Article

The Role of Environmental Background Processes in Determining Groundwater Level Variability— An Investigation of a Record Flood Event Using Dynamic Factor Analysis

Balázs Trásy [1,2,*]**, Norbert Magyar** [3]**, Tímea Havril** [1,4]**, József Kovács** [1,2] **and Tamás Garamhegyi** [1]

[1] Department of Geology, Eötvös Loránd University, 1117 Budapest, Hungary; timea.havril@geology.elte.hu (T.H.); kevesolt@geology.elte.hu (J.K.); tamas.garamhegyi@gmail.com (T.G.)

[2] Institutional Excellence Program, Eötvös Loránd University, 1117 Budapest, Hungary

[3] Department of Methodology for Business Analysis, Budapest Business School-University of Applied Sciences, 1054 Budapest, Hungary; Magyar.Norbert@uni-bge.hu

[4] József and Erzsébet Tóth Endowed Hydrogeology Chair and Fundation, Department of Geology, Eötvös Loránd University, 1117 Budapest, Hungary

* Correspondence: trasybalazs@gmail.com

check for updates

Abstract: Since groundwater is a major source of water for drinking and for industrial and irrigation uses, the identification of the environmental processes determining groundwater level fluctuation is potentially a matter of great consequence, especially in light of the fact that the frequency of extreme climate events may be expected to increase, causing changes in groundwater recharge systems. In the recent study, data measured at a frequency of one hour were collected from the Szigetköz, an inland delta of the Danube. These were then used to determine the presence, or not, and magnitude of any hidden environmental background factors that may be causing groundwater level fluctuations. Through the application of dynamic factor analysis, it was revealed that changes in groundwater level are mainly determined by (i) the water level of neighboring rivers and (ii) evapotranspiration. The intensity of these factors may also be estimated spatially. If the background factors determined by dynamic factor analysis do indeed figure in the linear model as variables, then the time series of groundwater levels can be said to have been accurately estimated with the use of linear regression. The accuracy of the estimate is indicated by the fact that adjusted coefficient of determination exceeds 0.9 in 80% of the wells. The results, via an enhanced understanding of the reasons for changes in the fluctuation of groundwater, could assist in the development of sustainable water management and irrigation strategies and the preparation for varying potential climate change scenarios.

Keywords: background factors; common trends; dynamic factor analysis; evapotranspiration dependent groundwater; hydrogeology; integrated evapotranspiration; managed aquifer recharge; multiple linear regression

---

Highlights:

Diverted river's inland delta assessed.
Dynamic factor analysis (DFA) and multiple linear regression used.
Driving factors of groundwater level determined.
Groundwater levels were estimated using linear models.

1. Introduction

The quality of surface water has a high degree of sensitivity [1], and this is especially true in the case of river water, thanks in part to its multifaceted roles (i.e., habitat for wildlife, source of drinking water, hydroelectric power, and a place for recreation) [2]. Regarding the water quality of rivers, riparian ecosystems play a crucial role [3] and also clearly determine the quality of subsurface waters, their nutrient content, as well as affecting the groundwater (GW) level as well.

Consequently, changes in the quality of surface waters may also have a major influence on groundwater systems. Furthermore, extreme meteorological events may well magnify these effects, as the uppermost layer of the riverbed—the location for biotectonic filtration—may be subject to disruption by a flood [4,5]. These processes can affect not only the urban environment and agriculture [6] but also endanger the drinking water supply system [7] and thus the quality of life in the region [8,9]. The water level fluctuations of a river play a key role in determining the nature of its broader environment.

The frequency of extreme meteorological events is expected to increase in the future [10]; indeed, the three most significant historic floods (i.e., characterized by the largest yield) of the second largest river in Europe, the Danube, have occurred in the last 15 years. The development of strategies to cope with changing environmental conditions predicted for the future requires a detailed knowledge of both surface and subsurface environmental processes; these, in turn, can be clearly determined via the examination of the groundwater itself. River and groundwater levels are important, however, not only from the perspective of flood protection and inland water protection but also on account of the sensitivity of many ecosystems to them. Taken together, these facts make an understanding of the interactions between surface and subsurface water crucial. Surface water can interact with GW in three basic way, depending on the hydraulic conditions of the subsurface zone, the water level of the surface water body, the hydraulic conductivity of the porous media, and on the meteorological settings. If water level of the surface water is lower than the GW level, surface water taps the GW as the latter flows through the riverbed into the river. In the opposite case, river water infiltrates through the sediments into the GW. The combination of the previous two situations can also occur, i.e., surface water can gain water in some part of the streambed from GW and lose it in other parts [11]. Moreover, direction of water flow between the two continua, as well as discharge and recharge processes, can be changed from place to place and from time to time. This dynamic system and its processes are well known and widely described in the literature by previous authors [11–13]. This strong connection between surface and groundwater is exploited by an intentional groundwater replenishment technique known as managed aquifer recharge (MAR) in order to compensate in cases of groundwater shortage. MAR is an increasingly important water management strategy in a world faced with changing climate and rising intensity of climate extremes [14].

The analysis and evaluation of complex datasets from the measurement of a large number of environmental variables represents one of the current challenges to earth and environmental sciences. The construction, development, and operation of an efficient and effective environmental monitoring system requires a detailed knowledge of any environmental processes which may be identified operating in the data. In the case of surface and subsurface water, the main questions are (i) which background factors (i.e., recharge or discharge) that have significant influence on the fluctuation of the water levels cause similar patterns (common trends) in the hydrographs, (ii) how strong are the different background factors, (iii) to what extent they contribute to water level change, i.e., how they affect the water balance. The natural environment cannot be protected and preserved in an appropriate condition without a proper monitoring network measuring water level and key water quality parameters. In order to be prepared in time for different potential environmental scenarios caused by climate change (such as the unequal distribution of precipitation and an increasing number of extreme events, such as floods and droughts), it is necessary to identify the environmental background factors which create and/or influence natural processes and their impact on the groundwater system [15].

Aims of the Study

The aim of the present research is to determine (i) which environmental background processes are foremost in the determination of groundwater level changes, (ii) whether the methodology applied can be reliably used to determine the background factors in the case of extreme events (such as floods) too, (iii) the accuracy of the estimation of time series of the background factors from the water level time series, (iv) the spatial distribution of the errors in these estimates.

Given the nature of the goals and the research site itself, instead of the generally applied deterministic hydrogeological approaches, a decision was taken to employ a stochastic method as more efficient, given that it deals with the random component. The identification of background processes requires the application of a certain kind of factor analysis modeling. In cases in which the data are not independent in time, factor analysis in its conventional form is not suitable. In such cases, dynamic factor analysis (DFA) is the proper method for the consideration of a lagged correlation structure [16], and provides a much deeper insight into the processes operating than ordinary deterministic approaches. In the course of DFA modelling common trends are sought (general patterns that can be observed over a significant part of the time series [17,18]). In other words, it is designed to identify underlying common trends and background factors or latent effects, especially in cases of multiple time series, while at the same time taking into account their mutual relationships [19,20]. In the case of the present study, dynamic factors were obtained on the basis of GW level time series. These were then compared to the water levels of rivers and meteorological data (potential background factors). Numerous studies have been published in recent years in the fields of hydrology and hydrogeology in which DFA was successfully applied to provide a solution to the two-fold problem of finding background (explanatory) processes and tying them to explored time series. For example, [21] used this method to quantify aquifer vulnerability, while the method has also been used to predict the intensity of latent effects governing changes in GW level in karstic [22] and other hydrogeological systems [23,24]. Ritter and Muñoz-Carpena [20] also used the time series of one variable (groundwater level) as response parameters at multiple sites in an agricultural area adjacent to the Everglades National Park. A further two of the more recent studies that should be mentioned are those of Kisekka, et al. [25] and Campo-Bescós, et al. [26].

2. Materials and Methods

The choice of the Szigetköz in NW Hungary as a pilot area came about as this area displays the consequences of numerous human-induced problems. The prime example is the diversion in October 1992 of a 58 km long section of the Danube in the area of the Szigetköz into a 27 km long concrete power canal for the Gabčíkovo power plant at river kilometer (rkm) 1851.75 (Figure 1) [27,28]. The diversion was carried out by Slovakia, because in 1989 Hungary had decided not to further participate in what had originally been intended as a bilateral project for the construction of a hydropower plant. Thus, Slovakia finished a version of the project, which circumvented Hungary and the need for cooperation between the two countries. The project was completed in 1992 [27]. Up until the diversion of the Danube occasioned by this project, the level of the GW had been uniform in the hydrogeologically quasi-homogeneous and isotropic aquifer with a thickness of several hundreds of meters [29]. This level was mainly determined by the natural water level fluctuation of the Danube [24,30]. Following the diversion, however, the majority of the Danube's flow was redirected to an artificial power plant canal that re-joins the original riverbed only at Palkovicovo (Szap), at rkm 1811. By this time, the flow along the original natural riverbed reached only approx. 20%, i.e., 250–350 $m^3\ s^{-1}$, of the pre-diversion yield. Due to river regulation and the construction of a dam, in both riverbeds (natural and artificial), a stable water flow system has developed, accompanied by a river with water level fluctuations measurable on scales of tens of centimeters over most of the year. However, occasional (and for a matter of some weeks annually in total) water level increase in both riverbeds is induced by the draining of the flood water from the upper sections of the Danube in Austria. In the original riverbed system, record flood events (which may be expected to be more frequent in the future) induce major—and therefore

well-traceable—changes that have a significant impact on GW monitoring systems in terms of both quantity [30,31] and water quality [31,32].
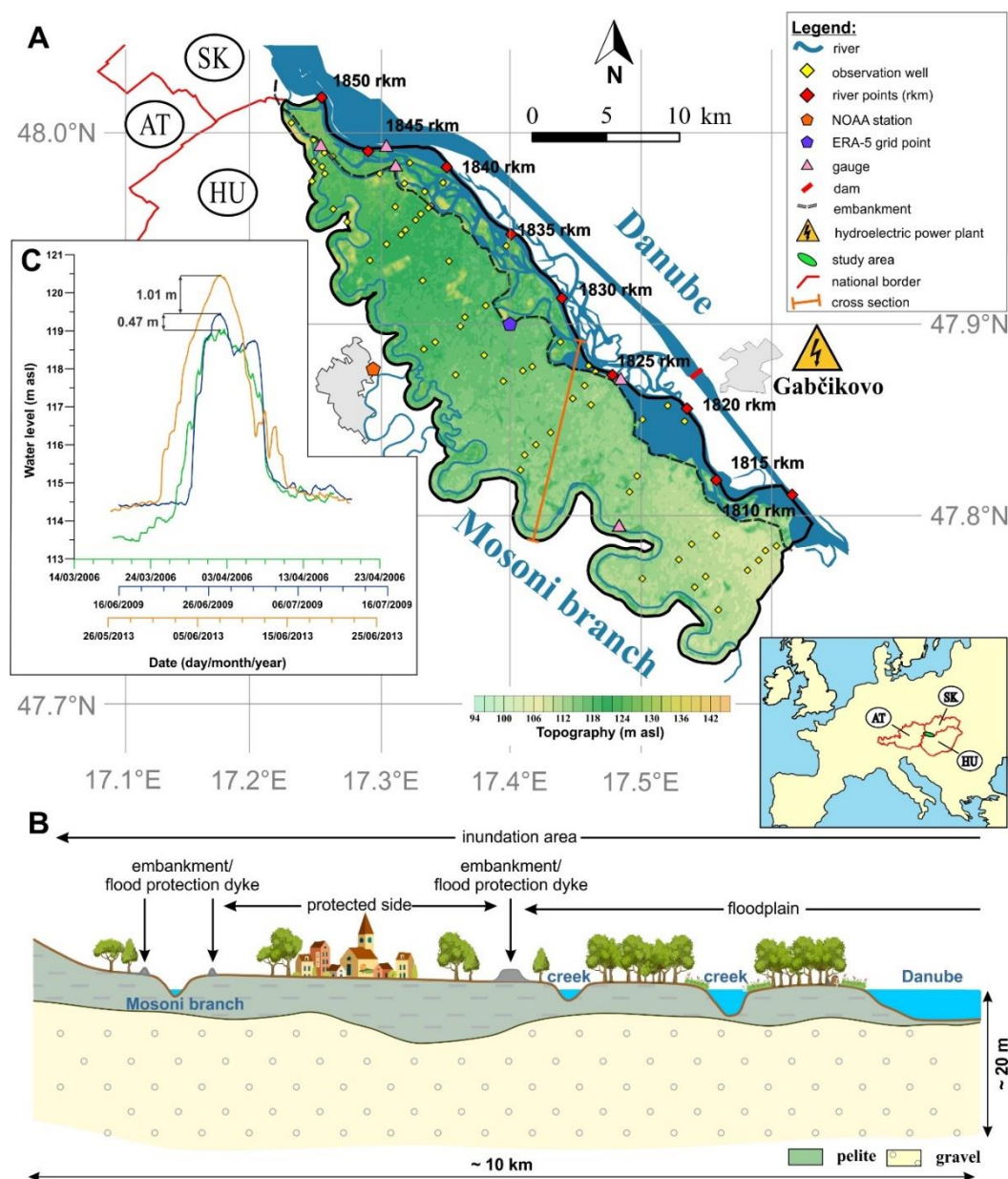

Figure 1. (**A**) Study area, the Szigetköz itself and its monitoring sites, along with the branches of the Danube and the "new" canal with the hydroelectric power plant at Gabčikovo. HU: Hungary, SK: Slovakia, AT: Austria. (**B**) Cross section summarizing the different parts of the broader environment of the river Danube. (**C**) Water levels during the investigated flood events of 2006, 2009, and 2013 [30].

The investigated part of Szigetköz (Figure 1) has an area of 342.7 km^2 and is regularly monitored by several wells. The average distance of the nearest neighbors (885.5 m) was calculated for all data points (i.e., for all monitoring wells), resulting in an average density of 1 well per 4 km^2. Geologically, the Szigetköz area is a part of the Little Hungarian Plain, which developed in the course of the Middle Miocene subsidence and the filling up of the alpine orogeny between the Eastern Alps and the Western Carpathians [33]. The uppermost 100–500 m sedimentary sequence of the Szigetköz is characterized by sand/pelite and gravel sediments. The surface gravel, characterized by high hydraulic conductivity (approx. a magnitude of 10^{-4}–10^{-3} m s^{-1}; [34]), is underlain by fine-grained sediments

with low conductivity, the sand, silt and clay deposits of the Upper Pannonian (approximate hydraulic conductivity is between 10^{-11} m s^{-1} and 10^{-5} m s^{-1} [34]. An overburden Holocene layer of 0–6 m thickness covers the surface of the alluvial fan and is characterized by a hydraulic conductivity of 1.2–2.3 m s^{-6} [35,36]. Isotope-hydrogeochemical methods have been widely applied in the study of the subsurface flow conditions of the Szigetköz. These have generally focused on the determination of the origin, residence time and flow velocity of the GW. On the basis of measurements of ^{36}Cl and Tritium/^{3}He [37,38], it has been documented that (i) the flow velocity within the Quaternary sandy gravel deposit series is 500–530 m year^{-1}, (ii) the main flow direction is NNW-SSE, and (iii) the groundwater is dominantly recharged by the River Danube [38].

Our study is focused on general patterns of water levels and not on the absolute values of them. However, to give an overview of the "normal state" water table and the change of the flow gradient during flood events, it presents potentiometric maps of the study area (Figure A1). For the summary table of the wells' metadata (location, screen elevation, depth, elevation at the wellhead; Table A1), see Appendix A. The spatial variance of the GW levels (i.e., the change in their geostatistical anisotropy) reflects changes in the topography of the water table as well [39]. In order to gain a better visualization of this phenomenon and show the field data, 2D kriged potentiometric maps were plotted. The flow conditions changed to a great extent after the arrival of the flood (see Figure 1) and major direction of water flow turned. In particular, along a 4–6 km corridor of the main branch, the NW flow direction rotated to N, while downstream of rkm 1830 it changed from NW to WNW. The reason behind this alteration in flow direction is the increase in hydraulic potential due to the extending flood front emanating from the Danube. This process overwrites the "normal" flow conditions prevailing in times of no flood.

2.1. Dataset Used for the Analysis

2.1.1. Groundwater Level Data

The location and spatial distribution of the monitoring wells is summarized in Figure 1. The GW levels are measured hourly by the North-Transdanubian Water Directorate using, in most cases, automatic loggers, DATAQUA sensors [40], in accordance with the Water Framework Directive [41] and based on the 30/2004 (XII 24) regulation of the Ministry of the Environment and Water, as it was then designated.

Data were available for the three largest flood events (2006, 2009, and 2013; Figure 1) occurring in the time since the beginning of continuous measurements, though their exact interval had to be determined. Thus, the first and the last points in time at which the water level changed at least 20 cm in 12 h were designated as the beginning and the end of each flood. In the past 25 years, the three highest recorded flood events (2006, 2009, and 2013) were chosen for investigation. The identification of background factors was carried out on the basis of the dataset of the high yield flood in 2013 (27 May–15 September 2013, dataset of 65 monitoring wells). Monitoring datasets of the floods in 2006 and 2009 were also integrated into the analysis. Due to the enormous energy of the flood event of 2013, the highest known historical yield of the Danube, the fine sediment of the riverbed was redeposited [30]. As a result, the clogged river section inhibits the flow of water toward the groundwater system, which can be interrupted by floods [32]. A flood with major environmental impact therefore became the subject of the present investigation. The inter-annual period (27 May–15 September 2013) was defined on the basis of previous results obtained by the same researchers. The results of the variogram analysis showed that depletion of the aquifer after a flood event finishes in autumn.

One fixed gauge is located in the middle of the Danube, one at the Mosoni branch, and others in three sluices located along to the upper section of the Mosoni branch. Data for the one daily water level measurement are available from each sluice.

2.1.2. Meteorological Data

The environmental (explanatory) surface parameters considered in this paper were the following: temperature, recorded hourly (°C), relative humidity (%) wind speed (m s^{-1}), air pressure (Pa), and dew point (°C), all of which were acquired from the National Centers for Environmental Information (NCEI) database of the National Oceanic and Atmospheric Administration (NOAA) for the location of Mosonmagyaróvár (station ID: 12815099999) in the year 2013 (between 1 January–31 December 2013).

Potential evapotrasniration (mm h^{-1}) was calculated using the FAO56-PM formula [42,43]. Net radiation [J] data from the Copernicus and ERA5 databases were combined with those from the NOAA database.

2.2. Methodology

2.2.1. Auto- and Cross-Correlation

Prior to the application of multivariate data analysis methods, the question of whether the data are dependent in time or not must be investigated. This may be achieved using autocorrelation, which is a general tool in environmental sciences for the analysis of a process' memory [44]. Autocorrelation represents the strength of the linear relationship between the time series and its lagged version of itself. The sample autocorrelation function is defined as $cor(x_t; x_{t+k})$, where x_t is the value of x at the time t, and k is the time lag. The value of the autocorrelation function is 1 at lag 0. The further a point is on the offset from the one before it, the greater the change in the process' autocorrelation [45]. The application of this method requires a stationary dataset.

Sometimes, the relationship between 2 different time series needs to be analyzed. For example, in riverside systems, as may be observed, the effects of a flood spread as a pressure surge within the aquifer. Therefore, increases in GW level—caused by the increased water level of the river—occur at different times in the different wells. This temporal relationship may be investigated using the cross-correlation method, which is a good tool for the assessment of delayed effects in datasets.

Cross-correlation is a measure of the similarity (i.e., the linear relation) of two variables as a function of the displacement of one variable by the time lag k relative to the other. Therefore, the sample cross-correlation function is defined as $cor(x_t, y_{t+k})$, where k is the time lag, x_t is the value of x at the time t, and y_{t+k} is the value of y at the time $t+k$. Similarly, the method may also be applied if dataset y is prior in time to dataset x. It is crucial to pay attention to which variable is shifted.

2.2.2. Dynamic Factor Analysis

If the number of time series is sufficiently high, it is possible to look for common trends (similar patterns) in the time series with the use of dynamic factor analysis (DFA) [46,47]. In place of general factor analysis, the application of DFA is suggested in those cases, in which the data of time series are not independent of each other [16,17,19,23,48]. The main idea of this method is to describe the variation in many observed time series as a linear combination of some common trends and factor loadings (weights).

These common trends are resulted by background factors (in this case, environmental variables, e.g., water levels of rivers, precipitation, evapotranspiration, etc., which can have an influence on the GW levels). The background factors can be identified by comparing the obtained common trends to the time series of potential factors. In this study, the sample correlation coefficient was used to measure the strength of the relationship. A high correlation coefficient indicates that, the given common trend is the result of the given background factor. Furthermore, the importance of the common trends, and thereby the identified background factors, can be measured by the factor loadings (weights of common trends in the linear model).

The MARSS (Multivariate Autoregressive State-Space Modelling) method was used to carry out DFA, allowing a set of common underlying trends among a relatively large set of water level time

series to be sought [49]. A detailed description of this method can be found in [50]; for a description adequate to the purposes of the present research, it is worth quoting at some length:

"The MARSS method provides maximum-likelihood parameter estimation for constrained and unconstrained linear multivariate autoregressive state-space models fit to multivariate time-series data. Fitting is primarily via an Expectation-Maximization (EM) algorithm (...). MARSS models are a class of dynamic linear models (DLM) and vector autoregressive (VAR) models (...). Parametric and innovations bootstrapping, Kalman filtering and smoothing, bootstrap model selection criteria (AICb), confidence intervals via the Hessian approximation and via bootstrapping and calculation of auxiliary residuals for detecting outliers and shocks are also applicable."

"Using DFA we are trying to explain temporal variation in a set of n observed time series using linear combinations of a set of m hidden random walks, where m << n. A DFA model is a type of MARSS model with the following structure (Equation (1)):

$$\begin{cases} x_t = x_{t-1} + w_t \; where \; w_t \sim MVN(0, Q) \\ y_t = Zx_t + a + v_t \; where \; v_t \sim MVN(0, R) \\ \qquad\quad x_0 = MVN(\pi, \Lambda) \end{cases} \tag{1}$$

where x_t is a vector containing the common random walks/hidden trends at time t, y_t represents the observed series, w_t is the process errors of the hidden trends, v_t is the observation errors, both are multivariate Gaussian distributed. Q and R are the covariance matrices, Z is the loading matrix, a is the offset, x_0 is the initial value of the random walk. The general idea is that the observations (y) are modelled as a linear combination of hidden trends (x) and factor loadings (Z) plus some offsets (a). The obtained hidden trends are denoted by x_1, x_2 ... x_n, where n is the number of the hidden trends.

As Harvey [49] discusses, there are multiple equivalent solutions to the dynamic factor loadings. We arbitrarily constrained Z in such a way to choose only one of these solutions, but fortunately the different solutions are equivalent, and they can be related to each other by a rotation matrix H. Let H be any $m \times m$ non-singular matrix. The following are then equivalent solutions (Equations (2) and (3)):

$$\begin{cases} y_t = Zx_t + a + v_t \\ x_t = x_{t-1} + w_t \end{cases} \tag{2}$$

and

$$\begin{cases} y_t = ZH^{-1}x_t^\dagger + a + v_t \\ x_t^\dagger = x_{t-1}^\dagger + w_t^\dagger \\ x_t^\dagger = Hx_t; w_t^\dagger = Hw_t \end{cases} \tag{3}$$

$x^\dagger$ are the rotated trends. There are many ways of doing factor rotations, but a common approach is the varimax rotation which seeks a rotation matrix H that creates the largest difference between loadings" (this description of MARSS is quoted in extenso from Harvey [49]).

2.2.3. Linear Regression

Linear regression generally applied to the estimation of the values of a dependent variable based on a set of independent variables. GW level time series (dependent variable) were estimated on the basis of the available hourly meteorological and surficial water level measurements (independent variables) using linear regression with ordinary least squares, as described below (Equation (4)):

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i \tag{4}$$

where Y_i is a value of the dependent variable, Y; β_0 is a constant; X_{1i}, X_{2i}, $\ldots$ X_{ki} are values of the independent variables, X_1, X_2, $\ldots$ X_k; ε_i is the random error, or residual; β_k is the partial regression coefficient belonging to X_k, which indicates how much Y changes for each unit of change in X_k, when all other variables in the model hold constant.

The degree of accuracy of the estimations thus arrived at was characterized by the adjusted coefficient of determination, which is well established in classical regression analysis [51,52]. The best estimates for the linear regression of any given well were also chosen on the basis of this indicator (Equation (5)):

$$R^2_{adj} = 1 - \left[\frac{\left(1 - R^2\right)(n-1)}{n-k-1}\right] \tag{5}$$

where *n* is the number of points in the data sample; and *k* is the number of independent regressors, i.e., the number of variables in the model, excluding the constant.

2.3. Software Used

The DFA was conducted in R 3.5.1, using the MARSS package, while the grids for mapping were obtained with the *sp* and *raster* packages. The mapping was performed using the *sp* package, Golden Software Surfer and QGIS 3.2.1.

3. Result: Which Background Processes Determines the Fluctuation Pattern of Groundwater Levels?

3.1. Time Frame of the Stochastic Relationships

To understand the stochastic relationship between the water level time series of wells and the Danube, cross-correlation analysis was conducted to examine the following: (i) time lag needed for the strongest possible linear connection between the water level time series of the flood observed at the Danube gauge and that of certain given wells, (ii) the strength of these relationships, and (iii) the spatial distribution of the values of correlation coefficients indicating the strongest linear connection. When the time lag of the water level time series of the Danube is observed in relation to wells, a higher correlation coefficient is seen than without the displacement (Figure 2).
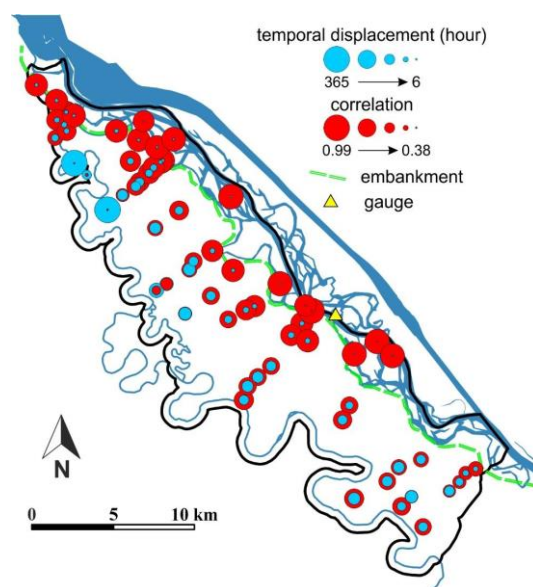

Figure 2. The results of cross-correlation between the water level time series of wells and the Danube gauge. As distance from the river increases, the maximum value of the correlation coefficient diminishes, while the degree of temporal displacement related to the given coefficients also increases.

The correlation coefficient reached its highest value (average: 0.81, Q1: 0.74, Q3: 0.91) in the course of shorter time lags within a zone 2–3 km wide along the Danube; further, it may also be observed that the lag associated with the highest correlation coefficient (average time lag in hours: 91 Q1: 41, Q3: 123)

increases in a direction perpendicular to the course of the Danube. Higher correlation coefficients were obtained in the floodplain area than in the wells further away from the Danube (Figure 2).

3.2. The Use of Dynamic Factor Models to Obtain Common Trends

Dynamic factor analysis was conducted to determine the presence, or not, of any pattern/common trends (CTs) in the time series of the GW levels. The application of the method is justified in cases in which a data series is dependent in time, so an autoregression function was applied in order to verify this. For each and every well, it turned out that even in the case of a 24-h displacement there is a strong linear connection between the displaced and the original time series (average r: 0.97, min: 0.93; max: 0.99); therefore, the application of DFA is well-founded.

The dynamic factor models were constructed using different parameters (number of identified CTs, different structures of the variance-covariance matrix denoted by *R*), and the best model was chosen on the basis of the Akaike information criterion (AICc, [53]). It is recommended that the model with the lowest AICc value be used [18]. The AICc values of the models differed greatly (Table 1), which only serves to emphasize the importance of finding the appropriate settings.

Table 1. The Akaike information criterion (AICc) values of the dynamic factor models relative to each other based on the structure of variance-covariance matrix (R matrix) and the number of common trends determined.

Number of Common Trends	Structure of R Matrix	ΔAICc
3	unconstrained	0
2	unconstrained	6006
1	unconstrained	14,077
3	equalvarcov	465,215
3	diagonal and unequal	472,695
3	diagonal and equal	580,897
2	equalvarcov	582,007
2	diagonal and unequal	606,329
2	diagonal and equal	706,953
1	equalvarcov	722,423
1	diagonal and unequal	788,822
1	diagonal and equal	884,828

It is apparent from the relative differences of the AICc values between the examined models that the best result can be obtained by means of three CTs, without putting any constraint on the elements of the variance-covariance matrix (*R*) and estimating them instead (unconstrained model). The estimates obtained using other models (different variances and no covariance, same variances, and no covariance) are considerably worse (Table 1).

The unconstrained model was applied to determine the most characteristic CTs of water level time series. In all three cases of the obtained CTs (Figure 3), the maximum absolute values can be found in the ±2 week time interval around the flood's maximum extent/level (7 June 2013).

The factor loadings show the weight of the given CT in the given well within the linear model obtained by DFA. These provide information with respect to the significance of the given background factor connected to the given CT, at the observation point. In Szigetköz, the factor loadings display a high degree of spatial distribution (Figure 4). The loadings of CT_1 display high values longitudinally, from the middle zone to the south-eastern border. CT_2 is crucial in the northern half of the Szigetköz, but it is dominant up to the Slovakian backflow of the Danube and throughout the floodplain zone. The loadings of CT_3 are highest in the strip running alongside the Mosoni branch, mainly in the upper part of the river, but the loadings (ranging from −0.16 to 0.2) are nonetheless considerably lower than those of the other two CTs.
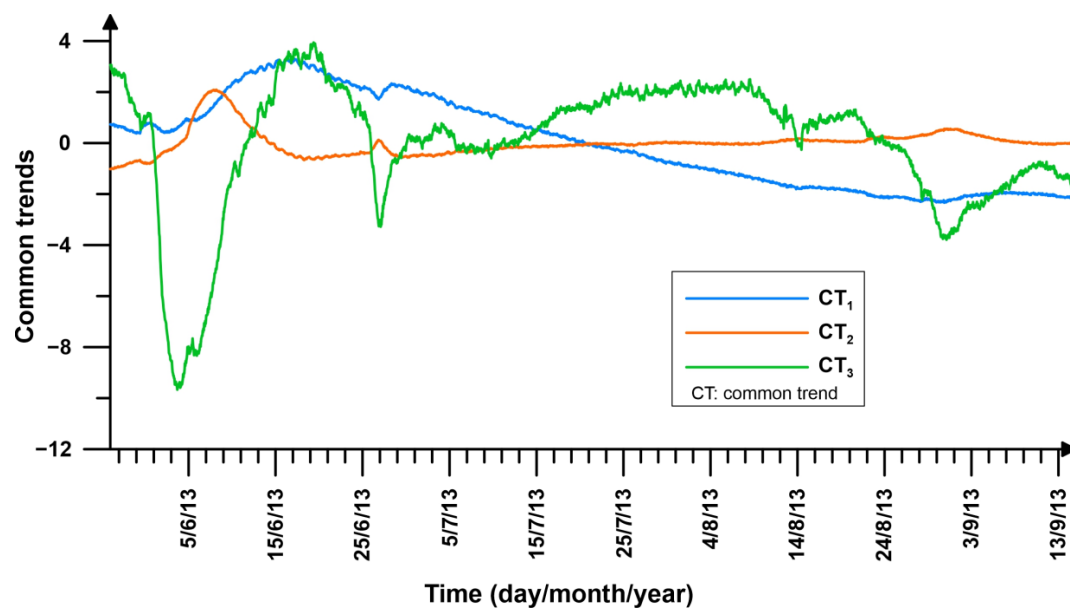
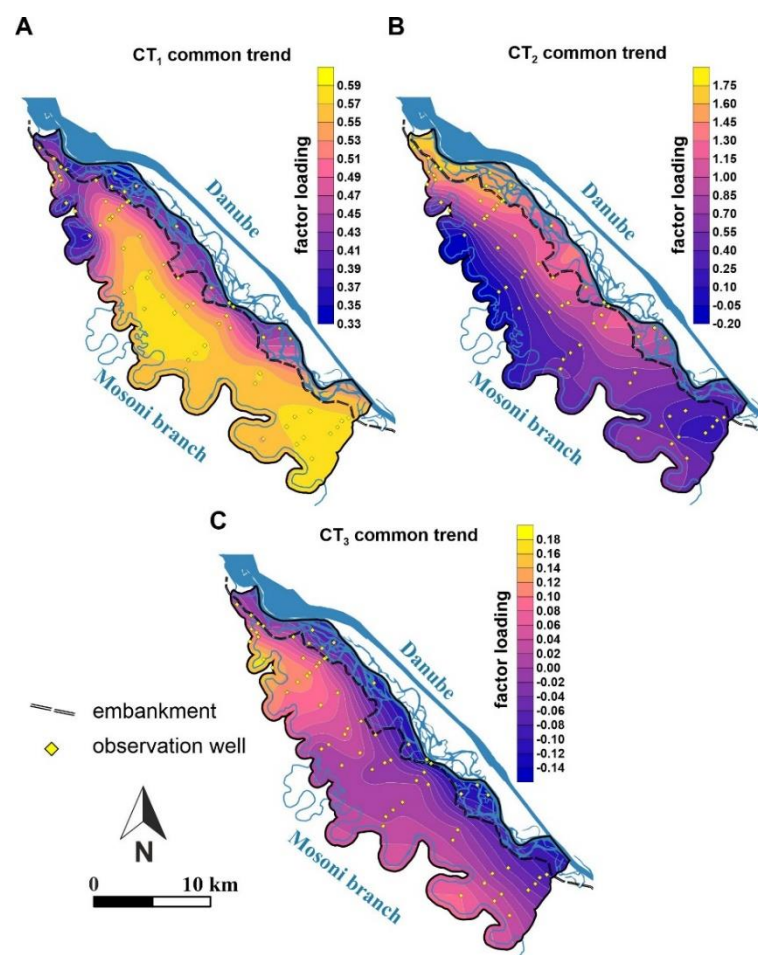**Figure 3.** The common trends of groundwater levels.

Figure 4. Spatial distribution of the loadings of common trends (CT_1, CT_2, and CT_3) determined for the investigated period.

The aim of investigating the relationship between a given CT and the time series of the potential background factors (e.g., meteorological parameters and surficial water levels) was to identify which

background processes cause the CTs arrived at in the model. The temporal and spatial nature of the impact of the "linear" factors of action (the rivers) on groundwater has already been recorded by cross-correlation analyses (Section 3.1). On this basis, it seems evident that the time series of potential background factors should be displaced temporally, too, to be consistent with the trends explaining a significant part of the different water levels' total variance. However, the degree of displacement differs from those presented in the previous chapter (Section 3.1), and this is inevitable, as CTs represent a general fluctuation pattern. In order to achieve consistency with the displacement that provided the best correlation coefficient, the three calculated CTs were cross-correlated with the time series of potential background factors. Based on the calculated correlation coefficients the three background factors were obtained (Figure 5).
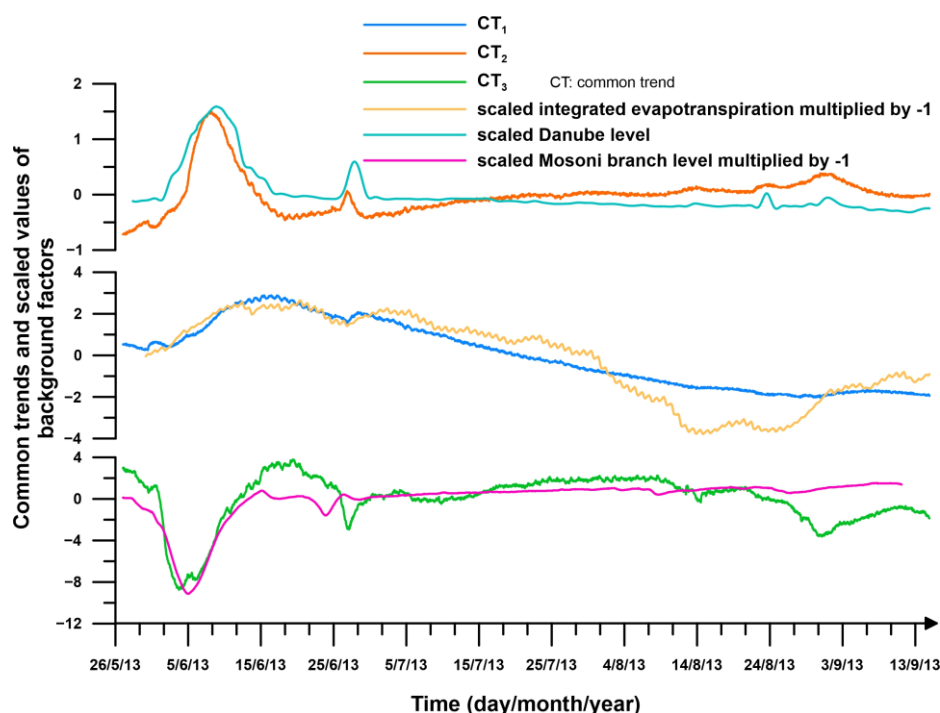

Figure 5. Common trends and the time series of background factors related to them in Szigetköz.

CT_1 was identified with the integrated evapotranspiration as a background factor, which is in close linear relation with CT_1 ($r = -0.896$), even without temporal displacement. A similar approach has been successfully applied in other studies [54–59] in which the integrated evapotranspiration played a significant role in the fluctuation of water levels. No major increase in the strength of the relationship was observed, not even when the CT_1 was displaced temporally—the correlation coefficient did not change significantly with a 4-day displacement either (Q1: -0.913, Q3: -0.905). Moreover, the highest $|r|$ value was 0.914, with a 76-h displacement.

CT_2 correlates well with the water levels of the Danube gauge, which means that the flow regime of the Danube determines behavior observed at wells where the weight of CT_2 is high. The correlation coefficient between the river water level and CT_2 is 0.60, but as a consequence of a 33-h displacement of the data from the Danube gauge, it improves to 0.67.

Making CT_3 consistent with a background factor proves to be considerably more complex. One possible reason for this may be the time series of the Mosoni gauge. The water level of the Mosoni branch is artificially regulated by the floodgates located near the reservoir of the hydroelectric power plant, so the flow regime of the river is almost entirely under anthropic control [60–62]. Among others, floodgates (NW part of the area) are used for the water flow regulation of the river, and they measure the water level only once a day [61]. The examination of data series showed that the operation of the floodgates determines whether it is the Mosoni branch or the supplementary system that drains

the waters from the reservoir. Consequently, the difference in the water level of the floodgates shows which channel (supplementary channels vs. the Mosoni branch) is draining the water at any given point in the investigated area. The strength of the linear relationship obtained from the daily data of the Mosoni gauge and the difference of the time series of the floodgates is marked: r = −0.96. The negative sign and the high value of |r| indicates that the two-time series are in close relation and in opposing linear connection. This close stochastic connection makes it possible to use the time series of the Mosoni gauge in place of that of the floodgates, the difference being that it is multiplied by (−1). In this way, CT_3 may be identified; the correlation coefficient with a 92-h displacement is 0.67. (Without displacement, the correlation coefficient is 0.41; Figure 5).

On the basis of the CTs and their loadings, the time series of wells were estimated based on the linear model obtained by DFA for the investigated time period to examine how efficiently the three CTs determined describe the variance of the water levels. The effectiveness of the estimations was characterized by the adjusted coefficient of determination (R^2). In the case of the DFA, the accuracy of the estimation could only have been assessed by taking all the CTs into account because the model adapts. On the basis of the results, it can be assumed that the three determined CTs did indeed prove to be efficient when it came to the estimation of the water level time series (average R^2: 0.94, Q1: 0.93, Q3: 0.99). Only in the case of a few wells (12.3%) was an R^2 value lower than 0.9 obtained.

4. Discussion

Groundwater is a geologic agent [63], i.e., its occurrence and movement affect a wide range of geologic and environmental processes [20,64], and continuously interacts with the ambient environment. The construction, development, and operation of an efficient and effective environmental monitoring system requires a detailed knowledge of any environmental processes that may be identified operating in the data. Therefore, to determine the factors affecting GW level data variance as precisely as possible has a key importance [20,48,65–67]. The application of methods capable of managing water level changes that take place due to extreme hydrometeorological events, as in the investigated period of time at Szigetköz, is of the highest priority [68,69]. In the case of the interactions between surface and GW, the fact that the effects of surface processes are usually felt underground with temporal delay [70] has to be taken into consideration in the estimations. In this study, the linear models were constructed with the use of linear regression, a common method in the profession, and were compared with DFA, using the GW level data of the time around a flood period [71].

On the basis of research carried out previously in the area, it may be stated that the Danube has one of the greatest influences on GW levels in the Szigetköz [24,37,72]. Given this, the temporal delay with which any change in the water level of the river appears in GW was determined using cross-correlation, that is, the point at which the linear connection is the closest between the water level times series of the river and the wells. It was established that as the distance from the river increases, the maximum value of correlation coefficients calculated from the temporal displacements diminishes, which happens because drawing away from the river, its impact on the GW level decreases. The pressure surge starting from a water level change in the Danube (e.g., the water level increase during floods) is mainly perceptible right in the river basin in terms of GW water levels. Nevertheless, it may also be observed that as one moves further from the river, the highest correlation coefficients were obtained with an ever-increasing time lag in the line of wells perpendicular to the Danube, which can be explained by the spread of the pressure surge and its temporal delay (Figure 2). This, in turn, sheds light on the fact that the most accurate estimation of the water level time series of wells with linear regression can be obtained by taking into account the temporal displacement of the water levels of rivers.

4.1. Identification of Background Factors

Based on the assessment of the degree of autocorrelation in the dataset, DFA may be considered adequate to identify common trends and the background factors shaping them [17,18,73]. The calculated CTs were identified with (i) the integrated evapotranspiration as well as the flow regimes of the

(ii) Danube and (iii) the Mosoni branch. However, it was established that the loadings of CT_3 are considerably lower than the other two (Figure 4), so this factor has an impact on data variance, that is, it does not contribute to the water level fluctuation in the area to such a great extent. In the course of the identification of background factors, the correlation coefficients calculated increased (in the cases of CT_2 and CT_3 by 0.12, on average), provided that the temporal delay was taken into account. As a special case, CT_1 did not yield better results, not even when the time series of the background factor (the integrated evapotranspiration) was displaced temporally (Δr 0.018 with a 72-day displacement). This indicates that the evapotranspiration process produces its effect laterally in the GW level virtually directly, and in a very short time; this had been confirmed in previous studies as well [74–77]. On the other hand, due to the temporality of the pressure caused by a rising water level, the impact of the river is delayed. The integrated evapotranspiration—considered a "surface factor" on account of the location of its impact—is crucial in areas where the impact of linear explanatory factors (rivers) is lower.

The average discharge of the River Danube (CT_2) is an order of magnitude higher (300–400 m^3 s^{-1}) than that of the Mosoni branch (20–30 m^3 s^{-1}). It is this factor—the river water level of all times—that affects the water level time series of wells in the floodplain area to the virtual exclusion of others (Figure 4). During the flood (early June), the high-energy water flow was capable of fracturing and remobilizing the fine fraction sediment in the clogged section of the riverbed [32,78]; this is where fine sediments settle down, causing low hydraulic conductivity on the riverbed reducing interaction between surface and GW. The degree of connection between the two continua therefore improved. Due to its higher hydraulic potential, the Danube could induce the pressure surge at every point, and it spread perpendicular to the river. Trásy et al. established that the phenomenon of water pouring from the river into the GW occasioned by the fracture of riverbed clogging changes the water flow temporarily, perhaps only for a few months, but nonetheless drastically [30]. As a result of this process, from the start of the flood, the connection of the river and the GW improved, so the river was able to affect the entire floodplain zone in the investigated area. However, in the upper sections of the area, where no significant riverbed clogging had taken place [32,78], the impact of the Danube proved to be more significant, because the infiltration rate of the riverbed remained good throughout the entire period. An additional factor in the significance of this impact is the fact that the thickness of the upper layers at the riverside is maximum 1–2 m, and the underlying strata are composed of gravel with high hydraulic conductivity [35,36]; this, naturally, does not stop infiltration [32]. The third CT (CT_3)—connected to the flow regime of Mosoni branch—also acts as a "linear" factor in terms of affecting the water levels.

The value of the coefficient of determination (R^2) was high practically without exception at every observation point between that obtained from the water level time series estimated based on the dynamic factor model and measured well data. On this basis, it may be assumed that by using the determined dynamic factor model, measured GW levels can be estimated with just a minor loss of information (R^2 exceeded 0.9 in the case of 57 wells, 87%). It can be stated that these three CTs—linked to different background processes (evapotranspiration, water level changes in rivers)—explain the majority of variance in water levels in the wells.

Over the course of the four months examined, no other pattern (common trends) was traceable over a considerable part of the time series of water levels. Two out of the three identified CTs (CT_2 and CT_3) are linked to local water flow systems. In terms of the space and time investigated, the impact of regional (basin-scale) groundwater flow systems was not traceable. One of the reasons for this may well be that the effect of a regional groundwater system cannot be detected over a time interval as short as that investigated.

4.2. Estimation of the Water Levels Based on Background Factors

It is crucial in the planning of monitoring systems to ensure they are able to describe the GW table properly on the basis of the water level time series of wells [79–81]. A framework for an optimized groundwater monitoring network and aggregated indicators is required for a comprehensive

environmental study. Hydrometeorological phenomena (e.g., precipitation, evapotranspiration, and temperature as driving factors) and surface waters usually affect the water levels of shallow aquifers to a significant extent [12,82]. In the absence of a better understanding of the effects of these driving factors, the representativity of monitoring systems cannot be guaranteed. For assessment, multiple linear regression, a commonly used estimation method, was employed [83,84]. The accuracy of the estimations was characterized by the adjusted R^2, which shows what percent of the total variance of the water levels is explained by the set of the independent variables (i.e., water levels of the rivers and the amount of evapotranspiration).

When it comes to the Szigetköz area, in the Danube [24], it is a determining factor [24] which water level is measured by the Danube gauge [31]. Measurements from this were used in the first instance as an independent variable to estimate the water level time series of wells. The average value of adjusted R^2 was 0.50 (Q1: 0.31; Q3: 0.72). In the cases of 6 wells (9% of the total), the value of adjusted R^2 exceeded 0.9, meaning that in those cases more than 90% of the total variance is explained by the water level of the Danube gauge. The insertion of the water level of the Mosoni branch into the linear model results in better estimations, especially on the protected side of the inundation area (Figure 1), that is, on the side which floodwaters cannot reach, because of the dam (average R^2: 0.61 Q1: 0.42; Q3: 0.83).

Cross-correlation assessment revealed that changes in the water level of the Danube appear in the hydrographs of wells with a certain delay (Section 3.1), and this delay depends on the distance from the river. With this in mind, it is worth carrying out the linear estimations while taking into account the temporal displacement of the background factors. Negative time lag is also necessary, because the Danube gauge is located in the middle section of the Danube, while the one on the Mosoni branch is in the lower one-third of the stream. Therefore, the effect of flood (i.e., water level increase) can be detected earlier in the upstream wells compared to the gauge located in the lower section of the river. The frontline of the flood in the examined area rolled away in less than 12 h [62]. In the case of all the wells, the linear model with the highest adjusted R^2 was chosen. In such a way that the variance inflation factor (the greater the value of the variance inflation factor (VIF), the less reliable the regression results will be. In general, a VIF above 10 indicates a high degree of correlation and is cause for concern) could not exceed 10 in any regression. Due to the displacement (meaning that of the time series of the gauges at the Danube and the Mosoni branch), the increase in the adjusted R^2 was rather large in several wells (average R^2: 0.83, Q1: 0.76, Q3: 0.94). In the case of 25 wells (38%), the adjusted R^2 exceeded 0.9. The displacement is especially noticeable in the cases of wells located on the protected side: when compared to cases without displacement, the most substantial improvement took place in this area (Figure 6).

Besides the rivers in the area, the groundwater level can be influenced enormously by evapotranspiration, as well [85,86], and so it is worth including it in the independent variables, as seen in the results of the dynamic factor analysis (Section 3.2). The inclusion of hourly potential evapotranspiration does not, however, increase the accuracy of the estimations significantly (average R^2: 0.83, Q1: 0.77, Q3: 0.94). Instead of evapotranspiration values, it might be better to use integrated potential evapotranspiration [87] for the estimation of the water levels. This parameter cumulates the deviation of the given potential evapotranspiration value from the average evapotranspiration of the whole time frame. A significant increase can be observed in the value of R^2 (average R^2: 0.91, Q1: 0.91, Q3: 0.96)—that is, there is a significant increase in the degree of accuracy in the estimation—principally on the protected side (Figure 7A). By shifting the time series of the Danube and Mosoni gauges and using integrated evapotranspiration, the value of R^2 proved to be higher than 0.9 in 52 wells (80% of wells), which means a remarkable 42% increase compared to the estimations based on the displaced gauge time series. Several descriptive statistics of R^2 related to the given set of independent variables can be seen in Figure 7B. Upon the examination of the spatial distribution of R^2 values, it can be observed that with the inclusion of the integrated evapotranspiration and the Mosoni branch the accuracy of the linear regression estimation increases considerably at wells on the protected side (Figure 7).
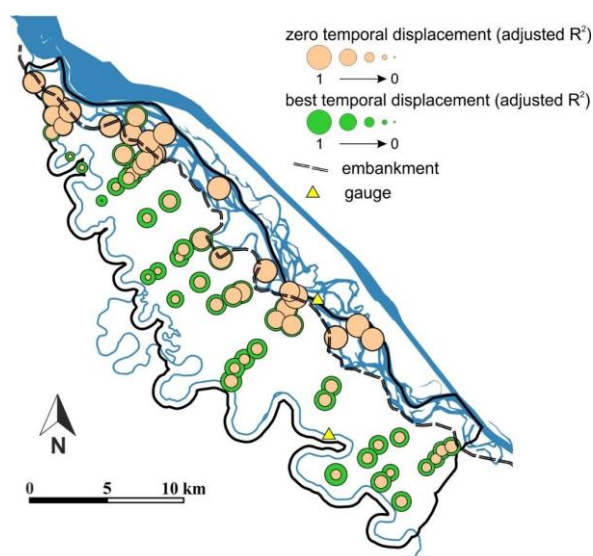
Figure 6. Spatial distribution of adjusted R^2 values, demonstrating that the temporal displacement of the water level is crucial, especially on the protected side, to achieve better modeling. The radius of the circles is in proportion to the adjusted R^2 values.
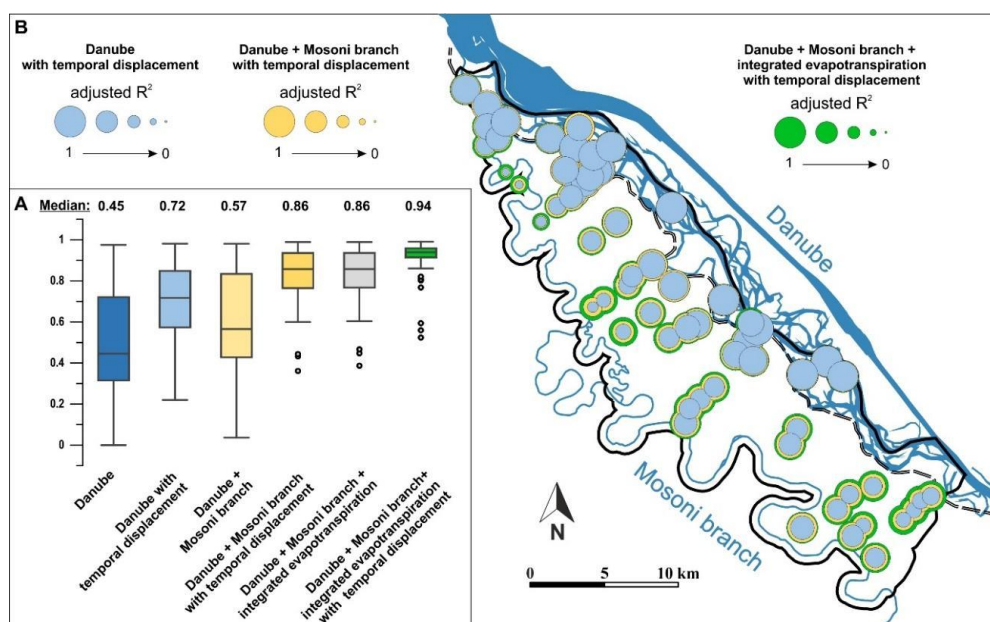

Figure 7. Box-and-whiskers plots of the R^2 values (**A**) and their spatial distribution (**B**) in relation to the given set of independent variables (blue: based on only the displaced Danube gauge, yellow: displaced Mosoni gauge included, green: integrated evapotranspiration also included).

4.3. Importance of Time Shift in Linear Regression Models

The water level of rivers is measured only at a number of points (gauges; see Figure 1A), so it might happen that the impact of a flood shows up sooner in the water level time series of a well further upstream from the gauge than in the water level time series of the gauge closest to the event as it occurs. In light of this, it is understandable why one has to apply the displacement of background factors in a negative direction in the case of regression. At times like this, it is worth considering displacement in opposing directions (the well's time series have to be displaced, as well as those of the gauge). It is for this reason that the best estimations were obtained by displacing the water level time series of wells in a negative direction, too. In addition to the flow regime of the Danube, the Mosoni branch and other

channels with similar flow regimes play a crucial role in the change of water levels. The estimations obtained by including the time series of the gauges on the Danube and the Mosoni branch in the model were enormously improved (average R^2: 0.83, Q1: 0.76, Q3: 0.94). However, the best estimations were obtained when, besides the water level of the rivers, the integrated evapotranspiration data was incorporated into the estimations (average R^2: 0.91; Q1: 0.91, Q3: 0.96); this would imply that evapotranspiration has a measurable role in affecting the water levels over the investigated time scale (Figure 7).

4.4. Comparison of the Results of Different Linear Models

The obtained dynamic factor model proved to be accurate thanks to two facts, namely:

- In most cases the observed water level time series can be estimated with a high degree of confidence on the basis of the linear combinations of the CTs and the factor loadings.
- The obtained CTs could be connected to those background processes (integrated evapotranspiration, water levels of rivers), which were significant independent variables in the linear regression model.

One of the main advantages of DFA is that the spatial distribution of the background factors' significance can be illustrated on maps on the basis of the loadings of CTs.

On the other hand, DFA cannot handle the fact that the effects of changes in surface water level can be observed only with a time lag in the different wells, and as a result, is capable of yielding only some general pattern of water levels on the basis of all the wells.

Taking the previous statements into consideration, the accuracy of estimations of GW levels using a linear model obtained from linear regression or from a dynamic factor model were assessed. The observed and estimated water levels of two wells can be seen in Figure 8.
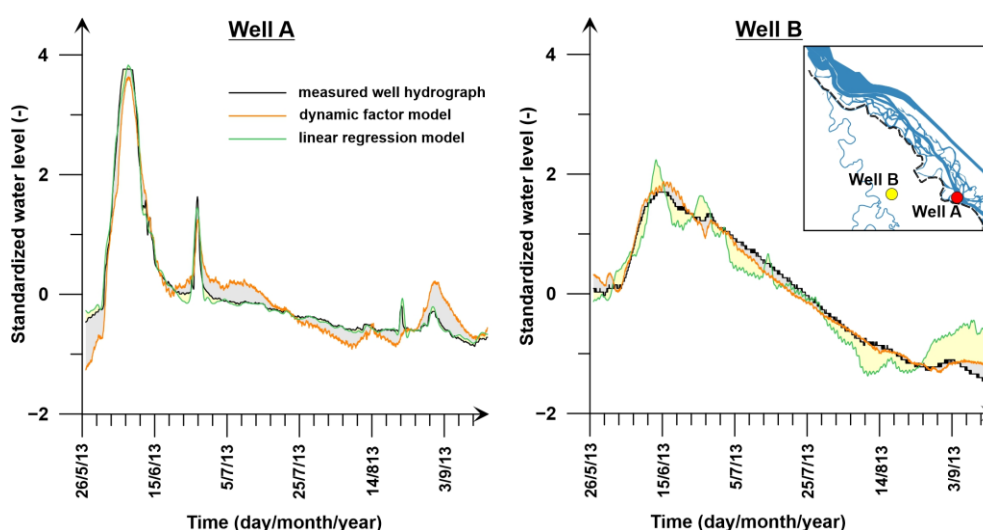

Figure 8. Hydrographs of two wells—"A", located in the floodplain and "B" on the protected side—with the results of the estimations obtained from different linear models and their difference colored gray and yellow.

As seen in Figure 7, the best estimations with linear regression can be achieved using the following set of independent variables: (i) the Danube gauge, (ii) the Mosoni gauge, and (iii) the integrated evapotranspiration.

Comparing the accuracy of estimations performed using DFA and the linear regression, the DFA gave better results in more than two-thirds (68%) of cases, despite the fact that the displacement of the background factors cannot be taken into account in this method. The spatial distribution of the differences at the highest achievable level of accuracy of estimation using the two methods is shown in

Figure 9. Better estimations were obtained in the floodplain area of the Danube using linear regression, a fact that can be explained by the very close linear connection between the wells and the river (typically r > 0.9). The phenomenon arises because only one background factor (i.e., the flow regime of the Danube) dominates the wells of the floodplain, so the delay with which the water level fluctuation can be observed is a more significant factor. Regression is capable of taking into account the temporal delay, a matter of some hours, with which the impact of the water level fluctuation of the river, may be observed in the well. In DFA it is not possible to include the time shift in the estimation of certain CTs. Latter is obtained from the time series of all the wells, while in the case of linear regression, the models were constructed well-by-well, based on the background factors.
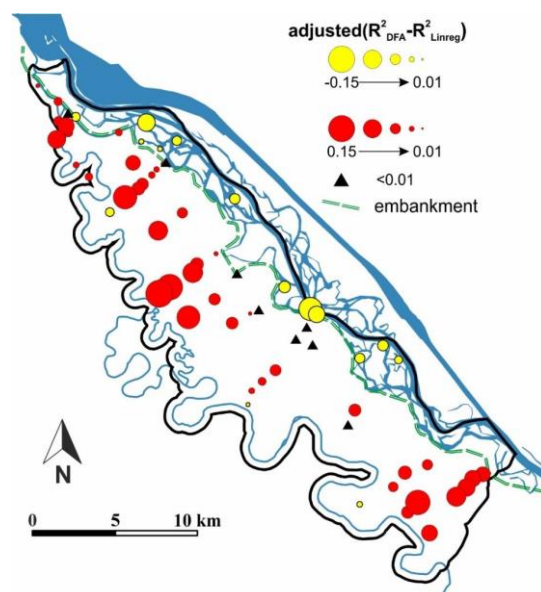

Figure 9. Differences in the value of adjusted R^2 in relation to the best models obtained from DFA and linear regression. Yellow circle: linear regression model is better; red circle: DFA model is better; triangle: difference is below 1%.

5. Conclusions

The diversion of the Danube in the Szigetköz area resulted in a significant change in the previously natural fluctuation of groundwater. The hidden background factors guiding these fluctuations were determined using dynamic factor analysis (Section 3.2). On the basis of the results, it was concluded that fluctuations in the groundwater level time series are mainly determined by the rivers and evapotranspiration (Figure 5). With the use of DFA, at each sampling point, the intensity of different background factors can also be estimated (Figure 4). As one of the background factors, evapotranspiration was identified, which plays the most significant role in the inner parts of the Szigetköz, mainly on the protected side. The water level fluctuation of the River Danube was identified as another background factor and was most powerful on the floodplain and in its immediate vicinity. The natural and artificial surface water paths of the area are considered the third background factor; the intensity of this factor is greatest in the southwestern and southeastern parts of the study area.

The time series of groundwater levels can be accurately estimated using linear regression if the background factors determined by dynamic factor analysis are used as independent variables in the linear model (Section 4.2). The accuracy of the estimate is indicated by the fact that in 80% of the wells, the adjusted coefficient of determination for the measured and estimated time series exceeds 0.9, which means that more than 90% of the total variance of the water level is explained by the background factors. The accuracy of estimation of the measured water level ranges improves significantly if the individual background factors are incorporated into the estimation process by determining in advance

the amount of time it takes for the given factor to reach the observation point (Figure 7). Thus, when applying linear regression, the time difference between the effects of various independent variables in the given water level time series is taken into account. In the case of "linear" background factors (rivers), this time lag is typically several days, while the delayed inclusion of a "surface" background effect (evapotranspiration) in time does not improve the accuracy of the estimation, i.e., it takes effect on the groundwater level within a very short time.

The results presented here highlight the need to compare the results of different model settings, as there may be significant deviations in their efficiency and/or appropriateness. In the case of floods, the commonly used numerical or semi-numerical models are typically less accurate at extreme levels. DFA also returned accurate data in the case of the special hydrological situation under examination, as well, and is therefore recommended to use this method for the modelling of these types of extreme events (floods and droughts). The results represent evidence that with the methods laid out here, it should be possible to model future high-magnitude flood events, which are likely to occur as a result of climate change.

Appendix A

Table A1. Metadata of wells assessed.

Well ID	WGS84 Latitude	WGS84 Longitude	Centre of Well Screen (m)	Well Depth (m)	Wellhead Elevation (m.a.s.l.)
62	47.783149	17.59044689	5.9	8.4	112.98
66	47.78833156	17.59564306	4.9	8.2	112.96
140	47.96504975	17.33023729	5.0	6.5	124.36
147	47.9445798	17.29621462	7.7	9.6	124.76
148	47.95223993	17.31065967	4.5	7.9	124.68
149	47.91354195	17.37445304	8.4	11.1	122.54
159	47.90248394	17.355061	5.8	7.0	122.03
3474	47.75682587	17.55996011	5.0	9.5	114.82
3478	47.96404833	17.32820611	4.8	8.6	125.18
3587	47.89388523	17.33576692	6.0	9.0	122.23
3591	47.95756653	17.31850121	4.5	8.9	124.71
3592	47.96082923	17.322411	4.5	9.2	124.10
3814	47.77779318	17.58199545	4.5	8.2	113.47
3816	47.79085741	17.60411415	5.5	9.2	113.55
3818	47.90727775	17.35842159	6.5	10.0	122.10
4063	47.94975274	17.30739154	5.5	9.0	124.67
4064	47.93587934	17.28357788	5.5	10.4	124.45
4066	47.89005148	17.327057	5.5	11.4	123.47
4123	47.87279472	17.45264943	5.5	10.4	118.06

Table A1. *Cont.*

| Well ID	WGS84 Latitude	WGS84 Longitude	Centre of Well Screen (m)	Well Depth (m)	Wellhead Elevation (m.a.s.l.)
4125	47.82792965	17.40419226	4.5	8.8	118.14
4126	47.9843638	17.24466649	5.5	8.8	128.10
4129	47.97663686	17.23703674	6.1	9.3	126.72
4322	47.84131643	17.41577716	5.5	8.0	118.14
4323	47.83584514	17.40713254	5.5	9.2	118.70
4328	47.86591327	17.44362108	5.2	8.9	117.59
110248	47.86281593	17.45788158	12.5	15.6	116.96
110250	47.88015579	17.46081051	13.0	16.3	120.67
110252	47.88215262	17.41164323	12.5	15.2	119.48
110413	47.9918023	17.24639689	7.2	9.3	126.58
110414	47.99167671	17.24617526	7.2	9.4	127.53
110610	48.00692551	17.2199182	7.5	12.0	129.59
110613	47.99828173	17.23751776	6.7	11.2	127.88
110617	47.98698464	17.23829643	6.5	12.8	127.61
110622	47.98075646	17.24687163	6.5	14.3	127.25
110623	47.98144785	17.28969306	7.5	10.8	126.57
110628	47.97652718	17.30898782	7.3	17.1	129.21
110629	47.97253934	17.32496921	5.9	15.6	129.61
110630	47.96425652	17.30232256	6.1	11.1	125.40
110636	47.96230796	17.25406413	12.0	17.9	125.57
110640	47.95573149	17.26510825	12.0	17.9	125.57
110660	47.93641205	17.34511936	10.5	16.0	123.21
110661	47.92596573	17.32491076	10.5	11.1	123.22
110673	47.9025542	17.39257339	12.0	16.1	121.44
110675	47.8741983	17.38934179	12.0	18.2	119.52
110682	47.84784264	17.42694192	4.0	9.5	118.98
110687	47.82601619	17.49478507	5.4	10.8	116.76
110688	47.8176137	17.48919283	5.0	10.4	116.03
110691	47.79571363	17.55711135	5.8	10.2	113.98
110692	47.79093894	17.53824284	5.9	11.7	114.11
110693	47.78273654	17.52853234	4.6	10.2	114.43
110695	47.77237445	17.50045653	4.1	9.8	115.78
110696	47.77412089	17.54958716	9.4	14.6	113.47
110697	47.76844817	17.54123952	9.8	15.2	114.45
110715	47.97742401	17.33912792	5.0	12.2	126.67
110724	47.89539812	17.43315366	5.4	14.5	123.32
110726	47.88316805	17.45542548	3.7	11.5	122.69
110749	47.87987304	17.40437303	7.4	13.0	118.31
110755	47.88758169	17.37400823	8.2	13.5	120.47
110756	47.8769739	17.35219003	7.3	12.4	120.83
110802	47.9450407	17.38940374	5.0	13.4	124.63
110804	47.8631916	17.51755044	6.5	14.8	119.75
110808	47.85515934	17.53094717	7.5	15.4	119.50
110811	47.85568808	17.49815723	14.5	19.6	119.94
111512	47.98975594	17.25283241	7.7	12.0	133.36
111530	47.98739806	17.31290189	5.5	10.0	128.39
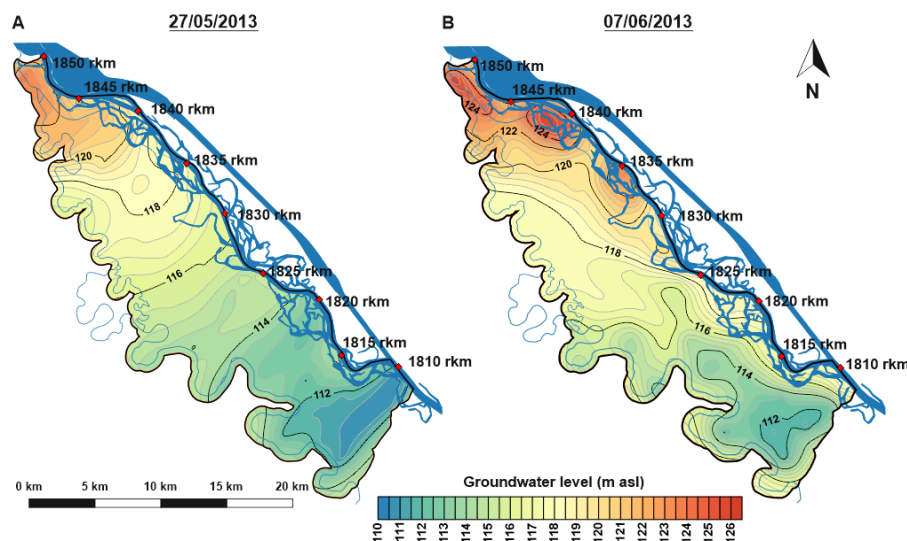
Figure A1. Potentiometric kriged maps of the study area (**A**) prior to and (**B**) during the flood of 2013, taking anisotropy into account [30].

References

1. Simeonov, V.; Stratis, J.; Samara, C.; Zachariadis, G.; Voutsa, D.; Anthemidis, A.; Sofoniou, M.; Kouimtzis, T. Assessment of the surface water quality in Northern Greece. *Water Res.* **2003**, *37*, 4119–4124. [PubMed]
2. Brismar, A. River systems as providers of goods and services: A basis for comparing desired and undesired effects of large dam projects. *Environ. Manag.* **2002**, *29*, 598–609. [CrossRef] [PubMed]
3. Kalbus, E.; Reinstorf, F.; Schirmer, M. Measuring methods for groundwater—Surface water interactions: A review. *Hydrol. Earth Syst. Sci.* **2006**, *10*, 873–887. [CrossRef]
4. Hubbs, S.A. *Riverbank Filtration Hydrology*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007.
5. Ray, C.; Melin, G.; Linsky, R.B. *Riverbank Filtration: Improving Source-Water Quality*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2003.
6. Zeng, R.; Cai, X. Analyzing streamflow changes: Irrigation-enhanced interaction between aquifer and streamflow in the Republican River Basin. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 493–502. [CrossRef]
7. Jaramillo, M. Riverbank filtration: An efficient and economical drinking-water treatment technology. *Dyna* **2012**, *79*, 148–157.
8. Winter, T.C. *Ground Water and Surface Water: A Single Resource*; DIANE Publishing Inc.: Darby, PA, USA, 1998.
9. Chapman, D.V.; Bradley, C.; Gettel, G.M.; Hatvani, I.G.; Hein, T.; Kovács, J.; Liska, I.; Oliver, D.M.; Tanos, P.; Trásy, B.; et al. Developments in water quality monitoring and management in large river catchments using the Danube River as an example. *Environ. Sci. Policy* **2016**, *64*, 141–154. [CrossRef]
10. Field, C.B. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2012.
11. Winter, T.C. Relation of streams, lakes, and wetlands to groundwater flow systems. *Hydrogeol. J.* **1999**, *7*, 28–45. [CrossRef]
12. Sophocleous, M. Interactions between groundwater and surface water: The state of the science. *Hydrogeol. J.* **2002**, *10*, 52–67. [CrossRef]
13. Brunner, P.; Therrien, R.; Renard, P.; Simmons, C.T.; Franssen, H.J.H. Advances in understanding river-groundwater interactions. *Rev. Geophys.* **2017**, *55*, 818–854. [CrossRef]
14. Dillon, P.; Stuyfzand, P.; Grischek, T.; Lluria, M.; Pyne, R.; Jain, R.; Bear, J.; Schwarz, J.; Wang, W.; Fernandez, E. Sixty years of global progress in managed aquifer recharge. *Hydrogeol. J.* **2019**, *27*, 1–30. [CrossRef]
15. Shah, N.; Nachabe, M.; Ross, M. Extinction depth and evapotranspiration from ground water under selected land covers. *Groundwater* **2007**, *45*, 329–338. [CrossRef] [PubMed]
16. Geweke, J. The dynamic factor analysis of economic time series. In *Latent Variables in Socio-Economic Models*; North-Holland: New York, NY, USA, 1977.

17. Zuur, A.F.; Fryer, R.; Jolliffe, I.; Dekker, R.; Beukema, J. Estimating common trends in multivariate time series using dynamic factor analysis. *Environmetrics* **2003**, *14*, 665–685. [CrossRef]

18. Zuur, A.; Tuck, I.; Bailey, N. Dynamic factor analysis to estimate common trends in fisheries time series. *Can. J. Fish. Aquat. Sci.* **2003**, *60*, 542–552. [CrossRef]

19. Muñoz-Carpena, R.; Ritter, A.; Li, Y. Dynamic factor analysis of groundwater quality trends in an agricultural area adjacent to Everglades National Park. *J. Contam. Hydrol.* **2005**, *80*, 49–70. [CrossRef] [PubMed]

20. Ritter, A.; Muñoz-Carpena, R. Dynamic factor modeling of ground and surface water levels in an agricultural area adjacent to Everglades National Park. *J. Hydrol.* **2006**, *317*, 340–354. [CrossRef]

21. Kovács, J.; Kovács, S.; Magyar, N.; Tanos, P.; Hatvani, I.G.; Anda, A. Classification into homogeneous groups using combined cluster and discriminant analysis. *Environ. Model. Softw.* **2014**, *57*, 52–59. [CrossRef]

22. Márkus, L.; Berke, O.; Kovács, J.; Urfer, W. Spatial prediction of the intensity of latent effects governing hydrogeological phenomena. *Environmetrics* **1999**, *10*, 633–654. [CrossRef]

23. Kovács, J.; Márkus, L.; Szalai, J.; Barcza, M.; Bernáth, G.; Kovácsné Székely, I.; Halupka, G. *Exploring Potentially Hazardous Areas for Water Quality Using Dynamic Factor Analysis*; Water Quality Monitoring and Assessment: Rijeka, Croatia, 2012; pp. 228–256. Available online: https://www.researchgate.net/publication/224829940_Exploring_Potentially_Hazardous_Areas_for_Water_Quality_Using_Dynamic_Factor_Analysis (accessed on 1 September 2019).

24. Kovács, J.; Márkus, L.; Szalai, J.; Kovács, I.S. Detection and evaluation of changes induced by the diversion of River Danube in the territorial appearance of latent effects governing shallow-groundwater fluctuations. *J. Hydrol.* **2015**, *520*, 314–325. [CrossRef]

25. Kisekka, I.; Migliaccio, K.; Munoz-Carpena, R.; Schaffer, B.; Li, Y. Dynamic factor analysis of surface water management impacts on soil and bedrock water contents in Southern Florida Lowlands. *J. Hydrol.* **2013**, *488*, 55–72. [CrossRef]

26. Campo-Bescós, M.A.; Muñoz-Carpena, R.; Southworth, J.; Zhu, L.; Waylen, P.R.; Bunting, E. Combined spatial and temporal effects of environmental controls on long-term monthly NDVI in the southern Africa Savanna. *Remote Sens.* **2013**, *5*, 6513–6538. [CrossRef]

27. Smith, S.E.; Büttner, G.; Szilagyi, F.; Horvath, L.; Aufmuth, J. Environmental impacts of river diversion: Gabcikovo Barrage System. *J. Water Resour. Plan. Manag.* **2000**, *126*, 138–145. [CrossRef]

28. Jansky, L.; Pachova, N.I.; Murakami, M. The Danube: A case study of sharing international waters. *Glob. Environ. Chang.* **2004**, *14*, 39–49. [CrossRef]

29. Völgyesi, I. A Kisalföld talajvíz-és rétegvíz helyzete. *Hidrológiai Közlöny* **1994**, *74*, 260–268.

30. Trásy, B.; Garamhegyi, T.; Laczkó-Dobos, P.; Kovács, J.; István Gábor, H. Geostatistical screening of flood events in the groundwater levels of the diverted inner delta of the Danube River: Implications for river bed clogging. *Open Geosci.* **2018**, *10*. [CrossRef]

31. Trásy, B. Assesment of the Szigetköz geological monitoring network (in Hungarian with English summary). *Dep. Phys. Appl. Geol.* **2012**, 118.

32. Trásy, B.; József, K.; István Gábor, H.; Timea, H.; Tibor, N.; Péter, S.; Csaba, S. Assessment of the interaction between surface-and sub-surface waters after the diversion in the inner delta of the River Danube using multivariate statistics. *Anthropocene* **2018**, *22*, 14. [CrossRef]

33. Sztanó, O.; Kováč, M.; Magyar, I.; Šujan, M.; Fodor, L.; Uhrin, A.; Rybár, S.; Csillag, G.; Tőkés, L. Late miocene sedimentary record of the danube/kisalföld basin: Interregional correlation of depositional systems, stratigraphy and structural evolution. *Geol. Carpathica* **2016**, *67*, 525. [CrossRef]

34. Domenico, P.A.; Schwartz, F.W. *Physical and Chemical Hydrogeology*; Wiley: New York, NY, USA, 1998.

35. Scharek, P. *The Geological Map Series of the Little Hungarian Plain–Mosonmagyaróvár*; Hungarian Geological Institute: Budapest, Hungary, 1991.

36. Don, G.; Kaiser, M.; Marsi, I.; Scharek, P.; Sıkhegyi, F.; Szeiler, R.; Tullner, T.; Zsambok, I. *Geological Setting of the Szigetköz*; Hungarian Geological Survey: Budapest, Hungary, 1993.

37. Balderer, W.; Synal, H.; Deak, J. Application of the chlorine-36 method for the delineation of groundwater infiltration of large river systems: Example of the Danube River in western Hungary (Szigetköz area). *Environ. Geol.* **2004**, *46*, 755–762. [CrossRef]

38. Stute, M.; Deák, J.; Révész, K.; Böhlke, J.; Deseö, É.; Weppernig, R.; Schlosser, P. Tritium/3He dating of river infiltration: An example from the Danube in the Szigetkoez area, Hungary. *Groundwater* **1997**, *35*, 905–911. [CrossRef]

39. Ahmadi, S.H.; Sedghamiz, A. Geostatistical analysis of spatial and temporal variations of groundwater level. *Environ. Monit. Assess.* **2007**, *129*, 277. [CrossRef]

40. DATAQUA. Available online: http://www.dataqua.hu (accessed on 1 September 2019).

41. EC. *Directive 2000/60/EC of the European Parliament and of the Council Establishing a Framework for Community Action in the Field of Water Policy*; Official journal of the European Communities: Brussels, Belgium, 2000.

42. Allen, R.G. Using the FAO-56 dual crop coefficient method over an irrigated region as part of an evapotranspiration intercomparison study. *J. Hydrol.* **2000**, *229*, 27–41. [CrossRef]

43. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO Rome* **1998**, *300*, D05109.

44. Kovács, J.; Hatvani, I.G.; Korponai, J.; Kovács, I.S. Morlet wavelet and autocorrelation analysis of long-term data series of the Kis-Balaton water protection system (KBWPS). *Ecol. Eng.* **2010**, *36*, 1469–1477. [CrossRef]

45. Hatvani, I.G. *Application of State-of-the-Art Geomathematical Methods in Water Protection*; Eötvös Loránd University: Budapest, Hungary, 2014.

46. Bánkövi, G.; Ziermann, M. Népgazdasági összefüggések dinamikus előrebecslésének kérdései. *Közgazdasági Szle.* **1973**, *20*, 1269–1286.

47. Ziermann, M.; Michaletzky, G. Factor analysis of time series. *Sigma J. Soc. Econ. Modeling* **1995**, *26*, 77–91.

48. Kuo, Y.-M.; Chang, F.-J. Dynamic factor analysis for estimating ground water arsenic trends. *J. Environ. Qual.* **2010**, *39*, 176–184. [CrossRef]

49. Harvey, A.C. *Forecasting, Structural Time Series Models and the Kalman Filter*; Cambridge University Press: Cambridge, UK, 1989.

50. Holmes, E.; Ward, E.; Scheuerell, M. Analysis of multivariate time-series using the MARSS package. *NOAA Fish. Northwest Fish. Sci. Cent.* **2014**, *2725*, 98112.

51. Nagelkerke, N. A note on a general definition of the coefficient of determination. *Biometrika* **1991**, *78*, 691–692. [CrossRef]

52. Sahoo, S.; Jha, M.K. Groundwater-level prediction using multiple linear regression and artificial neural network techniques: A comparative assessment. *Hydrogeol. J.* **2013**, *21*, 1865–1887. [CrossRef]

53. Akaike, H. Akaike's information criterion. In *International Encyclopedia of Statistical Science*; Springer: Berlin/Heidelberg, Germany, 2011; p. 25.

54. Knowles, L., Jr.; O'Reilly, A.M.; Adamski, J.C. Hydrogeology and simulated effects of the ground-water withdrawals from the Floridan aquifer system in lake county and in the Ocala National Forest and vicinity, north-central Florida. *Water Resour. Investig. Rep.* **2002**, *2*, 4207.

55. Butterworth, J.A.; Schulze, R.E.; Simmonds, L.P.; Moriarty, P.; Mugabe, F. Hydrological processes and water resources management in a dryland environment IV: Long-term groundwater level fluctuations due to variation in rainfall. *Hydrol. Earth Syst. Sci.* **1999**, *3*, 353–361. [CrossRef]

56. Ping, J.; Wei, X.; Nichol, C. Cumulative precipitation departure from average characterizing mountain system recharge in semi-arid North Okanagan, South Interior British Columbia, Canada. *J. Appl. Sci.* **2014**, *14*, 2156–2162. [CrossRef]

57. Xu, Y.; Van Tonder, G.J. Estimation of recharge using a revised CRD method. *Water SA* **2001**, *27*, 341–343. [CrossRef]

58. Garamhegyi, T.; Hatvani, I.G.; Szalai, J.; Kovács, J. Delineation of Hydraulic Flow Regime Areas Based on the Statistical Analysis of Semicentennial Shallow Groundwater Table Time Series. *Water* **2020**, *12*, 828. [CrossRef]

59. Magyar, N. Felszíni és Felszín Alatti Vizek Vizsgálata a Fertő Környezetében Többváltozós Adatelemző Módszerek Bevonásával. Ph.D. Thesis, Eötvös Loránd University, Budapest, Hungary, 2017.

60. Szabó, M. Vizes Élőhelyek Tájökológiai Jellemvonásai a Szigetköz Példáján. Ph.D. Thesis, Hungarian Academy of Sciences, Budapest, Hungary, 2005. Available online: http://real-d.mtak.hu/268/2/Szabo_Maria_tezisek.pdf (accessed on 1 September 2019).

61. TELEDAN. Slovakian-Hungarian Hydrometric Information System Of Danube River And Tributary. Available online: http://www.teledan.eu/eng/operativ-vizrajzi-adatok/ (accessed on 1 September 2019).

62. ÉDUVIZIG. Available online: http://edukovizig.hu/map/layout.html (accessed on 1 September 2019).

63. Tóth, J. Groundwater as a geologic agent: An overview of the causes, processes, and manifestations. *Hydrogeol. J.* **1999**, *7*, 1–14. [CrossRef]

64. Bredehoeft, J.; Norton, D. *The Role of Fluids in Crustal Processes*; National Academies Press: Washington, DC, USA, 1990.

65. Koerselman, W.; Van Kerkhoven, M.B.; Verhoeven, J.T. Release of inorganic N, P and K in peat soils; effect of temperature, water chemistry and water level. *Biogeochemistry* **1993**, *20*, 63–81. [CrossRef]

66. Tanos, P.; Kovács, J.; Székely, I.K.; Hatvani, I.G. Exploratory data analysis on the Upper-Tisza section using single and multi-variate data analysis methods. *Cent. Eur. Geol.* **2011**, *54*, 345–356. [CrossRef]

67. Hatvani, I.G.; Kovács, J.; Kovács, I.S.; Jakusch, P.; Korponai, J. Analysis of long-term water quality changes in the Kis-Balaton Water Protection System with time series, cluster analysis and Wilks' lambda distribution. *Ecol. Eng.* **2011**, *37*, 629–635. [CrossRef]

68. Moussa, R.; Voltz, M.; Andrieux, P. Effects of the spatial organization of agricultural management on the hydrological behaviour of a farmed catchment during flood events. *Hydrol. Process.* **2002**, *16*, 393–412. [CrossRef]

69. Westbrook, C.J.; Cooper, D.J.; Baker, B.W. Beaver dams and overbank floods influence groundwater–surface water interactions of a Rocky Mountain riparian area. *Water Resour. Res.* **2006**, *42*. [CrossRef]

70. Kollet, S.J.; Maxwell, R.M. Capturing the influence of groundwater dynamics on land surface processes using an integrated, distributed watershed model. *Water Resour. Res.* **2008**, *44*. [CrossRef]

71. Kovács, J.; Tanos, P.; Várbíró, G.; Anda, A.; Molnár, S.; Hatvani, I.G. The role of annual periodic behavior of water quality parameters in primary production–Chlorophyll-a estimation. *Ecol. Indic.* **2017**, *78*, 311–321. [CrossRef]

72. Erdélyi, M. A Kisalföld hidrogeológiája a vízlépcsők megépítése előtt és után. *Földrajzi Értesítő* **1990**, *39*, 7–26.

73. Molenaar, P.C. A dynamic factor model for the analysis of multivariate time series. *Psychometrika* **1985**, *50*, 181–202. [CrossRef]

74. Gerla, P.J. The relationship of water-table changes to the capillary fringe, evapotranspiration, and precipitation in intermittent wetlands. *Wetlands* **1992**, *12*, 91–98. [CrossRef]

75. Xu, C.Y.; Chen, D. Comparison of seven models for estimation of evapotranspiration and groundwater recharge using lysimeter measurement data in Germany. *Hydrol. Process. Int. J.* **2005**, *19*, 3717–3734. [CrossRef]

76. Timmermans, W.J.; Meijerink, A.M.J. Remotely sensed actual evapotranspiration: Implications for groundwater management in Botswana. *Int. J. Appl. Earth Obs. Geoinf.* **1999**, *1*, 222–233. [CrossRef]

77. Rosenberry, D.O.; Winter, T.C. Dynamics of water-table fluctuations in an upland between two prairie-pothole wetlands in North Dakota. *J. Hydrol.* **1997**, *191*, 266–289. [CrossRef]

78. Scharek, P.; Don, G.; Horváth, I.; Tóth, G. Result of the depositional process and hydrogeologic investigations in Szigetköz, Hungary. *Acta Geol. Hung.* **2000**, *43*, 85–106.

79. Bartram, J.; Ballance, R. *Water Quality Monitoring: A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes*; CRC Press: Boca Raton, FL, USA, 1996.

80. Yang, F.G.; Cao, S.Y.; Liu, X.N.; Yang, K.J. Design of groundwater level monitoring network with ordinary kriging. *J. Hydrodyn. Ser. B* **2008**, *20*, 339–346. [CrossRef]

81. Rentier, C.; Delloye, F.; Brouyère, S.; Dassargues, A. A framework for an optimised groundwater monitoring network and aggregated indicators. *Environ. Geol.* **2006**, *50*, 194–201. [CrossRef]

82. Freeze, R.A.; Cherry, J.A. *Groundwater*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1979; p. 604.

83. Adamowski, K.; Feluch, W. Application of nonparametric regression to groundwater level prediction. *Can. J. Civ. Eng.* **1991**, *18*, 600–606. [CrossRef]

84. Legates, D.R.; McCabe, G.J., Jr. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]

85. Sullivan, P.; Price, R.; Miralles-Wilhelm, F.; Ross, M.; Scinto, L.; Cline, E.; Dreschel, T.; Sklar, F. Precipitation and Groundwater Evapotranspiration as Hydraulic Drivers of Nutrient and Ion Accumulation in Everglades' Tree Islands, Florida. In Proceedings of the AGUFM, San Francisco, CA, USA, 13–17 December 2010; p. H33H-02.

86. Moreo, M.T.; Andraski, B.J.; Garcia, C.A. *Groundwater Discharge by Evapotranspiration, Flow of Water in Unsaturated Soil, and Stable Isotope Water Sourcing in Areas of Sparse Vegetation, Amargosa Desert, Nye County, Nevada*; US Geological Survey: Reston, VA, USA, 2017.

87. Szalai, J.; Varga, G.; Pappné, U.; A hidrometeorológiai és talajvízszintváltozások értékelése a Duna-Tisza közén az EU VKI szempontjainak tükrében. Magyar Hidrológiai Társaság XXV. *Országos Vándorgyőlés*. 2007. Available online: https://adtplus.arcanum.hu/hu/collection/HidrologiaiVandorgyules/ (accessed on 1 September 2019).