MGPB

The Mining-Geology-Petroleum Engineering Bulletin
UDC: 626/627:556.3
DOI: 10.17794/rgn.2017.1.4

Original scientific paper



# Cross-correlation Modelling of Surface Water – Groundwater Interaction Using the Excel Spreadsheet Application

**Kristijan Posavec[1]; Perica Vukojević[2]; Matija Ratkaj[3]; Tea Bedeniković[4]**

[1] *University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Pierottijeva 6, 10000 Zagreb, Associate Professor, PhD*
[2] *Geotehnički studio Ltd. Zagreb, Nikole Pavića 11, 10090 Zagreb, Mag.Ing.Geol.*
[3] *IHK Karlovac, Selce 25, 47000 Karlovac, Mag.Ing.Geol.*
[4] *University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Pierottijeva 6, 10000 Zagreb, Graduate student*

## Abstract

Modelling responses of groundwater levels in aquifer systems, which occur as a reaction to changes in aquifer system boundary conditions such as river or stream stages, is commonly being studied using statistical methods, namely correlation, cross-correlation and regression methods. Although correlation and regression analysis tools are readily available in Microsoft Excel, a widely applied spreadsheet industry standard, the cross-correlation analysis tool is missing. As a part of research of groundwater pressure propagation into alluvial aquifer systems of the Sava and Drava/Danube River catchments following river stages rise, focused on estimating groundwater pressure travel times in aquifers, an Excel spreadsheet data analysis application for cross-correlation modelling has been designed and used in modelling surface water – groundwater interaction. Examples of field data from the Zagreb aquifer system and the Kopački rit Nature Park aquifer system are used to illustrate the usefulness of the cross-correlation application.

## Keywords

cross-correlation, Excel spreadsheet application, surface water-groundwater interaction.

## 1. Introduction

Modelling of surface water – groundwater interaction, namely groundwater pressure propagation into alluvial aquifer systems of the Sava and Drava/Danube River catchments, as a response to changes in aquifer system boundary conditions i.e. following Sava, Drava and Danube River stage rise, has been conducted using the cross-correlation data analysis application. Cross-correlation is a process of comparing a time series, for example river stage and groundwater level time series, at successive lags in order to determine the lag or offset in time or distance between them at their position of maximum equivalence (**Davis, 2002**). Determining positions of pronounced correspondence in time series such as river stage and groundwater level time series, points to groundwater pressure travel times required for a certain part of the aquifer to respond, following river stage rise. Cross-correlation analysis is widely applied in studying interaction between groundwater systems and its boundary conditions, such as recharge through precipitation or recharge/discharge from rivers, lakes, streams or springs. Therefore, groundwater pressure travel times required to provoke responses of certain parts of the aquifers causing groundwater levels to rise or fall, are commonly calculated using cross-correlation analysis. **Fiorillo and Doglioni (2010)** studied the relation between rainfall and the discharge from two springs, located at the base of different karst massifs in southern Italy. They investigated a wide range of time windows of cumulative rainfall, due to the relation between spring discharge and the previous period of rainfall. Using cross-correlation analysis, they calculated the time required for water to flow through the karst aquifers, i.e. the main infiltration pathways in aquifers, and discharges at two spring sites. Since it is presumed that the pressure pulse is forced through phreatic conduits and almost instantaneously influences the spring discharge once the percolating water reaches the saturated zone (**Ford and Williams, 2007**), the time interval they calculated is likely to be related to the percolation time needed to travel through the vadose zone. **Crosbie et al. (2005)** evaluated the precipitation time parameter α from the partial cross correlation between the differenced water level signal and the rainfall signal, to exclude water level rises that are not caused by rainfall. The precipitation time parameter α is a parameter used in a methodology for filtering noise and non-rainfall-related impacts on water table fluctuations, when calculating groundwater recharge using the water table fluctuation method. **Larocque et al. (1998)** applied cross-correlation analysis to establish a link between

Corresponding author: Kristijan Posavec
*kristijan.posavec@rgn.hr*

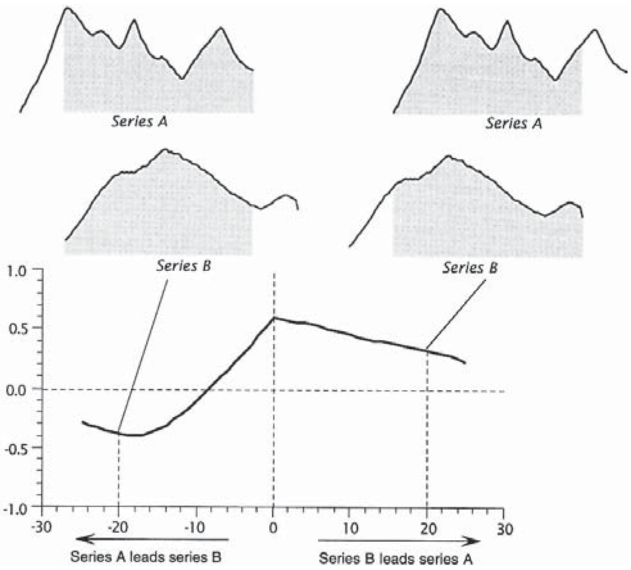


**Figure 1.** Cross-correlogram of two series, A and B. Shaded intervals indicate portions of A and B that are compared at two lag positions. Zero lag position indicates both series are aligned at their origin (**Davis, 2002**).

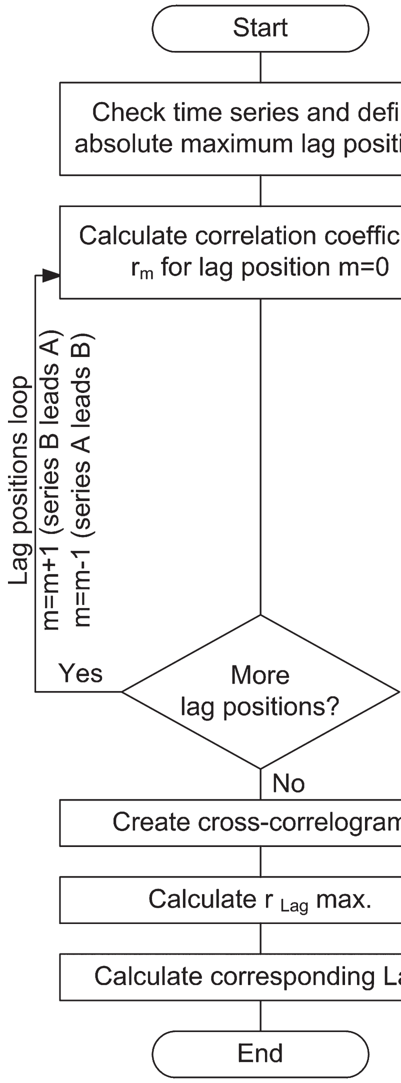


**Figure 2.** Program structure

Foulpougne and the Bandiat River, Tardoire and Bonnieure Rivers and between the Leche spring and the Echelle, Bandiat, Tardoire and Bonnieure Rivers. They established that river flow rates influence the spring flow rates. Similar analysis was performed by **Lee and Lee (2000)**, using cross-correlation functions to identify a characteristic relationship between the site rainfall and the corresponding natural water-level variations, where the rainfall was considered as an input and the water level as an output system variable. Time series solute data sets have been used by **Welch et al. (2013)** to determine travel times of river water into aquifers in losing environments. Besides direct interpretation of tracer arrival times, data interpretation has also relied on cross correlation analysis. **Cheng-Mau Wu et al. (2005)** used cross-correlation analysis to investigate effects of heterogeneity in aquifer parameters, transmissivity (T) and storage (S), on the analysis of traditional aquifer tests using the Theis analytic solution. They used cross-correlation analysis to represent distributions of the cross-correlations between the observed head and transmissivity field as well as distributions of the cross correlations between the observed head and storage coefficient field. They showed how the head perturbation at a given aquifer location at a given time is influenced by perturbation of aquifer parameters T and S at some different location. Interaction between stream water and groundwater by analysing their temperatures using cross-correlation analysis was studied in a work of **Lee et al. (2013)**. Their goal was to identify potential stream-groundwater exchange using vertical flow calculation. Their results revealed that groundwater temperatures show substantial correlation with the groundwater levels, which is indicative of somewhat long-term seasonal correlation. **Lee et al. (2003)** studied the interactive influence between groundwater flow and salinization in an underground liquid petroleum gas (LPG) cavern site using chemical analysis data and cross-correlation analysis between hydraulic head and operating pressure data. Their results revealed that a positive relationship between hydraulic head and operating pressure was far more conspicuous in the propane cavern area, and tidal change influenced the head variation in the butane cavern area. Cross-correlation analysis was also used in a study by **Seeboonruang (2015)**, where shallow groundwater and groundwater salinity fluctuation in the irrigated area was analysed. Dependency upon the surface water regulation and prior rainfall was shown, justifying the applicability of a multiple linear regression with lagged time dependence. A large number of different authors also used cross-correlation analysis in their studies of surface water – groundwater interaction, showing an extensive possibilities for the application of the method. In this study, surface water – groundwater interaction, i.e. responses of groundwater level fluctuations following river stages rise, was

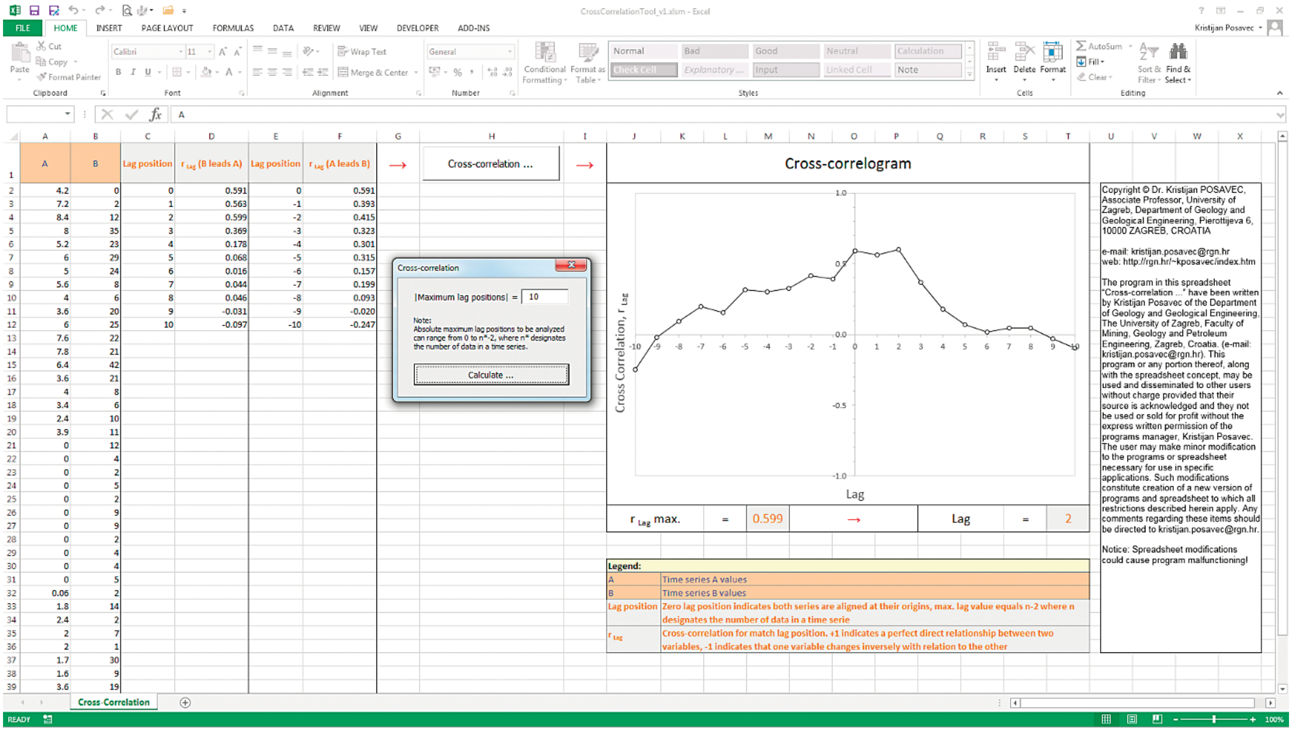


**Figure 3.** Excel spreadsheet layout showing calculated correlation coefficients for all match positions, cross-correlogram and maximum cross-correlation with corresponding lag time

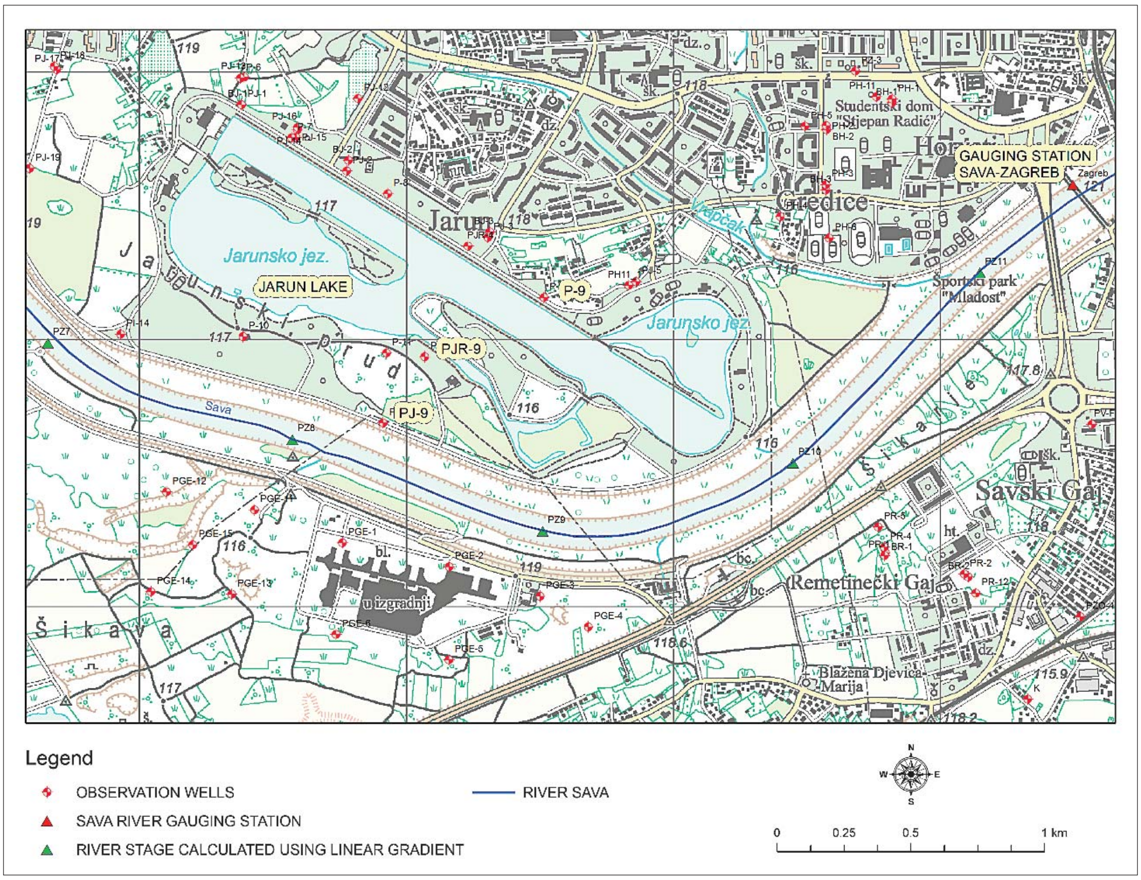


**Figure 4.** Situation map of observation wells and gauging station Sava–Zagreb – Site 1

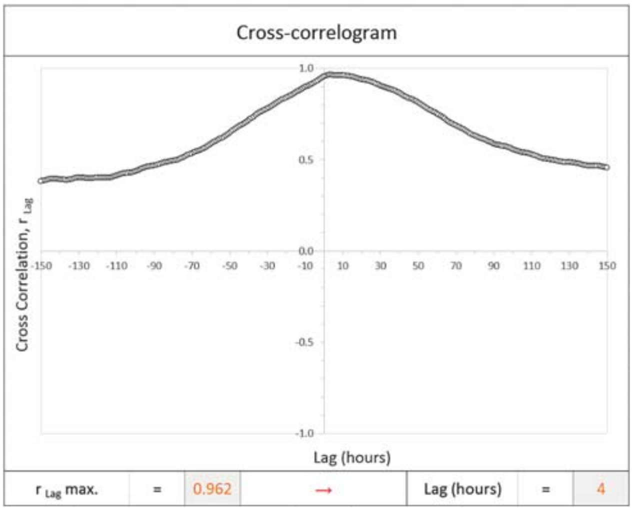


**Figure 5.** Cross-correlogram of the Sava River stage time series and groundwater level time series measured in observation well PJ-9

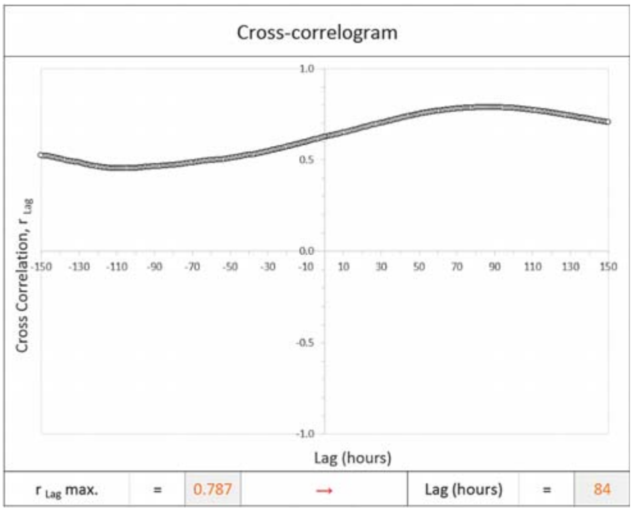


**Figure 7.** Cross-correlogram of the Sava River stage time series and groundwater level time series measured in observation well P-9

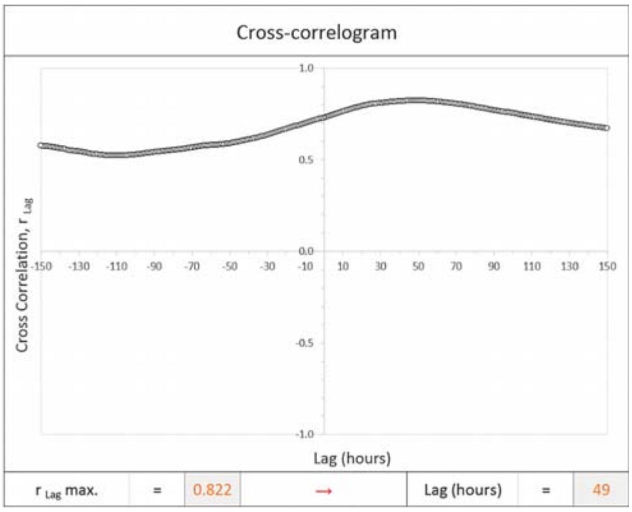


**Figure 6.** Cross-correlogram of the Sava River stage time series and groundwater level time series measured in observation well PJR-9

modelled using specifically designed cross-correlation Excel spreadsheet application in order to estimate groundwater pressure travel times. The cross-correlation analysis tool is not readily available in Microsoft Excel, although different commercial applications offering cross-correlation analysis are present in today's software market. Therefore, the Excel spreadsheet application with its open-access Visual Basic for Applications (VBA) code for cross-correlation analysis have been designed in order to process the data in research of groundwater pressure propagation into alluvial aquifer systems of the Sava and Drava/Danube River catchments following river stages rise, focused on estimating groundwater pressure travel times in aquifers. Analyses were conducted using hourly and daily measurements of the Sava, Odra, Drava and Danube River stages, Kopački rit Nature Park lake stages (Sakadaš Lake) and groundwater level measurements of the Zagreb aquifer system.

## 2. Method and Application Design

Modelling of surface water – groundwater interaction was studied using the cross-correlation statistical method. Generally, cross-correlation defines the measure of similarity of two series i.e. the degree to which two series are correlated with respect to lag of one relative to the other, and the process of their comparison at successive lag is called cross-correlation (**Davis, 2002**). It is most appropriately used in comparison of two series that have temporal dependence. In order to determine positions of pronounced correspondence of ground water level or lake level rise following river stage rise, time series were compared with each other at successive lags. Determined lag or offset in time between events of ground water level or lake level rise following river stage rise, at their position of maximum equivalence, defined groundwater pressure travel times. Besides, the strength of the relationship between the time series was defined through the correlation coefficient calculated for determined lag. In a cross-correlation analysis, time series measurements are required to be taken at corresponding times, for example, in analyses of the pressure propagation into the aquifer following river stage rise, river stages and groundwater levels are to be measured at the same time/date. The zero lag represent the alignment of two time series at their origins, namely at the same data measurement time, and comparison is made at successive lags, including both positive and negative lag positions where time series A leads time series B and vice versa, time series B leads time series A. Results of cross-correlation analysis are shown on a cross-correlogram (see **Figure 1**).

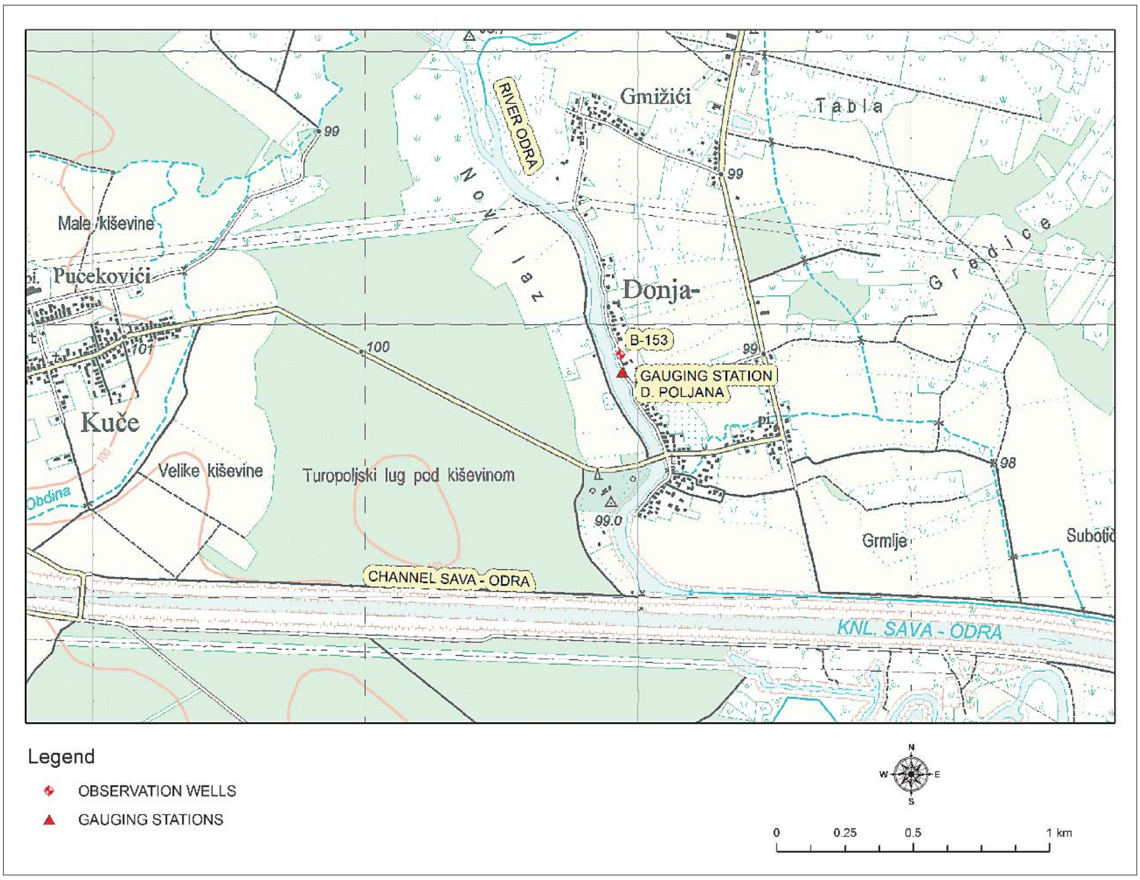


**Figure 8.** Situation map of observation well B-153 and gauging station D. Poljana on the river Odra– Site 2

The strength of the relationship between time series is determined through the correlation coefficient, using the same equation as for ordinary linear correlation given in **Equation 1** as follows:

$$r_m = \frac{n\sum Y_1Y_2 - \sum Y_1 \sum Y_2}{\sqrt{\left[n\sum Y_1^2 - \left(\sum Y_1\right)^2\right]\left[n\sum Y_2^2 - \left(\sum Y_2\right)^2\right]}} \quad (1)$$

Where:

$r_m$ – correlation coefficient for each match position m,

$n$ – the number of overlapped positions between the two time series,

$Y_1$ and $Y_2$ – two time series being compared.

The equation is applied only over the overlapped segments of the two compared time series, thus the absolute maximum number of overlapped positions $n$ between the two time series, equals $n^*$-2, where $n^*$ designates the total number of data in a given time series.

For the purpose of this research, the cross-correlation analysis tool was specifically designed as an application within the Excel spreadsheet, using VBA code, which is made open-access and freely available. An algorithm generally simulates comparison of two time series at successive lags, calculates the accompanying correlation coefficients and creates a cross-correlogram showing the maximum equivalence $r_{Lag}$ max., where $r_{Lag}$ max. designates the highest correlation coefficient for all overlapped positions between the two time series, as well as the match position $m$ of maximum equivalence i.e. the corresponding lag (see **Figure 2**).

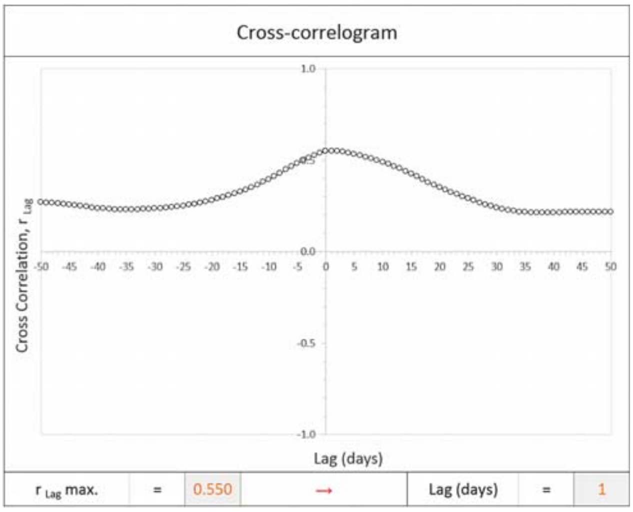


**Figure 9.** Cross-correlogram of the Odra River stage time series and groundwater level time series measured in observation well B-153

As input, a user-friendly and easy to use spreadsheet design, requires time series A and time series B values

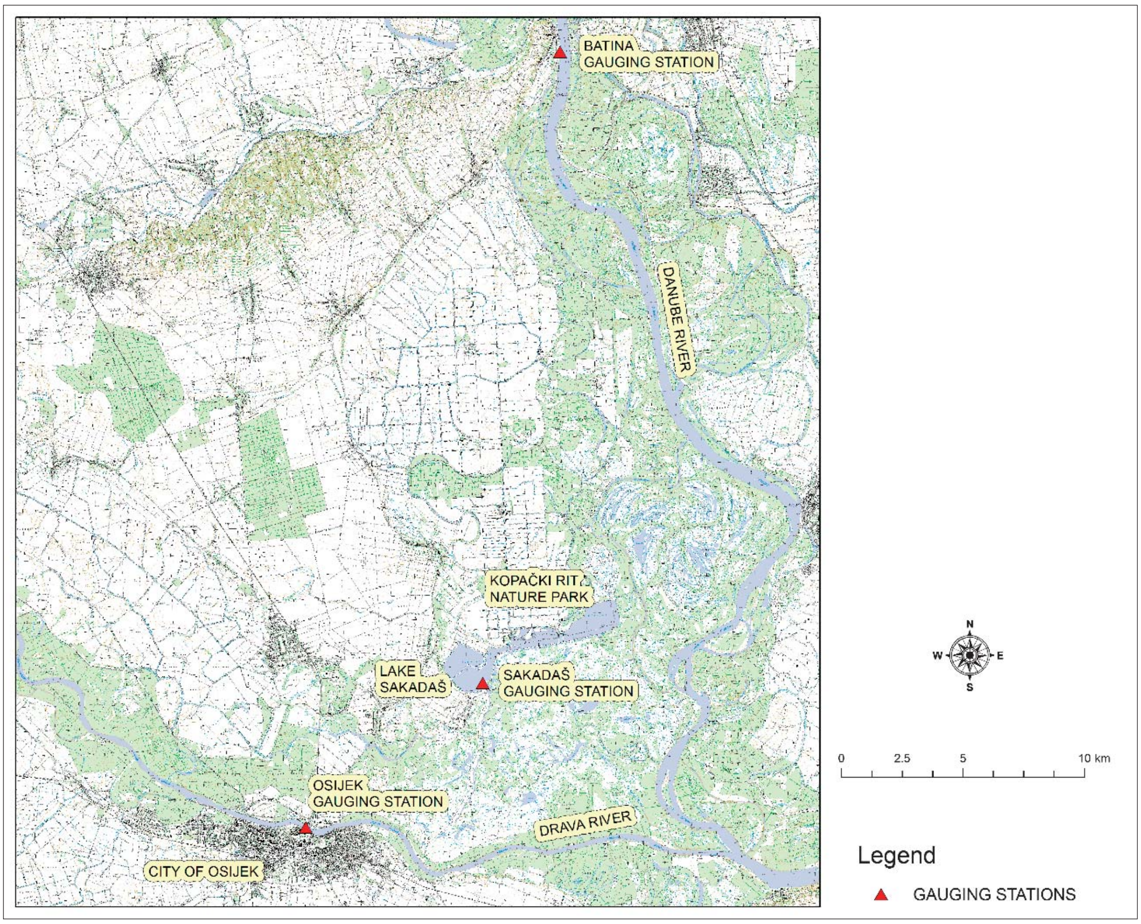


**Figure 10.** Situation map of gauging stations Osijek (Drava River), Batina (Danube River) and Sakadaš Lake (Kopački Rit Nature Park) – Site 3

measured on the same time/date. Calculation process begins with the "Cross-correlation …" button, where the user is required to set the absolute maximum lag positions to be analysed i.e. the absolute maximum number of overlapped positions to be compared and cross-correlated in the two analysed time series. The calculation process ends with the creation of a cross-correlogram and presentation of calculated maximum cross-correlation and corresponding lag time (see **Figure 3**).

## 3. Results

The city of Zagreb, Republic of Croatia's capital, lies on an unconfined alluvial aquifer whose groundwater levels are controlled by the Sava River, a perennial river connected to the aquifer water table. Cross-correlation modelling of surface water – groundwater interaction, namely pressure propagation into the Zagreb aquifer following the Sava River stage rise, was analysed using hourly measurements of the Sava River stages measured at the gauging station Sava – Zagreb and groundwater levels measured in observation wells PJ-9, PJR-9 and P-9 (**Ratkaj, 2014**), situated around the Jarun Lake, Site 1 (see **Figure 4**).

Results of cross-correlation analysis indicate that observation wells responded to the Sava River rise within the range of a few hours to 3.5 days, revealing that groundwater pressure propagated into the alluvial aquifer with pressure travel times equal to 4 hours for observation well PJ-9, 49 hours for observation well PJR-9 and 84 hours for observation well P-9 (see **Figures 5, 6 and 7**). High correlation coefficients ranging from 0.96 for observation well PJ-9 to 0.79 for observation well P-9 also reveal a very strong surface water – groundwater connection.

The second studied site analysed the small river Odra which runs through the eastern part of the Zagreb aquifer system situated within the Sava River catchment area, and its impact on groundwater levels of Zagreb aquifer. Daily measurements of the Odra River stage time series measured on gauging station D. Poljana were compared with groundwater level time series measured in observation well B-153, located in a near vicinity of the Odra River (see **Figure 8**).

Results of the cross-correlation analysis have shown a much weaker relationship of surface water – groundwater connection between the Odra River and groundwater levels of Zagreb aquifer, having low cross-correlation

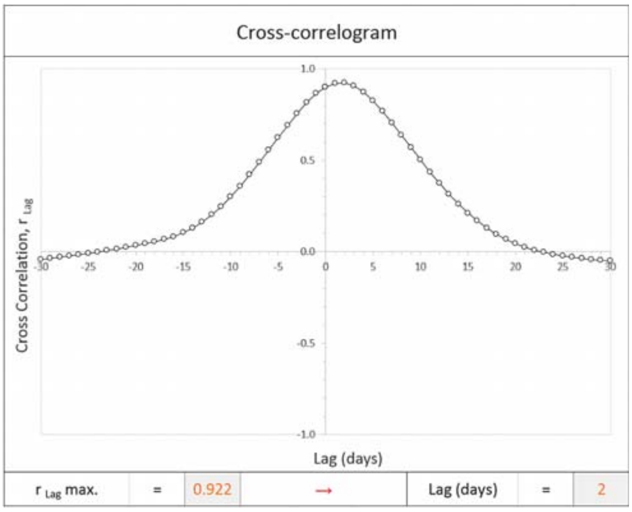


**Figure 11.** Cross-correlogram of the Drava River stage time series and the Sakadaš Lake stage time series

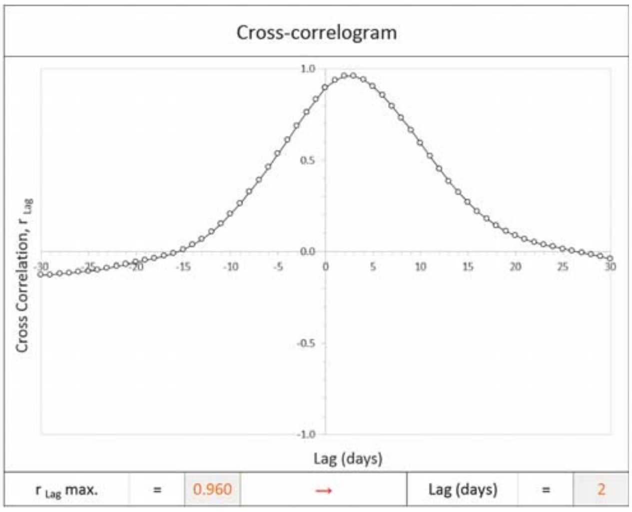


**Figure 12.** Cross-correlogram of the Danube River stage time series and the Sakadaš Lake stage time series

coefficient value of 0.55. Nevertheless, results showed a fast reaction i.e. fast pressure propagation, with pressure travel times equalling 1 day, which was expected due to the very small distance between the Odra River and observation well B-153 (see **Figure 9**).

On the third site, Kopački rit Nature Park, groundwater pressure propagation into the alluvial aquifer system and consequently Sakadaš Lake, following the Drava and Danube River stage rise, was analysed (**Vukojević, 2010**). Kopački rit Nature Park, also called the European Amazon, one of the largest fluvial-marshy lowlands in Europe, is situated in the central part of the Danube floodplain, between two important European rivers, the Drava and the Danube (see **Figure 10**).

Results of the cross-correlation analysis indicated that groundwater pressure propagated into the alluvial aquifer system and consequently Sakadaš Lake, following Drava and Danube River stage rise, with pressure travel times equalling 2 days (see **Figures 11 and 12**). Correlation coefficients ranging from 0.96 for Batina gauging station to 0.92 for Osijek gauging station showed similarly strong surface water – groundwater connection as on site 1, Zagreb aquifer.

## 4. Conclusions

Cross-correlation modelling of surface water – groundwater interaction in a research of groundwater pressure propagation into alluvial aquifer systems of the Sava and Drava/Danube River catchments following river stage rise, focused on estimating groundwater pressure travel times in aquifers, have been conducted using a specifically designed Excel spreadsheet application. Three sites were studied within the Zagreb alluvial aquifer and Kopački rit Nature Park, illustrating pressure propagation into alluvial aquifers as well as the usefulness of the cross-correlation spreadsheet application. Studies revealed very strong surface water – groundwater connection on two studied sites, site 1 and site 3 where large rivers Sava, Drava and Danube are interacting with aquifers, having cross-correlation coefficients larger than 0.9 and indicating pressure travel times ranging from a few hours to a couple of days. Site 2, where river stage time series of the small river Odra and groundwater levels of the Zagreb aquifer were compared, revealed a weaker surface water – groundwater connection, having a cross-correlation coefficient 0.55, respectively. Weaker interaction of the small river Odra and groundwater levels points to pronounced colmation processes in the river that lead to the clogging of the riverbed sediments. Due to far lesser energy of the water flowing in the river Odra, with regard to the river Sava, Drava and Danube, such cross-correlation was expected. The Excel spreadsheet application for cross-correlation analysis offers open-access VBA code and a simple and easy to use user interface. Further, it is offered free of charge in contrast to different commercial applications offering cross-correlation analysis in today's software market. Time series data processing is fast, reliable, and objective, which ensures consistency in the derivation and applicability for larger data sets.

## 5. Cross-correlation Application Availability

The Excel spreadsheet application for cross-correlation analysis with its open-access VBA code may be requested without charge via e-mail to the corresponding author.

## 6. References

*Papers*

Crosbie, R. S., P. Binning, and J. D. Kalma (2005): A time series approach to inferring groundwater recharge using the water table fluctuation method, Water Resour. Res., 41, W01008, doi:10.1029/2004WR003077.

Fiorillo, F. and Doglioni, A. (2010): The relation between karst spring discharge and rainfall by cross-correlation analysis (Campania, southern Italy). Hydrogeology Journal, 18, 8, 1881–1895.

Larocque, M., Mangin, A., Razack, M. and Banton, O. (1998): Contribution of correlation and spectral analyses to the regional study of a large karst aquifer (Charente, France). Journal of Hydrology, 205, 3-4, 217-231.

Lee, J.Y. and Lee, K.K. (2000): Use of hydrologic time series data for identification of recharge mechanism in a fractured bedrock aquifer system. Journal of Hydrology, 229, 3-4, 190–201.

Lee, J., Kim, R.H. and Chang, H.W. (2003): Interaction between groundwater quality and hydraulic head in an area around an underground LPG storage cavern, Korea. Environmental Geology, 43, 8, 901-912.

Lee, J.Y., Lim, H.S., Yoon, H.I. and Park, Y. (2013): Stream Water and Groundwater Interaction Revealed by Temperature Monitoring in Agricultural Areas. Water 2013, 5, 4, 1677-1698.

Seeboonruang, U. (2015): An application of time-lag regression technique for assessment of groundwater fluctuations in a regulated river basin: a case study in Northeastern Thailand. Environmental Earth Sciences, 73, 10, 6511-6523.

Welch, C., P. G. Cook, G. A. Harrington, and N. I. Robinson (2013): Propagation of solutes and pressure into aquifers following river stage rise, Water Resour. Res., 49, 5246–5259.

Wu, C.-M., T.-C. J. Yeh, J. Zhu, T. H. Lee, N.-S. Hsu, C.-H. Chen, and A. F. Sancho (2005): Traditional analysis of aquifer tests: Comparing apples to oranges? Water Resour. Res., 41, W09402, doi:10.1029/2004WR003717.

### *Books*

Davis, J.C. (2002): Statistics and Data Analysis in Geology, Third Edition. John Wiley & Sons, Inc., New York, 638 p.

Ford, D. and Williams, P. (2007): Karst Hydrogeology and Geomorphology. John Wiley & Sons, Chichester, UK, 562 p.

### *Reports written in non-English language*

Ratkaj, M. (2014): Identifikacija utjecaja vodostaja rijeke Save na vodostaj jezera Jarun (*Identification of impact of Sava river stage on lake Jarun stage*). Diplomski rad, Sveučilište u Zagrebu, Rudarsko-geološko-naftni fakultet, Zagreb. (*in Croatian*)

Vukojević, P. (2010): Identifikacija utjecaja vodostaja Dunava i Drave na vodostaj jezera Sakadaš (*Identification of impact of Danube and Drava river stage on lake Sakadaš stage*). Diplomski rad, Sveučilište u Zagrebu, Rudarsko-geološko-naftni fakultet, Zagreb. (*in Croatian*)

## SAŽETAK

## Kroskorelacijsko modeliranje interakcije površinskih i podzemnih voda primjenom aplikacije Excel

Modeliranje odaziva razina podzemne vode vodonosnih sustava, a koji se javlja kao reakcija na promjene rubnih uvjeta vodonosnih sustava kao što su vodostaji rijeka ili potoka, uobičajeno se provodi primjenom statističkih metoda kao što su korelacija, kroskorelacija i regresija. Iako su alati za korelacijsku i regresijsku analizu dostupni u tabličnome kalkulatoru Excel, široko korištenome industrijskom standardu tabličnih kalkulatora, alat za kroskorelacijsku analizu nedostaje. U sklopu istraživanja prijenosa tlaka podzemne vode u aluvijalnim vodonosnim sustavima slijeva Save i Drave/Dunava, koje je u fokusu imalo procjenu vremena prijenosa tlaka podzemne vode u vodonosnicima, aplikacija Excel za kroskorelacijsku analizu izrađena je i korištena u modeliranju interakcije površinskih i podzemnih voda. Primjeri analize terenskih mjerenja na području zagrebačkoga vodonosnog sustava i Nacionalnoga parka *Kopački rit* korišteni su za ilustraciju korisnosti aplikacije. Kroskorelacija je proces u kojemu se uspoređuju vremenski nizovi, kao što su to npr. vremenski nizovi vodostaja rijeka i razine podzemne vode, uzastopnim pomacima jednoga vremenskog niza, čime se određuje vremensko zaostajanje između vremenskih nizova temeljem određene pozicije maksimalne ekvivalencije (Davis, 2002). Određivanje pozicija izraženoga poklapanja vremenskih nizova, kao što su to vodostaji rijeka i razine podzemne vode, upućuje na vrijeme prijenosa tlaka podzemne vode koje je nužno da bi određeni dio vodonosnika reagirao na događaj porasta vodostaja rijeke. Istraživana su dva vodonosna sustava, zagrebački vodonosni sustav i vodonosni sustav Nacionalnoga parka *Kopački rit*, pri čemu su analizirani prijenosi tlaka u aluvijalne vodonosnike kao reakcija na porast vodostaja rijeke Save te Drave i Dunava. Također je dan i prikaz aplikacije za kroskorelacijsku analizu. Istraživanja su otkrila vrlo jaku vezu površinskih i podzemnih voda na dvama istraživanim lokalitetima, 1 i 3, na kojima dominiraju velike rijeke Sava, Drava i Dunav, a koje su u interakciji s vodonosnicima koja se, pak, koeficijentima kroskorelacije opisuje vrijednostima većima od 0,9, čija su posljedica vremena prijenosa tlaka od nekoliko sati do nekoliko dana. Lokalitet 2, na kojemu su uspoređivani vremenski nizovi vodostaja male rijeke Odre i razina podzemne vode zagrebačkoga vodonosnika, otkrio je slabiju vezu površinskih i podzemnih voda koja se koeficijentima kroskorelacije opisuje vrijednošću 0,55. Aplikacija Excel za kroskorelacijsku analizu nudi otvoreni VBA računalni kod te sučelje koje je jednostavno i lako za korištenje. Nadalje, aplikacija je besplatna za razliku od raznih komercijalnih aplikacija koje danas na tržištu nude programsku podršku za kroskorelacijsku analizu. Obrada vremenskih nizova podataka brza je, pouzdana i objektivna, što osigurava dosljednost u izvođenju obrada te primjenjivost na velike setove podataka.

**Ključne riječi**
kroskorelacija, tablični kalkulator Excel, interakcija površinskih i podzemnih voda