta, Nigeria. *Bioscience Methods*. <https://doi.org/10.5376/bm.2017.08.0001>.
- Krauskopf, K. B. (1955). Sedimentary deposits or rare metals. Economic Geology 50<sup>th</sup> anniversary volume, United States Department of the Interior Geological Survey, 411-463.
- Li, F., Huang, J., Zeng, G., Yuan, X., Li, X., Liang, J., Bai, B. (2013). Spatial risk assessment and sources identification of heavy metals in surface sediments from the Dongting Lake, Middle China. *Journal of Geochemical Exploration*, 132, 75-83. <https://doi.org/10.1016/j.gexplo.2013.05.007>
- Liu, Q., Wang, F., Meng, F., Jiang, L., Li, G., & Zhou, R. (2018). Assessment of metal contamination in estuarine surface sediments from Dongying City, China: Use of a modified ecological risk index. *Marine Pollution Bulletin*, 126, 293-303. <https://doi.org/10.1016/j.marpolbul.2017.11.017>
- Long, E. R., Macdonald, D. D., Smith, S. L., & Calder, F. D. (1995). Incidence of adverse biological effects within ranges of chemical concentrations in marine and estuarine sediments. *Environmental management*, 19(1), 81-97.
- Maigret, J. (1972). Campagne expérimentale de pêche des sardinelles et autres espèces pélagiques (Juillet 1970 - Octobre 1971). Observations concernant l'océanographie et la biologie des espèces. Laboratoire des Pêches.
- Müller, G. (1979). Distribution, Enrichment and Ecological Risk Assessment of Six Elements in Bed Sediments of a Tropical River, Chottanagpur Plateau: A Spatial and Temporal Appraisal. *Umsch. Wiss. Tech*, 79, 778-783.
- Nolting, R. F., Ramkema, A., & Everaarts, J. M. (1999). The geochemistry of Cu, Cd, Zn, Ni and Pb in sediment cores from the continental slope of the Banc d'Arguin (Mauritania). *Continental Shelf Research*, 19(5), 665-691.
- Ntakirutimana, T., Du, G., Guo, J., Gao, X., & Huang, L. (2013). Pollution and Potential Ecological Risk Assessment of Heavy Metals in a Lake. *Polish Journal of Environmental Studies*, 22(4).
- Oliveira, A., Palma, C., & Valença, M. C. de A. (2011). Heavy metal distribution in surface sediments from the continental shelf adjacent to Nazaré canyon. *Deep Sea Research Part II Topical Studies in Oceanography*, 58, 2420-2432.
- Ridgway, I. M., & N.B. PRICE. (1987). Geochemical associations and post-depositional mobility of heavy metals in coastal sediments: Loch Etive, Scotland. *Marine Chemistry*, 21, 229-248.
- Romankevich, E. A. (1984). Geochemistry of Organic Matter in the Ocean. *Clean Soil Air Water*, 13, 668.
- Sharifinia, M., Taherizadeh, M., Namin, J. I., & Kamrani, E. (2018). Ecological risk assessment of trace metals in the surface sediments of the Persian Gulf and Gulf of Oman: Evidence from subtropical estuaries of the Iranian coastal waters. *Chemosphere*, 191, 485-493. <https://doi.org/10.1016/j.chemosphere.2017.10.077>
- Sinex, S. A., & Helz, G. R. (1981). Regional geochemistry of trace elements in Chesapeake Bay sediments. *Environmental Geology*, 3(6), 315-323.

- Singovszka, E., Balintova, M., Demcak, S., & Pavlikova, P. (2017). Metal Pollution Indices of Bottom Sediment and Surface Water Affected by Acid Mine Drainage. *Metals*, 7(12), 284. <https://doi.org/10.3390/met7080284>
- Sutherland, R. A. (2000). Bed sediment-associated trace metals in an urban stream, Oahu, Hawaii. *Environmental geology*, 39(6), 611–627.
- Tijani, M. N., Jinno, K., & Yoshinari Hiroshiro. (2004). Environmental impact of heavy metals distribution in water and sediments of Ogunpa River, Ibadan area, southwestern Nigeria. *Journal of Mining and Geology*, 40, 73-83.
- Tomlinson, D. L., Wilson, J. G., Harris, C. R., & Jeffrey, D. W. (1980). Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. *Helgol-nder Meeresunters*, 33, 566-575.
- Turekian, K. K., & Wedepohl, K. H. (1961). Distribution of the Elements in Some Major Units of the Earth's Crust. *Geological Society of America Bulletin*, 72(2), 175. [https://doi.org/10.1130/0016-7606\(1961\)72\[175:DOTAIS\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1961)72[175:DOTAIS]2.0.CO;2)
- Uriah, L. A., & Shehu, U. (2014). Environmental risk assessment of heavy metals content of municipal solid waste used as organic fertilizer in vegetable gardens on the Jos Plateau, Nigeria. *American Journal of Environmental Protection*, 3(6-2), 1–13.
- USEPA. (2002). Onsite wastewater treatment systems manual. Office of Research and Development. US Environmental Protection Agency EPA/625/R-00/008.
- USEPA. (2014). Sediment Sampling United States Environmental Protection Agency, Washington, DC, USA. <https://www.epa.gov/sites/production/files/2015-06/documents/Sediment-Sampling.pdf> (Accessed 26, March 2017).
- Usero, J., Morillo, J., & Gracia, I. (2005). Heavy metal concentrations in molluscs from the Atlantic coast of southern Spain. *Chemosphere*, 59(8), 1175-1181. <https://doi.org/10.1016/j.chemosphere.2004.11.089>
- Vu, C. T., Lin, C., Shern, C.-C., Yeh, G., Le, V. G., & Tran, H. T. (2017). Contamination, ecological risk and source apportionment of heavy metals in sediments and water of a contaminated river in Taiwan. *Ecological Indicators*, 82, 32-42. <https://doi.org/10.1016/j.ecolind.2017.06.008>
- Wagne, M. M. (2013). *Contribution à l'étude de la qualité environnementale et sanitaire des eaux de la baie du Lévrier (Mauritanie)*. UNIVERSITE SIDI MOHAMED BEN ABDELLAH, Fès, Maroc.
- Wagne, M. M., BRAHIM, H. O., DARTIGE, A., & SÉFRIQUI, S. (2011). Contribution à l'étude du phytoplancton potentiellement nuisible de la baie du Lévrier (Mauritanie).
- Wedepohl, K. H. (1995). The composition of the continental crust. *Geochimica et cosmochimica Acta*, 59(7), 1217–1232.
- Yang, Q., Xu, Y., Liu, S., He, J., & Long, F. (2011). Concentration and potential health risk of heavy metals in market vegetables in Chongqing, China. *Ecotoxicology and Environmental Safety*, 74(6), 1664-1669. <https://doi.org/10.1016/j.ecoenv.2011.05.006>
- Yu, R., Hu, G., Lin, C., Yang, Q., Zhang, C., & Wang, X. (2017). Contamination of heavy metals and isotopic tracing of Pb in intertidal surface sediments of Jinjiang River Estuary, SE China. *Applied Geochemistry*, 83, 41-49. <https://doi.org/10.1016/j.apgeochem.2016.12.025>
- Zhang, G., Bai, J., Xiao, R., Zhao, Q., Jia, J., Cui, B., & Liu, X. (2017). Heavy metal fractions and ecological risk assessment in sediments from urban, rural and reclamation-affected rivers of the Pearl River Estuary, China. *Chemosphere*, 184, 278-288. <https://doi.org/10.1016/j.chemosphere.2017.05.155>
- Zoller, W. H., Hoffmann, G. L., & Duce, R. A. (1974). Atmospheric concentrations and sources of trace metals at the South Pole. *Science*, 183(4121), 198–200.

## **ROAD NETWORK BASED COMMUNITY DETECTION. CASE STUDY FOR AN EASTERN REGION OF AUSTRO-HUNGARIAN MONARCHY**

Zsolt *MAGYARI-SASKA*<sup>1</sup>

DOI: 10.21163/GT\_2019.141.06

**ABSTRACT.** The aim of the study was to find ways of measuring the relationship between clusters resulted on road networks topology and regional communities, before modern transport vehicles had widely spread. We have chosen as investigation area an eastern region of Austro-Hungarian Monarchy because at the beginning at the XX. century it had a modern transportation infrastructure for its time and a multicultural environment where different local or regional communities were present. Using data of that period we constructed a network model, attaching the populations' mother tongue as attribute data for settlements. After applying community detection algorithms considering only road network topology and later also the attribute data, we tried to compare the resulted clusters. We observed that in some cases mother tongue-based communities form sub-clusters for solely road network base clusters, while in other cases mother tongue-based communities restructure the road network based clusters. To quantify the differences, we identified three similarity/dissimilarity comparing methods, one based on clusters' spatial extends and two on statistical approaches.

**Key-words:** *Heterogeneous network, Modeling reality with graphs, Clustering, R, Spatial relation, Comparing communities*

### **1. INTRODUCTION**

In society from the very beginning, communities have appeared, that goes beyond the family, friends or acquaintances and even beyond local and settlement-level communities. These communities hold a local or regional identity that distinguishes them from the neighboring environment (Pohl, 2001).

The study of the positive and negative effects of these communities has inspired many studies (Semian & Chrmy, 2014; Chrmy & Janu, 2003), which all confirm the social importance of them. Most of these researches focus on the present, but at the same time also there also studies on past research and historical exploration (Baker & Biger, 1992).

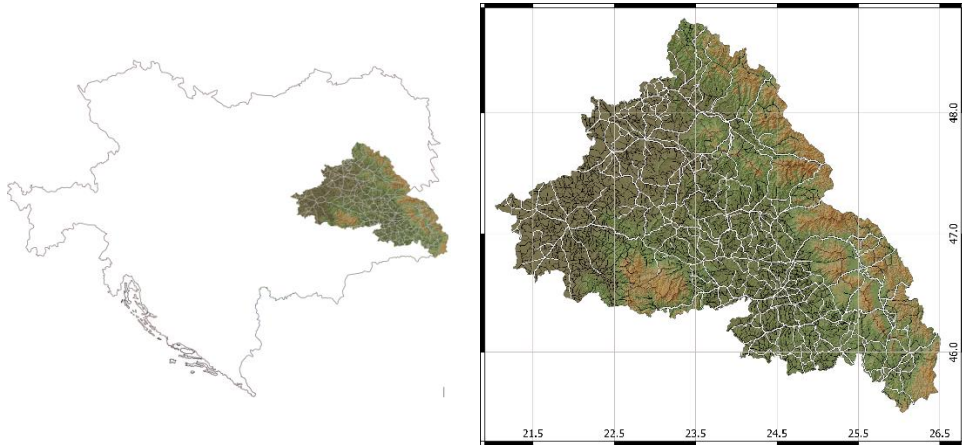
There are many types of communities which are based on different factors (McMillan & Chavis, 1986). This study attempts how can be explored the relationship between human communities and transport routes, and how these two concepts are related. As a test site an eastern region century of the Austro-Hungarian Monarchy (**Fig. 1**) at the beginning of the 19<sup>th</sup> century was selected, which had a relatively developed infrastructure at that time, and based on its geography and history had a multiethnic and multicultural society (Brie, 2014), which offered an opportunity for the appearance of several different regional communities.

Similar type of researches that explores spatial homogeneity and diversity and which studies populations clustering, is not new (McDoom & Gisselquist, 2015; York et al., 2011), but it is a constantly evolving trend where new methodological approaches have their place (Páez et al., 2012). The importance of networks is also highlighted in different studies as an important factor in analyzing social phenomena (Bobkova & Holesinska, 2017; Cadar et al., 2017).

---

<sup>1</sup> *Babes-Bolyai University, Faculty of Geography, 535500, Gheorgheni, Romania, zsmagyari@gmail.com*





**Fig. 1.** The study area within the Austro-Hungarian Monarchy (left) and its road network (right)

This research is made up of several interconnected processes. The basic tasks that model these processes are in fact answers to the following questions: how a weighted network model based on historical maps can be created; how the homogeneity of two connected settlements based on mother tongue can be expressed and measured; how the similarity of two different cluster map can be evaluated. The backbone of this research is answering these questions, projecting the results into the target area.

## 2. BUILDING THE DATA MODEL

### 2.1. Data sources and their limitations

Our first task was to create a network topology where nodes represent the settlements and the edges the existing directly connecting roads between them. A directly connecting road should not pass through other settlements but can go through multiple crossroads.

The input data we had was the point layer with settlements' position, and the line layer containing the road network, based on ten pieces of late XIX, early XX century Austro-Hungarian Empire map sheets. Their scale varied between 250.000 to 400.000. Another, but non-spatial data source was a tabular record of the 1910 census including the mother tongue of the population.

To build the desired network we had to properly snap the settlements on the road network. This step theoretically could have been done in the time of vectorization, we intentionally don't want to do this for two reasons. Firstly, not all settlements on the map are placed near roads, in some cases, there are several kilometers between the settlement and the nearest road. Secondly, the basic idea of the vectorization was to preserve as much as possible the original map layout and content and a forced snapping would have altered this principle. That's why we had to find post-processing techniques to achieve our goal.

Latter automatic snapping was not a possibility as snapping operation considers just vertices which is not necessarily the best solution as it disregards the minimum distance between the settlement point and the road. In addition, the automatic snapping would have neglected the distance between the settlement and the road (sometimes considerable).

Even in the first steps, it was easy to identify that in first phases we can build a heterogeneous network, which means that not all nodes have the same role. Some of them

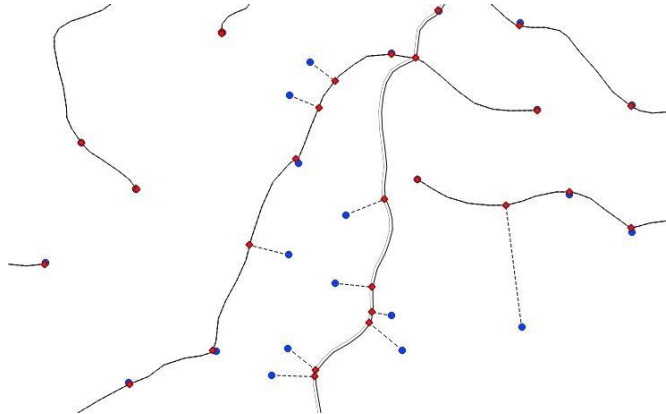
are crossroads with routing possibility while others representing the settlements which are the true data holder components of the network.

To build a usable network model, our processing algorithm used QGIS to create the heterogeneous network, Gephi for basic visualization and to eliminate non-connected components and R for homogenous network building, in which all nodes will represent settlements, maintaining the effective distance between them.

## 2.2. Building the heterogeneous network model

After the points representing the settlements were vectorized, the census data had to be geocoded based on the settlements name as no other uniquely identifying element exists. The used census data was available in electronic format for the study area from Varga E. Árpád database (<http://varga.adatbank.transindex.ro>).

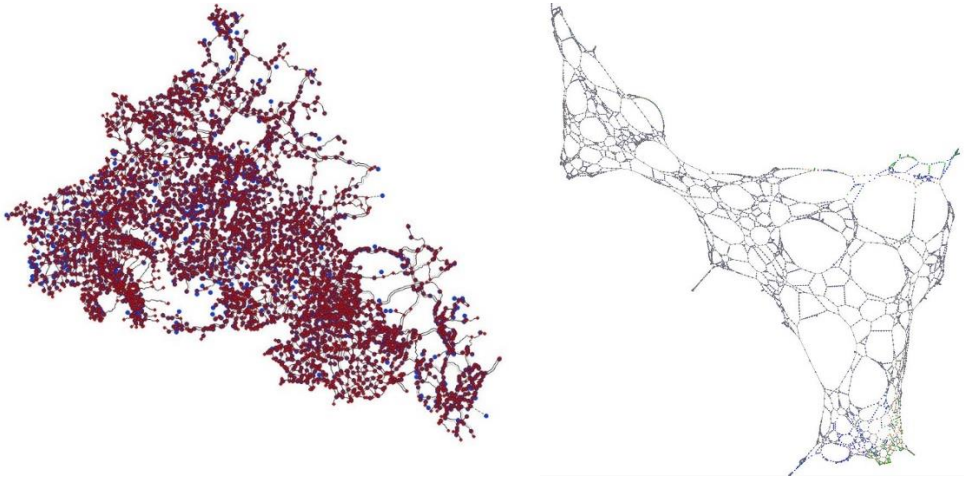
One of the first problems we had to overcome was that the road network was composed by arbitrary road sections, the junctions were in the middle of road segments disregarding any relationship between road network and settlements position. We had to transform in such a way that every road section to end at a crossroad. For this, the *Split with lines* command was used.



**Fig. 2.** Road network model in QGIS (red marks – junctions/crossroads, blue marks - settlements, dash lines – added segments to road network)

As mentioned the settlements position initially was not snapped to roads or crossroads. That means that every settlement had to be connected with a separate line to the road network. In QGIS the *Connect nodes to lines* from *Network* packages could do this. For every settlement a new road section is added to the nearest road, splitting the initial road segment into two and creating a junction (**Fig. 2**). After this operation, the *Network* package in QGIS is capable to build the node and edge tables representing the graph. The edges endpoints are the ending vertices of road sections. In such way, we had a network in which some of the vertices marks junctions or crossroads, while others vertices mark settlements, both without any data. The last step was to add the geolocated census data to the proper nodes with *Join attributes by location* command. The length of the road segments was also calculated and added to edges.

The resulted two layers (nodes and edges) were imported in Gephi (**Fig. 3**) where a component analysis was performed, eliminating 4 small parts of the graph which were not connected (as not the whole monarchy's map was used) to the main core. After this operation, we had 10266 edges and 9157 nodes, 3211 of them representing settlements.



**Fig. 3.** Settlements connectivity GIS data model (left) / network data model (right)

### 2.3. Building the homogeneous network model

The homogeneous network model, which should connect only settlements, if between them exists a direct road was built based on the following algorithm implemented in R.

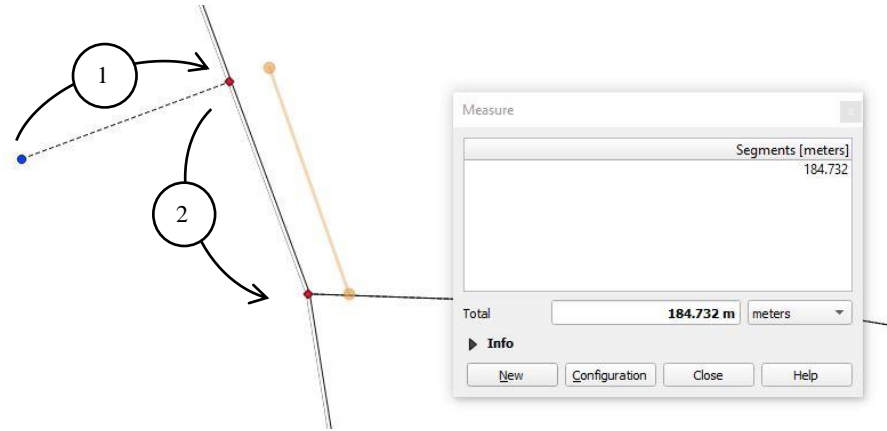
**Deleting of some singular nodes.** As not all nodes represent settlements, those which has only one connection and are not settlements could not participate in any routing sequence, they should be eliminated.

**Inserting the settlement nodes on the road network.** In fact, this operation is composed of three steps: 1) adding the edge's length which connects the settlement node and a junction node to both other edge of the junction node, 2) copying the settlement data to the junction node and 3) deleting the original settlement node.

**Putting some nodes in crossroads.** Now all settlement nodes are on roads but in many cases, the settlements are not positioned in crossroads but are very close to them. This situation causes incorrect routing data between settlements as it seems that one of them is connected directly to much more than in reality is. In this step, we overcame this situation by "rearranging" the characteristics of the network if a settlement node is closer to a junction node than a predefined distance. In such case, the settlement node is "moved" to the junction node adjusting the length of connecting edges, in such way that all implied edges to reflect the real length (**Fig. 4**).

The result so far still contains both types of nodes, therefore a new network should be built containing just the settlement nodes and the connecting weighted direct edges. In this step, the Dijkstra algorithm, implemented in R was used to find the shortest paths between settlement nodes.

In the end, we obtained a connected graph with 3211 nodes and 5279 edges were all nodes represents settlements having as attributes regarding the number of inhabitants with different mother tongues. An edge between two settlements represents the possibility to travel from one to other without crossing other settlements. Edges are weighted by travel distance.



**Fig. 4.** Settlement repositioning to a nearby crossroad in the network model  
(1 – inserting settlement on the road network; 2 – moving settlement to crossroad)

### 3. CLUSTER ANALYSIS – METHODS AND RESULTS

Cluster analysis refers to object grouping based on certain criteria. In connectivity models, represented by graph or networks, when the linkage between objects exists the clustering process is equivalent to community detection. This process is based on the similarity between nodes and can take account the topological characteristics of the network or beyond it also other attributes of the network components. In the first case, the most known similarity values are the common neighbors, cosine, Jaccard or min indexes (Fu et al., 2017) although particular indexes could be defined (Cheng et al., 2013).

In our case, we were interested to include in community detection values, related both to vertices having the mother tongue distribution attribute as for edges characterized by their spatial distance. There are several studies which includes such approaches, defining similarity, based on edge attributes (Steinhauser & Chawla, 2008), based on vertex attributes (Dang & Viennet, 2012; Boobalana et al., 2016) or even based on both types of attributes (Zhou et al., 2009; Chakrabarty et al., 2016).

The similarity value between two nodes (settlements) in this research was based on the Herfindahl concentration formulae which is often used in ethnic fractionalization analysis (Posner, D.N., 2004; Bossert et al., 2011; Fearon, D.J., 2013). In our case the original formula, which was developed to be a measure of fractionalization for each given location, was transformed in such a way to consider the sum of share differences of each mother tongue group for an edge connected settlement pair, as shown below:

$$edge_{AB} = 1 - \sum_{i=1}^k (A_i - B_i)^2 \quad (1)$$

where,

edge<sub>AB</sub> – edge attribute, based on mother tongue similarity

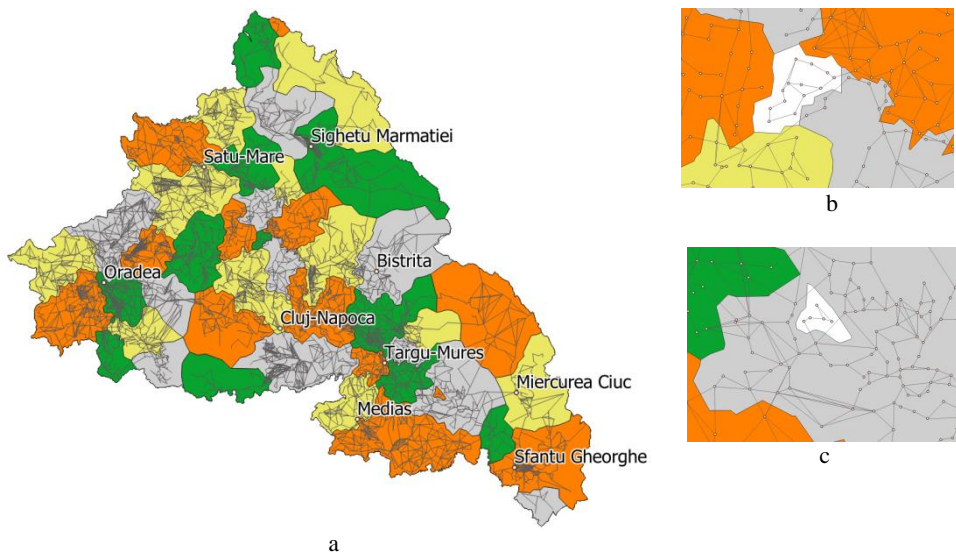
k – the number of different mother tongue categories

A<sub>i</sub>, B<sub>i</sub> – share of *i*. mother tongue in each settlement

It's easy to observe that in case of exact matches between mother tongue shares the edge attribute value will be 1, and as much as there are higher differences between shares

the value decreases to 0. After the similarity was defined, the cluster analysis had to come. There are many methods for community detection which differ in direction, complexity, resulted cluster shape (Neethu & Surendran, 2013; Bindiya et al. 2014). From the implemented clustering algorithms in igraph package of R we used the FastGreedy method. This algorithm uses a bottom-up hierarchical approach, merging smaller communities to maximize the modularity score, which is a measure of internal strength of a network component. When the modularity score could not become higher the algorithm stops.

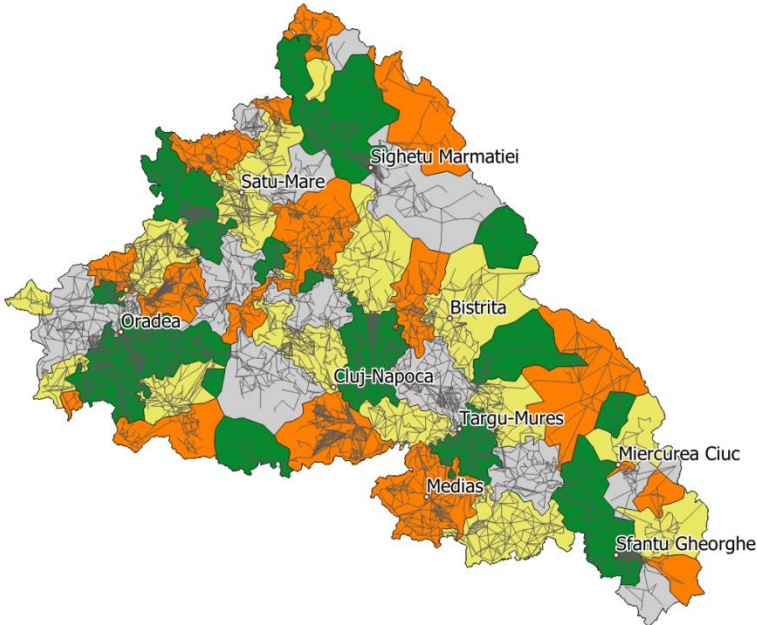
In our research, we started the cluster analysis on the created road network using different edge weight. In the first case we don't consider the mother tongue attribute, leaving the community detection algorithm to take account just the length of connecting roads. In this case, we obtained 46 structural communities as shown in **Fig. 5a**.



**Fig. 5.** Clusters resulted based on topological relation (a) with examples of isolated clusters (b,c) (lines represent the network model, colors has no special meaning just facilitate interpretation)

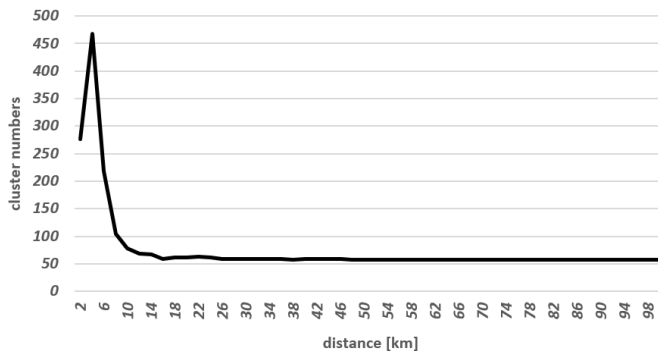
The cluster analysis was set up on a settlement network, each settlement receiving an id value corresponding to the cluster number to belongs to. The spatial extension of the clusters, represented as polygons were create with as Voronoi polygons, later merged by their cluster id attribute field. Due to this method, the area of clusters as it's shown on the map does not reflect the size of the cluster. Nonetheless, the spatial extent visualization helps to identify the individual cluster covered area and patterns of isolated clusters as presented in **Fig. 5b, 5c**. In the next phase of the analysis, we considered as edge weight only the similarity based on mother tongue as expressed in formula 1. In this case, we obtained 58 communities, with 26% more than the cluster number obtained at the previous step (**Fig. 6**).

The growth of cluster numbers theoretically suggests the division of initial clusters but that's not true for the whole study area. In the southeast region, we can observe sub-clusters of the road network based clusters, while in the west part we can observe clusters which aggregate partially or totally some road network based clusters.



**Fig. 6.** Clusters resulted based on mother tongue similarity  
(lines represent the network model, colors has no special meaning just facilitate interpretation)

Furthermore, we wanted to analyze if this cluster number and configuration changes dynamically if we try to limit the search radius at community detection, using a minor change in formula 1. We have repeated the cluster analysis based on mother tongue similarity gradually increasing the search distance from 2km to 100km. The evolution of cluster numbers stabilizes at 48km. The number and spatial extent of these stabilized clusters are exactly the same as those obtained when no distance limit was imposed. Beyond the 48km value, no changes appear in case of mother tongue-based community detection. We were also interested in the first distance value from which the final clustering result starts to appear. We considered a 10% limit against the 58 value representing the cluster numbers. In this case, starting from 16km the resulted clusters tend to be identical with the final ones (**Fig. 7**).



**Fig. 7.** The number of clusters evolution by considered search distance  
(cluster analysis based on road network topology and mother tongue similarity)

To measure the similarity of the resulted clusters we used three approaches: based on spatial extension, based on cluster size (settlement number) and based on mother tongue similarity. For the first one, based on the spatial extension we developed a comparison algorithm between two polygon layers, measuring the fragmentation with the following steps:

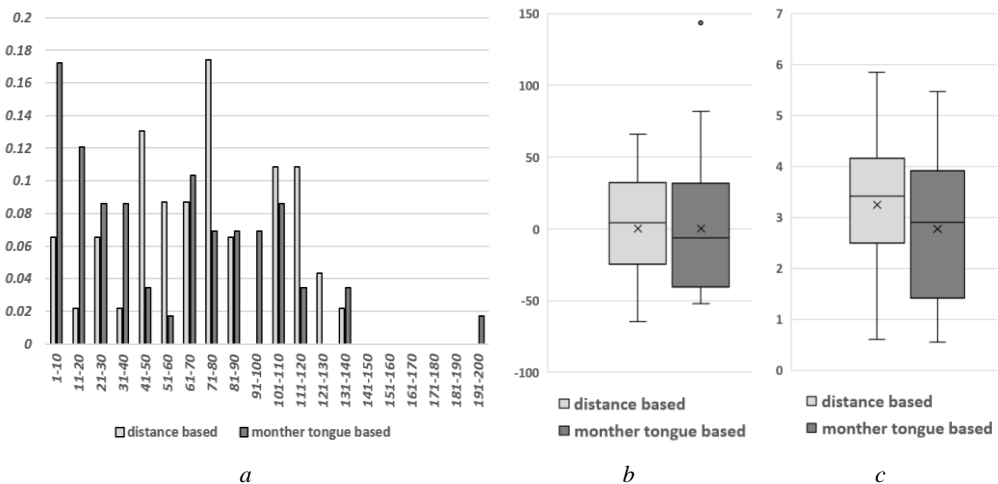
- one of the layers is considered the base layer, to which we will refer to. The other layer is considered the fragmenting layer.
- the fragmenting polygon layer is transformed to cutlines and it will divide the base layer in multiple polygons
- for each cluster value in the base layer, the ratio between the sum of the squared area with the same cluster value and the base layers' squared area is calculated as shown in formula 2, obtaining a value between 0 and 1. As higher the calculated value is there's a higher similarity.

$$cluster\ similarity_c = \frac{\sum_{i=1}^{n_c} A_i^2}{A_c^2} \tag{2}$$

where,

- c – cluster number for which the similarity is calculated
- n<sub>c</sub> – the number of polygons in the fragmented layer having c as cluster identifier
- A<sub>i</sub> – area on the fragmented layer
- A<sub>c</sub> – area on the base layer

By applying the mentioned algorithm, we have got 46 values, one for each road based cluster. The mean cluster similarity value of the road based clusters by the mother tongue base clusters was 0.6 with a standard deviation of 0.24. There were 6 clusters, 13% of the total, which had the same boundaries in both clusterings.



**Fig. 8.** Frequency distribution of cluster size (a), the statistical summary of the two clusters data series based on cluster size (b) and mother tongue relative standard deviation (c)

Secondly, we compared the resulted clusters' size. As the number of clusters differs in the two analyzed situations we do not compare the two data series directly, but their standardized values. The standardized values were calculated as differences between each value of a data series and its theoretical average size (settlement number divided by the

number of clusters). In this case even if in the percentile frequency distribution of cluster sizes there are differences, the two series statistical characteristics are similar (**Fig. 8b**). It's notable the presence of a huge cluster with 199 settlements in case of mother tongue-based community detection (**Fig 8a**), which has appeared by aggregating several road based cluster parts near Oradea. The last comparison method was to analyze the homogeneity of the resulted clusters after both clustering methods. For this we used the relative standard deviation (standard deviation divided by the mean) for all mother tongue classes, calculating its mean for each cluster. The statistical characteristics of the resulted two data series are similar with a notable lower mean value, representing a more homogeneous cluster, for mother tongue-based clustering (**Fig. 8c**), as it was expected.

#### 4. CONCLUSIONS

Our research followed two main directions: developing a study methodology for comparing two types of clustering results and applying it to a study region. Firstly, we presented an automated way to create a homogenous network, based on routes and settlements even if the two type of objects were vectorized independently with no initial spatial connection between them. The only parameter of the algorithm is the distance threshold for snapping a settlement to a crossroad. With all transformation, the algorithm maintains the real distance between the settlements. Another result of the research was a cluster analysis based on edge attached similarity values, which at data level was based on the share of a whole. For comparing the results of two clustering operations we presented three different approaches, each of them having a specific role in data evaluation.

Regarding the selected study area, the obtained cluster configuration becomes constant starting from the 48km distance limit in cluster analysis. This value can be considered close to the maximum distance that can be covered at that time without motorization. Comparing the clusters obtained with and without considering the mother tongue-based similarity between settlements the spatial similarity value indicates that in approx. 60% of the study area belongs to the same community in both clustering method. Both statistical comparison of the resulted clusters also indicates a notable similarity, especially regarding cluster size.

Considering the results of the three comparison types, we can affirm that road network topology was seriously related but not determinative in forming local or regional communities based on mother tongue till the I World War at the study area. Our study reinforces that human communities are formed considering the transportation network. However, the presented methodology had to be tested also in other location comparing the obtained results.

#### ACKNOWLEDGEMENT

The research was supported by the DOMUS scholarship program of the Hungarian Academy of Sciences.

#### REFERENCES

- Baker, A.R.H. & Biger, G. (1992) *Ideology and Landscape in Historical Perspective: Essays on the Meaning of Some Places in the Past*. Cambridge University Press, Cambridge.
- Bossert W., D'Ambrosio C. & La Ferrara E. (2011), A generalized index of fractionalization, *Economica*, 78, 723-750
- Bobkova, M. & Holesinska, A. (2017) Networking in a destination from the perspective of virtual relationships and their spatial dimension, *Geographia Technica*, 12(2), 10-19



- Brie M. (2014), Ethnicity and politics in the Romanian space. The case of northwestern Transylvania. Published in: No. Sorin Şipoş, Gabriel Moisa, Dan Octavian Cepraga, Mircea Brie, Teodor Mateoc (coord.), *From Periphery to Centre. The Image of Europe at the Eastern Border of Europe*, Editura Academia Română. Centrul de Studii Transilvane, Cluj-Napoca, 158-170.
- Cadar, R.D., Boitor, M.R. & Dumitrescu, M. (2017) Effects of the traffic volumes on accidents: the case of Romania's national roads, *Geographia Technica*, 12(2), 20-29
- Chakrabarty, A., Chelladurai, J. & Venkateswaran, S.K. (2016) Community detection in citation networks using attribute similarities, *International Journal of advances in cloud computing and computer sciences*, 2(5), 1-5
- Cheng, J., Xu, H., Gaybullaev, M., Leng, M. & Chen, X. (2013) Community Detection Algorithm based on Neighbor Similarity, *Telkommika*, 11(8), 4484-4490
- Chromý, P. & Janů, H. (2003) Regional identity, activation of territorial communities and the potential of the development of peripheral regions. *Acta Universitatis Carolinae – Geographica* 38, 105-117.
- Dang, T.A. & Viennet, E. (2012), Community Detection based on Structural and Attribute Similarities, *The Sixth International Conference on Digital Society*, 7-12
- Fearin D.J. (2013) Ethnic and cultural diversity by country, *Journal of Economic Growth*, 8(2), 195-222
- Fu, Y.H., Huang, C.Y. & Sun, C.T. (2017) A community detection algorithm using network topologies and rule-based hierarchical arc-merging strategies, *PLoS ONE* 12(11), 1-30
- McDoom O.S. & Gisselquist R.M. (2015) The Measurement of Ethnic and Religious Divisions: Spatial, Temporal, and Categorical Dimensions with Evidence from Mindanao, the Philippines., *Social Indicators Research*, 129(2), 863-891
- McMillan, D.W. & Chavis, D.M. (1986) Sense of Community, *Journal of Community Psychology*, 14, 6-23
- Neethu C.V. & Surendran, S. (2013) Review of Spatial Clustering Methods, *International Journal of Information Technology Infrastructure*, 2(3), 15-24
- Páez, A., Ruiz, M., López, F. & Logan, J. (2012) Measuring Ethnic Clustering and Exposure with the Q statistic: An Exploratory Analysis of Irish, Germans, and Yankees in 1880 Newark, *Annals of the Association of American Geographers. Association of American Geographers*, 102(1), 84-102.
- Pohl, J. (2001) Regional Identity, *International Encyclopedia of the Social & Behavioral Sciences*, 12917-12922
- Posner D.N., (2014) Measuring ethnic fractionalization in Africa, *American Journal of Political Science*, 48(4), 849-863
- Semian, M. & Chromý, P. (2014) Regional identity as a driver or a barrier in the process of regional development: A comparison of selected European experience, *Norsk Geografisk Tidsskrift - Norwegian Journal of Geography*, 68(5), 263-270
- Steinhaeuser, K. & Chawla, N.V. (2008) Community Detection in a Large Real-World Social Network, In: Liu H., Salerno J.J., Young M.J. (eds) *Social Computing, Behavioral Modeling and Prediction*, Springer, 168-175
- Varga E. Á. (2010) Ethnic and denominational statistics of Transylvania (1850-1992), in *hungarian* (Erdélyi etnikai és felekezeti statisztikája (1850-1992)). Erdélyi Magyar Adatbank. <http://varga.adatbank.transindex.ro>
- Varghese, B., Unnikrishnan, A. & Jacob, P.K. (2013) Spatial Clustering Algorithms- An Overview, *Asian Journal of Computer Science and Information Technology*, 3(1), 1-8
- York, A. M., Smith, M. E., Stanley, B. W., Stark, B. L., Novic, J., Harlan, S. L., Cowgill G.L. & Boone, C. G. (2011) Ethnic and Class Clustering through the Ages: A Transdisciplinary Approach to Urban Neighbourhood Social Patterns. *Urban Studies*, 48(11), 2399–2415
- Zhou, Y., Cheng, H., Yu, J.X. (2009) Graph Clustering Based on Structural/Attribute Similarities, *Proceedings of the VLDB Endowment*, 2(1), 718-729

## **REGIONAL DISPARITIES IN HUNGARIAN URBAN ENERGY CONSUMPTION – A LINK BETWEEN SMART CITIES AND SUCCESSFUL CITIES**

**Zoltán NAGY<sup>1</sup>, Tekla SEBESTYÉN SZÉP<sup>1</sup>, Dóra SZENDI<sup>1</sup>**

DOI: 10.21163/GT\_2019. 141.07

### **ABSTRACT:**

Cities account for 60–80% of global energy consumption, and based on projections the development of urban areas will be the main engine of energy use growth in the future. While it may seem that this topic plays only a marginal role in urban research, in energy economics more and more studies are focusing on the concept of the smart energy city and resilient city related to energy use as a possible way toward sustainability and human well-being. Our main objective is to examine the dimension of smart environment through residential energy use. We focus on the regional disparities of urban energy use (electricity use and natural gas consumption) in Hungary. The analysis covers 23 Hungarian cities and Budapest during the period from 2010 to 2015. The Theil Index and the area-based Gini index are calculated. We conclude that on the whole no significant inequalities or spatial differences were identified among the cities. The Theil Index components (within-group inequality component and between-group inequality component) draw attention to the within-group differences related to natural gas consumption. These disparities are more decisive than values of the between-group inequality components. It cannot be stated that belonging to the “elite” groups of cities causes significant changes in the urban electricity and natural gas consumption patterns.

*Key-words: AR-Gini, Theil index, Residential energy use, Disparities.*

### **1. INTRODUCTION**

Globally circa 210 million people without access to electricity live in urban areas and furthermore, approximately 500 million people still lack access to clean cooking facilities World Bank (2018). As Szlávik (2013) emphasizes, we cannot wait for an enlightened world government; a sustainable social, economic and political system can be achieved only at the local level with sustainable projects. Local initiatives and corporate social responsibility are highly important. There are no breakthroughs leading to sustainability (which is so popular nowadays), only small movements and shifts. There are economic decisions and objectives that can be achieved under the current circumstances, make a profit, improve well-being and serve the aims of sustainability as well. Actually, sustainable development can be interpreted as the long-term strategy of humanity. It exists not only globally but also on regional and local levels. With an increasing share of renewable energy sources and with local solutions to global problems, the role of decentralization, local approach to energy policy and local economic and community development are being reevaluated. The smart city concept can be the basis for these comprehensive social, economic and technological changes. This definition does not refer to a completed

---

<sup>1</sup> *Institute of World and Regional Economics, Faculty of Economics, University of Miskolc 3515 Miskolc-Egyetemváros, Hungary, nagy.zoltan@uni-miskolc.hu, regtekla@uni-miskolc.hu, regszdor@uni-miskolc.hu*

state/status, but to an operation logic and continuous development (Kulcsár and Szemerey 2016; Sáfián and Munkácsy 2015).

## 2. THEORETICAL BACKGROUND

The smart city concept is not novel in the economic literature. It first emerged in the 1990s (although the intensity of related research increased significantly after 2009, according to Jong et al. 2015) with regard to the sustainable development of urban areas and settlements and the reforms of the urban management systems (Kulcsár and Szemerey 2016). Hollands (2008) mentions San Diego, San Francisco, Amsterdam and Kyoto, as the first users of smart applications and technologies. However, Manchester, Southampton and Vancouver are presented for their best practices and good initiatives. As Egedy (2017) concludes, the three main pillars (or dimensions) of the smart city concept are sustainability, efficiency and wide participation. There are many other kinds of categories and definitions related to the smart city concept: eco city, sustainable city, low carbon city, knowledge city, intelligent city, digital city, resilient city, ubiquitous city, green city, information city, liveable city, hybrid city, creative city, humane city, learning city, wired city.

The borders often blur and there are significant overlaps; a detailed overview is given by Jong et al. (2015). The definition of smart city can be interpreted from two perspectives: there is a technical interpretation, which emphasizes the physical implementation of smart innovations and see the future in urban planning and decision-making based on algorithms (Baji 2017). From a social viewpoint, the main goal of these innovative solutions and improvements is a more democratic society and the active participation of citizens in community decisions that contribute to sustainable development and a higher quality of life (Jong et al. 2015).

The environmental dimension supplements this and it focuses on the “smart and green” technologies (Baji 2017). It targets the ecological improvement of urban areas. Important fields are urban water management, lighting, waste management, management of natural resources and energy management (these are relevant fields of environmental sustainability as well). It is worth examining how the performance of a selected city (related to these indicators) is compared to other urban areas. Are there social or spatial reasons for the differences? Can these differences be explained by the attitudes of the residents? We can get a clear picture about the environmental consciousness of residents then the intervention points become identifiable (Baji 2017). A well-functioning smart city may contribute to improving living standards, increasing urban competitiveness and overcoming obstacles such as poverty, social exclusion or environmental problems.

Our current study focuses on the environmental dimension of smart cities as a critical field of the smart city concept. Investigating the smart solutions that are running and under implementation, it can be stated that application of environmental solutions plays a significant role.

From our perspective efficient energy use is the most important. Most of the concepts emphasize it, such as Nam and Pardo (2011), Ladós (2011), Cohen (2014), Stankovic et al. (2017) and the ISO 37120 standard as well. Energy efficiency improvements can bear many positive external effects, such as direct effects (increasing the value of real estate, enhancing quality of life, intensifying tourism, development of the local smart business environment). Based on Giffinger (2015) the development of a smart environment can also

lead (directly or indirectly) to negative impacts, and this phenomenon is called the rebound effect. This suggests that the energy use of smart cities should be examined in more depth.

In this study a situation report is carried out that provides an energy efficiency overview for Hungarian cities (with county rights) to map the starting solution. We provide a review of the “elite category” of Hungarian medium-sized cities (based on Rechnitzer et al. 2014). In our view the most realistic chance of introducing smart apps and creating a smart city concept can be in this category (in some of these cities these initiatives have already started).

The research questions are: 1) Could be significant differences detected in the residential energy use of the selected cities (regarding the subcategory of “elite” cities)? 2) How these differences changed between 2010 and 2015?

### 3. DATA AND METHODOLOGY

In this study the regional disparities and the spatial distribution of Hungarian urban energy use (electricity use and natural gas consumption) are examined. The analysis covers 23 Hungarian towns with county rights and Budapest during the period of 2010-2015. The 23 towns are the following: Békéscsaba, Debrecen, Dunaújváros, Eger, Érd, Győr, Hódmezővásárhely, Kaposvár, Kecskemét, Miskolc, Nagykanizsa, Nyíregyháza, Pécs, Salgótarján, Sopron, Szeged, Székesfehérvár, Szekszárd, Szolnok, Szombathely, Tatabánya, Veszprém, Zalaegerszeg.

Annual data as listed below are applied in the calculations collected from the Hungarian Central Statistical Office (KSH): gross income (local currency unit LCU); resident population at the end of the year (data calculated further from finalised data of the population census); number of household electricity consumers; volume of electricity supplied to households (thousand kWh); total volume of electricity supplied (thousand kWh); total volume of piped gas supplied (not recalculated) (thousand m<sup>3</sup>); of total volume of gas supplied, volume of gas supplied to households (not recalculated) (thousand m<sup>3</sup>); of household gas consumers, number of those using gas for heating.

Based on these data we created the following indicators: residential gas consumption per household (m<sup>3</sup>), residential electricity consumption per household (kWh), natural gas consumption per capita (m<sup>3</sup>), electricity consumption per capita (kWh), income per capita (HUF). Application of these indicators enable us to compare the selected cities with different economic structure.

#### 3.1. Theil index

Conceição and Ferreira (2000) argue that inequality can be measured not only in the case of individuals (total inequality), but in the case of groups (between-group inequality) as well. Grouping individuals can be calculated on an area basis, or sex, qualifications, rural or urban population or even income deciles can serve as grouping criteria. If the groups are formed on a geographical basis, we intend to investigate and analyse the spatial difference of a selected indicator. The Theil index has become a very frequently used tool for regional differences (it used in e.g. Alcantara and Duro 2004, Zhang L. 2011).

The Theil index – similar to decomposition methods – consists of two components, i.e., a within-group inequality component and a between-group inequality component. Actually, the selected component shows – assuming the constancy/unchangingness of other indicators – the impact of the specific factor on the dependent variable. Based on Zhang et al. (2011) the formula of the Theil index is:

$$T(I) = \sum_i y_i \ln\left(\frac{\bar{I}}{I_i}\right)$$

where  $y_i$  is the gross income-share for city  $i$  for a given year;  $\bar{I}$  is average value of the measured (specific) energy data (considering the examined towns);  $I_i$  denotes the concerned indicator for city  $i$ . Similar to Zhang et al. (2011), in this study specific data (not total or absolute, but per capita or per household) are involved.

$$T(I) = T_B(I) + T_W(I)$$

where  $T_B(I)$  is the aggregate between-group variance component;  $T_W(I)$  is the aggregate within-group component.

$$T_B(I) = \sum_g y_g * \ln\left(\frac{\bar{I}}{\bar{I}_g}\right)$$

where  $y_g$  is the gross income share of group  $g$ ;  $\bar{I}_g$  is the average of the cities in group  $g$  (related to the selected specific energy data).

$$T_W(I) = \sum_g \sum_i y_g * y_{i,g} * \ln\left(\frac{\bar{I}_g}{I_{i,g}}\right)$$

where  $y_{i,g}$  is the gross income share associated with city  $i$  in group  $g$ ;  $I_{i,g}$  denotes the concerned indicator for city  $i$  in group  $g$ .

In the index zero indicates no inequality, while higher values of the index indicate greater disparity (the maximum value of the index is 1, indicating complete inequality).

Classification of the Hungarian towns with county rights has a significant literature (such as Lengyel 1999, Beluszky and Győri 2004). Here we cite Rechnitzer et al. (2014). The researchers, investigating the innovation clusters, conclude that in Hungary 17 towns (from the group of 23 Hungarian towns with county rights) can be grouped into an “elite” category (these are Pécs, Sopron, Miskolc, Szeged, Szolnok, Szekszárd, Kaposvár, Kecskemét, Szombathely, Debrecen, Eger, Zalaegerszeg, Veszprém, Nyíregyháza, Győr, Dunaújváros and Székesfehérvár), while 6 towns are not part of the group (Békéscsaba, Érd, Hódmezővásárhely, Nagykanizsa, Salgótarján and Tatabánya). The so-called “elite” category contains towns with outstanding features from the point of view of innovation and human development. These settlements have complex economic structures and are regional centres with great economic potential or with an economy that is highly oriented towards research, development and higher education (Rechnitzer et al. 2014). Calculation of the components for the Theil index is based on this classification.

### 3.2. AR-Gini index

The Gini coefficient is derived from the Lorenz curve. The Lorenz curve is especially suitable for the graphical representation of social inequalities, and has become a popular tool to illustrate not only income and the expenditure-related inequalities, but spatial variations as well (such as Dollman et al. 2015, Finn et al. 2009, Steinberger et al. 2010). The latter ‘shows the share of spending (or income) by households ranked by spending (or income). The further the curve is below the 45 degree line, the less equal the distribution. Correspondingly, the Gini coefficient is calculated as the area between the Lorenz curve

and the 45 degree line divided by the total area under the 45 degree line.’ (Dollman et al. 2015) The higher the coefficient, the more unequal the distribution is.

Application of the area-based Gini-index (called the AR-Gini index) can contribute to revealing the main reasons of the spatial inequalities. It is used in many research fields and is a frequently used tool for the examination of energy sources, energy use, different types of environmental damage (such as acidity, depletion of the ozone layer, emission, eutrophication, climate change), and for life-cycle- based (from cradle to grave) analysis of material flows (such as Steinberger et al. 2010).

The AR-Gini differs from the conventional Gini coefficient in two points (**Table 1**). First, the measure of disparity in terms of resource use is carried out using any physical unit (it is not monetarily expressed), second it is calculated on an area basis (not at the level of households or individuals) (Druckman and Jackson 2008).

**Table 1**

**Comparison of AR-Gini and the original Gini index**

	<b>AR-Gini</b>	<b>Gini</b>
<b>calculation basis</b>	calculated on an area basis	calculated on a per capita or household basis
<b>object of calculation</b>	calculated on a resource basis	income, wealth, expenditures (calculated on a monetary basis)

Source: own compilation based on Druckman and Jackson (2008)

Calculation of the Gini coefficient for income is carried out using the following equation (Druckman and Jackson 2008):

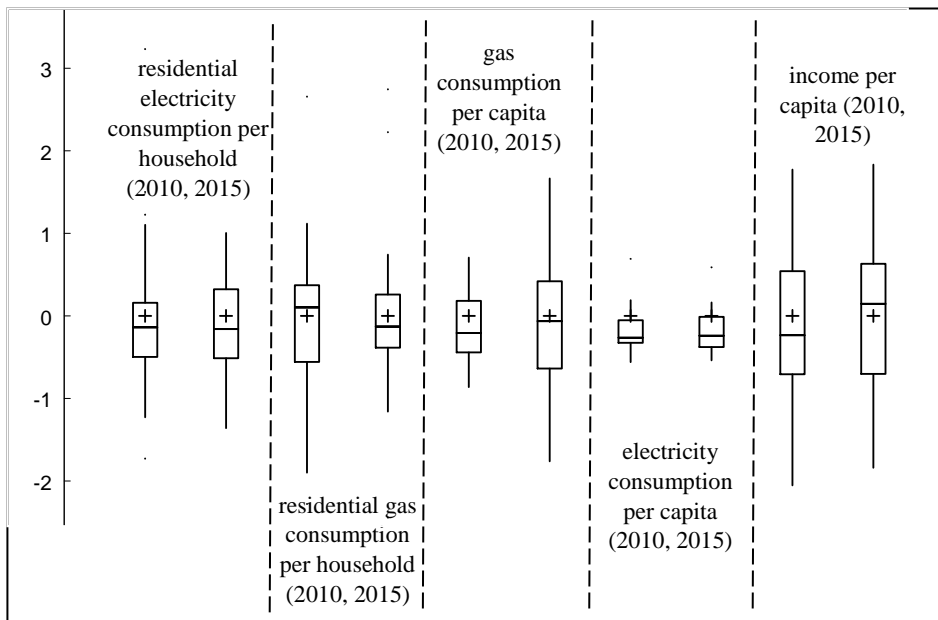
$$G = \frac{1}{2n^2\eta} \sum_{i=1}^n \sum_{j=1}^n |y^i - y^j|$$

where  $y^i$  and  $y^j$  are the incomes of the  $i$ th and  $j$ th household,  $\eta$  is the average income and  $n$  is the number of households. Adapting this formula to the calculation of AR-Gini, the explanation of the equation changes as well:  $y^i$  and  $y^j$  denote the average resource use in  $i$ th and  $j$ th area (in this study the resource use means the electricity consumption or natural gas use),  $\eta$  is the average resource use of each area,  $n$  is the number of output selected areas.

#### 4. RESULTS

As discussed above, no significant differences between the rural and urban (23 Hungarian towns with county rights and Budapest) energy use are found. According to the KSH (2018) database, while in 2015 37.5% of the Hungarian population lived in one of the 23 cities with county rights or in Budapest, 37.3% of the volume of electricity supplied to households and 40.4% of the volume of gas supplied to households were concentrated here. However, the question arises as to what differences can be observed among the energy use of the examined cities and whether there is any connection between their success (meaning belonging to the elite category based on Rechnitzer et al. 2014) and their energy consumption patterns. Or from the other side, whether the achieved level of energy efficiency and the decreasing energy use can contribute to urban development or to success. Does the society of the more developed or more successful cities consume (natural) resources more efficiently and more consciously?

To visualize the differences among the selected cities (regarding the selected variables) the Standard Deviation Method is applied and the distribution of data was displayed with a box plot chart (**Fig. 1**). Here the outliers are filtered out and the results slightly differ from the Gini coefficient values (**Table 2**). The latter highlights that the social disparities are the lowest among the selected cities in case of income per capita. The inequalities significantly decreased in the case of natural gas consumption (per household and per capita). The social differences regarding residential electricity consumption per household have hardly changed between 2010 and 2015. However, in the case of electricity consumption per capita a slight increase in Gini coefficient can be observed: the value of it is 0.24 in 2010 and 0.28 in 2015 (but this does not refer to an extremely high inequality).



**Fig. 1.** Box plot of selected variables  
Source: own compilation based on KSH (2019).

These trends are almost certainly related to the fact that between 2010 and 2012 a significant part of the society (the 1<sup>st</sup>-3<sup>rd</sup>-4<sup>th</sup>-5<sup>th</sup>-6<sup>th</sup>-8<sup>th</sup>-9<sup>th</sup> income deciles) decreased their energy expenditures. These households restrained their consumption and many switched energy sources, at least partly, from natural gas to cheaper sources (typically the expenditures on solid fuels, especially on wood, increased). In subsequent years the positive effects of energy efficiency improvements financially supported by the European Union funds could be realized. As a result of the utility cost reduction program (2013-2014), with special regard to the price effect, the social disparities related to the energy expenditure per household declined. However, residential energy consumption increased (this complex process is presented in detail in Sebestyén Szép 2018).

**Table 2**

**Gini coefficient values for the electricity consumption, natural gas use and the income per capita, in case of Budapest and Hungarian cities with county rights (2010, 2015)**

	Natural gas consumption per capita (m <sup>3</sup> )	Residential gas consumption per household (m <sup>3</sup> )	Electricity consumption per capita (kWh)	Residential electricity consumption per household (kWh)	Income per capita (HUF)
<b>Gini index (2010)</b>	0.22	0.16	0.24	0.1	0.07
<b>Gini index (2015)</b>	0.15	0.13	0.28	0.1	0.07

Source: own calculation.

The results of the AR-Gini index (**Table 2**) and the Theil index (**Table 3**) show similarities regarding electricity consumption. While spatial differences cannot be observed related to the residential electricity consumption per household, the results of the Theil index associated with electricity consumption per capita data indicate a small, but increasing tendency over time, where the within-group inequality component plays a greater role than the between-group inequality component.

The Theil indices associated with natural gas consumption (**Table 3**) confirm the results of the AR-Gini coefficient; a small and downward tendency can be identified in both categories (residential gas consumption per household (m<sup>3</sup>) and natural gas consumption per capita (m<sup>3</sup>)). The Theil index components (within-group inequality component and between-group inequality component) call attention to the differences in the within-group inequality component. These disparities are more important than the values of the between-group inequality components. It cannot be stated that the success, i.e. belonging to the “elite” category, causes significant changes in urban electricity and natural gas consumption patterns.

**Table 3**

**Theil index with regard to electricity and natural gas consumption for the Hungarian towns with county rights and Budapest (2010, 2015).**

Indicator	Year	Theil index	Between-group inequality component (T <sub>B</sub> (I))	Within-group inequality component (T <sub>w</sub> (I))
<i>natural gas consumption per capita (m<sup>3</sup>)</i>	2010	0.097	-0.001	0.098
	2015	0.018	-0.018	0.036
<i>residential gas consumption per household (m<sup>3</sup>)</i>	2010	0.104	0.037	0.068
	2015	0.047	0.040	0.007
<i>electricity consumption per capita (kWh)</i>	2010	0.111	-0.024	0.135
	2015	0.125	-0.030	0.155
<i>residential electricity consumption per household (kWh)</i>	2010	0.00	0.02	-0.02
	2015	0.00	0.01	-0.01

Source: own calculations

The results can be explained by the rebound effect and by high human development index (HDI) values. According to Sorrell et al. (2009) ‘the rebound effect is an umbrella term for a variety of mechanisms that reduce the potential energy savings from improved energy efficiency’ (Sorrell et al. 2009, p. 1457). The rebound effect shows as a result of



energy efficiency improvements, how much the additional residential energy use is, and what percentage of potential energy savings are lost. Madlener and Alcott (2009) conclude that the size of the rebound effect in the long run is between 10% and 30%. A previous study of ours (Sebestyén Szép 2013) identifies a lower result: the rebound effect associated with the residential energy consumption was 15% in East-Central Europe between 1990 and 2009. This finding is crucially important from the point of this study; this fact can explain why no significant difference can be detected among the residential energy use patterns in the selected cities. Probably in the richer, more successful cities (with higher income per capita) the residents can afford to buy modern (more energy efficient) household devices. They heat with better boilers (and probably they use gas for heating), but in the poorer and less successful towns many of the households switched the energy source for heating from electricity and natural gas to wood (in the context of rising energy prices). So the residents living in cities with higher gross income use more appliances and other energy-consuming devices, but their equipment is newer and more efficient, so its energy use is lower as well. By contrast the urban population in the poorer cities (with lower gross income per capita) is less able to afford to replace household appliances. The two processes actually balance each other: the richer use more, but in more efficient way, the poorer uses less, but with higher energy intensity.

Many studies focus on the relation between the degree of development of a country and its energy consumption development (this former is usually measured with HDI) (e.g. Arto et al. 2016; Steinberger and Roberts 2010; Martínez and Ebenhack 2008; Dias et al. 2006; Pasternak 2001). These papers emphasize that the energy use patterns of countries with HDI of 0.7-0.9 are similar, while significant growth in the energy use level is achieved in countries with HDI over 0.9. Pasternak (2001) concludes that there no country with annual electricity consumption below 4000 kWh per capita that has an HDI of 0.9 or higher (the examined time period in this case was 1980-1997). Above 5000 kWh per capita, no country has an HDI under 0.9 (cited in Arto et al. 2016). These publications describe this relation at national level, but analysis of the lower spatial levels (especially the urban analysis) is quite rare. This can be explained by limited data and methodological problems. Probably – if we start from the context described above – a similar process can be experienced at the urban level as well. So above a certain level of HDI – naturally it is achieved in all examined cities – increase in per capita energy consumption is no longer expected (the Hungarian HDI was 0.836 in 2015 according to UNDP (2018)). So development, or success, is separate from the additional energy use – and partly related to the Kuznets curve – the dematerialization is achieved. This is confirmed by Csák (2015). He concludes that the consumer culture is homogeneous – including also the opportunities as well – so the consumption of an area is determined by the intensity of its economic activities.

## 5. CONCLUSION

In this study we examined the regional disparities of urban energy use (electricity use and natural gas consumption) in Hungary. The analysis covered 23 Hungarian towns with county rights and Budapest during the period of 2010–2015. The Theil Index and the area based Gini-index were calculated as well. The basic research question is related to the size of inequalities based on the environmental components of smart cities. The following findings were made:

1. Significant differences between the rural and urban (23 Hungarian towns with county rights and Budapest) energy use were not experienced.

2. In the case of the examined cities significant inequalities and large spatial variances were not revealed with regard to the indicators of urban energy consumption (i.e. residential electricity consumption per household, residential gas consumption per household, natural gas consumption per capita, electricity consumption per capita). Furthermore, the already small territorial differences typically decreased between 2010 and 2015.
3. The Theil index components (within-group inequality component and between-group inequality component) call attention to the differences in within-group inequality component related to natural gas consumption (natural gas consumption per capita, and residential gas consumption per household) and electricity use per capita. It is evident that within-group disparities are currently the most important factor explaining the level variance in these energy indicators across the 24 cities involved in the study. These disparities are more important than values of the between-group inequality components.
4. It was not found that the success (i.e. belonging to the “elite” category) causes significant changes in urban electricity and natural gas consumption patterns.
5. There is a strong positive correlation between the electricity use per capita (kWh) and the income per capita, and between the natural gas consumption per capita (m<sup>3</sup>) and the income per capita.

In summary, focusing on the dimension of smart environment it can be concluded that significant spatial inequalities do not arise among the Hungarian cities with county rights in relation to the examined indicators of electricity use and natural gas consumption.

## ACKNOWLEDGEMENT

This research was supported by the project no. EFOP-3.6.2-16-2017-00007, titled Aspects on the development of intelligent, sustainable and inclusive society: social, technological, innovation networks in employment and digital economy. The project has been supported by the European Union, co-financed by the European Social Fund and the budget of Hungary. A part of this study was previously published in Hungarian (Nagy, Z., Sebestyén Szép, T., Szendi D. (2018): Területi különbségek a magyar megyei jogú városok energiafelhasználásában – I-II. rész. *Területi Statisztika*, 58(5): 447–461; 58(6): 551-566).

## REFERENCES

- Alcantara V., Duro J. A. (2004): Inequality of energy intensities across OECD countries: a note. *Energy Policy*, 32 (11), pp. 1257-1260
- Arto, I., Capellán-Pérez, I., Lago, R., Bueno, G., Bermejo, R. (2016): The energy requirement of a developed world. *Energy for Sustainable Development*, 33, pp.1-13.
- Baji P. (2017): Okos városok és alrendszereik – Kihívások a jövő városkutatói számára? (Smart cities and their domains – Future challenges for urban researchers?). *Tér és Társadalom*, 31 (1), pp.89-105.
- Beluszky P., Gyóri R. (2004): Fel is út, le is út... Városaink településhierarchiában elfoglalt pozícióinak változásai a 20. században. (Rise and fall... Changes in hierarchical position of the Hungarian cities in the XX. century). *Tér és Társadalom*, Vol. 1. pp. 1-41.

- Cohen, B. (2014): Estudio "Ranking de Ciudades Inteligentes en Chile"; Fundacion pais digital, Universidad del Desarrollo. <http://dg6223fhel5c2.cloudfront.net/PD/wp-content/uploads/2014/06/Ranking-Ciudades-Inteligentes-en-Chile.pdf>
- Conceição P., Ferreira P. (2000): The Young Person's Guide to the Theil Index: Suggesting Intuitive Interpretations and Exploring Analytical Applications. UTIP Working Paper Number 14. [https://utip.lbj.utexas.edu/papers/utip\\_14.pdf](https://utip.lbj.utexas.edu/papers/utip_14.pdf)
- Csák L. (2015): Energiapolitika: minden területi szinten. (Energy policy: on all territorial levels). Tér és Társadalom, 29.évf., 4.szám
- Dias R. A., Mattos C., R., Balestieri J. A. P. (2006): The limits of human development and the use of energy and natural resources. Energy Policy, Volume 34, Issue 9, pp.1026-1031.
- Dollman R., Kaplan G., La Cava G., Stone T. (2015): Household economic inequality in Australia. Research Discussion Paper 2015-15, Reserve Bank of Australia, Sydney. [https://gregkaplan.uchicago.edu/sites/gregkaplan.uchicago.edu/files/uploads/dollman\\_kaplan\\_lacava\\_stone.pdf](https://gregkaplan.uchicago.edu/sites/gregkaplan.uchicago.edu/files/uploads/dollman_kaplan_lacava_stone.pdf)
- Druckman A., Jackson T. (2008): Measuring resource inequalities: The concepts and methodology for an area-based Gini coefficient. Ecological Economics, Volume 65, Issue 2, pp. 242-252
- Egedy T. (2017): Városfejlesztési paradigmák az új évezredben – a kreatív város és az okos város. (Urban development paradigms in the new millennium – the creative city and the smart city). Földrajzi Közlemények, 141. 3. pp.254-262.
- Finn A., Leibbrandt M., Woolard I. (2009): Income & expenditure inequality: analysis of the NIDS Wavv I dataset. Discussion Paper 5. National Income Dynamics Study. <http://www.nids.uct.ac.za/publications/discussion-papers/wave-1-papers/96-nids-discussion-paper-no05/file>
- Giffinger, R. (2015): Smart City concepts: chances and risks of energy efficient urban development. Smartgreens Conference, IV. International Conference on Smart Cities and green ICT systems. Lisbon. [http://www.smartgreens.org/Documents/Previous\\_Invited\\_Speakers/2015/SMARTGREENS2015\\_Giffinger.pdf](http://www.smartgreens.org/Documents/Previous_Invited_Speakers/2015/SMARTGREENS2015_Giffinger.pdf)
- Hollands G. R. (2008): Will the real smart city please stand up? City, 12(3), pp. 303-320.
- Jong M., Joss S., Schraven D., Zhan C., Weijnen M. (2015): Sustainable-smart-resilient-low carbon-eco-knowledge cities; making sense of a multitude of concepts promoting sustainable urbanization. Journal of Cleaner Production, Volume 109, pp. 25-38
- KSH (2018): Dissemination Database. Hungarian Central Statistical Office. <http://statinfo.ksh.hu/Statinfo/themeSelector.jsp?&lang=hu>
- Kulcsár S., Szemerey S. (2016): Okos városok, intelligens városfejlesztés: Az intelligens városfejlesztés keretei. (Smart cities, intelligent urban development: Frames of the intelligent urban development) Falu Város Régió, 2, pp. 26-33
- Lados M. (editor) (2011): „Smart Cities” tanulmány. (Smart cities paper). IBM, MTA Regionális Kutatások Központja, Nyugat-magyarországi Tudományos Intézet, Győr. [http://www.rkk.hu/rkk/news/2011/smart\\_cities\\_tanulmany\\_IBM\\_RKK.pdf](http://www.rkk.hu/rkk/news/2011/smart_cities_tanulmany_IBM_RKK.pdf)
- Lengyel I. (1999): Mért a mérhetlent? A megyei jogú városok vizsgálata többdimenziós skálázással. (Measuring of immeasurable? Analysis of towns of county rank with multidimensional scaling). Tér és Társadalom Vol.1-2. pp. 53-73.
- Madlener R., Alcott B. (2009): Energy rebound and economic growth; A review of the main issues and research needs. Energy 34. pp.370-376.
- Martínez D. M., Ebenhack B. W. (2008): Understanding the role of energy consumption in human development through the use of saturation phenomena. Energy Policy, Volume 36, Issue 4, pp.1430-1435
- Nam, T., Pardo, T. A. (2011): Conceptualizing Smart City With Dimensions of Technology, People, and Institutions, from Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, ACM New York, NY.

- Pasternak, A.D. (2001). Global energy futures and human development: a framework for analysis. Global 2001 international conference on: "back-end of the fuel cycle: from research to solutions", France.
- Rechnitzer J., Páthy Á., Berkes J. (2014): A magyar városhálózat stabilitása és változása. (Stability and change in the Hungarian city network). *Tér és társadalom*, 28, pp.105-127.
- Sáfián F., Munkácsy B. (2015): A decentralizált energiarendszer és a közösségi energiatermelés lehetőségei a településfejlesztésben magyarországon. (The opportunities of decentralized energy systems and community energy production in settlement development in Hungary). *Földrajzi Közlemények* 2015. 139. 4. pp.257–272.
- Sebestyénné Szép T. (2013): Energiahatékonyság: áldás vagy átok? (Energy efficiency: malediction and benediction?) *Területi statisztika, Központi Statisztikai Hivatal*, 53 (1), pp.54-68.
- Sebestyénné Szép T. (2018): A hatósági árcsökkentés lakossági energiafelhasználásra gyakorolt hatásának vizsgálata indexdekompozícióval. (Analysing the effects of utility-cost reduction on household energy consumption, using index decomposition) *Közgazdasági Szemle. Vol. LXV.*, pp. 185-205, DOI:10.18414/KSZ.2018.2.185
- Sorrell S. (2009): Jevons' Paradox revisited: The evidence for backfire from improved energy efficiency. *Energy Policy* 37. pp.1456-1469 <https://doi.org/10.1016/j.enpol.2008.12.003>
- Stankovic, J., Dzunic, M., Dzunic, Z, Marinkovic, S. (2017). A multi-criteria evaluation of the European cities' smart performance: Economic, social and environmental aspects; *Zb. rad. Ekon. fak. Rij.* 35(2), pp. 519-550.
- Steinberger J. K., Roberts, J. T. (2010): From constraint to sufficiency: the decoupling of energy and carbon from human needs, 1975–2005. *Ecological Economics*, volume 70, issue 2, pp.425-433
- Szlávik J. (2013): Fenntartható gazdaság. (Sustainable economy) *CompLex Kiadó, Budapest*
- UNDP (2018): Human Development Indices and Indicators: 2018 Statistical Update. United Nations Development Programme. [http://hdr.undp.org/sites/all/themes/hdr\\_theme/country-notes/HUN.pdf](http://hdr.undp.org/sites/all/themes/hdr_theme/country-notes/HUN.pdf)
- World Bank (2018): SE4ALL Database  
<http://databank.worldbank.org/data/reports.aspx?source=sustainable-energy-for-all>
- Zhang L., Yang Z., Liang J., Cai Y. (2011): Spatial Variation and distribution of Urban Energy Consumptions from Cities in China. *Energies*, 4, pp.26-38.

## **GIS-BASED APPROACH TO IDENTIFY THE SUITABLE LOCATIONS FOR SOIL SAMPLING IN SINGAPORE**

*Mărgărit-Mircea NISTOR<sup>1</sup>, Harianto RAHARDJO<sup>1\*</sup>, Alfrendo SATYANAGA<sup>1</sup>, Eng-Choon LEONG<sup>1</sup>, Koh Zhe HAO<sup>1</sup>, Aaron Wai Lun SHAM<sup>2</sup>, Hongjun WU<sup>3</sup>*

DOI: 10.21163/GT\_2019.141.08

### **ABSTRACT:**

Shallow slope failures due to rainfall commonly occur in residual soil, especially in tropical areas like Singapore. Therefore, it is critical to understand the distribution of residual soil properties throughout Singapore Island. Proper procedures are required for selection of appropriate locations of soil sampling to obtain the representative soil properties. The aim of this study is to establish the necessary procedures and methods which are applicable to select suitable locations for soil sampling in Singapore. In this study, the Geographical Information System (GIS) with the incorporation of three layers: digital elevation model (DEM), slope angle, and the soil sampling locations from past studies were used to generate suitability map for determination of soil sampling locations in Singapore. Five suitability classes were implemented in each layer: very low, low, medium, high, and very high. Four types of spatial analyses such as “Analysis by weights”, “Inference matrix method”, “Fuzzy Overlay Gamma method”, and “per-cell statistic maximum method (PCSM)” were assessed to identify the appropriate method for determination of the suitable locations for soil sampling in Singapore. These analyses were carried out using Spatial Analyst Tools in ArcGIS environment. The results of this study indicated that the spatial analysis by weights and Fuzzy Overlay Gamma are suitable for determination of soil sampling locations up to 100 data points. The spatial analysis using the inference matrix is suitable for determination of limited number of soil sampling locations. The spatial analysis using the PCSM method is suitable for determination of large number of soil sampling locations. The findings from this study will benefit the practical engineers in surveying as well as other related researchers for determination of suitable locations of soil sampling.

**Key-words:** Spatial Analyst Tools, suitability map, soil sampling, digital elevation model, soil investigation.<sup>1</sup>

## **1. INTRODUCTION**

The earth's atmosphere is naturally composed of a number of gases that act like the glass panes of a greenhouse which are able to retain heat to keep the temperature of the Earth stable at an average temperature of 16 °C. The average surface temperature of the earth would be around -10 °C without the natural warming effect of these gases, carbon dioxide (CO<sub>2</sub>) (the most prolific among other gases), methane (CH<sub>4</sub>), nitrous oxide (NO<sub>2</sub>),

---

<sup>1</sup>Nanyang Technological University, School of Civil and Environmental Engineering, 50 Nanyang Avenue, Singapore. \*Corresponding author email: [chrahardjo@ntu.edu.sg](mailto:chrahardjo@ntu.edu.sg)

Co-authors e-mail: [margarit@ntu.edu.sg](mailto:margarit@ntu.edu.sg), [alfrendo@ntu.edu.sg](mailto:alfrendo@ntu.edu.sg), [zhkoh@ntu.edu.sg](mailto:zhkoh@ntu.edu.sg), [cecileong@ntu.edu.sg](mailto:cecileong@ntu.edu.sg)

<sup>2</sup>Building and Construction Authority, Special Functions Group, Enforcement & Structural Inspection Department, email: [aaron\\_sham@bca.gov.sg](mailto:aaron_sham@bca.gov.sg)

<sup>3</sup>Singapore Land Authority, Land and Estate Management, email: [wu\\_hongjun@sla.gov.sg](mailto:wu_hongjun@sla.gov.sg), +65 6790 5246

ozone (O<sub>3</sub>) and halocarbons. However, study by IPCC (2013) indicated the increase in the concentrations of these gases in the atmosphere which was resulted in the de-stabilizing effect on the global climate. This phenomenon, referred to as global warming is a major issue all around the world with its negatives effects on the climate, ecosystems, water resources, and human activities (Haerberli et al., 1999; Kargel et al., 2005; Oerlemans, 2005; Nistor & Petcu, 2015; Păcurar, 2015; Collins, 2008; Rahardjo et al., 2016).

Climate changes may cause extreme weather events in the world, such as the alteration in the rainfall pattern, amount, intensity and frequency. Previous studies indicated that rainfall events have become more intense in a number of countries (Fujibe, 2008; Groisman et al., 2005; Kuhn et al., 2011; Kristo et al., 2017). The increase in rainfall intensity may increase the impact of natural hazards, namely flooding and slope failures. Prolonged and heavy rainfalls in tropical countries have been observed in recent years. The report from IPCC (IPCC 2013) showed that it is expected that rainfall intensity in South East Asia, including Singapore will increase about 2 to 15 %, according to A1B scenario of the forth assessment report (AR4). The increase in rainfall intensity creates a risk that affects the environment sustainability. In this region, numerous slope failures commonly occur in steep residual soil slopes with a deep groundwater table during rainfalls (Toll et al., 1999; Rahardjo et al., 2013; Singh et al., 2008; Tohari, 2012; Jotisankasa et al., 2010).

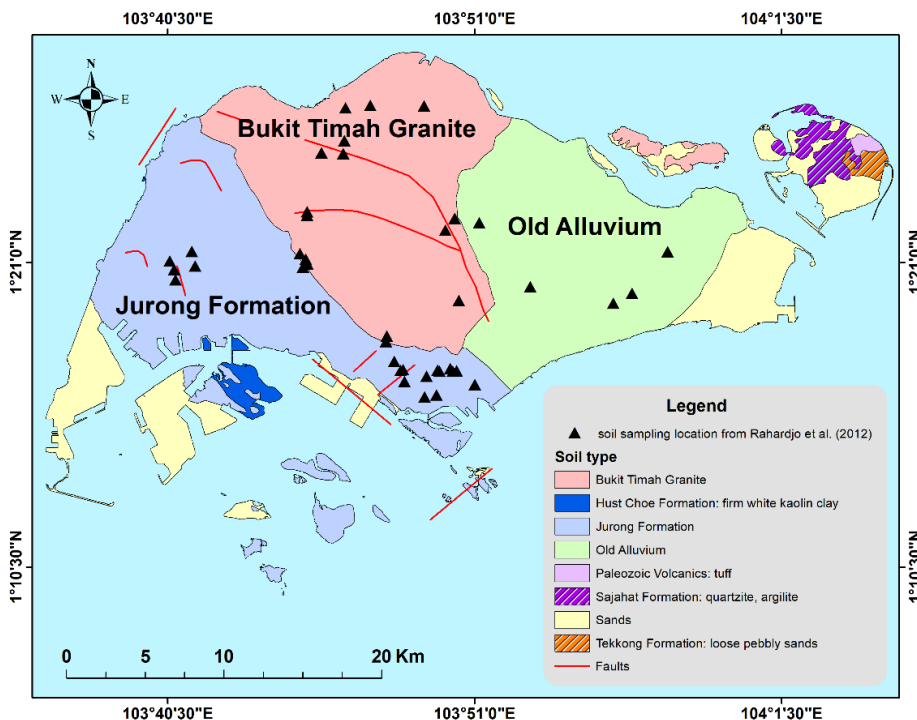
Residual soil slopes have heterogeneous characteristics. Due to weathering, the soil properties of residual soil vary with depths for different locations in Singapore (Rahardjo et al., 2004). Permeability and shear strength vary gradually with depth, controlling both local seepage response to rainfall infiltration and location of the shear surface (Rahardjo et al., 2012). The variability of residual soil properties must be taken into account in the seepage and stability analyses of slopes. In this case, the proper number of soil sampling is necessary in the development of spatial distribution of soil properties with high accuracy. However, a large number of soil sampling contributes to the high cost of soil analyses. Therefore, it is necessary to have an appropriate tool for spatial analyses of saturated and unsaturated soil properties.

GIS applications are commonly used for spatial analyses in the engineering and environmental fields. Special function in ArcGIS environment called Spatial Analyst Tools were used by many researchers for determination of suitable locations of data collection in the development of high accuracy spatial distribution of those particular data (McCoy & Johnston, 2002). Dezsi et al. (2015) used GIS approach to select the suitable locations for a winter resort in the Western Carpathians. Several researchers used Spatial Analyst Tools in ArcGIS for determination of the spatial distribution of groundwater and climatic data in different regions of South Europe (Baltas, 2007; Nistor et al., 2015; Nistor, 2016). Paul et al (2002) and Elshehaby & Taha (2009) adopted the GIS techniques for the mapping of the spatial distribution of glaciers and measurements of the environmental changes (i.e. hydrogeology, climate, glaciology) in Switzerland. Limited research works were performed using GIS technique for landslides investigations and slope hazard risk development (Dai & Lee, 2002).

Sambah & Miura (2014) determined the vulnerable areas to tsunami in the eastern coast of Japan using Spatial Analyst Tools based on pair-wise comparison matrix in GIS. Șerban et al. (2016) used unmanned aerial vehicle (UAV) technology and GIS to delineate the flood-prone area in the North-East of Apuseni Mountains, Transylvania. The remote sensing and the Spatial Analyst Tools in ArcGIS were used to determine the multi criteria evaluation for development of landslide susceptibility map in the Haraz Watershed, Iran (Pourghasemi et al., 2012).

The soil quality and variability of salinity at spatial scale in Dawling National Park from southwestern Mauritania was checked and analyzed using GIS by Abidine et al. (2018). Cervi et al. (2010) used the statistical and deterministic methods for the development of landslides susceptibility map in the Northern Apennines, Italy. Bouajaj et al. (2016) utilized GIS applications for slope stability analysis in the development of slope susceptibility map. Based on the GIS, Murekatete & Shirabe (2018) applied spatial and statistical analyses to determine the variability of least-cost paths using raster artificial landscape data. In the city of Palermo, Dardanelli et al. (2017) determined the acoustic map of the vehicular traffic of an urban agglomeration using GIS technology. In their study, the cartography analysis including the terrain morphology, traffic flow integration, and resident population in the area were the main data used for the acoustic map for vehicular traffic. Bilaşco et al. (2018) used GIS technology to determine the accessibility of agricultural lands in the territory of Teiuş, Alba-Iulia, and Sebeş.

The objective of this study is to investigate an appropriate method for the development of the suitability map for soil sampling locations in Singapore. The scope of works include spatial analyses using Spatial Analyst Tools based on four methods in ArcGIS environment.



**Fig. 1.** Main soil formation in Singapore and the location of the soil sampling from Rahardjo et al. (2012). Background image source: Esri, Garmin, GEBCO, NOAA NGDC, and other contributors.

## 2. MATERIALS AND METHODS

### 2.1. Residual soils in Singapore

Residual soil is the product of the in-situ mechanical and chemical weathering of underlying rocks, which have lost their original rock fabrics. The most important characteristic of residual soils is the low strength due to the destruction of the bonds and the cementation of the material from the weathering processes. In certain cases, they appear to have high shear strength, but as they reach saturation the shear strength reduces significantly with zero or very small effective cohesion (Lumb, 1965). Residual soils in Singapore are often a problem during heavy rainfall events since the flow of rainwater into the unsaturated zone results in an increase in the pore-water pressure and a decrease in the shear strength of residual soil (Fredlund & Rahardjo, 1993). As a result, rainfall-induced slope failures frequently occur in tropical areas that are covered mainly with residual soils (Rahardjo et al, 2013).

Three main soil types exist in Singapore Island. The western part of Singapore is predominantly covered by the residual soil from Jurong Formation, the north and central parts of Singapore are covered by the residual soil from Bukit Timah Granite, whereas the eastern sides of Singapore are covered by the residual soil from Old Alluvium (Oliver & Gupta, 2017). Rahardjo et al (2012) carried out soil sampling in 40 locations throughout Singapore island. **Figure 1** shows the main soil formations in Singapore and 40 locations of the soil samplings from previous studies by Rahardjo et al (2012).

### 2.2. Data preparation

Four (4) different methods were incorporated in Spatial Analyst Tools of ArcGIS to study an appropriate method for the development of the suitability map for the soil sampling locations in Singapore. Prior to spatial analyses, the data input in terms of GIS layer must be prepared correctly to ensure the reliability of the results of the analyses. In this study, five (5) different classes were used to determine the suitability of each GIS layer, i.e. very low, low, medium, high and very high. Three (3) GIS layers were incorporated in the Spatial Analyst Tool for spatial analyses using four different methods, i.e. Analysis by Weights, Inference Matrix, Fuzzy Overlay Gamma and Per-Cell Statistic Method (PCSM). The first layer is the digital elevation model (DEM) layer which provides the elevation of the slope. The second layers is the slope layer, which is referring to the inclination of the slope. The third layer is associated with the location of the previous soil sampling in Singapore based on the study from Rahardjo et al (2012). Forty (40) locations of soil sampling from Rahardjo et al (2012) were used in this study since all soil samplings were conducted within the slope surface and they were distributed in different formations in Singapore. It is important to use this data since the purpose of this study is to determine the suitability of the soil sampling location for the purpose of development of the slope susceptibility map in the future. The laboratory test results of soil samples from the selected locations will be used in this study for the analyses of rainfall-induced slope failures in Singapore. Toll (1999) observed that the slope failures in Singapore are related to shallow slip surface. Therefore, the determination of soil sampling location in this study only incorporated soil layer. Geological layer is not relevant for the selection of soil sampling location in this study.

The classification of GIS layers with respect to DEM, slope angle and distance to the previous soil sampling locations are presented in **Figure 2**.



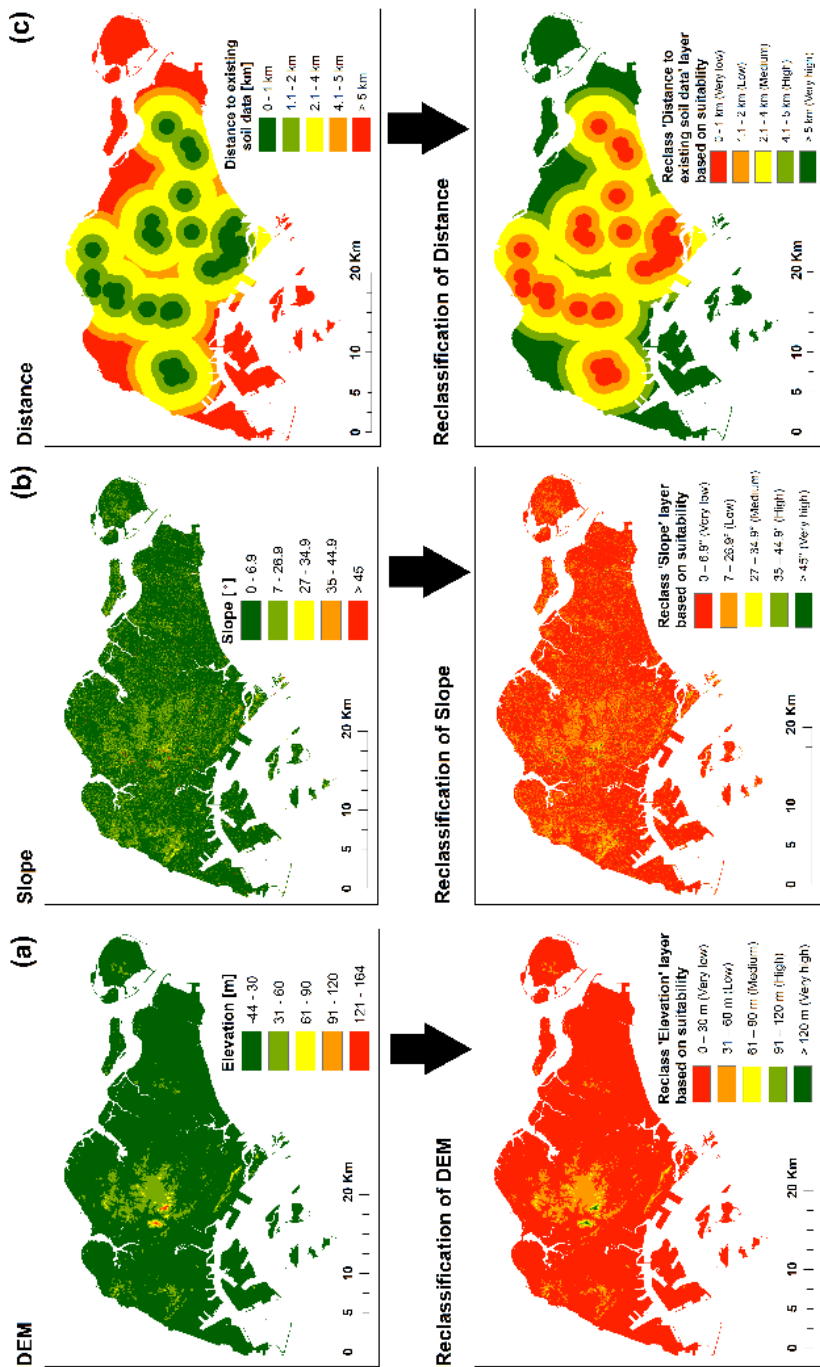


Fig. 2. Classification and reclassification of (a) DEM, (b) Slope angle, and (c) Distance to the previous soil sampling locations

**Figure 2** shows that the elevation of Singapore island varies from -44 m to 164 m. The DEM layer was reclassified based on 5 suitability classes with the same interval in each class. Slope located in the high elevation is considered to be more prone to environmental conditions (i.e. cracks due to high temperature, high infiltration due to rainfall). Hence, it is very important to have soil sampling in the high elevation. The slope layer was divided into 5 classes as follows:  $0^{\circ}$ - $6.9^{\circ}$ ,  $7^{\circ}$ - $26.9^{\circ}$ ,  $27^{\circ}$ - $34.9^{\circ}$ ,  $35^{\circ}$ - $44.9^{\circ}$ ,  $>45^{\circ}$ . This classification was based on the condition of slope angle in Singapore. The ground surface is commonly inclined to  $7^{\circ}$  to drain the rainwater properly into the main drainage especially during high rainfall periods. The engineered slopes in Singapore are commonly designed with  $27^{\circ}$  inclination. Slopes with an inclination higher than  $45^{\circ}$  are commonly associated with steep slopes in Singapore. The slope layer was also reclassified into 5 suitability classes with steep slope (angle higher than  $45^{\circ}$ ) is considered very suitable for the locations of the soil sampling. The GIS layer related to the distance to the previous soil sampling location was classified into 5 classes with the same interval of 1 km in each class. 1 km was adopted in this study since this is the lowest resolution of Singapore raster map.

**Table 1** illustrates the classification and the reclassification of each GIS layers based on the suitability areas for soil sampling. The reclassification was necessary because the datasets should have the same range of numbers and the same impact to the final results. Reclassification was conducted by normalizing the raster data between 0 and 1. The reclassification of the raster data was completed with 'Reclassify' function in ArcGIS. The spatial resolution of the layers was set at  $10 \text{ m}^2$  (**Table 1**).

**Table 1. Values of each parameter used in the spatial analysis and the classes**

Suitability for the test sites	Reclassification		Reclassification		Distance to the previous test sites (km)	Reclassification of the Distances (integer numbers)
	DEM (m)	n of DEM (integer numbers)	Slope (Radian degrees)	n of Slope (integer numbers)		
Very low	0 - 30	1	0 – 6.9	1	0 - 1	1
Low	31 - 60	2	7 – 26.9	2	1 - 2	2
Medium	61 - 90	3	27 – 34.9	3	2 - 4.	3
High	91 - 120	4	35 – 44.9	4	4 – 5	4
	121 -	5	> 45	5	> 5	5
Very high	164					

### 2.3. Spatial analysis using weights

The suitability map based on spatial analysis using weights was developed by applying certain weightage into DEM, slope and distance to previous soil sampling location layers in the "Raster Calculator" function in ArcGIS. The weightage of 0.3, 0.3 and 0.4 was applied

on the reclassification maps of all layers in "Raster Calculator" using Equation 1. Those weightages were selected based on the importance of each layer in the spatial analyses (Deszi et al, 2015; Nistor et al, 2015).

$$\text{Suitability} = (\text{DEM} \times 0.3) + (\text{Slope angle} \times 0.3) + (\text{Distance} \times 0.4) \tag{1}$$

The 30% weightage for DEM and slope layers were selected since both factors provided similar contribution into slope stability analyses. The 40% weightage for distance to previous soil sampling location layer was selected since this layer had the highest impact in the spatial analyses. This layer was significantly required to avoid the overlapping between the new soil sampling location in this study and the previous soil sampling location from Rahardjo et al (2012).

### 2.4. Spatial analysis using inference matrix method

The suitability map based on spatial analysis using inference matrix method was generated using predefined overlay procedure with 5 x 5 matrices (i.e. very low, low, medium, high and very high). Two steps were required in the development of the suitability map based on three required GIS layers, DEM, slope and distance to previous soil sampling location layers.

The first step is to establish the classification for the suitability of the terrain for soil sampling based on relationship between elevations and slope layers.

The second step is to generate the suitability map for the new soil sampling in this study based on the relationship between the suitability of the terrain for soil sampling and the distance to the previous soil sampling location layer.

The inference matrix to establish the suitability of the terrain for soil sampling is presented in **Figure 3**. The inference matrix to develop the suitability map of the new soil sampling location in this study is presented in **Figure 4**.

		Slope (degree)				
		0 – 6.9	7 – 26.9	27 – 34.9	35 – 44.9	> 45
Suitability		Very low	Low	Medium	High	Very high
Elevation (m)	≤ 30	Very low	Very low	Low	Low	Medium
	31 - 60	Low	Very low	Low	Medium	Medium
	61 - 90	Medium	Low	Low	Medium	High
	91 - 120	High	Low	Medium	Medium	High
	> 120	Very high	Medium	Medium	High	High

Terrain suitability for test sites location				
Very low	Low	Medium	High	Very high

**Fig. 3.** The inference matrix to establish the suitability of terrain for soil sampling.

			Terrain data				
			Class 1	Class 2	Class 3	Class 4	Class 5
Suitability			Very low	Low	Medium	High	Very high
Distance to previous sites (km)							
Close nearest	0 - 1	Very low	Very low	Very low	Low	Low	Medium
Nearest	1 - 2	Low	Very low	Low	Low	Medium	Medium
Moderate nearest	2 - 4	Medium	Low	Low	Medium	Medium	High
Far	4 - 5	High	Low	Medium	Medium	High	High
Very far	> 5	Very high	Medium	Medium	High	High	Very high

Integrated suitability for test sites location				
Very low	Low	Medium	High	Very high

**Fig. 4.** The inference matrix to develop the suitability map for the new location of the soil sampling.

## 2.5. Spatial analysis using Fuzzy Overlay Gamma

The suitability map based on Fuzzy Overlay Gamma spatial analysis was developed using an automatic approach in ArcGIS. The 'Fuzzy Overlay Gamma' was selected because the method yields the values between 0 and 1. The required variables in the analysis were 'Fuzzy Product' and 'Fuzzy Sum' (ESRI, 2013). Both of these variables are dependent to the power of gamma. The general function incorporating both variables is presented in Equation 2. The general equation used in Fuzzy Overlay Gamma spatial analysis is presented in Equation 3.

$$\mu(x) = (\text{FuzzySum})^\gamma \times (\text{FuzzyProduct})^{1-\gamma} \quad (2)$$

$$\text{fuzzyGammaValue} = \text{pow}(1 - ((1 - \text{arg1}) \times (1 - \text{arg2}) * \dots), \text{Gamma}) \times \text{pow}(\text{arg1} \times \text{arg2} * \dots, 1 - \text{Gamma}) \quad (3)$$

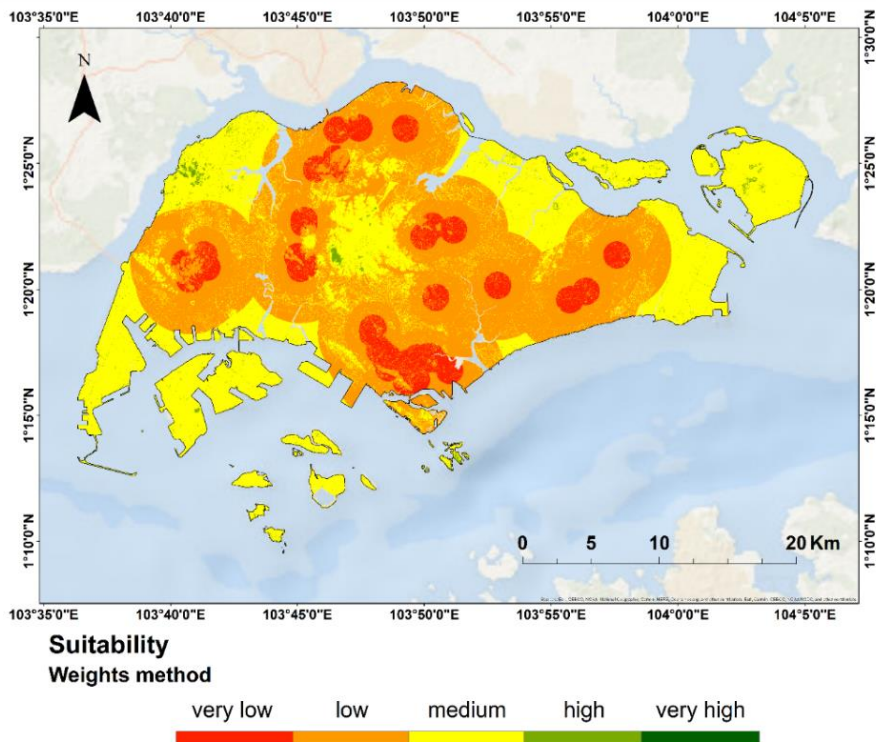
The gamma value of 1 produces an output equal to 'fuzzy Sum'. The gamma value of 0 provides an output equal to 'Fuzzy Product'. The gamma value between 0 and 1 generates different output of 'Fuzzy Overlay Gamma' as compared to the output of other fuzzy methods, e.g. 'fuzzy Or' or 'fuzzy And'. The Fuzzy Overlay Gamma can be used when the expected values should be greater than fuzzy Product and less than fuzzy Sum. The final suitability map for determination of the new soil sampling location in this study can be classified into five classes similar to the suitability map obtained from other methods of spatial analyses.

## 2.6. Spatial analysis using per-cell statistic maximum method

The suitability map based on spatial analysis using per-cell statistic maximum (PCSM) method was produced by selecting the maximum values from each GIS layer of DEM, slope and distance to the previous soil sampling location layers. Prior to the analysis, all three layers should be superimposed correctly to ensure the accurate calculation of maximum value of each cell from all layers. All cells in the suitability map should be independent of all cells in each GIS layer.

### 3. RESULTS

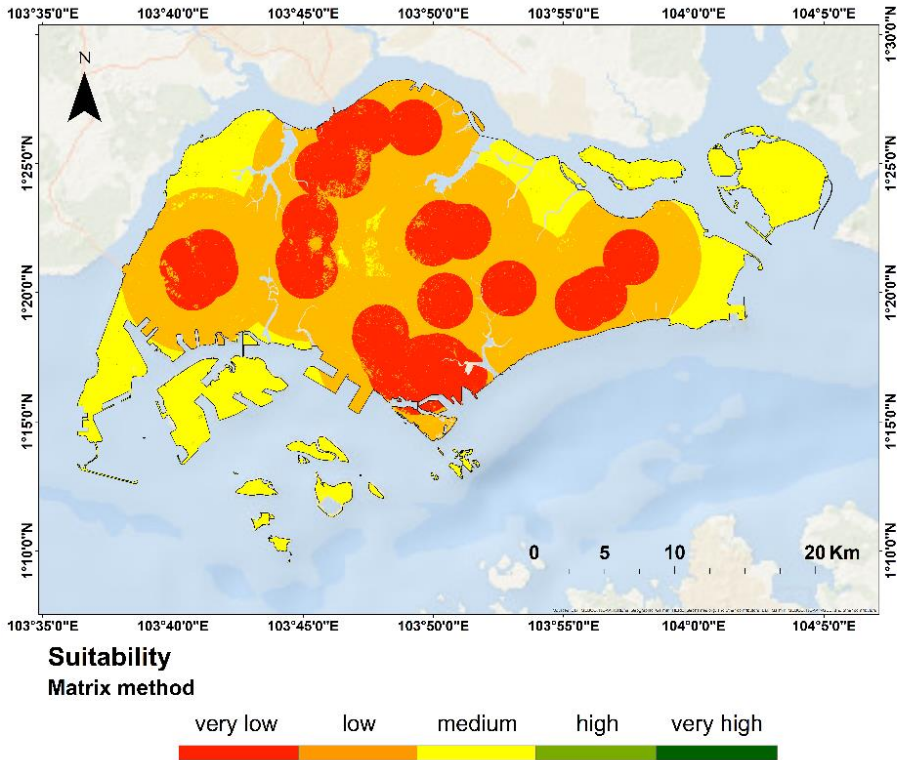
The suitability map for the soil sampling locations in Singapore was developed based on four methods using Spatial Analyst Tools in ArcGIS. The results from analysis using the spatial analysis by weights indicated that the high and very high suitability for the soil sampling are located in the small areas in the central, north western and north eastern parts of Singapore. The central parts of Singapore was very suitable for soil sampling since this area was located in high elevation and majority of slopes in this area were steep. Few locations with high suitability for soil sampling could also be found in the western, southern, and eastern parts of Singapore. However, these locations are close (within 1-1.5 km) to the previous soil sampling locations. The areas with medium suitability extended mostly in the north western, eastern, north eastern and south western parts of Singapore with a distance longer than 5 km to the previous soil sampling locations. In general, majority of areas have low suitability for new soil sampling locations based on the spatial analysis by weights. All cells in these areas were located within a distance of 1 to 5 km to the previous site investigation locations. **Figure 5** shows the suitability map for soil sampling locations obtained from spatial analysis by weights. It can be seen that the central, the north western, north eastern and eastern parts of Singapore are suitable for new soil sampling locations.



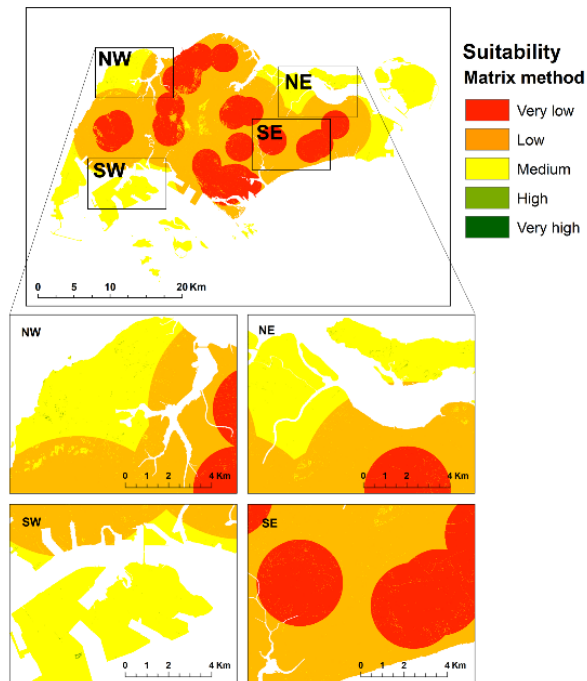
**Fig. 5.** Suitability Map for soil sampling locations in Singapore obtained from spatial analysis using weights. Background image source: Esri, Garmin, GEBCO, NOAA NGDC, and other contributors.

The results of spatial analysis using the inference matrix method indicated that only very few areas were associated with high and very high suitability for new soil sampling locations. Majority of areas produced from the spatial analysis using this method were associated with very low, low and medium suitability for soil sampling. The areas with very high and high suitability for soil sampling were located in the north western, north eastern and south western parts of Singapore (Fig. 6).

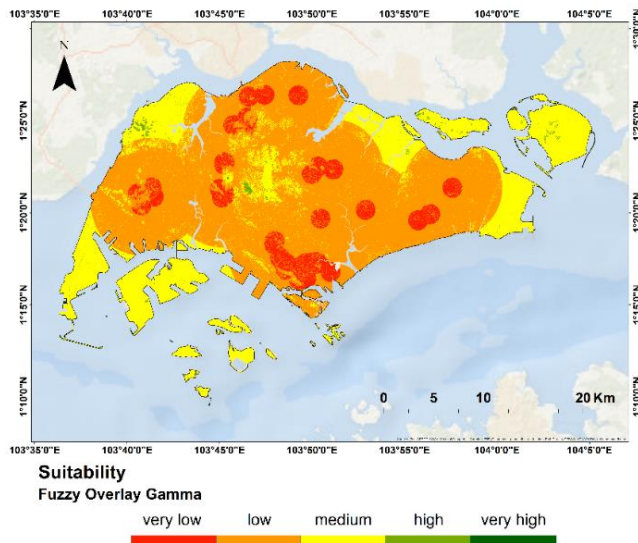
The spatial analysis results using this method also shows that the areas with a distance up to 1 km to the previous soil sampling locations were classified as very low suitability for new soil sampling locations. The areas located at the distance between 1 and 5 km to the previous soil sampling locations were classified as low suitability for new soil sampling locations. The areas located at a distance higher than 5 km to the previous soil sampling locations were classified as moderate suitability for new soil sampling locations. **Figure 6** shows that the eastern, north western and north eastern of Singapore are suitable for new soil sampling locations.



**Fig. 6.** Suitability Map for soil sampling locations in Singapore obtained from spatial analysis using the inference matrix. Details of this map could be seen in the **Figure 7**. Background image source: Esri, Garmin, GEBCO, NOAA NGDC, and other contributors.



**Fig. 7.** Details of Suitability Map carried out using the inference matrix. Details of this map could be seen in the Supplementary material 1. Background image source: ESRI, Garmin, GEBCO, NOAA NGDC, and other contributors.



**Fig. 8.** Suitability map for soil sampling locations in Singapore obtained from spatial analysis using Fuzzy Overlay Gamma. Background image source: ESRI, Garmin, GEBCO, NOAA NGDC, and other contributors.

The analysis using Fuzzy Overlay Gamma method indicated that majority of areas in Singapore was classified as low suitability for new soil sampling locations. **Figure 8** shows that very few areas with high and very high suitability for new soil sampling locations were located in the north western and central parts of Singapore. The areas with medium suitability were located in the northeastern and southwestern parts of Singapore. The areas with very low suitability were located within a distance of less than 1 km to the previous soil sampling locations in Singapore. The suitability map from this method shows that the eastern, north western and north eastern of Singapore are suitable for new soil sampling locations.

The suitability map based on PCSM indicated that areas with very high and high suitability were distributed in the northern, western, southwestern, eastern and central parts of Singapore (**Fig. 9**). The areas with a distance of 1 km to the previous soil sampling locations were classified as low suitability. The areas with a distance between 1 and 4 km to the previous soil sampling locations were classified as medium suitability. Very few areas with very low suitability were observed in the southern part of Singapore. The suitability map from this method shows that only the southern part of Singapore is not suitable for new soil sampling location.

#### 4. DISCUSSIONS

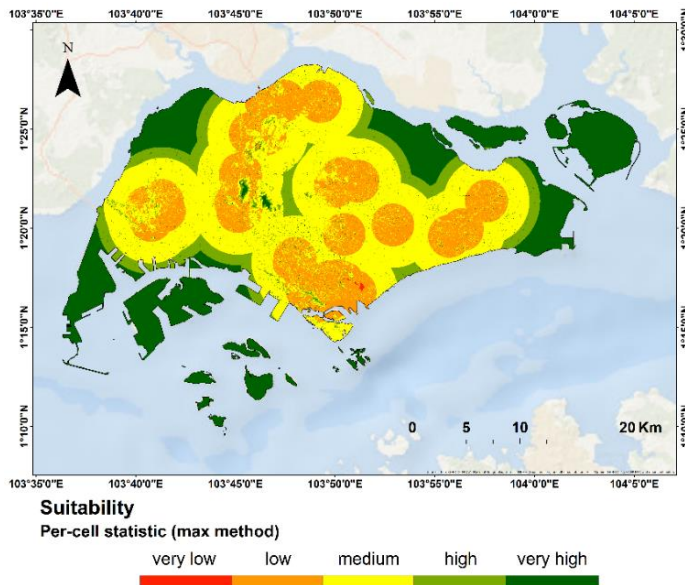
The results of the spatial analyses using four different methods in the Spatial Analyst Tools of ArcGIS indicated that the suitability map from the spatial analysis by weights is very close to that from the spatial analysis using Fuzzy Overlay Gamma. This is attributed to the similar algorithm used in both methods which depend on the weightage of each layer and the summation of all layers.

The areas with high and very high suitability were easily defined from the suitability map obtained from both methods. Eleven (11) to one hundred (100) soil sampling locations can be selected from the suitability map. The suitability map based on spatial analysis using the inference matrix method provided only few suitable locations for soil sampling. This limited results were attributed to the reduction of the number of cells with high suitability as compared to the cells in the suitability map from spatial analysis using other methods. The spatial analysis using the inference matrix was useful to determine the new soil sampling locations when the previous data were very limited (less than 10 data points).

The spatial analysis using the PCSM method produced the suitability map for a large areas with high and very high suitability, especially in the central, northeastern and northwestern parts of Singapore. These areas were located at a distance of greater than 5 km to the previous soil sampling locations. These areas were also located in the high elevation and with steep slope angle.

The suitability map from the spatial analysis using the PCSM method provided higher distinct differences between each suitability classes as compared to that from the spatial analysis using other methods since the PCSM method extracted the maximum values from each GIS layer of DEM, slope and distance to the previous soil sampling locations. The PCSM method is very suitable for determination of high number of soil sampling locations.





**Fig. 9.** Suitability Map for soil sampling locations in Singapore obtained from spatial analysis using per-cell statistics maximum (PCSM). Background image source: Esri, Garmin, GEBCO, NOAA NGDC, and other contributors.

## 5. CONCLUSIONS

Based on the data in this study, the following conclusions can be withdrawn:

- Spatial analyses for development of suitability map for soil sampling locations using four methods in the Spatial Analyst Tools of ArcGIS, i.e. spatial analysis by Weights, spatial analysis using the inference matrix, spatial analysis using Fuzzy Overlay Gamma and spatial analysis using per-cell statistic maximum have been presented and discussed.
- The spatial analysis by weights and Fuzzy Overlay Gamma are suitable for determination of soil sampling locations up to 100 data points.
- The spatial analysis using the inference matrix is suitable for determination of limited number of soil sampling locations (up to 10 data points).
- The spatial analysis using the PCSM method is suitable for determination of large number of soil sampling locations (more than 100 data points).
- Based on four methods used in the spatial analyses, the northwestern, the southwestern, central and the eastern parts of Singapore are considered as the suitable locations for new soil sampling.

## ACKNOWLEDGEMENT

The authors would like to acknowledge the funding support from Building Construction Authority and the sharing of the data from Singapore Land Authority, who are the collaborator of "The Development of Slope Management and Susceptibility Geographical Information System" project.

## REFERENCES

- Abidine, M.M.O., Aboudi, A.E., Kebd, A., Alouimine, B.B., Dallahi, Y., Soulé, A. & Vadel, A. (2018) Modeling the spatial variability of the electrical conductivity of the soil using different spatial interpolation methods: case of the Dawling National Park in Mauritania. *Geographia Technica*, 13(2), 1–11.
- Baltas, E. (2007) Spatial distribution of climatic indices in the northern Greece. *Meteorological applications*, 14, 69–78.
- Bilaşco, Ş., Roşca, S., Păcurar, I., Moldovan, N., Vescan, I., Fodorean, I. & Petrea, D. (2018) Roads accessibility to agricultural crops using GIS technology. Methodological Approach. *Geographia Technica*, 13(2), 12–30.
- Bouajaja, A., Bahia, L., Ouadifa, L. & Awa, M. (2016) *Slope stability analysis using GIS. The International Archives of the Photogrammetry. Remote Sensing and Spatial Information Sciences*, Volume XLII-2/W1, 3rd International GeoAdvances Workshop 16–17 October 2016, Istanbul, Turkey.
- Cervi, F., Berti, M., Borgatti, L., Ronchetti, F., Manenti, F. & Corsini, A. (2010) Comparing predictive capability of statistical and deterministic methods for landslide susceptibility mapping: a case study in the northern Apennines (Reggio Emilia Province, Italy). *Landslides*, 7, 433–444.
- Collins, D.N. (2008) Climatic warming, glacier recession and runoff from Alpine basins after the Little Ice Age maximum. *Annals of Glaciology*, 48(1), 119–124.
- Dai, F.C. & Lee, C.F. (2002) Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology*, 42, 213–228.
- Dardanelli, G., Marretta, R., Santamaria, A.S., Strega, A., Lo Brutto, M. & Maltese, A. (2017) Analysis of technical criticalities for GIS modelling an Urban noise map. *Geographia Technica*, 12 (2), 41–61.
- Dezsi, Şt., Nistor, M.M., Man, T.C. & Rusu, R. (2015) The GIS assessment of a winter sports resort location. Case study: Beliş District, Western Carpathians. *Carpathian Journal of Earth and Environmental Sciences*, 10(1), 223–230.
- Elshehaby, A.R. & Taha, L.G.E. (2009) A new expert system module for building detection in urban areas using spectral information and LIDAR data. *Applied Geomatics*, 1, 97–110.
- ESRI. (2013) *The language of spatial analysis*. Esri Press, 380 New York Street, Redlands, California 92373-8100.
- Fredlund, D.G. & Rahardjo, H. (1993) *Soil mechanics for unsaturated soils*. John Wiley and Sons, Inc., New York.
- Fujibe, F. (2008) Long-Term Changes in Precipitation in Japan. *Journal of Disaster Research*, 3(1), 51–52.
- Groisman, P.Y., Knight, R.W., Easterling, D.R., Karl, T.R., Hegerl, G.C. & Razuvaev, V.N. (2005) Trends in intense precipitation in the climate record. *Journal of Climate*, 18(18), 1326–1350.
- Haeberli, W.R., Frauenfelder, R., Hoelzle, M. & Maisch, M. (1999) On rates and acceleration trends of global glacier mass changes. *Physical Geography*, 81A, 585–595.
- Intergovernmental Panel on Climate Change. (2013) *Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Final draft underlying scientific-technical assessment*. Stockholm, Sweden. 205 pp.
- Jotiskansa, A. & Mairaing, W. (2010) Suction-monitored direct shear testing of residual soils from landslide-prone areas. *Journal of Geotechnical and Geoenvironmental Engineering ASCE*, 136(3).
- Kargel, J.S., Abrams, M.J., Bishop, M.P., Bush, A., Hamilton, G., Jiskoot, H., Käab, A., Kieffer, H.H., Lee, E.M., Paul, F., Rau, F., Raup, B., Shroder, J.F., Soltesz, D., Stainforth, S., Stearns, L. & Wessels, R. (2005) Multispectral imaging contributions to global land ice measurements from space. *Remote Sensing of Environment*, 99(1), 187–219.
- Kuhn, N.J., Schütt, B. & Baumhauer, R. (2011) Managing the impact of Climate Change on the Hydrology of the Gallocanta Basin, NE-Spain. *J. Environ. Manage.*, 92, 275–283.

- Kristo, C., Rahardjo, H. & Satyanaga, A. (2017) Effect of variations in rainfall intensity on slope stability in Singapore. *International Soil and Water Conservation Research*, 5, 258–264.
- Lumb, P. (1965) The residual soils of Hong Kong. *Géotechnique*, 15, 180–194.
- McCoy, J. & Johnston, K. (2002) *Using ArcGIS™ Spatial Analyst*. [Kopp S, Borup B, Willison J, Payne B (contributors)]. Printed by ESRI, Redlands, CA, USA.
- Murekatete, R.M. & Shirabe, T. (2018) A spatial and statistical analysis of the impact of transformation of raster cost surfaces on the variation of least-cost paths. *International Journal of Geographical Information Science*, DOI: 10.1080/13658816.2018.1498504.
- Nistor, M.M. (2016) Spatial distribution of climate indices in Emilia-Romagna region. *Meteorological applications*, 23, 304–313.
- Nistor, M.M. & Petcu, I.M. (2015) Quantitative analysis of glaciers changes from Passage Canal based on GIS and satellite images, South Alaska. *Applied Ecology and Environmental Research*, 13(2), 535–549.
- Oerlemans, J. (2005) Extracting a Climate Signal from 169 Glacier Records. *Science*, 308, 675–677.
- Oliver, G.J.H. & Gupta, A. (2017) *A Field Guide to the Geology of Singapore*. Published by Lee Kong Chian Natural History Museum, Republic of Singapore.
- Păcurar, A. (2015) The climate change and its impact on international dimension of tourism. *Carpathian Journal of Earth and Environmental Sciences*, 10(2), 281–292.
- Paul, F., Kääb, A., Maisch, M., Kellenberger, T. & Haerberli, W. (2002) The new remote-sensing-derived Swiss glacier inventory: I. Methods. *International Glaciological Society*, 34(1), 355–361.
- Pourghasemi, H.R., Pradhan, B., Gokceoglu, C. & Moezzi, K.D. (2012) *Landslide Susceptibility Mapping Using a Spatial Multi Criteria Evaluation Model at Haraz Watershed, Iran*. In 'Terrigenous Mass Movements' [Pradhan B, Buchroithner M (eds.)]. Springer-Verlag Berlin Heidelberg, 400 p, doi 10.1007/978-3-642-25495-6\_2.
- Rahardjo, H., Satyanaga, A., Hoon, K., Sham, W.L., Ong, C.L., Huat, B.B.K., Fasihnikoutalab, M.H., Asadi, A., Rahardjo, P.P., Jotisankasa, A., Thu, T.M. & Viet, T.T. (2016) *Slope Safety Preparedness in Southeast Asia for Effects of Climate Change*. In 'Slope Safety Preparedness for Impact of Climate Change' [Ho K, Lacasse S, Picarelli L (eds.)] CRC Press, Taylor and Francis Group, 13 December, ISBN: 9781138032309, pp. 572.
- Rahardjo, H., Satyanaga, A. & Leong, E.C. (2013) Effects of Flux Boundary Conditions on Pore-water Pressure Distribution in Slope. *Engineering Geology, Special Issue on "Unsaturated Soils: Theory and Applications"*, 165, 133–142.
- Rahardjo, H., Satyanaga, A., Leong, E.C. & Ng, Y.S. (2012) Variability of Residual Soil Properties. *Engineering Geology*, 141–142, 124–140.
- Rahardjo, H., Lee, T.T., Leong, E.C. & Rezaur, R.B. (2004) A Flume for Assessing Flux Boundary Characteristics in Rainfall-induced Slope Failure Studies. *Geotechnical Testing Journal, ASTM International*, 27(2), 145–153.
- Sambah, A.B. & Miura, F. (2014) Integration of Spatial Analysis for Tsunami Inundation and Impact Assessment. *Journal of Geographic Information System*, 6, 11–22.
- Șerban, G., Rus, I., Vele, D., Brețcan, P., Alexe, M. & Petrea, D. (2016) Flood-prone area delimitation using UAV technology, in the areas hard-to-reach for classic aircrafts: case study in the north-east of Apuseni Mountains, Transylvania. *Natural Hazards*, 82, 1817–1832.
- Singh, H., Huat, B.K. & Jamaludin, S. (2008) Slope assessment systems: A review and evaluation of current techniques used for cut slopes in the mountainous terrain of West Malaysia. *Electronic Journal of Geotechnical Engineering*, 13, 1–24.
- Tohari, A. (2012) *Impact of Climate Change on Landslides Hazard in West Java*. Proceeding of Workshop on Slope Stability. Parahyangan Catholic University, Bandung.
- Toll, D.G., Rahardjo, H. & Leong, E.C. (1999) *Landslides in Singapore*. Proc. 2nd International Conference on Landslides. Slope Stability and the Safety of Infra-Structures, Singapore, July 27–28, pp. 269–276.

## **DROUGHT EVALUATION WITH NDVI-BASED STANDARDIZED VEGETATION INDEX IN LOWER NORTHEASTERN REGION OF THAILAND**

*Tanutdech ROTJANAKUSOL<sup>1</sup>, Teerawong LAOSUWAN<sup>1\*</sup>*

DOI: 10.21163/GT\_2019. 141.09

### **ABSTRACT:**

Drought is a natural phenomenon that usually occurs in various areas of the northeastern part of Thailand and it has a major impact on agriculture. The main objective of this study was to apply remote sensing technology and the Standardized Vegetation Index (SVI) for the evaluation of drought in the aforementioned areas. For the implementation, data of the Normalized Difference Vegetation Index (NDVI) from the Terra/MODIS satellite was utilized for the analysis in order to examine the drought areas due to the change of the vegetation conditions by the SVI within a period of three years: 2014, 2015, 2016. According to the study, it was found that the worst drought was in 2016, followed by the years 2014 and 2015, respectively. Furthermore, for the reliability of this technique, the analysis results from the SVI to find the statistical relationship with the rainfall in the target areas was completed. The results of both sets of the data showed a high relationship in all three years (2014,  $R^2 = 0.83$ ; 2015,  $R^2 = 0.88$  and 2016,  $R^2 = 0.97$ ). Thus, it could be concluded that examining the drought situation by remote sensing technology and the SVI at different periods in the areas of the Lower Northeastern region of Thailand could identify the forms of drought effectively.

*Key-words: Drought Evaluation, Remote Sensing, NDVI, SVI.*

### **1. INTRODUCTION**

The current issue found in Thailand the most is drought on account of less rain than usual or because of a dry spell. That is to say, the amount of continuous rainfall is less than 1 mm over a period of 15 days. This has a direct impact on agriculture and most water sources; therefore, drought causes major losses; e.g., dry soil, blocked plant growth, reduction of products, low quality products, and low product quantities (Gomasathit et al., 2013; Nistor et al., 2018). Generally, a drought arises from the lack of seasonal rain, but there are also other causes: for example, natural causes owing to a dry spell and less rainfall: low moisture storage capability of the soil: deforestation: lower amounts of water in reservoirs due to water exploitation: greenhouse effects, and industrial development. Thus, drought is not a result of a single cause (Laosuwan et al., 2016; Uttaruk & Laosuwan, 2017).

In Thailand, the drought as mentioned appears annually, especially in winter to summer, starting from the middle of October onward. The rainfall reduces gradually until the rainy season returns again in mid-May of the following year. Then, it appears again in the middle of the rainy season due to a dry spell, around the end of June to July (Mongkolsawat et al., 2001). Rainfall is, a key factor that should be taken into

---

<sup>1</sup> *Department of Physics, Faculty of Science, Khamriang Sub-District, Kantarawichai District, Maha Sarakham 44150, tanutdech.r@msu.ac.th. \*corresponding author: teerawong@msu.ac.th*

consideration and examined to find out the relationship with the vegetation index, as well as to explore the rainfall periods affecting vegetation. The relationship between the vegetation index and rainfall is a crucial variable to determine the drought areas (Wattanakij & Mongkolsawat, 2008). The application of remote sensing technology based on data from satellites for examining irregular drought areas is another technique that reflects vegetation change. That is because recorded data from satellites can be continuously repeated and can monitor real-time spatio-temporal change of land cover (Mongkolsawat et al., 2009; Seekaw et al., 2014; Kogan, F., & Guo, 2015). When that data is brought for digital image processing together with mathematical statistics, the issue to be studied will be clarified; e.g., the Normalized Difference Vegetation Index (NDVI) calculation (Costea and Haidu, 2010; Costea et al., 2012). The results of the NDVI perfectly exhibit changes in the condition of the vegetation in each weather period (Tucker, 1979; Malo & Nicholson, 1990; Peters et al., 2002; Gessner et al., 2013; Furtuna et al., 2015; Furtuna et al., 2018; Wang et al., 2013; Wu, et al., 2013; Vrieling et al., 2014).

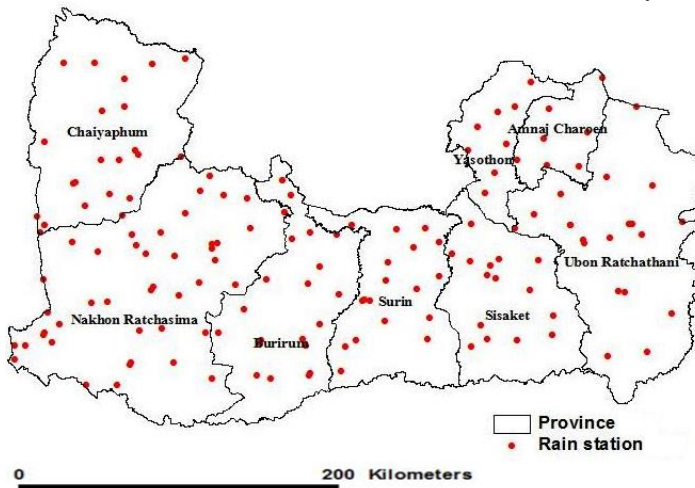
Therefore, drought is a spatial issue, and data from satellites for examining drought in risky areas can raise better effectiveness of displaying the status of the issue. Hence, this study aimed to apply remote sensing technology and the Standardized Vegetation Index (SVI) for the evaluation of drought in areas of the Lower Northeastern region of Thailand.

## 2. STUDY AREAS

This study selected the implementation areas in the Lower Northeastern region of Thailand (**Fig.1**). They were divided into:

(1) The Lower Northeastern provincial cluster 1 consisting of four provinces; i.e., Nakhon Ratchasima, Chaiyaphum, Buri Ram, and Surin. The size of the area was 52,389.89 km<sup>2</sup> or 30.83 % of the total area of the Northeast, and 7.70 % of the country.

(2) The Lower Northeastern provincial cluster 2 comprising four provinces; i.e., Ubon Ratchathani, Sisaket, Amnat Charoen, and Yasothon. The size of the area was 31,907.74 km<sup>2</sup> or 18.89 % of the total area of the Northeast, and 6.2 % of the country.



**Fig. 1.** Study area.

## 2.1. Data

### 2.1.1. Data from Terra/MODIS

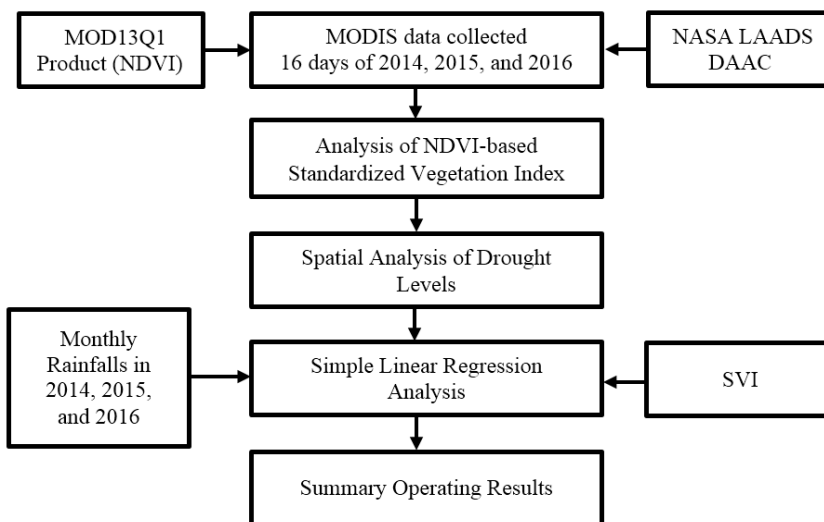
The Terra/MODIS satellite was designed for monitoring and examining natural resource data. The width of the swath was approximately 2,330 km., with the resolutions from 250 to 1000 m. and a 36-band record system. Data from the satellite, therefore, was most suitable for monitoring spatial change of vegetation. That was why the MOD13Q1 package (NDVI) was utilized in the study.

### 2.1.2 Data from rainfall

The data of the average monthly rainfall were collected from the rainfall station of the Thai meteorological station (Fig.1 shows 154 station located in the Lower North-eastern region) for the period 2014-2016.

## 3. MATERIAL AND METHOD

The implementation procedures were set as demonstrated in **Fig. 2**.



**Fig. 2.** Implementation procedures.

### 3.1. Analysis of Terra/MODIS

The Terra/MODIS satellite was designed for monitoring and examining natural resource data. The width of the swath was approximately 2,330 km., with the spatial resolutions from 250 to 1000 m. and a 36-band record system. The data from the satellite was most suitable for monitoring spatial change. That was why the MOD13Q1 package (NDVI) was utilized in the study. The data were collected from the Terra/MODIS satellite (January-December) during 2014-2016. It was downloaded from LAADS DAAC via <https://ladsweb.modaps.eosdis.nasa.gov/> and using MODIS Conversion Toolkit (plugin for ENVI program) for georeference data.

### 3.2. Rainfall analysis

The data of the average monthly rainfall were collected from the rainfall station of the Meteorological Department, Thailand (particularly the station located in the Lower Northeastern region) for the period 2014-2016.

### 3.3. SVI analysis

Because the SVI depended on a z-score in each pixel of the data from the Terra/MODIS, the Z-score was to find the standard deviation (SD) from the mean in the unit of the SD, which was calculated from the NDVI of each pixel in each season. The study was separated into the three seasons of Thailand; namely, summer (17 February-16 May); rainy season (17 May-16 October), and winter (17 October-16February(. The z-score was analyzed as shown in Equation 1 (Peters et al., 2002; Laosuwan et al., 2016).

$$Z_{ijk} = \frac{NDVI_{ijk} - \overline{NDVI_{ij}}}{\sigma_{ij}} \quad (1)$$

Where;

$Z_{ijk}$  = the z-value for pixel i during week j for year k

$NDVI_{ijk}$  = the weekly NDVI value for pixel i during week j for year k

$\overline{NDVI_{ij}}$  = the mean NDVI for pixel i during week j over n years and

$\sigma_{ij}$  = the standard deviation of pixel i during week j over n years

According to Equation 1,  $Z_{ijk}$  was the hypothesis value in order to conform with the standard normal distribution ( $\bar{x} = 0$ ,  $SD = 1$ ) for the test of the hypothesis from every pixel in each season for the period of 2014-2016. There was the probability of  $SVI = P(Z_{ijk})$  (of the z-score from the NDVI. To reflect the probability of the possible vegetation conditions, the SVI analysis could be done as shown in Equation 2 (Peters et al., 2002; Park et al., 2008; Zhang et al., 2013).

$$SVI = \frac{(Z_{ijk} - Z_{ijMIN})}{Z_{ijMAX} - Z_{ijMIN}} \quad (2)$$

Where;

$Z_{ijk}$  = z-value for pixel i during week j for year k;

$Z_{ijMAX}$  = maximum of z-value for pixel i during week j and

$Z_{ijMIN}$  = minimum of z-value for pixel i during week j

According to Equation 2, the probability of each pixel was displayed as the SVI, which helped view the vegetation verdancy in terms of the probability of each pixel in all seasons in the different periods. The study focused on a period of three years (2014-2016) in order to present the comparison between the high and low levels of the drought within the fixed periods of the seasons. This was to estimate the probability of the current vegetation based on the past. The SVI was more than 0 but less than 1 ( $0 < SVI < 1$ ). 0 was the minimum z-score of the NDVI at the pixel of a certain period whereas 1 was the maximum z-score of the NDVI at the pixel of a certain period (Peters et al., 2002; Zhang et al., 2013).

### 3.4. Spatial analysis of the drought areas

This was the classification based on monthly vegetation density in the aforementioned period. From the SVI, the study classified the drought into five levels (Table 1). 0.00-0.05 represented the very poor vegetation density (worst drought) whereas 0.95-1.00 represented a very good vegetation density (minimum drought) (Peters et al., 2002).

**Table1.**

**The vegetation levels.**

SVI level	Vegetation density
0.95– 1.00	very good
0.75 – 0.95	good
0.25 – 0.75	average
0.05 – 0.25	poor
0.00 – 0.05	very poor

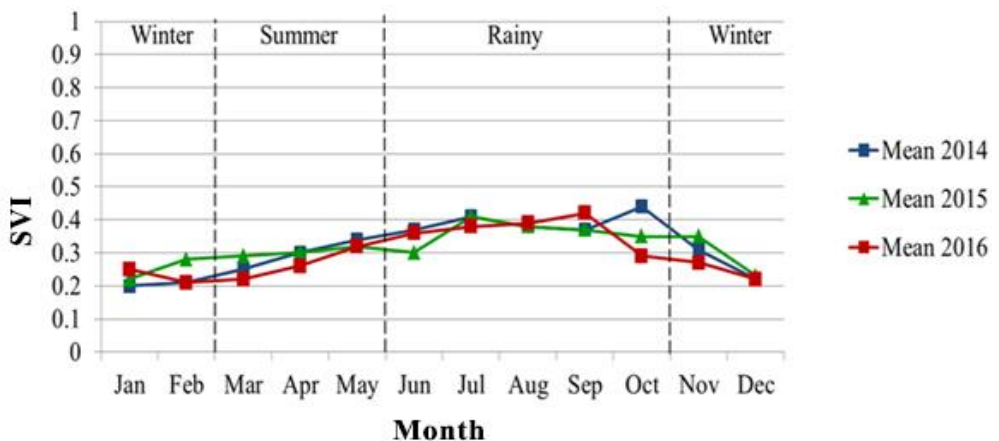
### 3.5. Statistical relationship analysis

For the reliability of the data analyzed from the SVI, the study brought the analysis results from the SVI to find the statistical relationship in the form of a linear regression analysis with monthly rainfalls in the three-year period.

## 4. RESULTS AND DISCUSSIONS

### 4.1. Variation of SVI

For the results of the monthly SVI variation in the three seasons; i.e., summer, rainy, and winter, the means of the SVI in all three seasons (2014-2016) are shown in Fig. 3. The graphs signified the SVI variance in each period of each year.



**Fig. 3.** Comparison of SVI in 2014-2016.



From **Fig. 3**, the average indicated vegetation or drought condition in each period and each season. The result displayed that the year 2014 had the maximum average of 0.44 in October in the rainy season, and the minimum in 0.22 in December at the cold season. The year 2015 had the maximum average of 0.41 in July in the rainy season, and the minimum in 0.23 in December at the cold season. The year 2016 had the maximum average of 0.42 in September in the rainy season, and the minimum t in 0.22 in December at the cold season.

#### **4.2. The drought areas**

According to the examination of drought by Terra/MODIS based on the SVI technique, it was found that the SVI variance in each period was caused by the amount of rainfall that was the determinant of the drought. The comparison of the SVI data of the different periods in each year revealed various changing vegetation conditions. The SVI spatial data obviously mirrored the different levels of the vegetation conditions. The spatial analysis of the drought areas in the three years is shown in **Figs. 4-6**.

According to **Fig. 4** with respect to the areas in 2014, the condition of the vegetation was found to have a minimum during summer in March (SVI average = 0.20) with the maximum distribution in May (SVI average = 0.34) ; had a minimum during the rainy season in June (SVI average = 0.37) and increased to its maximum in September (SVI average = 0.44), and had a high spatial distribution during winter in November (SVI average = 0.31) and started to decline until December (SVI average = 0.22).

According to **Fig. 5** with respect to the areas in 2015, the condition vegetation was found to have a minimum during summer in April (SVI average = 0.30) with the maximum distribution in May (SVI average = 0.32); had a minimum during the rainy season in June (SVI average = 0.30) and increased to its maximum in July (SVI average = 0.41), and had a high spatial distribution during winter in November (SVI average = 0.35) and started to decline until December (SVI average = 0.23).

According to **Fig. 6**, with respect to the areas in 2016, the condition of the vegetation was found to have a minimum during summer in March (SVI average = 0.22) with the maximum distribution in May (SVI average = 0.26); had a minimum during the rainy season in June (SVI average = 0.30) and increased to its maximum in September (SVI average = 0.42), and had a high spatial distribution during winter in November (SVI average = 0.27), and started to reduce until December (SVI average = 0.22).

For this study, the annual spatial SVI analysis was conducted for the annual spatial result displayed by overlaying the based on monthly data. The annual SVI results gave a more vivid expression of the spatio-temporal drought than the monthly ones. **Fig. 7** clearly demonstrates that there was the maximum spatial drought (SVI annual averages data from 12 months) in 2016 , followed by 2014 and 2015, respectively.

#### **4.3. Statistical relationship analysis**

The analysis of the monthly means of the SVI in all three periods and the rainfall illustrated the conformity of the SVI. The relationship of the statistics, the SVI, and the rainfall was shown in **Fig. 8-10**. In **Fig. 8**, focusing on 2014, the statistical relationship analysis between the SVI and the rainfall generated the relationship equation as  $y = 1242.97x - 261.55$ , and  $R^2 = 0.83$ . In **Fig. 9**, focusing on 2015, the statistical relationship analysis between the SVI and the rainfall generated the relationship equation as  $y = 1441.9x - 349.08$ , and  $R^2 = 0.88$ . In **Fig. 10**, focusing on 2016, the statistical relationship analysis between the SVI and the rainfall generated the relationship equation as  $y = 1420.54x - 306.57$ , and  $R^2 = 0.97$ .

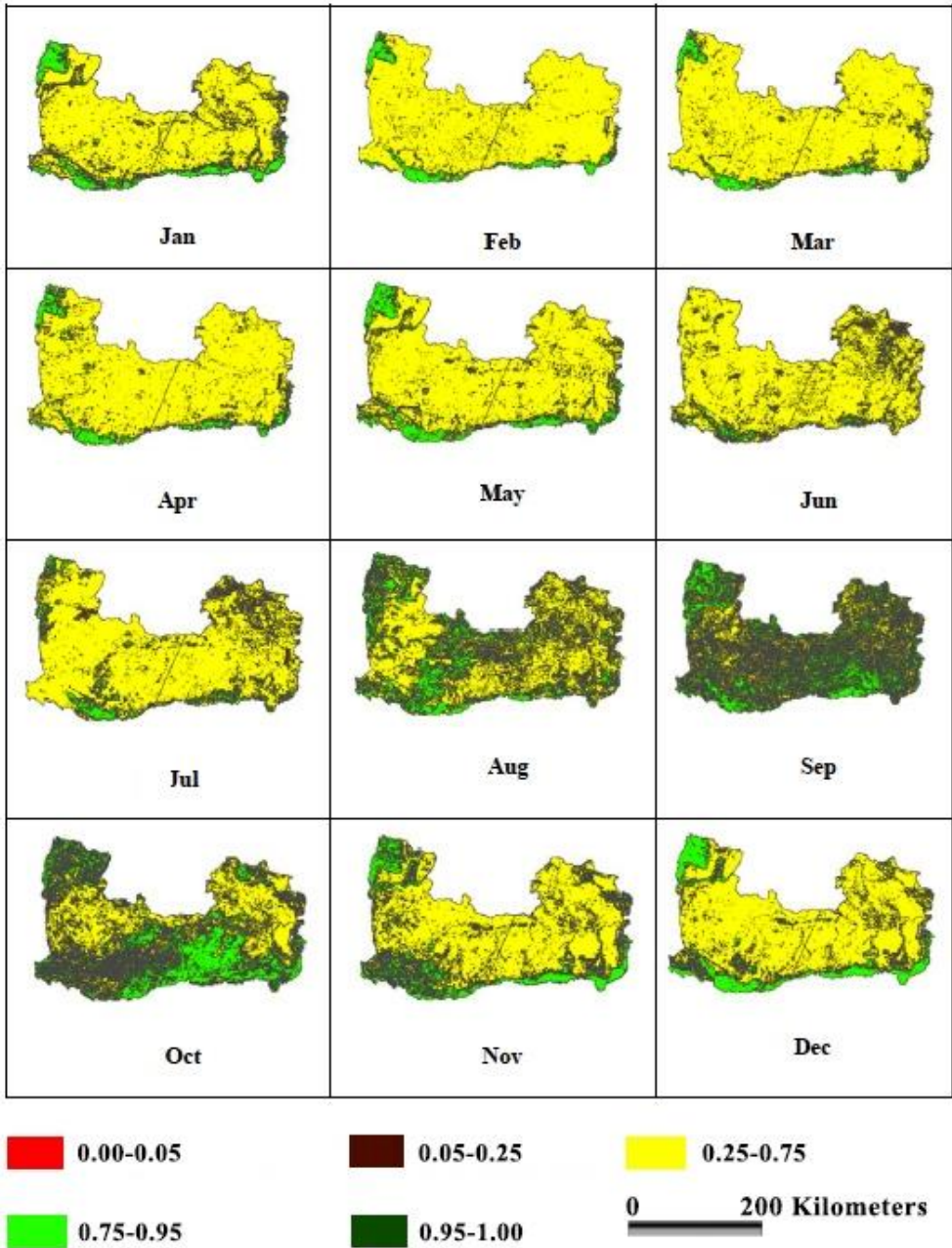


Fig. 4. Spatial analysis of vegetation conditions in 2014.

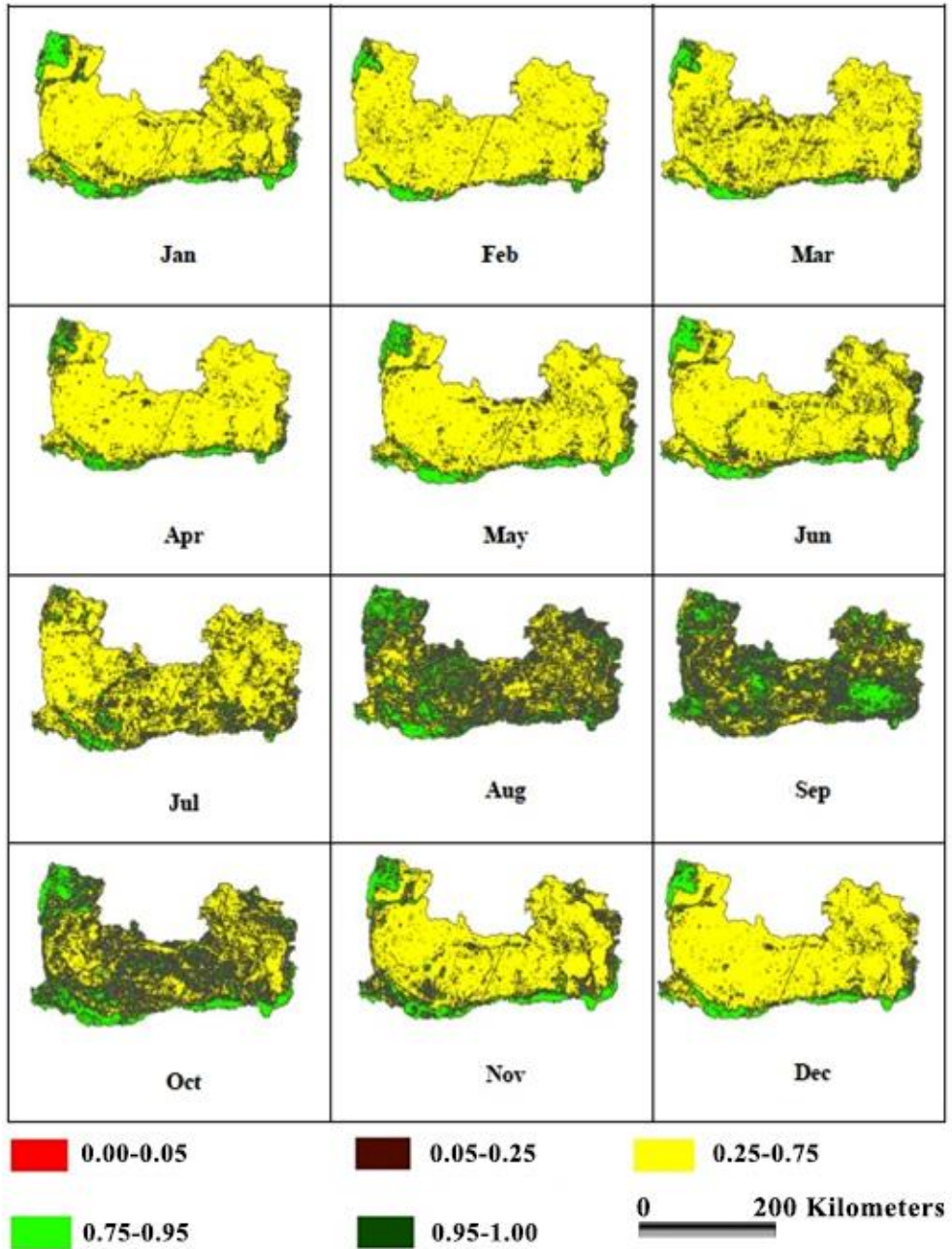


Fig. 5. Spatial analysis of vegetation conditions in 2015.

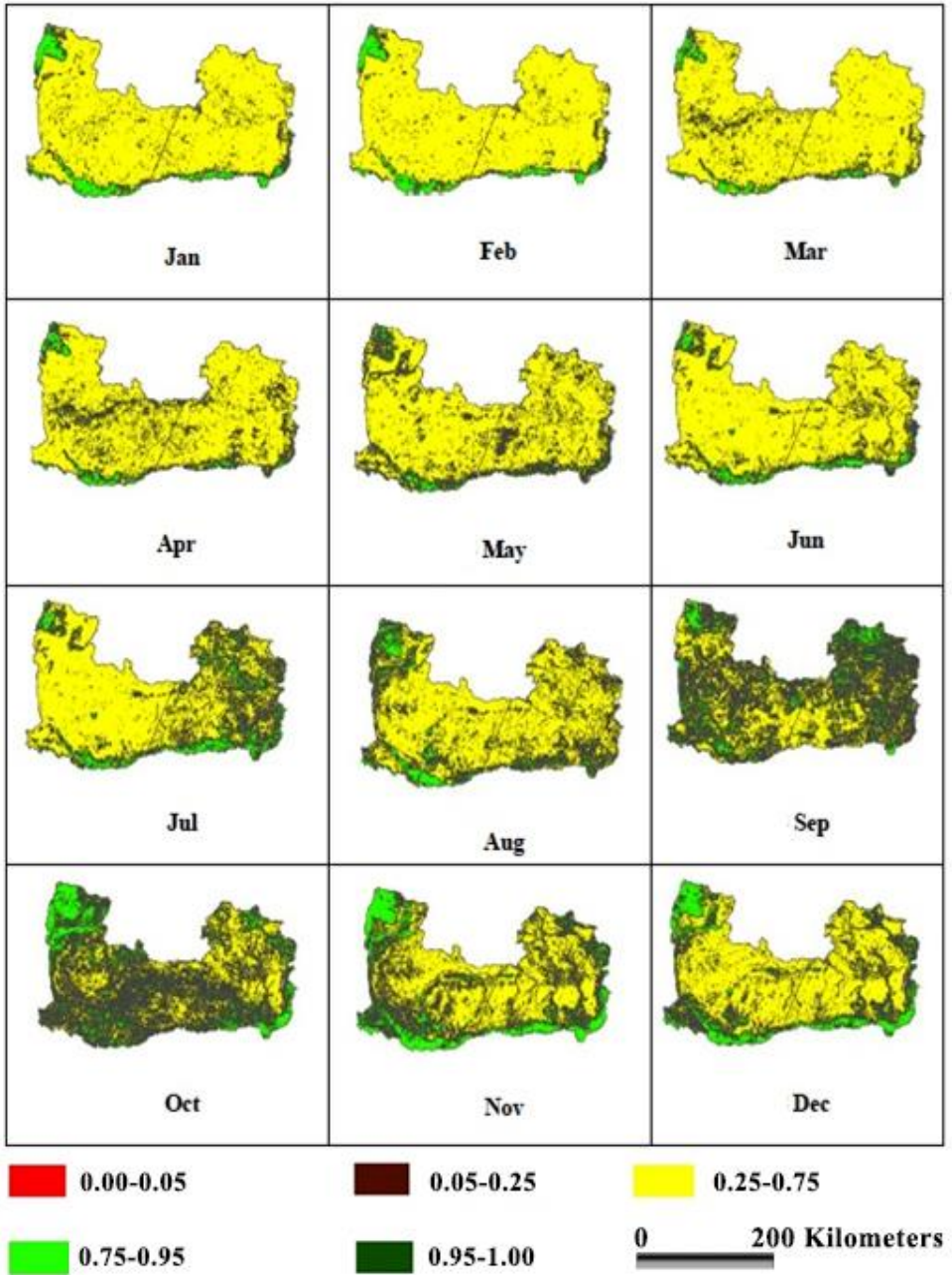


Fig. 6. Spatial analysis of vegetation conditions in 2016.

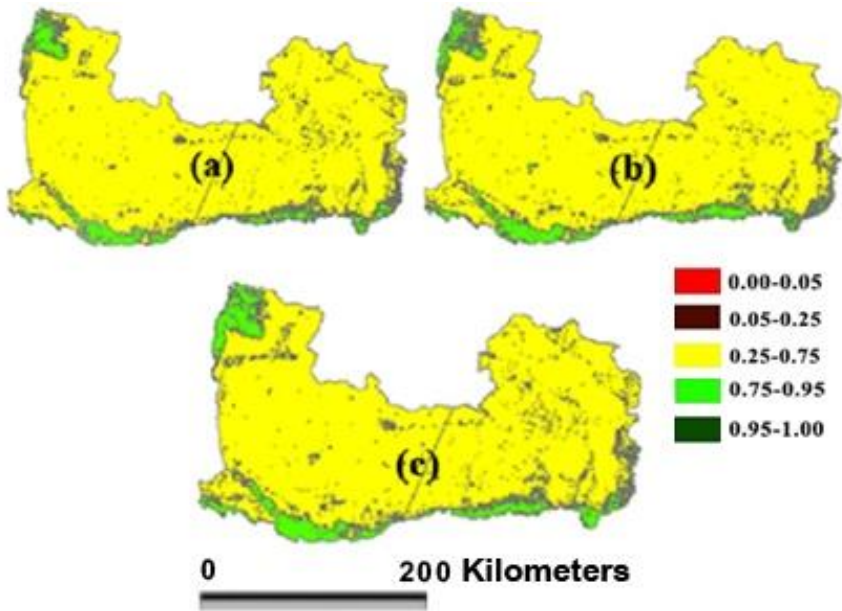


Fig. 7. Spatial analysis (a) 2016 (b) 2014 and (c) 2015.

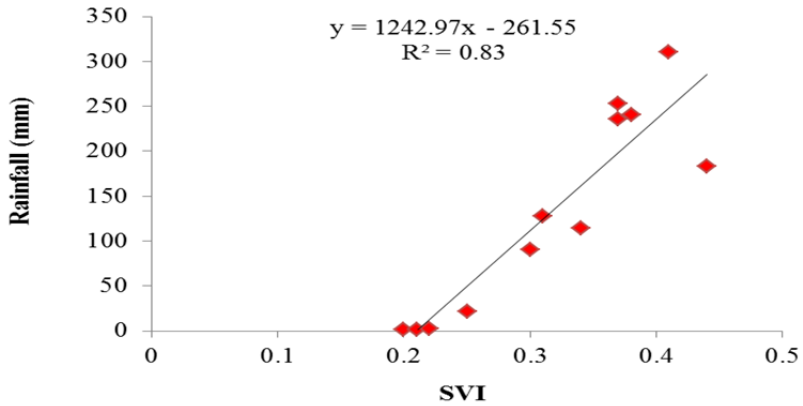
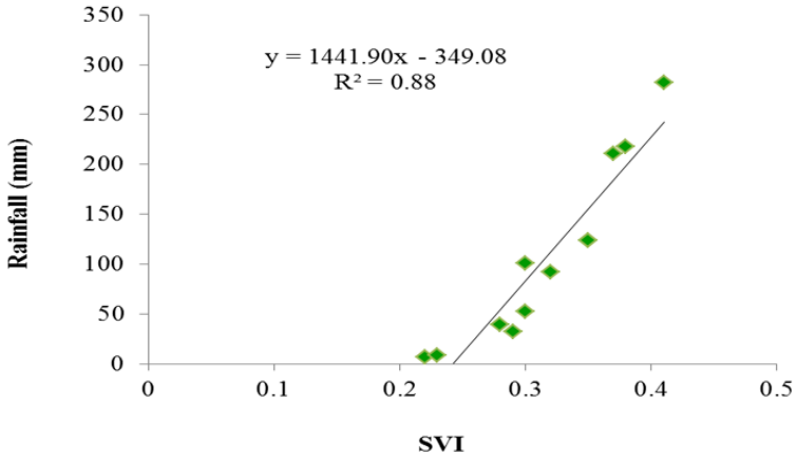
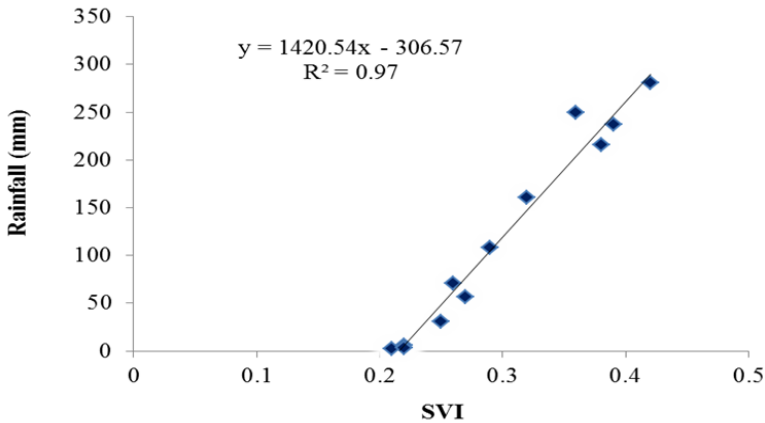


Fig. 8. Relationship of SVI and rainfall data in 2014



**Fig. 9.** Relationship of SVI and rainfall data in 2015



**Fig. 10.** Relationship of SVI and rainfall data in 2016

## 5. CONCLUSIONS AND DISCUSSION

According to the analysis of the means of the SVI in this study, it was discovered that the closer the SVI accessed to 0, the more severe the drought. In contrast the closer the SVI accessed to 1, the higher the moisture increased. When the analysis results of the SVI in all three years were calculated into the percentages of the areas, it was found that the maximum drought was in 2016, with the drought areas at 91.29 % or 76,955.30 km<sup>2</sup>. The second one was in 2014, with the drought areas being 90.99 % or 76,702.41 km<sup>2</sup>, followed by 2015, in which the drought areas was 89.21 % or 75,201.91 km<sup>2</sup>, respectively. The statistical relationship analysis between the SVI and the rainfall unveiled a high conformity. This could be noticed from the coefficient of determination ( $R^2$ ). All three years possessed the coefficient of determination ( $R^2$ ) values close to 1. To clarify, in 2014,  $R^2 = 0.83$ ; in 2015,  $R^2 = 0.88$ ; and in 2016,  $R^2 = 0.97$ . Furthermore, it was proved that high rainfall

affected the high SVI. On the contrary, low rainfall affected a low SVI. However, the SVI change might be slightly slower than the rainfall because of the development of vegetation after there was sufficient water for growth.

In addition, the results of this research were compared with other similar research articles. For example

1) A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States by Gu et al., 2007. The results showed that MODIS data, after analysis by Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) and Normalized Difference Drought Index (NDDI), were able to identify grassland drought.

2) A combination of meteorological and satellite-based drought indices in a better drought assessment and forecasting in Northeast Thailand by Thavorntam et al., 2015. The results showed that Terra MODIS VIs Product after analysis by Vegetation Condition Index (VCI) can be applied to assess drought severity and drought-affected areas for efficient drought management and planning.

3) Drought Detection by Application of Remote Sensing Technology and Vegetation Phenology by Uttarak & Laosuwan, 2017. This research indicates that, an analysis Landsat 8 satellite data with Vegetation Condition Index (VCI) can be identified drought area from vegetation phenology.

The results of these three articles were similar to this present study of which the technique and the results should be regarded as the logical analysis and examination criteria of the drought areas in the Northeastern region of Thailand. Moreover, they can be applied for quick and reliable examination as well as future estimation of drought areas. Furthermore, the related government agencies and private sector can utilize this technique to analyze any expected drought areas, so to obtain the results for writing future sustainable drought prevention and mitigation plans in other areas of Thailand.

## REFERENCES

- Costea, G., Haidu, I. (2010). Detection of recent spatial changes regarding landuse in small basins from the Apuseni Natural Park. *Geographia Technica*, 5(2), 11-27.
- Costea, G., Serradj, A., Haidu, I. (2012). *Forest cartography using Landsat imagery, for studying deforestation over three catchments from Apuseni mountains, Romania*. In: Advances in Remote Sensing, Finite Differences and Information Security. Proceedings of the 8th WSEAS International Conference on Remote Sensing, (REMOTE' 12), Prague, Czech Republic, September 24-26, 2012, WSEAS PRESS. ISBN: 978-1-61804-127-2, 109-114.
- Gessner, U., Naeimi, V., Klein, I., Kuenzer, C., Klein, D. & Dech, S. (2013). The relationship between precipitation anomalies and satellite-derived vegetation activity in Central Asia. *Global and Planetary Change*, 110, 74-87, 2013.
- Gomasathit, T., Laosuwan, T., Sangpradit, S. & Rotjanakusol, T. (2015). Assessment of Drought Risk Area in Thung Kula Rong Hai using Geographic Information System and Analytical Hierarchy Process. *International Journal of Geoinformatics*, 11 (2), 21-27.
- Gu, Y., Brown, J.F., Verdin, J. P. & Wardlow, B. (2007). A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophysical Research Letters*, 34(6), 1-6
- Kogan, F., & Guo, W. (2015). *Agricultural Drought Detection and Monitoring Using Vegetation Health Methods*. In Remote Sensing of Water Resources, Disasters, and Urban Studies; Thenkabail P.S., Ed.; CRC Press: Boca Raton, FL, USA, 339-348, 2015.

- Laosuwan, T., Sangpradid, S., Gomasathit, T. & Rotjanakusol, T. (2016). Application of Remote Sensing Technology for Drought Monitoring in Mahasarakham Province, Thailand. *International Journal of Geoinformatics*, 12(3), 17-25.
- Malo, A. & Nicholson, S. E. (1990). A Study of Rainfall and Vegetation Dynamics in the African Sahel Using Normalized Difference Vegetation Index. *Journal of Arid Environments*, 19, 1-24.
- Mongkolsawat, C., Thirangoon, P., Suwanwerakamtorn, R., Karladee, N., Paiboonsak, S. & Champathet, P. (2001). An evaluation of drought risk area in Northeast Thailand using remotely sensed data and GIS. *Asian Journal of Geoinformatics*, 1, 33-44.
- Mongkolsawat, C., Wattanakij, N., Kamchai, T., Mongkolsawat, K. & Chuyakhai, D. (2009). Exploration of Spatio-Temporal Drought Patterns using Satellite-Derived Indices for Crop Management in Northeastern Thailand. *Proceedings of the 30th Asian Conference on Remote Sensing China*. 18-23 October 2009. Beijing, China.
- Nistor, M.M., Man, T.C, Benzaghta, M.A., Nedumpallile Vasu, N., Dezsil, Ş. & Kizza, R. (2018). Land Cover and Temperature Implications for the Seasonal Evapotranspiration in Europe. *Geographia Technica*, 13(1), 85-108.
- Park, J., Kim, K. & Choi, Y. (2008). Application of Vegetation Condition Index and Standardized Vegetation Index for Assessment of Spring Drought in South Korea. *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, Boston, MA.
- Furtuna, P., Haidu, I., Alexe, M., Holobaca, I. (2015) Change detection in the Cluj forest district, using remote sensing and GIS application. *Environmental Engineering and Management Journal*, 15 (6), 1361-1367.
- Furtuna, P., Haidu, I., Maier, N., (2018) Synoptic processes generating windthrows. A case study in the Apuseni mountains (Romania). *Geographia Technica*, 13(2), pp. 52-61.
- Peters, J.A., Walter-Shea, E.A., Ji, L., Andres, V., Michael, H. & Svoboda, M.D. (2002). Drought Monitoring with NDVI-Based Standardized Vegetation Index. *Photogrammetric Engineering & Remote Sensing*, 68(1), 71-75.
- Seekaw, A., Mongkolsawat, C. & Suwanwerakamtorn, R. (2014). Using standardized vegetation index to assess drought areas in Northeast Thailand. *Journal of Remote Sensing and GIS Association of Thailand*, 15 (2), 25-38.
- Thavornnam, W., Tantemsapya, N. & Armstrong, L. (2015). A combination of meteorological and satellite-based drought indices in a better drought assessment and forecasting in Northeast Thailand. *Natural Hazards*, 77(3), 1453-1474.
- Tucker, C.J. (1979). Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sensing of Environment*, 8(2), 127-150.
- Uttarak, Y. & Laosuwan, T. (2017). Drought Detection by Application of Remote Sensing Technology and Vegetation Phenology. *Journal of Ecological Engineering*, 18(6), 115-121.
- Vrieling, A., Meroni, M., Shee, A., Mude, A. G., Woodard, J., de Bie, C. A. J. M. & Rembold, F. (2014). Historical extension of operational NDVI products for livestock insurance in Kenya. *International Journal of Applied Earth Observation and Geoinformation*, 28, 238-251.
- Wang, J., Rich, P.M. & Price, K.P. (2013). Temporal Responses of NDVI to Precipitation and Temperature in the Central Great Plains, USA. *International Journal of Remote Sensing*, 24(11), 2345-1364.
- Wattanakij N. & Mongkolsawat C. (2008). Drought Detection in Northeast Thailand using Standardized Vegetation Index of Multi-Temporal Satellite Data, In: *Proc. 4th Environment Naresuan Conference*, Naresuan Univeristy, Thailand, 206-215.
- Wu, J., Zhou, L., Liu, M., Zhang, J., Leng, S. & Diao, C. (2013). Establishing and Assessing the Integrated Surface Drought Index (ISDI) for Agricultural Drought Monitoring in Mid-eastern China. *International Journal of Applied Earth Observation and Geoinformation*, 23, 397-410.
- Zhang, Y., Peng, C., Li, W., Fang, X., Zhang, T., Zhu, Q., Chen, H. & Zhao, P. (2013). Monitoring and estimating drought-induced impacts on forest structure, growth, function, and ecosystem services using remote-sensing data: Recent progress and future challenges. *Environmental Reviews*, 21, 103-115.



## **DECISION SUPPORT FOR CYCLING INFRASTRUCTURE PLANNING: A CASE STUDY IN PRĚROV, CZECHIA**

*Aleš RUDA<sup>1</sup>*

DOI: 10.21163/GT\_2019.141.10

### **ABSTRACT:**

Speaking about strategical planning, proposing new cycle paths is based on both quantitative evaluating and qualitative respecting strategic intent of the city. The main goal of the paper is to extend currently accepted approaches in modelling of cycling infrastructure towards geostatistical and multicriterial analysis. The research documents a comparative study of different interpolation techniques producing raster surfaces for further processing together with Euclidean distance surface of points of interest. Choosing an appropriate interpolation method, IDW, EBK, RBF, Ordinary Kriging (OK) with spherical variogram, Ordinary Kriging (OK) with linear variogram, Simple Kriging (SK) with spherical variogram and Simple Kriging (SK) with linear variogram) were considered based on state of the art analysis and further examined. EBK and OK (spherical variogram) achieved very similar values bringing the most accurate predicts and small error estimation. Considering two variants (safety and suitability for construction) determining weights, two sets of proposals were presented. Finally, it was find out that none of examined variants itself can be the ideal solution but mixture of both. Using topological overlay missing segments (connections between existing cycle paths) were recommended for new cycle paths construction.

*Key-words: Cycling, Multicriterial evaluation, Spatial analysis, Strategical planning.*

### **1. INTRODUCTION**

Building sustainable urban transport has been a matter for several years. With success, cities are improving the environment and increasing the safety of cyclists by creating sustainable urban mobility (Ruda, 2016). This trend is confirmed by the growing importance of the Association of Cyclists in Czechia, the greater participation of cities in the bicycle competition and the improvement of the possibilities of drawing funds from the European funds, the State Fund for Transport Infrastructure. The basic prerequisite for the use of support from subsidy programs is the construction and maintenance of cycling infrastructure, which will increase the safety of transport and its accessibility to persons with reduced mobility. The primary objective of the funding was to eliminate cycling traffic in the extravilan from roads of the 1st, 2nd and 3rd classes and move it into a parallel network of cycle paths (STIF, 2013). The National Cycle Development Strategy (MT and TRC, 2015) accepts the possibility of building cycling lanes and other integration measures in urban areas on roads of the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> classes if the roads meet the width-appropriate parameters, rather than the construction of expensive bicycle paths, because the crossing of the newly built cycling paths increases the danger. The transport policy of the Czech Republic for the period 2014-2020 with a view to 2050 (MT, 2013), where the significance of cyclotourism is mentioned as an important attribute for the development of tourism and

---

<sup>1</sup> *Mendel University in Brno, Department of Environmentalistics and Natural Resources, Czech Republic, ales.ruda@mendelu.cz*

it does not neglect its importance in commuting to schools and also a concept of ecotourism (Paul, 2013). The National Road Safety Strategy (Besip, 2011) places great emphasis on the support of cyclotourism in connection with the drafting of legislative measures, emphasizes the obligation to use the protective elements in the legislation. In the White Paper on Transport (European Commission, 2011) cyclotourism is addressed in the context of prioritizing transport safety and promoting sustainable behavior. In the Green Book (Commission of the European Communities, 2007) cycling is presented in the context of its importance for environmentally friendly modes of transport and construction of infrastructure aimed at increasing the safety of cyclists, as in the Czech Republic's National Reform Program (2014). In other strategic documents of regional development of the Czech Republic (MPO CR, 2010), Policy of Territorial Development of the CZ (MRD, 2015) and a whole series of resolutions of the Government of the Czech Republic, the significance of cycling is emphasized as a support for improvement of the lifestyle and health of the population.

The aim of the paper is to illustrate the methodology using geostatistical GIS tools and multicriterial analysis and its possible outputs in order to evaluate necessary indicators (according to public authorities) and recommend suitable solution for planning the construction of cycle paths.

## **2. LITERATURE REVIEW**

The prediction of cyclist movement is crucial to the planning and allocation of the cost of building a new infrastructure, which is based on the collection of primary data (González and De Lázaro, 2013; Papafragkaki and Photis, 2014). Approaches to the methodology of counting cyclists are addressed in numerous studies (Lindsey et al., 2014; Nordback et al. 2013; Lowry et al., 2013; Schneider et al., 2005; Zaki et al., 2013). Basically, the census can be organized along edges or in nodal points (intersections) of the network in the short or long term (Lowry et al., 2013). Using cyclists' counts as a valuable source of basic data is considered in many studies and various techniques to count bicyclists and pedestrians are described (Nordback et al. 2013). Counts can be conducted at mid-segment or at intersections, short-term or long-term. Nordback et al. (2013) also conclude that counts can either be done manually or by automated technologies. Manual counts are usually binned in time periods of 2 hours, while continuous counts 24-hours per day, 365 days per year are gathered by automated counters usually using video counts, active infrared, passive infrared, inductive loops or pneumatic tubes and provide mid-segment screen line. Manual counting is financially less demanding and is usually carried out at irregular moments or continuously 24 hours a day with the support of many bicycle counters. At intersections, bike counting is usually realized manually in short periods of time and uses four different counting techniques: absolute sum of passing cyclists, four-way sum when leaving the intersection, four-way sum on arrival at the intersection, or 12-way sum of arrival and following exit from intersection (Lowry et al., 2013) - the number of paths must be adjusted to the intersection geometry. However, the observed values can be influenced by many factors: weather patterns, demographic characteristics, distance from points of interest etc. (Cardoso et al., 2012; Di Piazza et al., 2011). The accuracy of these results must consider manual counting-related error. Some recent studies (Esawey et al., 2013; Nordback et al., 2013; Nosal 2014; Roll, 2013) combine short-term manual and long-term automatic censuses to achieve the most accurate results.

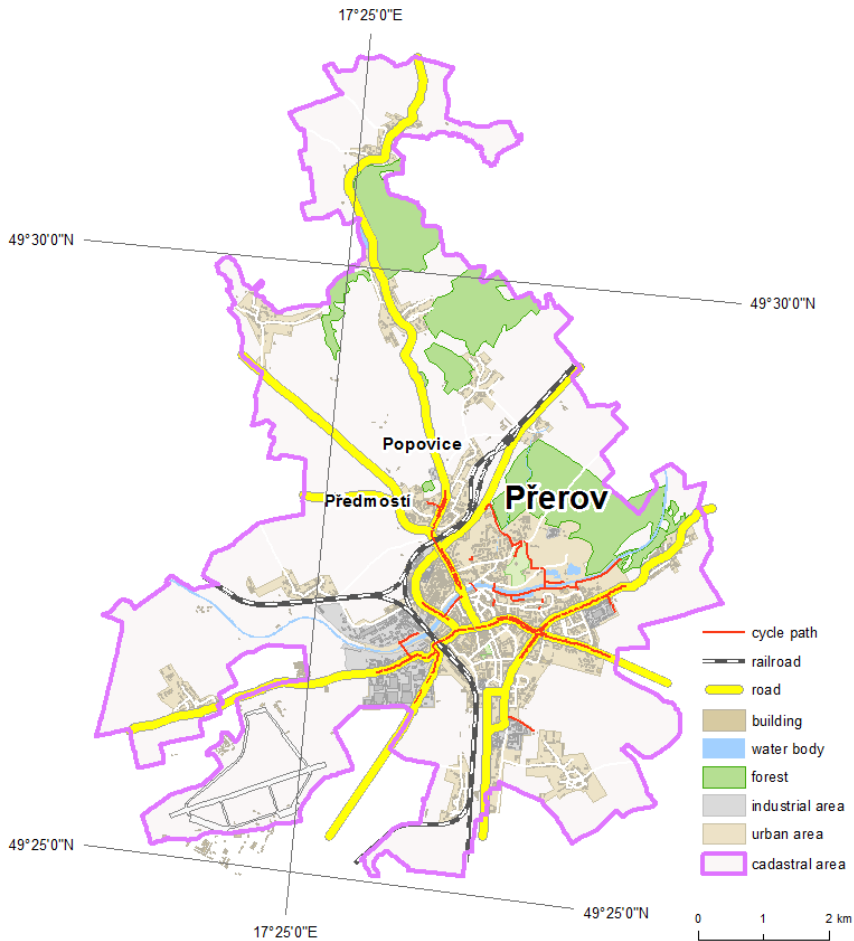
However, the original input data obtained from the primary data collection may not be statistically representative; therefore, different statistical prediction methods are used to densify the point field. Straus and Miranda-Moreno (2013) created a spatially regression model that estimates the resulting values at unknown intersections in relation to land use, demographic characteristics, intersection geometry or weather conditions. The least squares linear regression application was used by Griswold et al. (2011) and, similarly to Straus and Miranda-Moreno (2013), a variety of explanatory variables (average slope, presence of cycling markings, land use and traffic characteristics - intersections density, percentage of intersections, etc.). Studies on safety analyses and identification of obstructions in cycling infrastructure (Brüde and Larsson, 1993; Strauss et al., 2015) are also used. To estimate the daily cycling intensity based on short-term traffic surveys, it is appropriate to set a number of different variations in the intensity of cycling: according to the share of the morning intensity between 5:00 am and 9:00 am to all-day (24h) intensity, according to the share of the afternoon intensity between 2:00 pm and 6:00 pm to all-day (24h) intensity, according to both of the above factors and the size and time of the peak intensity (STIF, 2015). With the progressive improvement of tracking technology, global positioning systems can be used for this purpose in combination with GPS tracking devices (Dardanelli et al., 2017; Strauss et al., 2015). In addition to cycling, this data also provides information for assessing the behavior of cyclists (Broach et al., 2012; Donkwook et al., 2014, Xu et al., 2016). In addition to the contribution of quantitative methods of cycling assessment, qualitative studies are equally important. Rybarczyk and Wu (2010) have proposed a model based on the evaluation of supply and demand for cycling infrastructure services. Participatory approach to decision models has also been incorporated (Lundberg and Weber, 2016; Milakis and Athansopoulos, 2014; Shafizadeh and Niemeier, 1997). Application of spatial aspects in the planning process can be found at Evans et al. (2007), who integrated land use (land use) and the characteristics of the transport system to a planning tool Transit Development System, or by Singh et al. (2014) who extended their work with other spatially significant criteria. Modelling a cycling infrastructure load is one of the current challenges which transport infrastructure planning is inherently associated with. A number of studies (Barnes and Krizek, 2005; Cardoso et al., 2012, Nordback et al., 2013, Porter et al., 1999, Turner et al., 1999) show, that neither the prediction of the number of cyclists passing the intersections cannot be generalized. We usually work with not too large data set and even so results with highly accurate estimation are expected. Intersection with passing cyclists represents points with a quantified attribute and two possible solutions are available: estimation of density and interpolation of fictitious (artificial) surface. Creating predictive models of bicycle volumes in urban areas is challenging as it is considered extremely complex issue. But they are widely used for safety analysis and calculation of crash risks, identifying priority locations for facility improvements or estimating changes in volumes following up infrastructure changes and new projects (Griswold et al., 2011). Modelling the load on the territory by cycling is one of the current challenges that transport infrastructure planning is inherently associated with. Many studies (Barnes and Krizek, 2005; Cardoso et al., 2012; Nordback et al., 2013; Porter et al., 1999; Turner et al., 1999) show that neither the prediction of the number of cyclists passing through individual intersections cannot be generalized but we usually have to work with not too large data set, so that modelling must be adapted to the results with high accuracy. Data on passing cyclists through intersections can be primarily represented as points with a quantified attribute, and two processing options are available: estimation of density and interpolation of fictitious (artificial) surface. Density calculation can be solved by the core estimation

method, with the Kernel Density function (Silverman, 1986) being used in GIS software, resulting in a smoothed surface. If the counting positions are equally distributed and real clusters cannot be considered as random locations suitable for density determination, interpolation is more appropriate. In the context of interpolation methods, the use of local, exact/approximate and stochastic interpolation methods (eg. IDW - Inverse Distance Weighted, RBF - Radial Base Function, Kriging, etc.) has been used, even though some authors have used in their studies global interpolation methods based on the principle of "classical" regression analysis (Griswold et al., 2011; Strauss and Miranda-Moreno, 2013).

### 3. RESEARCH METHODOLOGY

#### 3.1. Study area

The city of Přeřov is located in the east of the Czech Republic on the river Bečva (**Fig. 1**). Its geographical location makes it an intersection of ways and heart of Moravia. Nowadays Přeřov has approximately 44 000 inhabitants and is hospitable place for visitors.



**Fig. 1.** Cadastral area of Přeřov

Traditional cultural and sport events contribute to the development of tourism as well as comfortable accommodation and broad offer of restaurants and pubs. Currently, Přerov is the social, administrative and cultural centre of the District with developing economics. Cyclotourism is an appropriate mode of transport in Přerov, which is also illustrated by the newly built bicycle tower (storage for bicycles) at the city transport hub (bus and train station). Bicycle moves on marked cycling trails, on cycle paths or out of them. The absence of efficient and complete cycling infrastructure leads to the dispersal of cyclists across the city in an effort to find their safe route. There have been more than 25 km of cycle paths and cycling lanes in Přerov and its local parts, and more than 20 years have been invested only in the city budget of about 800 000 Euros. Support for cycle paths is also stated in the Strategic Plan for Territorial and Economic Development of the Statutory City of Přerov for the period 2014-2020. In addition, a Sustainable Urban Mobility Plan is currently being developed to address transport in a comprehensive manner in line with the needs of residents, city visitors and the business community to improve the quality of city life. With regard to the creation of materials for decision-makers, a methodology including both the collection of primary and secondary data editing, as well as data processing and the interpretation of results was proposed (Fig. 2).

### 3.2. Data collection and analysis

For the case study in the city of Přerov, data was primarily collected at 30 pre-selected counting points (intersections). Bicyclists were counted on two business days during the rush hour in the morning, 5:30 – 9:00 a. m. and in the afternoon, 2:00 – 5:00 p. m. The counting day was set Tuesday (10.5. 2016 and 21.5. 2016) as a business day, that does not precede neither follow the weekend days. The counting occurred twice in two following weeks in May. Volunteer counters marked numbers of arrivals and departures of cyclists in  $n$ -paths intersections, then partial average numbers were calculated. This way of counting provides not only the total number of cyclists at intersections but also cycling intensity for each direction. Input dataset completion (road and street network) was dealt with data provided by the Municipal Government of city of Přerov in the scale of 1:10 000 using ArcGIS for Desktop 10.5. Further data processing was realized in S-JTSK Krovak East North coordinate system.

The analysis and data evaluation stage included the identification of the criteria to be assessed. For the purpose of this study, we also considered the **surface of the cyclist's estimation of motion**, the **mapping error of the interpolated surface** and the **Euclidean distance from the POIs** (Points Of Interest). To choose an appropriate interpolation method local, exact/approximate method IDW (*Inverse Distance Weighted*), RBF (*Radial Base Function*), Simple Kriging, Ordinary Kriging, Universal Kriging and Empirical Bayesian Kriging were considered based on state of the art analysis as useful and examined. For selecting a suitable interpolation technique, the cross validation was taken into account comparing Mean Error (ME), Root Mean Square Error (RMSE), Root Mean Square Standardized Error (RMSSE) and Average Standard Error (ASE). For interpolation of bicycle movement estimation, data from counting points were used. This data was averaged so that an average number of cyclists per 1 hour during rush hour were calculated. Six methods were tested when choosing the most suitable interpolation method and significant statistical parameters of their results were compared. With regard to the two intended scenarios (traffic safety and suitability for construction), the parameters of the transport network were determined as weight and used in map algebra computation within multicriterial evaluation.

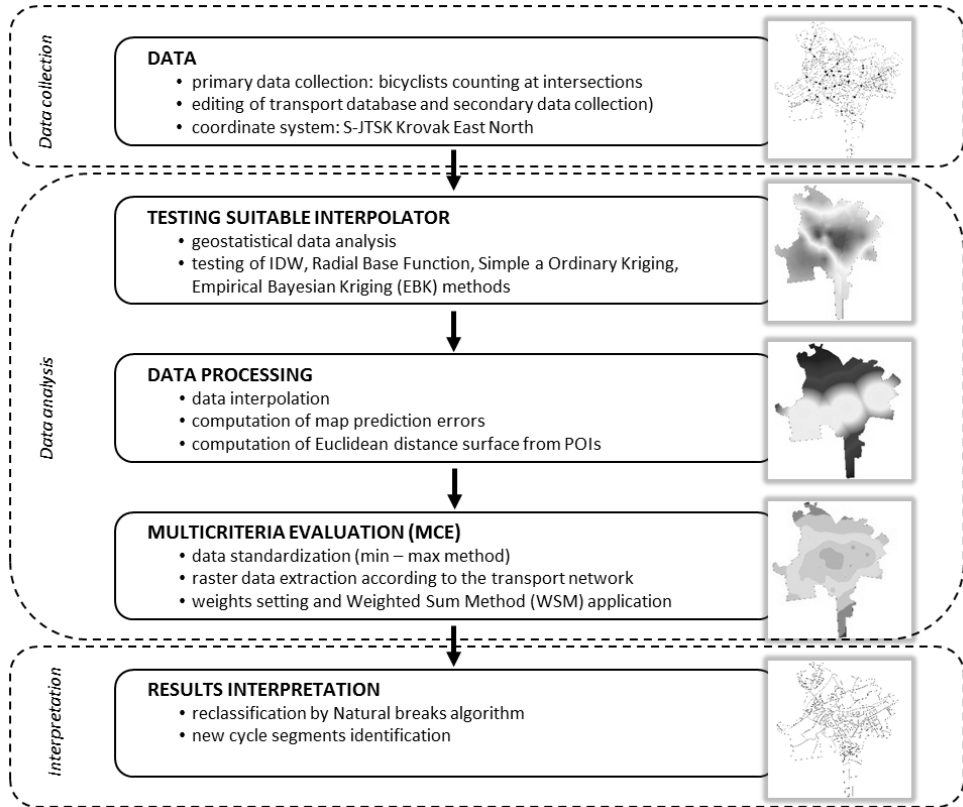


Fig. 2. Methodology flow chart

Data conversion from raster data model to vector data model was also applied. Data standardization was subsequently implemented using the linear range method (min - max), when the maximizing (1) and minimizing (2) criteria were differentiated using the calculation:

$$x'_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}}, \quad (1)$$

$$x'_{ij} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}}, \quad (2)$$

where  $x'_{ij}$  is the adjusted value for the  $i$ -th variant of the  $j$ -th criterion and  $x_{ij}$  is the input value.

The adjusted (standardized) value becomes from 0 to 1. For the two designed scenarios the ranking method was applied to set the weights. The most important criterion is the  $k$  number ( $k$  = number of criteria), the second  $k-1$ , the least important criterion refers to 1. Generally, the  $b_i$  value is assigned to the  $i$ -th criterion. The weight of the  $j$ -th criterion is calculated by (3):

$$v_i = \frac{b_i}{\sum_{i=1}^k b_i}, \quad \sum_{i=1}^k b_i = \frac{k(k+1)}{2}. \quad (3)$$

Weighted Sum Method (WSM) was used for calculation of the weighted result within each scenario individually (4).

$$WSM = v_i \sum_{j=1}^n b_j, \quad (4)$$

where  $v_i$  represents the  $i$ -th weight and  $b_i$  is the standardized value of the  $j$ -th criterion.

The output surface from MCE calculation was statistically examined and according to data distribution Natural breaks algorithm was used for reclassification into 5 classes representing the range from the less suitable to the most suitable candidates.

## 4. RESULTS

### 4.1. Intensity of cyclists' movement

An ordinary geostatistical data processing necessary for testing interpolation techniques comprises four parts (1. data analysis, 2. structural analysis, 3. applying interpolation technique and 4. evaluation) and is being understood as an iterative procedure, especially between parts 2 - 4). During data analysis basic exploratory statistic data analysis was completed. Relatively diagonal direction with lightly significant deviation in Q-Q plot and together with featureless asymmetry (skewness: 0,35), close mean (76,1) and median (78,1) values localized approximately between minimum (24,1) and maximum (139,5) indicate normal distribution. It was also confirmed by Shapiro-Wilk normality test that the data came from a normally distributed population.

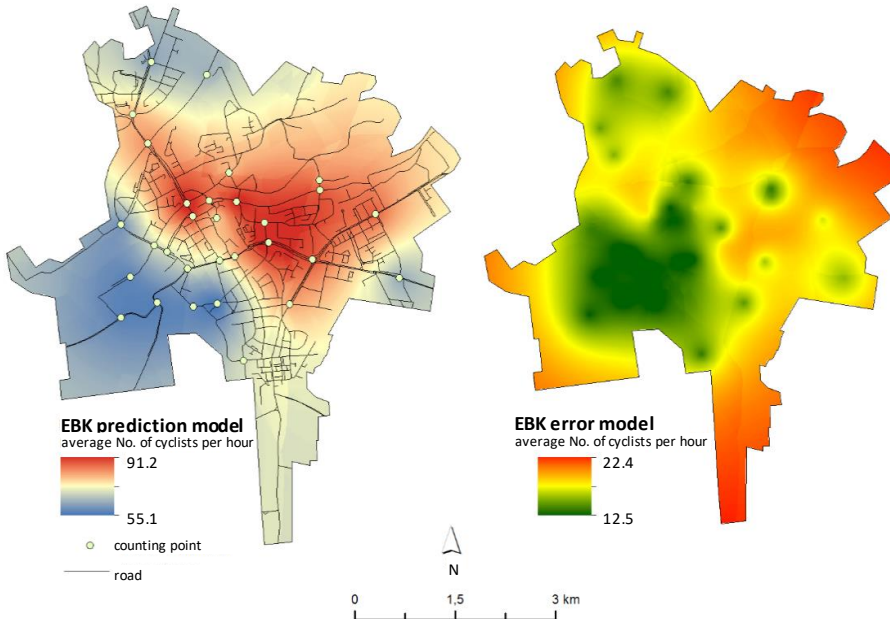
For comparative analysis, seven settings (IDW, EBK, RBF, Ordinary Kriging (OK) with spherical variogram, Ordinary Kriging (OK) with linear variogram, Simple Kriging (SK) with spherical variogram and Simple Kriging (SK) with linear variogram) for data set interpolation have been considered based on the state of the art. Although Universal Kriging (with spherical or linear variogram) is recommended as a useful tool, generated surfaces did not reflect movement of cyclists and this technique has not been applied in the study. Cross-validation was followed in order to compare measured errors and two residues in two farthest points (ID 3 and ID 25; ID means identifier) (**Table 1**). Two best results in evaluated criteria were typed in bold. The evaluation shows that IDW and RBF have worst results although IDW was the most accurate in prediction of point ID 25 which is caused by computing average values in predicted surface. EBK and OK (spherical variogram) achieved very similar values. Both have smallest ME and ASE is relatively close to RMSE which declares suitable variability of predicted values. These two methods will be compared in detail. Other methods gave satisfactory results but with worse errors than EBK and OK (spherical variogram). Cross-validation results for EBK and OK (spherical variogram) showed that both methods overestimate smaller values and underestimate higher values. The same was determined at both Simple Kriging methods. Opposite EBK, OK (spherical variogram) has regression line closer to theoretically needed diagonal line (see more in Ruda and Floková, 2017).

Table 1

Cross-validation results for tested interpolators

<i>method / error</i>	<i>ME</i>	<i>RMSE</i>	<i>RMSSE</i>	<i>ASE</i>	<i>ID 3 residue</i>	<i>ID 25 residue</i>
<i>IDW (coefficient= 2)</i>	3.8673	32.9553	-	-	48.9	<b>0.05</b>
<i>RBF (multiquadric)</i>	2.8621	45.7472	-	-	48.91	<b>-38.4</b>
<i>EBK</i>	<b>0.7395</b>	32.767	<b>0.9684</b>	<b>33.9832</b>	-23.72	53.69
<i>OK (spherical var.)</i>	<b>0.894</b>	32.0088	<b>0.9483</b>	<b>33.9183</b>	-21.53	43.58
<i>OK (linear var.)</i>	0.9536	32.1266	0.9475	34.0735	-25.49	42.27
<i>SK (spherical var.)</i>	2.1803	<b>31.8536</b>	0.9095	34.5406	<b>-19.88</b>	46.06
<i>SK (linear var.)</i>	2.1522	<b>31.286</b>	0.9052	34.6201	<b>-20.27</b>	46.39

Deeper analysis of individual variograms and generated prediction surfaces revealed significant differences. Variogram of EBK was set as a result of 100 simulations with standard circular neighborhood type. Its empirical variogram oscillates along theoretical variogram with higher amplitude which can increase the prediction. The EBK prediction map (**Fig. 3**) estimates higher values in the middle of interpolated area and creates graded zones of smaller values southward from the centre of the city. The surface is relatively smooth and does not estimate enclaves of unexpected values. Farthest areas are estimated without significant changeability. Prediction surface is less smoothed but estimates farthest input points (see ID 3 and ID 25 in **Table 1**) with higher accuracy. Densest area of input points is realistically graded in order to cyclists' movement. Partial uncertainty of predictions can be seen in southwest corner where values without input points are estimated not quite good for cyclists' movement.



**Fig. 3.** Prediction map of EBK and map of prediction errors



Actual model slightly overestimates the variability of estimated values, but provides the most accurate estimate opposite to other interpolators. From the results of the prediction model, it can be seen that the most intense movement of cyclists is estimated in the city centre with a gradual decline towards the east (Meopta's headquarters – the biggest employer) and a slower downward trend to the south (industrial zone). At the northern (the Předmostí suburb area) and the southwestern outskirts (Precheza's – chemical industry – and Dalkia's industrial zone), the decline is even more pronounced.

The biggest estimation error is indicated especially in the marginal parts and in the eastern half of the area outside the entry points. The benefit of the error model is to compensate for a higher error estimate, as we can see for example in the southern tip of the territory.

Significant points of interest (POIs) have been chosen with the premise that during days of counting cyclists are heading mostly to work and for services these days: the business centre, bicycle storage, the area of important employers – Dalkia (thermal power plant), Precheza (chemical industry), Meopta (optical industry) and declining Přerov engineering industry. In a simplified form, the surface of the Euclidean distance was calculated from these points to favour those communications that are closer.

#### 4.2. Multicriterial evaluation

The suitability of the individual segments was realized using a weighted sum of the standardized values of the above mention criteria (derived surfaces). The calculation was carried out over the pixels above which considered traffic segments pass. Only those communications that could be used for the development of the cycle paths were included (1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> class roads, roads with cycling parameters, residential connections and non-differentiated roads).

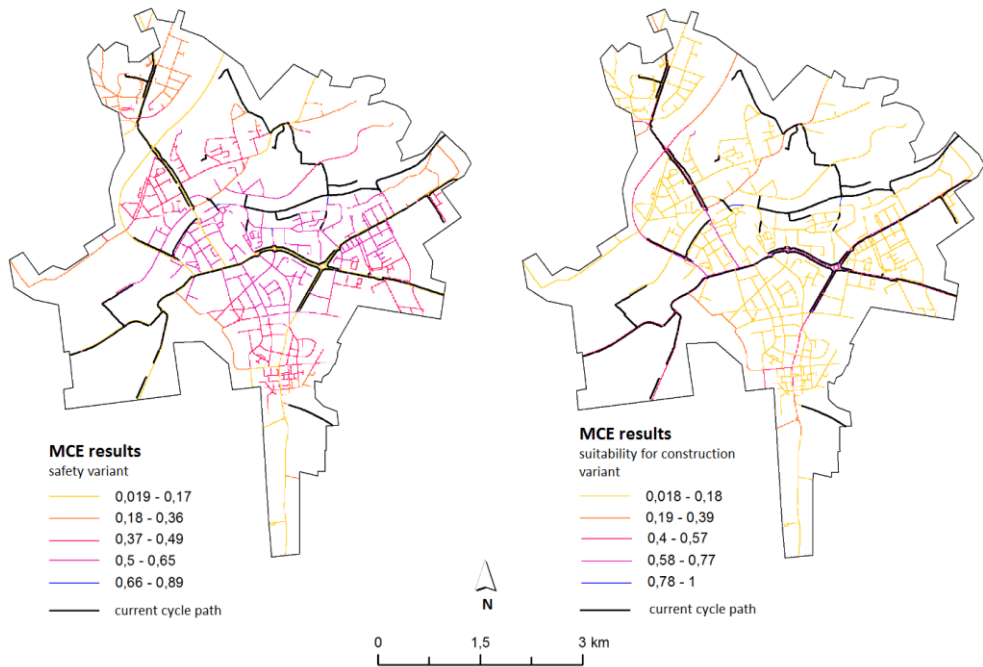
Pavements with inappropriate parameters, security routes, connections with staircases and roads intensively used for lorry transport were excluded. The type of these communications served to determine the weights for the safety and suitability for construction by the ranking method (Table 2).

Table 2

Weights setting for types of communications

<i>criteria</i>	<i>safety variant</i>			<i>suitability for construction variant</i>		
	<i>k value</i>	<i>calculation</i>	<i>weight</i>	<i>k value</i>	<i>calculation</i>	<i>weight</i>
1 <sup>st</sup> class road	1	1/56	0.02	4	10/35	0.285
2 <sup>nd</sup> class road	2	3/56	0.05	4	10/35	0.285
3 <sup>rd</sup> class road	3	6/56	0.11	3	6/35	0.17
residential connect.	5	15/56	0.27	1	1/35	0.03
cycle paths	6	21/56	0.37	5	15/35	0.43
nondifferentiated	4	10/56	0.18	2	3/35	0.085

The **safety** variant aims to identify the potential of each segment for the safety of riding, while **suitability for construction** variant favours segments with inexpensive construction (Fig. 4).



**Fig. 4.** MCE results.

While the safety variant favours segments of communications that are not used by cars and preserves rather the importance of the cyclists' movement prediction, the suitability for construction variant is based on the assumption that current communications will be also used for cycle path construction.

For comparison, the results of both variants were reclassified by the same algorithm (Fig. 5). At first glance, it is clear that safety variant defines much more moderate sections than suitability for construction variant, which separates the least suitable sections from the most appropriate ones. For the correct interpretation, only segments rated as satisfactory or very satisfying (the last two highest rated categories) were selected for further selection. These segments were identified by a sequence number, and it is clear from the results of both variants that segments 1, 3, 5 and 7 were identically identified in both variants.

Another beneficial basis for decision making is also a diagram map representing the number of cyclists arriving at intersections (counting points).

Fig. 6 illustrates the sum of arriving cyclists (during one day) from all examined directions. It is clear from the map, that the busiest parts are already covered by existing cycling infrastructure. In a more detailed perspective, it would be advisable to recommend not only the construction of cycle paths in the segments mentioned above, but also the extension of segments 6, 9 and 13 to join the city center, and segments 14-16 which would connect the southern industrial part of the city to the existing cycle network. Only segments 10 – 11 should be excluded for now, because they run along the frequented first class road and possible cycle path constructions would include would involve a major extensive reconstruction of the road.

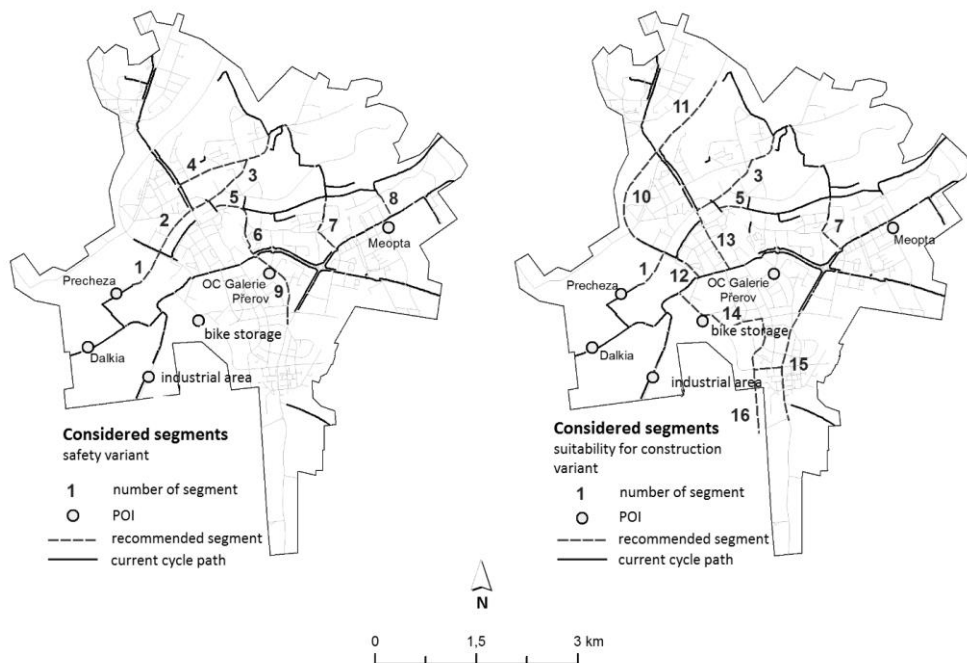


Fig. 5. Recommended segments for cycle path construction.

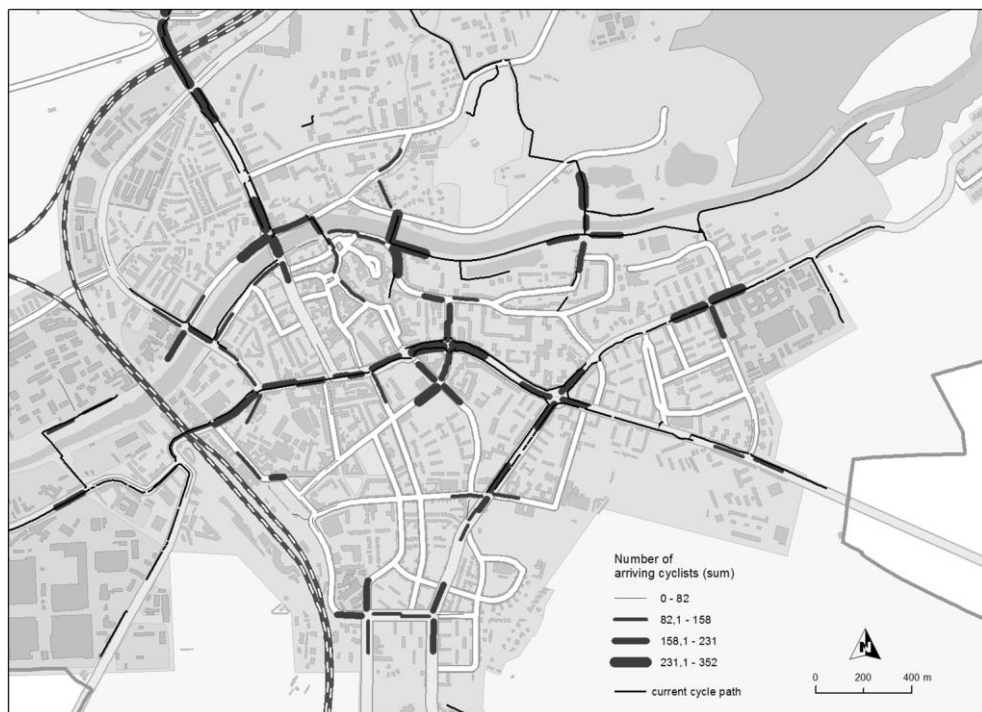


Fig. 6. Loading of intersections by arriving cyclists in the city center.

## 5. CONCLUSIONS

Predicting cyclist's movement as an example of socioeconomic activity is very difficult but also very challenging task. This kind of activity is associated with attributes describing both the interest of cyclists and the existing infrastructure of cycle paths or any other line transport infrastructure. Counting points providing input data are irregularly distributed with prevailing clusters and vast unmeasured areas. Therefore, estimating needed values is loaded by a high uncertainty. The construction of cycling infrastructure has been solved in Přerov city for a long time. There is no clear unambiguous methodology from the currently available studies. The design of factors, individual criteria and their indicators, including preferences, cannot be determined in advance with regard to the consensual agreement of city representatives. In many ways, it would help to identify the factors that affect the development of cycling. It is also appropriate to involve participatory mapping that takes into account citizens' needs and qualitatively oriented studies that will allow individual preferences to be better evaluated. Spatial decision-making support in GIS will enable decision makers to have a more detailed and visually beneficial view of the solved problem and facilitate work that requires spatial planning and decision making (González, M.J. and Pascual M.E., 2016).

The use of the above mentioned methods is very beneficial in relation to the solved problems, but it is also characterized by some degree of uncertainty. The creation of surface of cyclists' movement provides the required information, but does not sufficiently take into account a number of socio-economic characteristics that may affect the movement of cyclists (population deployment, all barrier movements, etc.). The solution could be the use of geographically weighted regression, which both determine the leading factors and consider their impact in the final output. The linear regression model is recommended in applications where precious estimates of average daily traffic are desired, the logistics regression model for applications where level of traffic is desired as it enables predicting the probability of a road belonging to one of traffic volume thresholds. In spite of the fact that non-motorised traffic is influenced by many other factors than motorised there is a matter of correcting generally used methods for non-motorised count. But this approach is data-intensive. Lowry and Dixon (2012) presented own GIS tool to estimate average annual daily traffic based on ordinary least square regression employing functional classification, number of lanes and connectivity importance index as explanatory variables which were found significant for the case study of Moscow, Idaho presented by them. Log linear least square regression was used in the study of the Alameda County, California (Griswold et al., 2011). Dependent variables were transformed using natural logarithm before creating the model. Comparing to negative binomial form of regression model which was also tested in this study similar model coefficient was obtained but log linear form is easier to apply and interpret. Besides bicycle counts there were variables referring to intersection site characteristic (average slope, bicycle lanes and markings), surrounding land use (number and network distance to certain points of interests) and surrounding transportation system (intersection density, connectivity) included in the modelling process.

Another uncertainty lies in the design of individual scenarios, which were built on the priorities of city representatives. Even in this case, it is very important to choose the individual factor scales and to perform a sensitivity analysis at the same time. In both cases, the above-mentioned scenario proposals and the construction of new segments were based on the data from the census (frequency, dominant directions used by cyclists, intersections profile) and the distribution of key points of interest.

However, it is clear that, with regard to city policy, only one design of suitable segments for the construction of cycling infrastructure cannot be created, but a list of possible solutions, which the city council will subsequently determine or refuse to build, is needed.

## REFERENCES

- Barnes, G. & Krizek, K. (2005) Estimating Bicycling Demand. *Transportation Research Record, Journal of the Transportation Research Board*, 1939, 45-51.
- Besip. (2011) *National Road Safety Strategy 2011-2020*. Available from: <http://www.ibesip.cz/cz/besip/strategicke-dokumenty/narodni-strategie-bezpecnosti-silnicniho-provozu/nsbsp-2011-2020> [Accessed September 2018].
- Broach, J., Dill, J. & Gliebe, J. (2012) Where do cyclists ride? A route choice model developer with revealed preference GPS data. *Transportation Research Part A*, 46, 1730-1740.
- Brüde, U. & Larsson, J. (1993) Models for Predicting Accidents at Junctions Where Pedestrians and Cyclists are Involved. How Well Do They Fit? *Accident Analysis and Prevention*, 25 (5), 499-509.
- Cardoso, O.D., García-Palomares, J.C. & Gutiérrez, J. (2012) Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Applied Geography*, 34, 548-558.
- Dardanelli, G., Marretta, R., Santamaria, A.S., Strega, A., Lo Brutto, M. & Maltese, A. (2017) Analysis of technical criticalities for GIS modelling an urban noise map. *Geographia Technica*, 12 (2), 41 – 61.
- Di Piazza, A., Lo Conti, F., Noto, L.V., Viola, F. & La Loggia, G. (2011) Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. *International Journal of Applied Earth Observation and Geoinformation*, 13 (3), 396-408.
- Donkwook, L., Jinsul, K. & Minsoo, H. (2014) Density Map Visualization for Overlapping Bicycle Trajectories. *International Journal of Control and Automation*, 7 (3), 327-332.
- Esawey, M., Limel, C., Sayed, T. & Mosa, A.I. (2013) Development of daily adjustment factors for bicycle traffic. *Journal Transportation Engineering*, 139 (8), 859–871.
- European Commission. (2011) *White paper*. Available from: <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52011DC0144> [Accessed September 2018].
- Evans IV, J.E., Pratt, R.H., Stryker, A. & Kuzmyak, J.R. (2007) *Transit-Oriented Development -- Traveler Response to Transportation System Changes*. Transit Cooperative Research Program (TCRP) Report 95. Washington, Transportation Research Board.
- González, M.J. & Pascual M.E. (2014) Strategic spatial thinking and change. *European Journal of Geography*, 7 (1), 24-35.
- González, M.J. & De Lázaro, M.L. (2013) Strategic planning and sustainable development in Spanish cities. *European Journal of Geography*, 4 (1), 48-63.
- Griswold, J.B., Medury, A. & Schneider, R.J. (2011) Pilot Models for Estimating Bicycle Intersection Volumes. *Transportation Research Record, Journal of the Transportation Research Board*, 2247 (1), 1-7.
- Griswold, J.B., Medury, A. & Schneider, R.J. (2011) *Pilot Models for Estimating Bicycle Intersection Volumes*. Available from: <http://escholarship.org/uc/item/380855q6>
- Lindsey, G., Nordback, K. & Figliozzi, M.A. (2014) *Institutionalizing Bicycle and Pedestrian Monitoring Programs in Three States, Progress and Challenges*. Available from: [http://www.pdx.edu/ibpi/sites/www.pdx.edu/ibpi/files/14-4181\\_in\\_Compendium.pdf](http://www.pdx.edu/ibpi/sites/www.pdx.edu/ibpi/files/14-4181_in_Compendium.pdf) [Accessed September 2018].

- Lowry, M. & Dixon, M. (2012) *GIS Tools to Estimate Average Annual Daily Traffic*. Available from: [http://ntl.bts.gov/lib/45000/45900/45948/KLK725\\_N12-03.pdf](http://ntl.bts.gov/lib/45000/45900/45948/KLK725_N12-03.pdf) [Accessed September 2018].
- Lowry, M., McGrath, R., Cool, S., Cook, R. & Skiles, M. (2013) Data Collection and Spatial Interpolation of Bicycle and Pedestrian Data. Available from: [http://depts.washington.edu/pactrans/wp-content/uploads/2016/01/UI\\_KLK853\\_BikePed\\_FinalReport\\_Lowry.pdf](http://depts.washington.edu/pactrans/wp-content/uploads/2016/01/UI_KLK853_BikePed_FinalReport_Lowry.pdf) [Accessed September 2018].
- Ludnberg, B. & Weber, J. (2016) Non-motorized transport and university populations, an analysis of connectivity and network perceptions. *Journal of Transport Geography*, 39, 165-178.
- Milakis, D. & Athansopoulos, K. (2014) What about people in cycle network planning? applying participative multicriteria GIS analysis in the case of the Athens metropolitan cycle network. *Journal of Transport Geography*, 35, 120-129.
- Ministry of Industry and Trade (MIT). (2010) Strategic Framework for Sustainable Development of the Czech Republic (*Strategický rámeč udržitelného rozvoje České republiky*). Available from: <http://www.mpo.cz/dokument71639.html> [Accessed September 2018].
- Ministry of Regional Development (MRD). (2015) Territorial Development Policy of the Czech Republic (*Politika územního rozvoje České republiky*). Available from: <http://www.strukturalni-fondy.cz/getmedia/fcf05cb6-a4b9-4af8-ac00-896bfd1b7e32/Politika-uzemniho-rozvoje-aktualizace-c-1.pdf?ext=.pdf> [Accessed September 2018].
- Ministry of Transport (MT), Transport Research Centre (TRC). (2013) Transport Policy of the Czech Republic for the period 2014-2020 with a view to 2050. (*Dopravní politika ČR pro období 2014-2020 s výhledem do roku 2050*). Available from: <https://www.databaze-strategie.cz/cz/md/strategie/dopravni-politika-cr-pro-obdobi-2014-2020-s-vyhledem-do-roku-2050> [Accessed September 2018].
- Nordback, K., Marshall, W.E. & Janson, B.N. (2013) Development of Estimation Methodology for Bicycle and Pedestrian Volumes Based on Existing Counts. Available from: <http://www.coloradodot.info/programs/research/pdfs> [Accessed September 2018].
- Nosal, T. (2014) Improving the accuracy of bicycle AADT estimation, temporal patterns, weather and bicycle AADT estimation methods. Available from: [http://digitool.library.mcgill.ca/webclient/StreamGate?folder\\_id=0&dvs=1472369098397~876&usePid1=true&usePid2=true](http://digitool.library.mcgill.ca/webclient/StreamGate?folder_id=0&dvs=1472369098397~876&usePid1=true&usePid2=true) [Accessed September 2018].
- Papafragkaki, A. & Photis, Y.N. (2014) GIS-based location analyses of administrative regions. Applying the median and covering formulations in a comparative evaluation framework. *European Journal of Geography*, 5 (3), 37-59.
- Paul, S. (2013) Analysis of tourism attractiveness using probabilistic travel model, A study on Gangtok and its surroundings. *European Journal of Geography*, 4 (2), 46-54.
- Porter, C., Suhrbier, J. & Schwartz, W.L. (1999) Forecasting Bicycle and Pedestrian Travel State of the Practise and Research Needs. *Transportation Research Record*, 1674, 94-101.
- Roll, J.F. (2013) Bicycle Traffic Count Factoring, An Examination of National, State and Locally Derived Daily Extrapolation Factors. Available from: [http://pdxscholar.library.pdx.edu/open\\_access\\_etds/998](http://pdxscholar.library.pdx.edu/open_access_etds/998) [Accessed September 2018].
- Ruda, A. & Floková, L. (2017) Interpolation techniques for predicting the movement of cyclists. In *Lecture Notes in Geoinformation and Cartography*, eds. Ivan, I., Horák, J. and Inspektor, T., 383-398. Cham, Springer.
- Ruda, A. (2016) Exploring tourism possibilities using GIS-based spatial association methods. *Geographia Technica*, 11 (2), 87-101.
- Rybarczyk, G. & Wu, C. (2010) Bicycle Facility Planning Using GIS and Multi-criteria Decision Analysis. *Applied Geography*, 30, 282-293.

- Schneider, R., Patten, R. & Toole, J. (2005) Case Study Analysis of Pedestrian and Bicycle Data Collection in U.S. Communities. *Transportation Research Record, Journal of the Transportation Research Board*, 1939 (1), 77-90.
- Shafizadeh, K. & Niemeier D. (1997) Bicycle Journey-to-Work Travel Behavior Characteristics and Spatial Attributes. *Transportation Research Record, Journal of the Transportation Research Board*, 1578, 84-90.
- Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*. New York, Chapman and Hall.
- Singh, Y.J., Fard, P., Zuidgeest, M., Brussel, M. & Van Maarseveen, M. (2014) Measuring transit oriented development, a spatial multi criteria assessment approach for the City Region Arnhem and Nijmegen. *Journal of Transport Geography*, 35, 130-143.
- State Transport Infrastructure Fund (STIF). (2013) *Cyklostrategie 2013*. Available from: <http://www.cyklodoprava.cz/finance/statni-fond-dopravni-infrastruktury> [Accessed September 2018].
- Strauss, J. & Miranda-Moreno, L.F. (2013) Spatial Modeling of Bicycle Activity at Signalized Intersections. *The Journal of Transport and Land Use*, 6 (2), 47-58.
- Strauss, J., Miranda-Moreno, L.F. & Morency, P. (2015) Mapping cyclist activity and injury risk in a network combining smartphone GPS data and bicycle counts. *Accident Analysis and Prevention*, 83, 132-142.
- Turner, S., Hottenstein, A. & Shunk, G. (1999) Bicycle and pedestrian travel demand forecasting, Literature review. Available from: <http://d2dtl5nnlpfr0r.cloudfront.net/tti.tamu.edu/documents/1723-1.pdf> [Accessed Sept. 2018].
- Xu, Y., Shaw, S., Fang, Z. & Yin, L. (2016) Estimating Potential Demand of Bicycle Trips from Mobile Phone Data—An Anchor-Point Based Approach. *ISPRS International Journal of Geo-Information*, 5 (8), 131-154.
- Zaki, M.H., Sayed, T. & Cheung, A. (2013) Automated Collection of Cyclist Data Using Computer Vision Techniques. Available from: <http://docs.trb.org/prp/13-0745.pdf> [Accessed September 2018].

## **CHARACTERIZATION OF HEAT WAVES: A CASE STUDY FOR PENINSULAR MALAYSIA**

*Wayan SUPARTA<sup>1\*</sup> and Ahmad Norazhar Mohd YATIM<sup>2</sup>*

DOI: 10.21163/GT\_2019.141.11

### **ABSTRACT:**

The present work aims to investigate the characteristics of heat waves in Peninsular Malaysia based on the Excess Heat Factor (EHF) Index. This index was calculated based on the daily maximum and minimum temperatures over nine meteorological stations in Peninsular Malaysia during the period 2001 to 2010. The selected station is representing all of the states in Peninsular Malaysia. Statistical analysis found that the highest of the EHF happened at the Kuala Lumpur station in 2002 with an index of  $9.1^{\circ}\text{C}^2$  and the lowest was in Alor Setar in 2006 with an index of  $0.1^{\circ}\text{C}^2$ . The EHF moderate was found at Kuantan with an index of  $4.2^{\circ}\text{C}^2$ . Moreover, the longest heat wave with 24 days has happened in Ipoh, Perak with amplitude of  $29.4^{\circ}\text{C} - 33.0^{\circ}\text{C}$ . Most of the heat wave characterized in Malaysia occurred during the El Nino events especially moderate El Nino in 2002 until 2005, and 2010. The Southeast, northeast and west part of Malaysia experience the highest average heat wave activity. These results indicated that the heat wave conditions in Peninsular Malaysia are anxious and this requires immediate investigation because it has a direct impact on agriculture, particularly health, economic, and human being.

**Keywords:** *Heat wave, Excess heat factor, Climate change, Characteristics, Peninsular Malaysia.*

## **1. INTRODUCTION**

Heatwave is one of the most threatening natural hazards and can adversely affect ecosystem, infrastructure, human health, and social life (Zuo et al., 2014). Populations are very vulnerable to changes in heat wave attributes and extreme heat wave events can increase human morbidity and mortality rates (Anderson & Bell, 2011). Previous studies show that heat wave is responsible for more deaths than all other natural hazards in Australia (Coates et al., 2014) and climate risk in Romania (Bocancea, 2018). Although there is no standard definition of the heat wave, it can be referred as a period of consecutive days of abnormal weather. Heat waves are increasing globally due to the effects of climate change. Numerous studies have indicated that climate change is expected to exacerbate the heat wave events, particularly more frequent in their duration and intensity (Coumou & Rahmstorf, 2012). Asia region also experiences adverse effects of climate change. The Intergovernmental Panel on Climate Change (IPCC, 2013) states that South Asian countries, will be at the greatest risk for the emergence of heat waves. The heat wave impacts could be enlarged and significant from region to region, for example in developing countries such as Malaysia.

The heat wave event is one of the uncommon natural hazards observed and has had a significant impact on Malaysia. However, the heat wave mitigation has not been taken seriously by most governments and non-governmental organization (NGOs) working on

---

<sup>1</sup> *Pembangunan Jaya University, Department of Informatics, South Tangerang City, Banten 15413, Indonesia, Corresponding author: wayan.suparta@upj.ac.id*

<sup>2</sup> *Universiti Malaysia Sabah, Kota Kinabalu, 88400 Sabah, Malaysia, norazhar@ums.edu.my*



monitoring and disaster risk reduction. The impact of heat wave in Malaysia is currently under-reported and it is difficult to assess information such as morbidity, mortality, and economic consequences. Although the heat wave events in Malaysia have not been extensively investigated, in recent years there have been a number of studies focusing on climate change and temperatures. Wai et al. (2005) used 50 years (1951-2001) of temperature data set to study the global warming trend in Malaysia and found that the annual mean temperature increased from 0.99 to 3.44°C/100 years. A study by Makerami et al. (2012) found that an acceptable thermal condition in outdoor spaces of a hot and humid climate of Malaysia is less than 34°C where the comfort condition is in the early morning (9-10 am) and late afternoon (4-5 pm). Mohd Salleh et al. (2015) also found that most of the stations across Peninsular Malaysia showed an inclination toward a temperature above the annual mean surface temperature of 26°C to 28°C and with high annual precipitation values (1200–2400 mm). From the preliminary studies above provide clues that the investigation of heat wave and their characteristics will be useful to improve human understanding and awareness.

This study will focus on the analysis of heat wave characteristics in Peninsular Malaysia from 2001 until 2010. We choose an Excess Heat Factor (EHF) method across different climates, which is an ideal method to normalize the climatology variation in heat wave from a hazard point of view (Perkins & Alexander, 2013). This method is introduced to create a universal method for heat wave definition and measurement (Nairn et al., 2009). The EHF index capable provides results which more focused on the pattern of heat wave frequency, duration, amplitude, and a number of days in the past decade. The results of the investigation of heat wave will provide a better understanding and importance of climatology events and early warning systems especially during extreme weather in Malaysia.

## 2. DATA AND METHODS

Section 2.1 describes details about the data set used in this study and Section 2.2 describes the method conducted to determine the heat wave events and characterization.

### 2.1 DATA

The daily maximum temperature ( $T_{\max}$ ) and minimum temperature ( $T_{\min}$ ), both in °C units for the period 2001-2010 were obtained from the Malaysian Meteorological Department (MetMalaysia). Data on temperature were based on the availability of the meteorological data at nine stations across Peninsular Malaysia. The meteorological station in Johor Bahru and Malacca is represent the Southwest of Peninsular Malaysia; Kuala Lumpur and Perak is for the West; Penang and Perlis is for the Northwest; Pahang, Kelantan, and Terengganu is for the East. All the selection stations represent the urban, suburban, and industrial area. In this study, we focused on Peninsular Malaysia for the year 2001-2010 where the previous study has shown that there were El Nino events recorded globally during this study period. **Table 1** compiling the detail of each station and Figure 1 depicts the location of the stations.

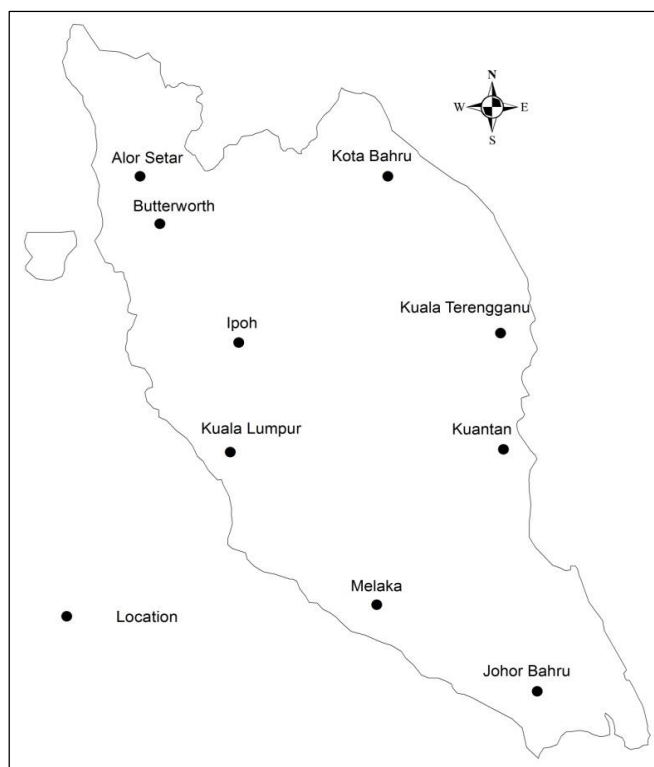
The figure shows that Peninsular Malaysia or West Malaysia is located in Southeast Asia between 1°N - 6°N and 101°E - 105°E, which covers an area of 130,598 km<sup>2</sup>. The climate of Peninsular Malaysia is characterized by two monsoon regimes namely the Southwest Monsoon (SWM) and the Northeast Monsoon (NEM). The SWM usually influenced by low-level southwesterly winds begins in May and ends in August. For NEM,

it is dominated by northeasterly winds that cross over the South China Sea. This season usually begins in November and usually lasts between 3-4 months the following year in February. During NEM, the exposed areas on the eastern part of Peninsular Malaysia receive heavy rainfall while the SWM is a drier period for the whole country, particularly for the other states of the west and north coast of Peninsular Malaysia. For a larger view, Peninsular Malaysia surrounded by two large oceans which are the Pacific Ocean to the east and the Indian Ocean to the west. This situation makes Peninsular Malaysia climate strongly influenced by natural climate variability associated with these oceans (Tangang et al., 2012).

**Table 1.**

Nine selected stations in Peninsular Malaysia

No.	Station	Region	Classification	Longitude	Latitude
1	Johor Bahru	Southwest	Industrial	103° 40' E	1° 38' N
2	Malacca	Southwest	Suburban	102° 15' E	2° 16' N
3	KLIA	West	Urban	101° 42' E	2° 43' N
4	Ipoh	West	Urban	101° 05' E	3° 18' N
5	Butterworth	Northwest	Suburban	100° 16' E	5° 18' N
6	Alor Setar	Northwest	Suburban	100° 24' E	6° 12' N
7	Kota Bharu	East	Urban	102° 18' E	6° 10' N
8	Kuala Terengganu	East	Suburban	103° 06' E	5° 23' N
9	Kuantan	East	Industrial	102° 15' E	3° 30' N

**Fig. 1.** Locations of nine selected stations in Peninsular Malaysia.

## 2.2 METHODS

There are a number of indices to determine the heat wave that has been developed based on the difference parameter. In this study, we chose the EHF index method to determine the heat wave events, which was developed by Nairn and Fawcett (2013). The EHF was determined based on the combined effect of two excessive heat indices, i.e. excess heat index significance ( $EHI_{sig}$ ) and excess heat index acclimatization ( $EHI_{accl}$ ) (Nairn and Fawcett, 2015). The  $EHI_{sig}$  defined as an unusually high heat arising from a high daytime temperature to unusually high overnight temperature. It can be measured by comparing directly three days period (TDP) of daily mean temperature (DMT) against a climate reference annual temperature (95<sup>th</sup> percentile). For the calculation 95<sup>th</sup> percentile, we used 10 years period climate reference at each particular location. From the results, if the TDP average is higher than the climate reference, then each day in this period is marked as unusual warm or excess heat event. The units of  $EHI_{sig}$  are degree Celsius ( $^{\circ}C$ ) and the equation is given by (Nairn & Fawcett, 2013).

$$EHI_{sig} = \frac{(T_i + T_{i+1} + T_{i+2})}{3} - T_{95} \quad (1)$$

where  $T_i$  is the DMT on day  $i$  and  $T_{95}$  is the 95<sup>th</sup> percentile of DMT for the climate reference period of 2001–2010. In addition to  $EHI_{sig}$ , DMT is the average of  $T_{max}$  and  $T_{min}$  as given by

$$T = (T_{max} + T_{min})/2 \quad (2)$$

The second heat index is  $EHI_{accl}$  which defined as a period of heat that is warmer than the recent past (on average). Acclimatization to higher temperatures depends on the characteristics of human physical adaptation that takes between two and six weeks, which involved adjustment of physiological cardiovascular, endocrine, and the renal systems (Nairn & Fawcett, 2015). In this index, 30 days has been used as the period required for determining acclimatization.  $EHI_{accl}$  can be measured by determining the difference between the same three-day-averaged DMT compared against the average DMT over the preceding 30 days. The units of  $EHI_{accl}$  is also in  $^{\circ}C$  and the index is given by (Nairn & Fawcett, 2013).

$$EHI_{accl} = \frac{(T_i + T_{i+1} + T_{i+2})}{3} - \frac{(T_{i-1} + \dots + T_{i-30})}{30} \quad (3)$$

where  $T_i$  is the DMT on day  $i$ .  $EHI_{accl}$  is an anomaly of three day of DMT with respect to the previous 30 days. Positive  $EHI_{accl}$  means the three days are warmer than the recent past (on average). Then, the EHF as in equation (4) is calculated based on the combination of these two indices, i.e.  $EHI_{sig}$  and  $EHI_{accl}$ , which provides a comparative measure of frequency, duration, amplitude, and spatial distribution of a heat wave event. The unit of EHF is  $^{\circ}C^2$ .

$$EHF = EHI_{sig} \times \max(1, EHI_{accl}) \quad (4)$$

where 1 in equation 4 is basically necessary to make a small positive value at least for a short heat wave. EHF calculation incorporates the effects of humidity on heat tolerance

indirectly by using the mean, rather than the maximum daily temperature. It provides a comparative measure of intensity, load, duration and spatial distribution of a heat wave event, and has a strong signal-to-noise ratio (Guyton & Hall, 2000). Universal understanding the impact of a heat wave on human health are varied, but mostly shows that vulnerable population for the following three days is more sensitive and affected (Keggenhoff et al., 2014). As a result, the heat wave is defined as a period of at least three days with  $\text{EHF} > 0$  (positive value) with the combination of excess heat and heat stress with respect to the local climate<sup>13</sup>. Then, we use the EHF index results to calculate annual values of four heat wave characteristics based on Fischer and Schar (2006) as below.

1. The heat wave amplitude (HWA) is the peak of EHF value from the hottest heat wave event of the year.
2. The heat wave number (HWN) is the annual number of heat wave events.
3. The heat wave frequency (HWF) is the annual frequency of days contributing to the heat wave events (the sum of participating heat wave days per year).
4. The heat wave day (HWD) is the duration of the longest annual heat wave event (must be  $\geq 3$  days).

To characterize a spatial distribution of each heat wave in Peninsular Malaysia, we used the inverse distance weighted (IDW) interpolation method in ArcGIS version 10.3 software. ArcGIS software provides tools like spatial analysis tools for spatial data analysis either raster or vector data that apply statistical theory and technique. Figure 2 shows the overall methodology that has been conducted in this study. The next section discusses the result and discussion of the finding.

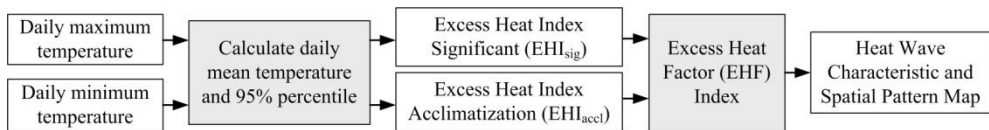
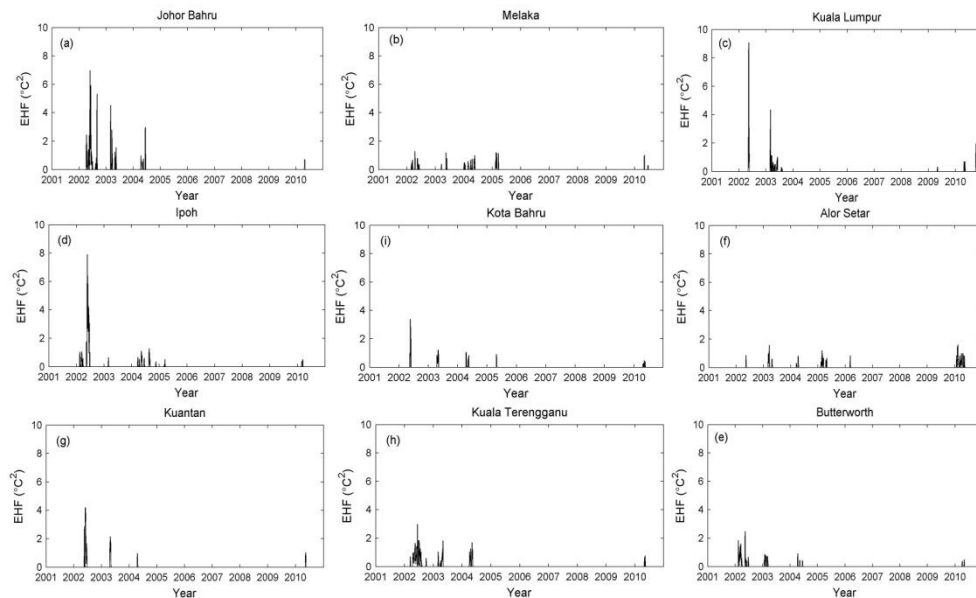


Fig. 2. Summary of research methods.

### 3. RESULT AND DISCUSSION

Heat wave conditions exist when the EHF is positive at least three consecutive days while a single or double positive EHF value does not define a heat wave (Perkins & Alexander, 2013). Based on the EHF index, the heat wave events for all stations from 2002 to 2006 and 2010 have been identified to determine the characteristics of HWN, HWF, HWD, and HWA. Figure 3 illustrates the variation of the EHF index across Peninsular Malaysia during the period 2001-2010. Referring to Figure 3(a), the EHF index in Johor Bahru shows the heat wave events happened from 2002 to 2004 and 2010 with the highest EHF index in May 2002 with  $7.0^{\circ}\text{C}^2$ . Figure 3(b) presents the EHF index for Melaka with heat wave event happened from 2002 to 2005. The highest EHF index is  $1.3^{\circ}\text{C}^2$  in April 2002. Meanwhile, Figures 3(c) and 3(d) show the EHF index for KLIA and Ipoh with heat wave event happened in the year 2002, 2003, 2009 and 2010 for KLIA while from 2002 to 2005 and 2010 for Ipoh. These two locations show the highest EHF index compared to others with EHF index  $9.1^{\circ}\text{C}^2$  for KLIA and  $7.9^{\circ}\text{C}^2$  for Ipoh and both happen in May 2002. For Northwest area in Figures 3(e) and 3(f), the heat wave event in Butterworth was happened from 2002 to 2004, and 2010 with the highest EHF index is  $3.5^{\circ}\text{C}^2$  in May 2002 while Alor Setar shows heat wave event happened from 2002 to 2006 and 2010 with the highest index is  $1.8^{\circ}\text{C}^2$  in March 2010. Meanwhile, the heat wave events over East

Peninsular Malaysia for Kuantan and Kuala Terengganu in Figures 3(g) and 3(h) indicated the heat wave event happened from 2002 to 2004, and 2010. It slightly differences for Kota Bahru in Figure 3(i) with the heat wave event was happened from 2002 to 2005, and 2010. The highest of the EHF index for Kuantan is  $4.2^{\circ}\text{C}^2$  and Kuala Terengganu is  $3.0^{\circ}\text{C}^2$  which both happen in June 2002 while for Kota Bahru is  $3.4^{\circ}\text{C}^2$  happened in May 2002.



**Fig. 3.** Variation of EHF index (a) Johor Bahru, (b) Melaka, (c) Kuala Lumpur, (d) Ipoh, (e) Butterworth, (f) Alor Setar, (g) Kuantan, (h) Kuala Terengganu and (i) Kota Bahru.

From the result, the heat wave event almost happened during southwest monsoon (SWM) which is between March and July. Major heat wave events with the highest EHF index happened in 2002 for all locations except Alor Setar in 2010. The critical value recorded during the study period may be due to severe dry spells in East Malaysia that have been recorded during the El Niño events which are three driest years (1963, 1997, and 2002) for Peninsular Malaysia (MMD, 2009). We can conclude that the variation of heat wave events for the most area in Peninsular Malaysia is affected by El Niño during the period 2001-2010. Abul and Gazi (2016) compiled El Niño event and classified during 1952-2010 globally which indicated moderate El Niño during 2002-2003 and 2009-2010 while weak El Niño event happened during 2004-2005 and 2006-2007. The result of this study also shows the difference of EHF index refers to the urban, industrial and suburban area. Urban and industrial areas such as Kuala Lumpur, Ipoh, Kuantan, Kota Bahru, and Johor Bahru recorded higher EHF index than the suburban area (Alor Setar, Kuala Terengganu, Melaka, and Kota Bahru). For urban and industrial regions, it is exposed to more heat wave events compared to the suburban or rural area. The effect of heat in urban areas is probably due to the result of the interaction of synergies such as surface moisture deficiency in urban areas, low wind speed, and also a difference in ambient temperature between the two areas (Fisher et al., 2007).

**Table 2** shows the summary of the heat wave characteristics over Peninsular Malaysia during the period 2001-2010. From the Table, the standard annual mean surface

temperature for this country is ranged from 26°C to 28°C<sup>22</sup>. For the value of 95th percentile in each location, Kuala Lumpur and Ipoh were the highest with 30.2°C and the lowest was in Kuala Terengganu and Melaka at 29.5°C. For HWN, the most heat wave cases happened in Johor Bahru with 1.8 events/year from April until June 2002 with 8 cases recorded. The lowest of HWN were observed in Kuantan with 0.6 events/year where on April 2004 is one case recorded. For HWF, it happens at a similar location like HWN, where the highest value is 14.5 days/year and the lowest value is 6.6 days/year. For HWD, the highest sum of participating heat wave days happened in a similar location as HWN and HWF, where during 2002 happened with 74 days from April until June and the lowest is 3 days in April 2004 for Johor Bahru dan Kuantan, respectively. The longest duration was at Ipoh from 7 May until 30 May 2002 (24 days) with an average of 4.7 days/year and the shortest was at Kuantan with 1.4 days/year (3 April until 5 April 2004). Over the considered period of HWA, the highest was recorded in Kuala Lumpur with 1.6°C<sup>2</sup>/year and the lowest was in Alor Setar with 0.5°C<sup>2</sup>/year. Based on the EHF calculation, the strongest and the weakest amplitude heat wave peaks were recorded in Kuala Lumpur on 17 May until 21 May 2002 with index range 2.7°C<sup>2</sup> - 9.1°C<sup>2</sup> and in Alor Setar from 3 March until 6 March 2004 with index range 0.1°C<sup>2</sup> - 0.2°C<sup>2</sup>, respectively.

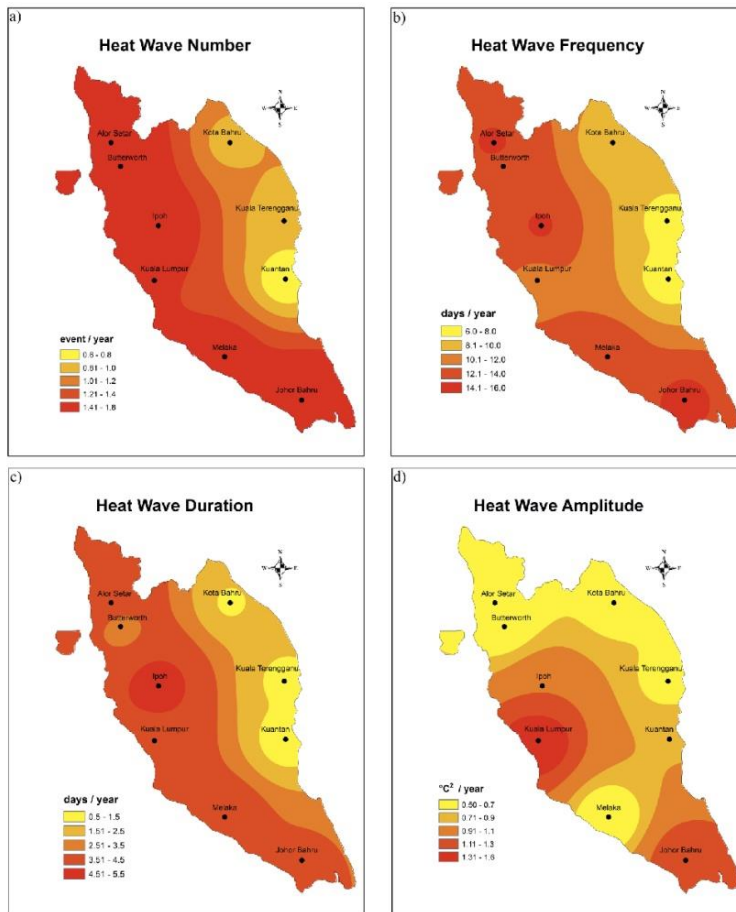
**Table 2.**

Peninsular Malaysia heat wave characteristic for the period 2001-2010:  
95% Percentile, HWN, HWD, HWF, HWA

Station	95 <sup>th</sup> Percentile (°C)	HWN (no. of event/year)	HWF (days/year)	HWD (days/year)	HWA (°C <sup>2</sup> /year)
Johor Bahru	29.4	1.8	14.5	3.1	1.3
Melaka	29.5	1.6	13.9	3.7	0.6
KLIA	30.2	1.4	11.6	3.7	1.6
Ipoh	30.2	1.7	14.3	4.7	1.1
Alor Setar	30.0	1.7	14.2	4.0	0.5
Butterworth	30.0	1.6	13.8	2.6	0.7
Kuantan	29.6	0.6	6.6	1.4	0.9
Kuala Terengganu	29.5	0.8	7.2	1.7	0.7
Kota Bharu	29.8	0.9	8.1	1.8	0.7

Figure 4 shows the general spatial distribution of heat wave characteristic for Peninsular Malaysia. HWN Figure 4(a) showed the highest average value which concentrated in the southwest and northwest of Peninsular Malaysia (1.5-1.9 event/year) and the lowest value is observed in the east part of Peninsular Malaysia (0.50-0.70 event/year). Similar to HWN, the highest average value for HWF (>14 days/year) can be found in the southwest and northwest and the lowest 6.0-8.0 days/year) is in East Peninsular Malaysia (Figure 4(b)). In Figure 4(c), HWD shows the longest average duration of 4 to 7 days per year which predominantly located in the west and northwest of Peninsular Malaysia. The shortest heat waves (1-2 days/year) are observed in the east part of Peninsular Malaysia. Figure 4(d) shows the HWA with the highest (1.3°C<sup>2</sup>-1.5°C<sup>2</sup>/year) and the lowest (0.5°C<sup>2</sup>-0.7°C<sup>2</sup>/year) is located in west part and some part of the Northeast (Alor Setar) and southeast (Melaka) of Peninsular Malaysia. From these spatial distributions, the heat wave events are dominated by seasons in the southeast, northeast, and west parts of Malaysia where the majority of heat waves has happened during the

southwest monsoon season. This characteristic is also similar to what happened in Iran (Esmailnejad, 2016).



**Fig. 4.** Spatial pattern of heat wave characteristic over Peninsular Malaysia (a) heat wave number (b) heat wave frequency (c) heat wave duration, and (d) heat wave amplitude.

Based on the result represented by 95 percentile, HWD, HWF, and HWN, most of the station are resulting in longer duration of heat wave events which more than a week. This condition of the heat wave in Peninsular Malaysia can induce to affect public health especially during SWM and El Nino events. More than that, the higher value in HWA indices will have a greater impact on agriculture, which usually due to high values of evapotranspiration associated with the lack of precipitation (Croitoru et al., 2016). The yields can be severely affected when crop exposed to stressful condition during heat waves, mainly due to the lack of water in the irrigation system during the drought season. Over the study period, the range of heat wave amplitude for all locations is between 29.4°C and 33.0°C. Compared to the optimum annual temperature for the major economic crops in Malaysia (NC2, 2000) in **Table 3**, we can assume that the heat wave events during the study period affected some agriculture plants, especially crop production.

**Table 3.**

Optimum annual temperature of the major economic crop in Malaysia  
(Malaysia Second National Communication, 2000).

Crop	Optimum Annual Temperature
Oil Palm	22°C - 32°C
Rice	24°C - 34°C
Rubber	23°C - 30°C
Cocoa	25°C - 32°C

## 5. CONCLUSIONS

This paper has successfully characterized the heat wave events in Peninsular Malaysia based on daily maximum and minimum temperature during the period 2001-2010. The combined effect of excess heat (EHI<sub>sig</sub>) and heat stress (EHI<sub>acc</sub>) were employed to obtain the EHF index. The result showed that the highest EHF index has happened at Kuala Lumpur in 2002 and the lowest was at Alor Setar in 2006. The most of the heat wave event was recorded during El Nino events especially during moderate El Nino in 2002 until 2005 and 2010. Based on HWD, HWF, HWN, and HWA, locations with the highest climatological values has been characterized and found that the southwest (Johor Bahru) and the west (Kuala Lumpur) of Peninsular Malaysia are the highest for HWN and HWF, and HWA, respectively. On the other hand, the longest duration of the heat wave (HWD) was found in Ipoh with 24 days. The characteristics of heat wave from EHF index was also compared to spatial distribution maps which shows that southeast, northeast and west part of Malaysia experience a more heat wave event during the study period.

The results of our investigation showed that the heat wave conditions in Peninsular Malaysia are anxiously, and therefore, critical study and exploration is needed because this event has a direct impact on agriculture, economic, and human health. Under these circumstances, further research needs to be undertaken in order to produce standard definition and threshold heat wave for Malaysia. This information would be useful for health policymakers to enable them to better plan for future climate change impacts. The use of longer-term period would possible to give more accurate information about the occurrence of heat wave indices and events overall Malaysia.

## ACKNOWLEDGMENTS

We would like to thank the Malaysia Meteorological Department for providing the temperature data. The second author is PhD candidate and supported by the Faculty Science and Natural Resources, University of Malaysia Sabah, 88400, Kota Kinabalu, Sabah.

## REFERENCES

- Abul, Q. A. & Gazi, M. A. (2016) The impact of El-Niño on agro-economics in Malaysia and the surrounding regions: An analysis of the events from 1997-98. *Asian J. Earth Sci.*, 9, 1-8.
- Anderson, G. B. & Bell, M. L. (2011) Heat wave in the United States: Mortality risk during heat wave and effect modification by heat wave characteristics in 43 U.S. communities. *Environmental Health Perspectives*, 119, 210-218.
- Bocancea, R. S. (2018), "Heat waves frequency: a study case of Iasi City, Romania (1961-2016). *Geographia Technica*, 13(1), 10-19.
- Coates, L., Haynes, K., O'Brien, J., McAneney, J., & De Oliveira, F. D. (2014) Exploring 167 years of vulnerability: an examination of extreme heat events in Australia 1844-2010. *Environmental Science & Policy*, 42, 33-44.



- Coumou, D. & Rahmstorf, S. (2012) A decade of weather extremes. *Nature Climate Change*, 2(7), 491–496.
- Croitoru, A. E., Piticar, A., Ciupertea, F. A. & Rosca, C. F. (2016) Changes in heat wave indices in Romania over the period 1961–2015. *Global and Planetary Change*, doi: 10.1016/j.gloplacha.2016.08.016.
- Esmailnejad, M. (2016), The spatial analysis of heat waves in southeast of Iran a case study: Sistan and Baluchistan Province. *Geographia Technica*, 11(2), 50-60.
- Fischer, E. M. & Schar, C. (2010) Consistent geographical patterns of changes in high-impact European heatwaves. *Nature Geoscience*, 3, 398–403.
- Fischer, E. M., Seneviratne, S. I., Luthi, D. & Schar, C. (2007) Contribution of land-atmosphere coupling to recent European summer heat wave. *Geophys. Res. Lett.*, 34, L06707.
- Guyton, A. C. & Hall, J. E. (2000) Textbook of medical physiology 10th ed. Philadelphia, PA: W. B. Saunders Company.
- Intergovernmental Panel on Climate Change (IPCC) in Climate Change 2013: Summary for Policymakers (Eds Stocker, T.F. et al.) (Cambridge University Press, 2013).
- Keggenhoff, I., Elizbarashvili, M. A. & King, I. (2014) Heat Wave Events over Georgia since 1961: Climatology, Changes and Severity. *Climate*, 3, 308-328.
- Makerami, N., Salleh, E., Jaafar, M. Z. & Ghaffarian, H. A. (2012) Thermal comfort conditions of shaded outdoor spaces in hot and humid climate of Malaysia. *Building and Environment*, 48, 7-14.
- Malaysia's Second National Communication (NC2) Vulnerability & Adaptation Assessment. (Ministry of Natural Resources and Environment Malaysia 2000).
- Malaysia Meteorological Department (MMD), (2009) Climate change scenarios for Malaysian 2001-2099. (Malaysian Meteorological Department, Scientific Report January 2009).
- Mohd Salleh, N. H., Hasan, H. & Kassim, S. (2015) Trends in Temperature Extremes across Malaysia. *Advances in Environmental Biology*, 9 (27), 174-181.
- Nairn, J. R., Fawcett, R. G., & Ray, D. (2009) Defining and predicting excessive heat events, a national system. In: *Proc. Third Cent. Aust. Weather Clim. Res. Underst. High Impact Weather*.  
[http://www.cawcr.gov.au/events/modelling\\_workshops/workshop\\_2009/papers/NAIRN.pdf](http://www.cawcr.gov.au/events/modelling_workshops/workshop_2009/papers/NAIRN.pdf).  
 Accessed on 20 February 2019.
- Nairn, J. R. & Fawcett, R. G. (2013) Defining heatwaves: heatwave defined as a heat-impact event servicing all community and business sectors in Australia.  
[http://www.cawcr.gov.au/publications/technicalreports/CTR\\_060.pdf](http://www.cawcr.gov.au/publications/technicalreports/CTR_060.pdf). Accessed on 20 February 2019.
- Nairn, J. R. & Fawcett, R. G. (2015) The excess heat factor: a metric for heatwave intensity and its use in classifying heatwave severity. *Int. J. Environ. Res. Public Health*, 12, 227–253.
- Peng, R. D., Boob, J. F., Tebaldi, C., McDaniel, L., Bell, M. L., & Dominici, F. (2011) Toward a quantitative estimate of future heat wave mortality under climate change. *Environmental Health Perspectives*, 119(5), 701-706.
- Perkins, S. E. & Alexander, L. V. (2013) On the measurement of heat wave. *Journal of Climate*, 26, 4500–4517, doi: 10.1175/JCLI-D-12-00383.1.
- Tangang, F. T., Juneng, L., Salimun, E., Sei, K. M., Le, L. J., & Muhamad, H. (2012) Climate change and variability over Malaysia: Gaps in science and research information. *Sains Malaysiana*, 41(11), 1355-1366.
- Tertre Le, A., Lefranc, A., Eilstein, S., Declerc, C., Medina, S., Blanchard, M., Chardon, B., Fabre, P., Filleul, L., Juso, J. F., Pascal, L., Prouvost, H., Cassadou, S. & Ledrans, M. (2006) Impact of the 2003 heatwave on all-cause mortality in 9 French cities. *Epidemiology*, 17, 75-9.
- Wai, N. M., Camerlengo, A., and Wahab, A. K. A. (2005) A Study of Global Warming in Malaysia. *Jurnal Teknologi*, 42, 1-10.
- Zuo, J., Pullen, S., Palmer, J., Bennetts, H., Chileshe, N. & Ma, T. (2014) Impacts of heat wave and corresponding measures: a review. *Journal of Cleaner Production*, 92 (1–12).  
<http://dx.doi.org/10.1016/j.jclepro.2014.12.078>.

## **MOUNTAIN TRACKS DEVELOPMENT METHODOLOGY FOR ADVENTURE RECREATION ACTIVITIES IN GURGHIU MOUNTAINS, ROMANIA**

*Adrian TORPAN<sup>1</sup>, Mihai VODA<sup>2</sup>*

DOI: 10.21163/GT\_2019.141.12

### **ABSTRACT:**

The focus of this study is mountain recreation trails evaluation for an effective implementation as organized leisure activity in Gurghiu Mountains. Here we show that our tracks classification model can be applied for any Romanian forest. Our research was developed using GIS and geospatial environmental datasets from GoogleEarth combined with field work on every proposed trail in order to assess the morphological features and the ecologically sensitive areas for future preservation. We found that gravel roads and forest exploitation tracks network constitutes the best premises for mountain bikes access in a forested area. We structured them in categories based on difficulty level, surface characteristics and climatic conditions. Furthermore we found that minimum investments are needed to mark and manage these tracks in a cost-effective manner, for the natural forest benefits and protection. Our results show how mountain biking recreation popularity can sustainably contribute to the local communities' development.

*Key-words: Mountain biking, Gurghiu, Romania, Forest, Trails.*

### **1. INTRODUCTION**

This paper examines the mountain forest trails potential for organized leisure activities development in Gurghiu Mountains. Studies show that mountain biking will increase in the future due to the new technological advances in electric propulsion and smartphone/GPS orientation devices (Pröbstl-Haider et al., 2018; Voda & Montes, 2018; Ernawati et al., 2018; Gupta et al., 2018).

This article contributes to understanding why the ecologically sensitive forest areas are more protected if the mountain gravel roads are opened to the public access for recreation activities and properly controlled (Ballantyne et al., 2014). Importantly, it outlines the environmental factors which need to be monitored in future studies and argues that forest exploitation tracks represent the best network for mountain biking expansion. However, recent research suggests that landowners, hunters and the National Forest Administration Authorities are wishing to restrict tourism development in the natural woodlands (Voda et al., 2017; Forest Law 133, 2005; Pröbstl-Haider et al., 2018).

As mountain biking is increasingly seen as a remedy for summer economic ills in other countries with mountainous areas, it is worth examining its potential in Romania which is a pioneer in this process. Switzerland, Germany, Italy and Western Austria opened their forests for mountain bikers, encouraging summer tourism development as a feasible alternative to winter season sports (Pröbstl-Haider et al., 2018; Hardiman & Burgin, 2013).

This innovative approach makes it possible to analyse the entire Gurghiu Mountains gravel roads network as a functional Geosystem with common and unique attributes given

---

<sup>1</sup>Folkuniversitetet, Jönköping, Sweden, [adriantorpan@yahoo.com](mailto:adriantorpan@yahoo.com);

<sup>2</sup>Dimitrie Cantemir University, 540546 Targu Mures, Romania, [mihaivoda@cantemir.ro](mailto:mihaivoda@cantemir.ro).

by the various morphometric characteristics, environmental factors and changing climate conditions (Gupta et al., 2018; Voda & Montes, 2018) and by a quantitative approach (Haidu, 2016). Further, this article contributes to the forest trails classification system, elaborating an efficient assessment methodology that can be securely applied in any natural environment for recreational activities development.

## 2. METHODOLOGY

Given the complexity of natural phenomena that are influencing the adherence on the analysed trails as well as the hazards that can influence the access to the mountain routes, we developed a complex scale (Torpan Scale) to properly quantify the trails' difficulty grades for recreational activities (Voda et al., 2017).

The Torpan Scale (TS) was issued to quantify the risks and difficulties of going through the mountain trails for adventure mountain biking. The Torpan Scale uses three main criteria for grading the mountain trails. These are: the morphometric criterion, the adherence criterion and the criterion given by the weather-climatic factors.

The primary index considered in the morphometric criterion assessment was the one given by the slope. In terms of mountain slopes, we took into account the values in the **Table 1**.

**Table 1.**  
Slope grades under dry track conditions

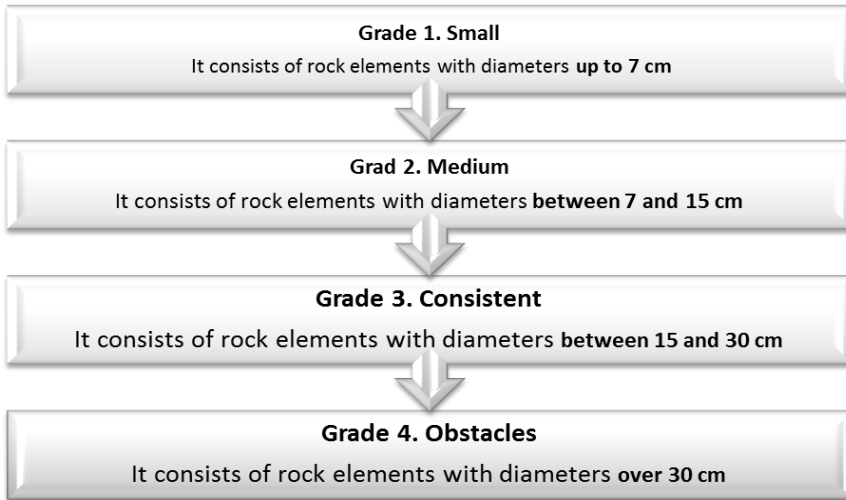
Difficulty Degree	Degree of slopes difficulty in dry road conditions
Grade 1	Small slopes: up to 10 <sup>0</sup>
Grade 2	Average slopes: up to 10-20 <sup>0</sup>
Grade 3	Steep slopes: up to 20-30 <sup>0</sup>
Grade 4	Very steep slopes: over 30 <sup>0</sup>

The adherence criterion is the most sensitive in defining the difficulty classes of the analysed routes because it directly changes the adhesion coefficient that influences the safety on the mountain trails. The classification of the trail elements stability composed of rock fragments after the Torpan Scale has the following values (**Table 2**):

**Table 2.**  
Difficulty grades given by the non-fixed elements percentage

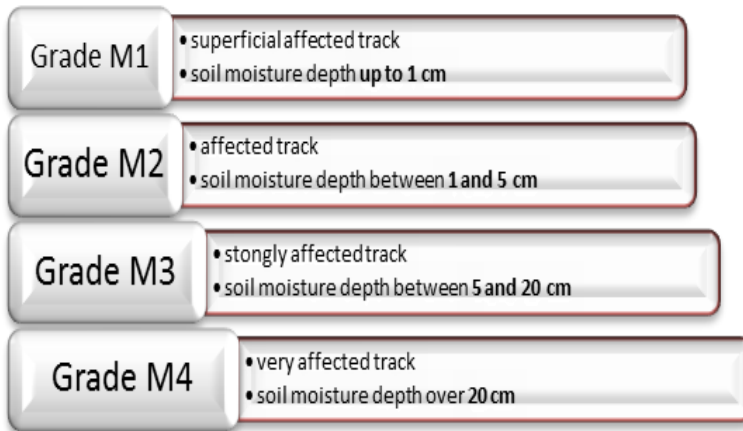
Grade	Percentage of solid non-fixed elements in the macadam component
Grade 1	Has a percentage of solid non-fixed elements <b>up to 20%</b>
Grade 2	Has a percentage of solid non-fixed elements <b>between 20% and 50%</b>
Grade 3	Has a percentage of solid non-fixed elements <b>over 50%</b>

The rocks diameter has a strong limitative factor and presents real risks in the following cases: in combination with slope, increased humidity, sinuous routes, sudden changes in trail's direction or in combination with wet soils (**Fig.1**):



**Fig. 1.** Rocks diameter difficulty grades.

The climatic factors taken into account for difficulty classes' determination after the Torpan Scale were rainfall, wind and temperature. Atmospheric precipitations are impacting the running surfaces increasing soil moisture depth, stream flow, generating torrents and rock falls that can directly affect the surface layer of mountain trails (Sarpe & Haidu, 2017; Voda et al., 2018). For the soil trails with or without consistent rocks we have the following degrees of difficulty under various humidity/moisture level (**Fig.2**):



**Fig. 2.** Gurghiu Mountains MTB routes moisture depth grades.

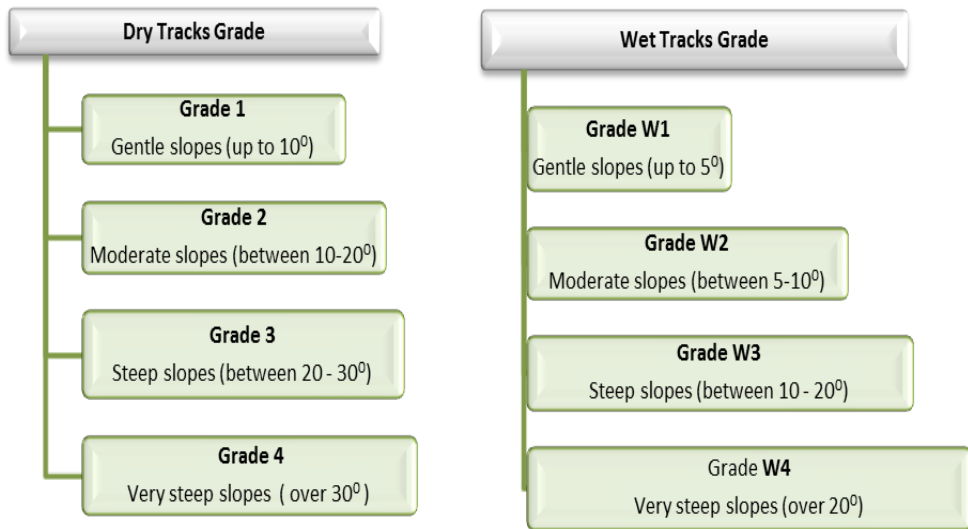
However, the main limiting factor is given by the slope, which in combination with humidity, changes the slope difficulty classification and presents the following values (**Table 3**):

**Table 3.**

**Slope grades under wet track conditions**

<b>Difficulty Degree</b>	<b>Degree of slopes difficulty under wet route conditions</b>
Grade 1	Small slopes: up to 5 <sup>0</sup>
Grade 2	Average slopes: up to 5-10 <sup>0</sup>
Grade 3	Steep slopes: up to 10-20 <sup>0</sup>
Grade 4	Very steep slopes: over 20 <sup>0</sup>

Grade M1 may occur under the conditions of morning condensation or light rain. It does not affect the grading except in the presence of slopes over 10 degrees. Hence, the evolved erosion processes or the medium and large non-fixed rocks along the route, determine lower slope values for fitting to the same difficulty class level (**Fig.3**).



**Fig. 3.** Wet and dry tracks condition influence on slope grading.

Sudden temperature drops may have a strong restriction role especially in late autumn when early frost phenomena and certain solid precipitation may occur. These are phenomena with an important impact on the trail by radically reducing the adherence along the mountain routes. Lower temperatures associated with strong winds are affecting body thermal comfort, especially for two-wheeled sports practitioners, who are directly exposed to temperature and wind fluctuations. Strong winds can cause forest vegetation collapses along the mountain trails. They can also directly affect paths in high ridge sectors exposed to rough air currents. Especially mountain biking sports enthusiasts can be affected by climate driven trail disruptions.

The combination of all analysed criteria generated the Torpan Scale that includes 10 difficulty classes of mountain routes. It covers the entire range of adventure recreation tracks. The Torpan Scale defines four major categories of difficulty listed in the following tables: **Table 4** and **Table 5**.

**Table 4.****General difficulty categories and the corresponding classification**

Easy/ low-grade routes	Class I, II, III, IV.
Medium-difficulty routes	Class V, VI, VII.
Difficult routes	Class VIII, IX.
Extreme routes	Class X.

Under heavy rainfall, considerable temperature drops, and solid precipitations, class VII (medium difficulty routes) may change the category to overall risk class VIII, IX or even X depending on the degree of damage done to the trail surface (snow or ice presence).

**Table 5.****Torpan scale of difficulty categories, classes and grades**

Difficulty category	Difficulty class	Slope grades/ dry track condition	Slope grades / wet track condition	Solid non-fixed elements grades	Rocks diameter difficulty grades	Routes moisture depth grades
Easy	Class I	1	1	1	1	1
	Class II	1/2	1	1	1	1
	Class III	2	1	1/2	1	1
	Class IV	2	1	2	2	1
Medium	Class V	2	1	2	2	2
	Class VI	2/3	2	2	2/3	2
	Class VII	2/3	2	2/3	2/3	2/3
Difficult	Class VIII	2/3	2/3	2/3	3	2/3
	Class IX	3/4	3	2/3	2/3/4	3/4
Extreme	Class X	4	3/4	3	3/4	4

**3. RESULTS AND DISCUSSIONS**

The analysed area in Gurghiu Mountains is represented by an erosion witness of a volcanic plateau, located around 1300 –1415 m altitude. Spectacular andesitic cliffs are bordering the adjacent river valleys. Fluvial incision and stratigraphic succession are responsible for the high fragmentation depth (500 m on average) and the cliffs development. The columns area shows a recent dynamic with gravity detachments and debris slopes at their base. The predominantly soils are andesite soils, developed on the plateau. Lito series soils are found on the debris talus and steep slopes which border the plateau.

The forest exploitation gravel road network is well maintained in Gurghiu Mountains like everywhere in Romania and represents the best way to access and use MTB for recreational purposes (Voda et al., 2014). We found that there are no ecologically sensitive areas on these routes with unchanged morphological features along the years. Crossing the mountains from East to West, the forest exploitation tracks are following the main rivers and their tributaries. We considered them legal tracks and structured the gravel roads network in categories based on slope difficulty level, surface characteristics and climatic conditions.

The illegal random tracks were also evaluated and monitored because our researches on field revealed more than 25 routes across Gurghiu Mountains that are affecting the natural ecosystems. Used for motorized recreation purposes by hard enduro practitioners, illegal tracks follow small creeks upstream, dense forest deer trails or pasture routes on the volcanic plateau. Those illegal mountain tracks were not considered for our MTB recreational network model (Fig.4).

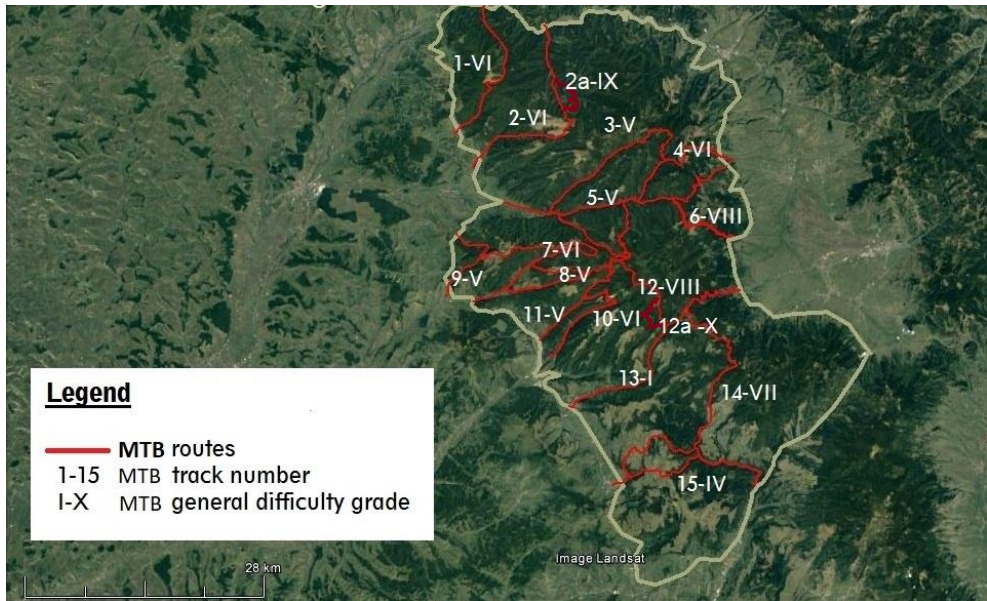


Fig. 4. Gurghiu Mountains tracks classification map.

In order to correctly evaluate the Gurghiu Mountains MTB routes, we analysed two categories of risks that are directly affecting the tracks: natural risks and anthropogenic origin risks. In the natural risks category, the geomorphological risks and weather related risks were included.

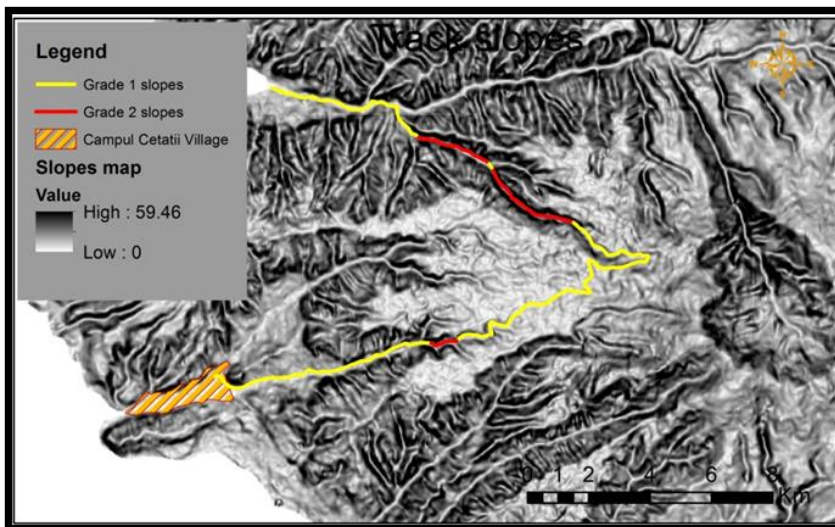
The geomorphological risks are closely correlated with the morphometric features of specific volcanic substrate composition with external influences from the hydric and thermic regime, weather-related phenomena (often with a trigger role) and geological risks caused by the major movements of the crust. The most dynamic phenomena that can affect MTB routes today are represented by mass movements, rolling stones, cave-ins, flow hillside, subsidence, landslides and other processes represented by ravines, torrential floods, stream runoffs that can occur on the slope under the gravity direct impact on the gravel roads infrastructure used for MTB activities (Gupta et al., 2018).

The analysis of weather related risks for the Gurghiu Mountains MTB recreational routes required special attention due to their destructive power and the unpredictability factor, which cannot be controlled by man. In practice, they often have a major role in unleashing geomorphological phenomena. Giving rise to a domino effect, the weather phenomena are expanding the various natural risks with direct impact on MTB gravel roads infrastructure (Voda et al., 2014).

The anthropogenic origin risks are caused by human intervention for deforestation purposes in the Gurghiu Mountains, where the MTB recreational routes were analysed. In most cases human intervention creates phenomena and events with high risk by disrupting natural systems relative equilibrium. The most common human activities impacting on Transylvanian mountains routes are represented by excessive deforestation leading to severe imbalances on local natural woods ecosystems. Mountain biking development will increase transparency diminishing illegal logging activities (Voda and Montes, 2018).

Along the criteria related to MTB routes adherence, climate, natural risks and anthropogenic related risks, the main criteria considered in the MTB accessibility routes assessment is given by track surface morphometric characteristics and route slopes as the main limiting factor analysed through GIS. Also their influence is enhanced or diminished depending on rainfall amount, the presence of freezing temperatures and the slope orientation. GIS analysis revealed that the presence of erosion linear deployment on routes running surfaces (runoff, ravination and torrents) is strongly influenced by the slope degree in which they develop, the slope being the main factor accelerating the expansion of these phenomena.

In general, we found that the easier wet track grades tend to be harder than expected compared to the normal dry condition grades. The remoteness and poor volcanic rock quality in Gurghiu Mountains mean conditions are highly variable. There are some very committing *Grade M2-W2* climbs because often, the challenge of the route comes from the fluctuating mountain weather conditions, rather than technical gravel road gradients. Thus the moderate slope characteristics of *Grade2* are becoming as difficult as the steep slopes when the moisture *GradeM2* is affecting the tracks. As a case study we evaluated the scenic route of Gurghiu Valley – Sirod Valley - Niraj Valley leading to Campul Cetatii Village. According to the Gurghiu Mountain MTB tracks classification, the route is easily accessible for MTB practitioners in normal dry conditions and positive temperatures. Thus we differentiated the sectors that present moderate *Grade2* slopes, from the gentle variation of the gradient along the route (*Grade1*) that is less difficult for MTB activities (**Fig.5**).

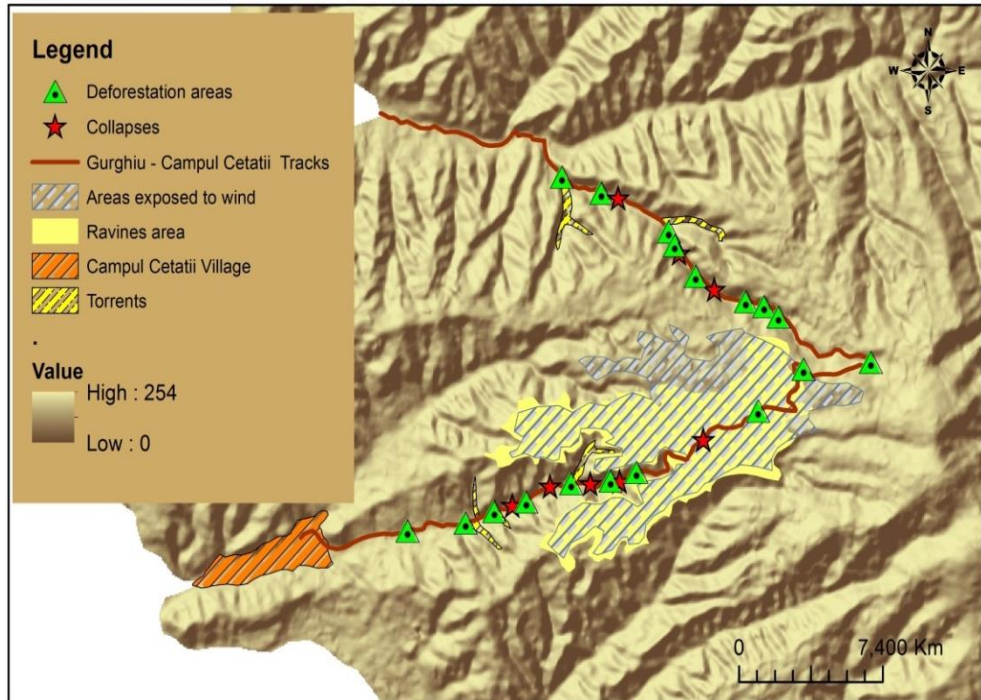


**Fig. 5.** Gurghiu Valley-Campul Cetatii Village route slope grades map.



Using GIS and geospatial environmental datasets from GoogleEarth combined with field work, we analysed all the route sections that are exposed to natural and anthropogenic risks (**Fig. 6**).

Our field research identified eighteen deforestation areas and seven collapses of soil-structure interaction along the Gurghiu Valley - Sirod - Niraj – Campul Cetatii Village track. After this MTB recreational route evaluation, we came to the conclusion that previous uncontrolled deforestation activities determined torrents and ravines development, increased the surfaces of wind exposed areas and led to soil structure collapses.



**Fig. 6.** Mountain bike recreational route risks map.

Romanian National Forest Administration Authorities did not manage deforestation activities accordingly in terms of ecosystem balance maintenance. As a result, in time, the dynamic equilibrium was affected. We assessed the morphological features and the ecologically sensitive areas of the Gurghiu Valley - Sirod - Niraj – Campul Cetatii Village MTB recreational route for future preservation.

We strongly believe that the inclusion of this forest exploitation track in a MTB recreational network constitutes the best premises for Gurghiu mountainous region ecosystems protection and local villages' future sustainable development. We found that National Forest Administration is constantly investing in gravel roads maintenance and cooperates with tourist trails responsible authority for the tracks effective marking.

#### 4. CONCLUSIONS

This article examined a network of more than 1000 km length. Our study identified 40 off road mountain tracks during six years of field research and proposed an MTB recreation frame model for Gurghiu Mountains. The analysis reveals the recreational potential of the tracks selected from the National Forest Administration gravel roads network which were evaluated, classified and designed following MTB practitioners needs.

Further, the study confirms what MTB recreational tours are offering: the sensation of freedom (that comes especially riding on less crowded mountain routes), pure air to breathe (there are no polluted areas) and beautiful landscapes (spectacular sceneries constitute the main attraction).

Our study results are providing a protocol for the methodological steps used to evaluate forest tracks and create adventure recreational networks. The present research validated 409.8 km of forest tracks from Gurghiu Mountains including 15 beautiful routes of different general difficulty categories (GDC) with classes from I to X. Most of the tracks have a moderate to low difficulty level, the easiest being number 13 (Trans-Gurghiu) of 30 km long with I-GDC and the hardest being track number 12 (Fancel-Bucin) with VIII-GDC and 24 km length.

According to analysis presented in this article, the best MTB recreational track from Gurghiu Mountains is represented by Poieni route (7-VI). Starting in Orsova Village and passing through Gurghiu Meadow- Obarsia Glade-Prislop Glade- Copriana Glade- Breditel Meadow, it follows Niraju Mare Valley downstream to the ending point in Campu Cetatii Village. The longest MTB route has 50 km of gravel road, forming a IV grade circuit from Corund traditional village to Zetea storage reservoir. The shortest (9-V) has only 3.7 km length but it is a very important linkage track between Gurghiu and Nirajului river valleys.

Further research is needed to identify the local communities' leisure infrastructure and accommodation possibilities for the mountain biking practitioners. The present study did not examine the bike rental opportunities and bike service availability in the analysed area.

The findings presented in this research give ground for any forest trail evaluation for the inclusion in future outdoor recreational areas. The mountain biking recreation popularity can actively contribute to the sustainable tourism development not only in Gurghiu forests but on any woodland.

#### REFERENCES

- Ballantyne, M., Gudes, O., & Pickering, C.M. (2014) Recreational trails are an important cause of fragmentation in endangered urban forests: A case-study from Australia. *Landscape and Urban Planning*, 130, 112–124.
- Ernawati, N.M., Torpan, A. & Voda, M. (2018) Geomedia role for mountain routes tourism development. Mesehe and Pisoiu Waterfall comparative study. *Geographia Technica*, 13 (1), 41–51.
- Forest Law 133. (2015) *Motorised public acces into the National Forest Fund*. [Online]. Available from: [http://legea\\_133\\_2015\\_modificare\\_lege\\_46\\_2008\\_codul\\_silvic.php](http://legea_133_2015_modificare_lege_46_2008_codul_silvic.php) [Accessed December 2018].
- Google Earth. (2018). *Google Earth*. [Online]. Available from: [www.google.com/earth/](http://www.google.com/earth/) [Accessed August 2018].

- Gupta, S.K., Negru, R., & Voda, M. (2018) The Indian Himalaya`s unique attributes: Hemkund Sahib and The Valley of Flowers. *Geographia Technica*, 13 (2), 62-73.
- Haidu, I. (2016) What Is Technical Geography. *Geographia Technica*, 11(1), 1-5. DOI: 10.21163/GT\_2016.111.01.
- Hardiman, N., & Burgin, S. (2013) Mountain biking: downhill for the environment or chance to up a gear? *International Journal of Environmental Studies*, Vol. 70, No. 6, 976–986.
- Pröbstl-Haider, U., Lund-Durlacher, D., Antonschmidt, H., & Hödl, C. (2018) Mountain bike tourism in Austria and the Alpine region – towards a sustainable model for multi-stakeholder product development, *Journal of Sustainable Tourism*, 26:4, 567-582, DOI: 10.1080/09669582.2017.1361428
- Sarpe, C. A., & Haidu, I. (2017). Temporal sampling conditions in numerical integration of hydrological systems time series. *Geographia Technica*, 12 (1), 82- 94.
- Voda, A.I, Sarpe, C.A, & Voda, M. (2018) Methods of maximum discharge computation in ungauged river basins. Review of procedures in Romania. *Geographia Technica* 13(1), 130–137.
- Voda, M., Moldovan, L., Torpan, A., & Henning, A. (2014) Using GIS for mountain wild routes assessment in order to qualify them for tourism valorisation. *Geographia Technica* 9 (1), 101–108.
- Voda, M., Torpan, A. & Moldovan, L. (2017) Wild Carpathia Future Development: From Illegal Deforestation to ORV Sustainable Recreation. *Sustainability*, 9(2254), 1-11.
- Voda, M., & Montes, Y.S. (2018) Descending mountain routes future: the North Yungas and Fagaras Geosystem`s comparative study. *Geographia Technica*, 13 (2), 152-166.

## Aims and Scope

**Geographia Technica** is a journal devoted to the publication of all papers on all aspects of the use of technical and quantitative methods in geographical research. It aims at presenting its readers with the latest developments in G.I.S technology, mathematical methods applicable to any field of geography, territorial micro-scalar and laboratory experiments, and the latest developments induced by the measurement techniques to the geographical research.

**Geographia Technica** is dedicated to all those who understand that nowadays every field of geography can only be described by specific numerical values, variables both of time and space which require the sort of numerical analysis only possible with the aid of technical and quantitative methods offered by powerful computers and dedicated software.

Our understanding of **Geographia Technica** expands the concept of technical methods applied to geography to its broadest sense and for that, papers of different interests such as: G.I.S, Spatial Analysis, Remote Sensing, Cartography or Geostatistics as well as papers which, by promoting the above mentioned directions bring a technical approach in the fields of hydrology, climatology, geomorphology, human geography territorial planning are more than welcomed provided they are of sufficient wide interest and relevance.

### Targeted readers:

The publication intends to serve workers in academia, industry and government. Students, teachers, researchers and practitioners should benefit from the ideas in the journal.

## Guide for Authors

### Submission

Articles and proposals for articles are accepted for consideration on the understanding that they are not being submitted elsewhere.

The publication proposals that satisfy the conditions for originality, relevance for the new technical geography domain and editorial requirements, will be sent by email to the address [editorial-secretary@technicalgeography.org](mailto:editorial-secretary@technicalgeography.org).

This page can be accessed to see the requirements for editing an article, and also the articles from the journal archive found on [www.technicalgeography.org](http://www.technicalgeography.org) can be used as a guide.

### Content

In addition to full-length research contributions, the journal also publishes Short Notes, Book reviews, Software Reviews, Letters of the Editor. However the editors wish to point out that the views expressed in the book reviews are the personal opinion of the reviewer and do not necessarily reflect the views of the publishers.

Each year two volumes are scheduled for publication. Papers in English or French are accepted. The articles are printed in full color. A part of the articles are available as full text on the [www.technicalgeography.org](http://www.technicalgeography.org) website. The link between the author and reviewers is mediated by the Editor.

### Peer Review Process

The papers submitted for publication to the Editor undergo an anonymous peer review process, necessary for assessing the quality of scientific information, the relevance to the technical geography field and the publishing requirements of our journal.

The contents are reviewed by two members of the Editorial Board or other reviewers on a simple blind review system. The reviewer's comments for the improvement of the paper will be sent to the corresponding author by the editor. After the author changes the paper according to the comments, the article is published in the next number of the journal.

Eventual paper rejections will have solid arguments, but sending the paper only to receive the comments of the reviewers is discouraged. Authors are notified by e-mail about the status of the submitted articles and the whole process takes about 3-4 months from the date of the article submission.

Indexed by: **CLARIVATE ANALYTICS**  
**SCOPUS**  
**GEOBASE**  
**EBSCO**  
**SJR**  
**CABELL**

**ISSN: 1842 - 5135 (Print)**  
**ISSN: 2065 - 4421 (Online)**

