A New Approach in Determining the Decadal Common Trends in the Groundwater Table of the Watershed of Lake “Neusiedlersee”

Norbert Magyar 1, István Gábor Hatvani 2,* 1, Miklós Arató 3,4, Balázs Trásy 5, Alfred Paul Blaschke 6,7 and József Kovács 5

Abstract: Shallow groundwater is one of the primary sources of fresh water, providing river base-flow and root-zone soil water between precipitation events. However, with urbanization and the increase in demand for water for irrigation, shallow groundwater bodies are being endangered. In the present study, 101 hydrographs of shallow groundwater monitoring wells from the watershed of the westernmost brackish lake in Europe were examined for the years 1997–2012 using a combination of dynamic factor and cluster analyses. The aims were (i) the determination of the main driving factors of the water table, (ii) the determination of the spatial distribution and importance of these factors, and (iii) the estimation of shallow groundwater levels using the obtained model. Results indicate that the dynamic factor models were capable of accurately estimating the hydrographs (avg. mean squared error = 0.29 for standardized water levels), meaning that the two driving factors identified (evapotranspiration and precipitation) describe most of the variances of the fluctuations in water level. Both meteorological parameters correlated with an obtained dynamic factor (r = −0.41 for evapotranspiration & r = 0.76 for precipitation). The strength of these effects displayed a spatial pattern, as did the factor loadings. On this basis, the monitoring wells could be objectively distinguished into two groups using hierarchical cluster analysis and verified by linear discriminant analysis in 98% of the cases. This grouping in turn was determined to be primarily related to the elevation and the geology of the area. It can be concluded that the application of the data analysis toolset suggested herein permits a more efficient, objective, and reproducible delineation of the primary driving factors of the shallow groundwater table in the area. Additionally, it represents an effective toolset for the forecasting of water table variations, a quality which, in the view of the likelihood of further climate change to come, is a distinctive advantage. The knowledge of these factors is crucial to a better understanding of the hydrogeological processes that characterize the water table and, thus, to developing a proper water resource management strategy for the area.

Keywords: dynamic factor analysis; hydrograph time series; shallow groundwater; water resource management; watershed of Lake “Neusiedlersee”
1. Introduction

Shallow groundwater (SGW) is one of the main sources of freshwater on Earth [1] affecting all terrestrial ecosystems by providing river base-flow and root-zone soil water in the lack of precipitation [2]. Thus, SGW serves as an important source for irrigation [3] and the maintenance and restoration of ecosystems (for details, see Fan et al. [4]). By the early 2010s, it was the source of one third of freshwater withdrawals world-wide [5]. The scale of this makes clear the necessity of a deeper understanding of the complex spatiotemporal dynamics of SGW, dynamics which are driven by hydrological processes [6]. The variability of the SGW level is driven by both anthropogenic and natural factors [6]. In the case of the former, this is mainly related to irrigation (~70% of global freshwater withdrawals at the turn of the century [7]) and the needs of industry. These together cause drops in the water table. The natural factors are related to local meteorological conditions (precipitation and temperature [8]) and subsurface re/discharge [9], both of which affect the groundwater reservoir budget.

Net recharge is determined by rainfall, surface water, soil and aquifer properties and topography [10]. These factors should all be considered in determining groundwater vulnerability, since, with recharge water, pollutants enter the subsurface water bodies and can travel laterally within the aquifer or vertically to the water table [11], rendering higher risks to higher recharge rates [10]. The SGW table plays an important role in groundwater-dependent ecosystems [12], and has an effect on surface vegetation and yield of agricultural crops [13,14]. Consequently, a decline in the water table can reduce well yield and increase pumping costs, having serious economic impact on areas where groundwater is used as the major source of irrigation [15–17]. While in temperate climates diurnal fluctuations in the groundwater table are driven in large measure by the water consumption of vegetation [18], it can also be affected by other factors, such as the depth of the water table, soil moisture, and so on. It is worth mentioning that withdrawals from, or interaction with, a river (via pressure propagation) can have an even more direct effect on the groundwater level and its fluctuation.

With increased agricultural and industrial activity and urbanization, SGW outtake began to surpass the rate of natural replenishment in certain regions of the world, causing a drop in water tables (for reviews, see the works of Fan et al. and Taylor et al. [4,5]). Moreover, with climate change, the spatial pattern of water demand for irrigation has changed and will continue to do so, rendering the indirect effects of climate on groundwater smaller than the direct impacts of climate on recharge [5]. Climate and land cover are in constant interaction with the subsurface water table, since the former determines the amount of precipitation and evapotranspiration, while the potential amount of water reaching the subsurface water table is regulated by local soil and geology [5]. Therefore, changes in local climate patterns will inevitably induce changes in the SGW table [19], together with changes in soil moisture and surface water [5]. Due to the increased probability of extreme climate events (e.g., droughts and floods) which are both more severe and longer in duration, the importance of SGW for irrigation purposes in agriculture is likely to grow further. Therefore, determining the relative strengths and interplay of the main drivers of SGW level fluctuation is critical in water resource management [20].

The area selected for the study was the north-westernmost part of the Danube Basin adjacent to the Viennese Basin on the border of Austria and Hungary (Figure 1) covering a part of Burgenland (AT) and a small part of the Little Hungarian Plain (HU) (~4000 km²; Figure 1). As a vulnerable agricultural region composed of diverse landscapes, this was considered a suitable setting in which to determine the main drivers of the SGW table. The region has several characteristics that make it particularly interesting with regard to the fluctuation of the SGW table: (i) as in the whole of Austria, SGW is the main source for irrigation [21], and in the studied area 72% of SGW withdrawal is used for agricultural purposes, mostly for irrigation; [22]. (ii) it includes the watershed of Lake “Neusiedlsee” in northern Burgenland, the largest steppe (endorheic) lake in Europe [23,24] and the largest (surface area 309 km²) shallow lake in Austria [25,26]; (iii) about two-thirds of the area
is under Ramsar Convention protection, and is an International Union for Conservation of Nature protection zone [27]; it is also a UNESCO biosphere reserve [28,29], and has even been designated as a World Heritage site [30,31]; (iv) certain parts of the Seewinkel region (E part of the study area; Figure 1) have an exceptionally high ecological value due to the presence of saline lakes, which form a unique biotope [30] requiring a high groundwater table in order to sustain the stabilization of the salt content and ensure the capillary uptake of salt from the soil [32]; (v) the climate of the area is under the influence of semi-arid conditions [33], with an annual average precipitation amount of ~670 mm year\(^{-1}\) (1997–2012) and its climate can be classified as temperate without a dry season, but with cold summers denoted by Cfb [34].

In arid and semi-arid regions, groundwater serves as the most reliable and sustainable water source for drinking, agriculture, industry, ecosystem services and even recreational purposes [35]. Thus, sustaining subsurface water bodies in good quality and quantity is of utmost importance [36]. Unfortunately, natural and anthropogenic influences can negatively alter the quantity and quality of groundwater rendering it at higher vulnerability than surface water in arid and semi-arid regions [37].

Although the annual average precipitation is not below 500 mm year\(^{-1}\), one of its sub-regions, the Seewinkel (Figure 1), has previously been classified as a semi-arid region [32]. Such regions of the world are under constantly growing stress from the negative effects of climate change [38], and these are further aggravated by local land surface processes taking place due to human activities [39].

Lastly, (vi), as in the Everglades National Park (FL, USA) [40], active water resource management measures are being taken in the area to retain the surface waters, prevent flooding and minimize SGW extraction to the level of only that which is absolutely necessary.

In the interests of efficient water resource management strategies for such vulnerable SGW aquifers, the countervailing and/or reinforcing influence of the driving characteristics of the above-mentioned phenomena on the SGW levels must be understood and quantified. Several multivariate analysis methods, such as principal component analysis (PCA) and factor analysis (FA), have long proven useful in obtaining a picture of the drivers of the SGW table [41]. Both the named approaches aim to decompose the total variance of the examined variables into orthogonal vectors, which can subsequently be examined in terms of their relation to potential driving phenomena. Such studies have been conducted on SGW systems all around the world, from the North China Plain, where precipitation and water outtake for irrigation have been found as the main driving factors of SGW table variations [42], to the Thessaly Basin (Greece) [43], from the watershed of four lakes in the USA [41] to Korea [44], where sets of hydrographs with similar behavior were determined for more efficient water resource management.

It should be noted that, although the application of PCA (and FA) have proved successful in both the determination of driving factors in sets of hydrograph time series and in increasing the effectiveness of water resource management, in these cases the lagged serial correlation structure of the data was not taken into account, rendering the results inaccurate [45,46]. Thus, when analyzing the common trends in time series with serial correlation, it is suggested that dynamic factor analysis (DFA) be used, since it allows the estimation of common patterns and interactions in several time series, as well as studying the effect of explanatory time-dependent variables [47,48].
The question of determining the driving factors of the subsurface water table using DQA has been addressed in a few studies in recent decades, e.g., for aquifer vulnerability studies [50], for predicting the intensity of common trends governing the changes in groundwater level in karstic [45] and other hydrogeological systems [47,51,52], or for determining common trends in order to explore the driving factors of the temporal variation in soil and bedrock water content [53].

The specific objectives of the present study were therefore to determine (i) the main driving factors of the SGW table in the area, (ii) the spatial distribution and importance of these factors, and (iii) the possibility of estimating SGW levels using the derived dynamic factor model in the agriculturally important border region between Austria and Hungary by means of dynamic factor analysis.

2. Materials and Methods

2.1. Hydrogeological Characteristics of the Study Area

In the studied area, most of the SGW wells are located in an unconfined (~95%) mixed gravel and sand aquifer complex, generally 5 to 25 m thick [32]. Only a small part can be considered confined or semi-confined, that which lies in the upper northern section. The conductivity is generally low but does vary widely over the aquifer [54,55]. The relief of the aquitard is undulating and sometimes interaction between the groundwater and the small salt lakes in the area east of Lake“Neusiedlersee”may be observed [30]. Due to the interaction between groundwater and surface water, groundwater protection will, in the long run, become an even bigger priority in the protection of surface water ponds [56]. However, comprehensive studies employing state-of-the-art statistical tools to explore the SGW characteristics of the area are extremely scarce.

The amount of water taken for irrigation purposes varies from season to season. Nearly nothing is pumped during winter, while during summer pumping depends mainly on the weather situation and can vary between only 1–2 L s⁻¹ up to 80–100 L s⁻¹ for a sole irrigation well. The wells generally used for irrigation in the area tap the same aquifer. It...
should be noted that pumping is constrained by regulations when the groundwater falls below a certain level.

2.2. Subsurface Water Levels (Response Parameters)

In the course of the research, the monthly time series (1997–2012) of 101 SGW level monitoring wells were assessed, 15 from the Hungarian and 86 from the Austrian part of the study area (Figure 1). The former set was provided by the Hungarian North-TransDanubian Water Directorate, while the latter came from the database of the Austrian Federal Ministry of Agriculture, Regions and Tourism. The average distance between the SGW monitoring wells is 2469 m (min: 37.2 m, max: 9939.8 m), with a mean screening depth of ~7.4 m. The aquifer thickness is 2 to 25 m in the Seewinkel area and is very heterogeneous due to the relief of the aquiclude.

2.3. Environmental Explanatory Parameters

For reasons which should become clear, it was reasonable to suspect that precipitation and evapotranspiration would be the environmental parameters driving SGW levels. Daily precipitation data were retrieved from 38 meteorological stations (Figure 1) and were downscaled to integrated monthly precipitation time series (Prec; Equations (1) and (2)), then, lastly, an areal average was arrived at:

\[ y_t = x_t - \hat{m} + y_{t-1}, \text{ if } t > 0 \]  

(1)

where \( x_t \) is the original time series, \( y_t \) the integrated time series, \( \hat{m} \) the average of \( x_t \) and:

\[ y_0 = x_0 - \hat{m} \]  

(2)

Similar integrated time series were used from the potential evapotranspiration data (pET) estimated using the Penman equation [57] from the synoptic meteorological data of the Andau station (Figure 1; monthly average of the daily minimum and maximum temperatures, relative humidity, wind speed, air pressure, vapor pressure, cloud cover and daily bright sunshine hours, height of the flora and latitude).

2.4. Data Preprocessing

As the first preprocessing step in any data analysis procedure [58], the typos and outlying records of the assessed hydrographs were filtered manually by inspecting the time series of the SGW levels at each sampling site and consulting with the authorities (data owners) if there were questionable values. Missing data were imputed using multivariate regression analysis of the neighboring wells within a 20 km search radius correlating at least to a value of \( r > 0.7 \) and when the number of complete cases was larger than 36. The maximum number of consecutive values imputed was six. If an SGW well still had missing data after the previous steps had been taken, it was excluded from the analysis. Since the temporal sampling frequency was not uniform in the two countries, monthly averages were formed from the data to make them comparable in the analyses.

2.5. Applied Methodology

Dynamic factor analysis (DFA) was the method applied to find the background factors driving the common trends of the SGW levels while taking their lagged autocorrelation structure into account [48,59,60]. DFA is a specialized time series technique developed originally to study economic time series models [60]; it has, however, been applied in various fields of earth sciences, e.g., fisheries [61], limnology [46], oceanography [62], karst-hydrology [50], and hydrogeology [45,63]. In the present study it was applied to the monthly average SGW level time series. The factor loadings determined the weight of the different SGW levels of the wells in a given factor. Following a commonly used methodology [46,50], the obtained DF time series were correlated with the possible explanatory variables determining which background factor is driving a set of SGW monitoring
wells’ hydrographs. In the following step, the SGW level time series were estimated and the estimations’ accuracy determined employing mean squared error (MSE), using the obtained factor time series and weights.

The spatial distribution of the determined background factors was assessed using hierarchical cluster analysis (Wards’ method, squared Euclidean distance [64]), and validated with the use of linear discriminant analysis (LDA; [65]). The groups of SGW monitoring wells thus obtained were then characterized to achieve a better understanding of their SGW levels.

All computations were performed in an R statistical environment [66] with MARSS [67], mass, and stats packages, while for the visualizations MS Excel and Golden Software Surfer 17 were used.

Dynamic Factor Analysis

The principle of DFA is very similar to traditional factor analysis, in which multidimensional vectors are replaced by vectors with a much smaller number of dimensions. The important difference, however, is that neither the original vectors nor the replacement vectors are independent. A general formulation for a model with \( p \) common trends can be described as follows.

\[
\eta_t = \eta_{t-1} + w_t, \quad w_t \sim N_p(0, Q) \tag{3}
\]

\[
\xi_t = Z\eta_t + \alpha + v_t, \quad v_t \sim N_n(0, R) \tag{4}
\]

\[
\eta_0 \sim N_p(\pi, \Lambda) \tag{5}
\]

The vector \( \xi_t = \begin{pmatrix} \xi_{t,1} \\ \vdots \\ \xi_{t,n} \end{pmatrix} \) contains the \( n \) time series, and \( \eta_t = \begin{pmatrix} \eta_{t,1} \\ \vdots \\ \eta_{t,p} \end{pmatrix} \) is a vector of dimension \( p \) containing common random walks (hidden trends) at time \( t \). Meanwhile, \( w_t, v_t \) are multivariate Gaussian distributed white noises with \( Q \) and \( R \) covariance matrices. \( Z \) is the loadings matrix and \( \alpha \) is the offset [67].

Following common practice, it was assumed that matrix \( Q \) is an identity matrix, thus the coordinates of \( w_t \) are independent. The time series \( \xi_{1,}, \ldots, \xi_{n,} \) were standardized and thus vectors \( \alpha \) and \( \pi \) appeared as null vectors. To make the model identifiable, \( z_{i,j} (j > i, i < p) \) was set to zero.

The covariance matrix \( R \) can be estimated using the following four settings:
- a diagonal matrix with equal variances in the diagonal
- a diagonal matrix with unequal variances in the diagonal
- a non-diagonal matrix with equal variances in the diagonal and equal covariances
- an unconstrained covariance matrix

As in many dimension reduction techniques (i.e., PCA and FA), the determination of the number of common trends is crucial. The use of more common trends results in a better fit; in such cases, however, more parameters must be estimated. The small-sample corrected Akaike Information Criterion (AICc) was used, and this may be defined as twice the difference between the log likelihood function and the number of parameters to determine the number of trends (\( p \)) and the type of the covariance matrix \( R \) [48].

Finally, since dynamic factor analysis provides multiple solutions, varimax rotation was applied to the factors in order to determine the \( H \) matrix in the case in which the largest difference is obtained between loadings.

3. Results

3.1. Estimation of Common Trends and Driving Factors of the SGW Levels and Their Spatial Distribution

The question of which external influences drive the common trends of the SGW levels, and to what extent, in the area was investigated. Since the SGW levels’ monthly average time series showed persistence over time (average \( r_1(192) = 0.90, p = 0.00085 \) for the 101
SGW monitoring wells), the application of conventional dimension reduction methods (e.g., PCA, FA) was rejected, since these do not account for the lagged autocorrelation structure of the given time series [46,48,50,68].

Thus, DFA was applied, with different model criteria. The best result was obtained in the case of three factors (AICc = −178.5) in such a way that the elements of the $R$ covariance matrix were not constrained, but instead variances and covariances are estimated. The error matrix indicated that the less complex models performed the least well (Table A1).

After obtaining the three common trends of the SGW levels, the dynamic factor time series were correlated with the possible explanatory variables (see Sect. Environmental, explanatory parameters). The first dynamic factor (DF1) was found to be driven by pET (Figure 2A; $r(192) = −0.41$, $p = 0.0597$), the second (DF2) by precipitation (Figure 2B; $r(192) = 0.76$, $p = 0.028$), while no external driving factor was identified for the third dynamic factor, which indicated the least variability. The areal integrated precipitation time series was also characterized through the use of DFA. Dynamic factors of the 38 meteorological stations’ integrated precipitation time series were obtained and correlated with DF2. The first factor showed a similar correlation coefficient ($r(192) = 0.77$, $p = 0.029$) as calculated from the monthly averages ($r(192) = 0.76$, $p = 0.028$).

![Figure 2](image-url)

**Figure 2.** The time series of the first dynamic factor and integrated potential evapotranspiration data (pET) (A); the time series of the second dynamic factor and integrated precipitation (Prec) (B). The inverted values of the integrated pET values are shown (A), due to its water level decreasing effect. The Pearson correlations between the dynamic factor time series and the environmental variables’ time series are reported in the upper right corner of the panels and both values were significant ($p < 0.06$).

By determining the factor loadings for each of the SGW monitoring wells, it becomes clear how important their previously identified driving factors are at any given location (Figure 3). In order to explore the spatial pattern of SGW level fluctuations in the area, the dynamic factor loadings belonging to the DFs were grouped using hierarchical cluster analysis. Based on the dendrogram obtained (Figure A1), two groups could be delineated...
(GR1 and GR2; Figure 3) and the ratio of the original grouped cases were 98.01% as validated by LDA, rendering it a perfectly acceptable result.

The two sets of SGW monitoring wells were distributed mainly following the relief of the region (Figure 3). Most of the sampling sites of GR1 (circles on Figure 3) can be found in the more elevated areas (mean surface elev. of the wells: ~177.1 m asl.), while the sampling sites of GR2 (triangles on Figure 3) are located in lower elevated areas (mean surface elev. of the wells ~124.8 m asl.). If the loadings themselves are plotted together with the groupings, it can be concluded that the SGW monitoring wells did not display different characteristics in the two groups regarding DF1 (Figure 3A). However, there was a clear distinction between the groups regarding DF2 (Figure 3B), which had previously been found to be related to Prec (Figure 2B), i.e., the loadings of DF2 were higher in GR2, than in GR1 (Figure 3B). This implies that the background processes behind DF2 play a more important role in separating the cluster groups than those of DF1. This pattern in large part resembles the pattern indicated by the loadings, with DF2 dominating the north-eastern part of the study area, and almost fully overlapping with the Seewinkel area (Figure 1).

3.2. Estimation of SGW Levels from the Dynamic Factor Models

For each SGW well the water level time series was estimated based on the dynamic factor model obtained (Equations (3)–(5)) and then compared to observed values. The goodness of these estimations was evaluated using MSE ranging between 0.05–0.96 (Table 1, Figure A2), resulting in a spatial pattern (Figure 4A). The best estimations were obtained in the SGW time series of GR2, dominated by DF2 and located mostly in the Seewinkel area (Figure 4A). In the case of the wells belonging to GR1, the interquartile range of the MSE values calculated between the observed and estimated water level time series was more than two times higher than those in GR2 (Figure A2). With the DF models it became possible to estimate the SGW time series of numerous wells almost perfectly—these are shown in white in Figure 4A. For example, SGW levels of well 305524 was estimated almost perfectly (MSE: 0.07) using the DFA model (Figure 4B). The poorest estimations were obtained along the Wulka Stream (Figure 4A, mean MSE: 0.46, max: 0.96).
Table 1. Group statistics of mean squared error (MSE).

| Statistics          | GR1  | GR2  | All  |
|---------------------|------|------|------|
| Mean                | 0.451| 0.195| 0.286|
| Median              | 0.414| 0.178| 0.224|
| Standard deviation  | 0.217| 0.098| 0.194|
| Range               | 0.785| 0.349| 0.911|
| Minimum             | 0.173| 0.046| 0.046|
| Maximum             | 0.958| 0.396| 0.958|
| Number of wells     | 36   | 65   | 101  |

Figure 4. Spatial distribution of the MSE obtained from the measured and modelled shallow groundwater (SGW) time series (1997–2012) (A); and an example of the original and estimated time series by the dynamic factor (DF) model presented for SGW levels of well 305524 (B).
4. Discussion

The groundwater of the studied area is recharged primarily by precipitation, while the only natural small river (Wulka) is fed mainly from the aquifer [69]. The area studied is often characterized by pronounced fluctuations in climatic conditions, which manifest themselves in longer periods of occasionally wet but more usually periods of dry-to-drought conditions—both of which affect the groundwater body. The expected changes in climate (higher temperatures and changes in precipitation intensity and distribution) make the study area especially vulnerable [70]. However, the changes forecast in precipitation and shifts in seasonal patterns are still very uncertain due to the uncertainties involved in deriving small-scale statements from the climate models for this region [71,72]. Taken together, these factors underline the importance of water resource management measures. Results concur with general knowledge that, at the level of interannual variability, changes in SGW levels are driven primarily by evapotranspiration and the interannual distribution of precipitation [73]. Therefore, finding significant correlations between precipitation and SGW level variations can help assess aquifer vulnerability [17,74].

The common trends of the SGW level time series were determined using different model criteria as suggested by Zuur et al. [48], linking the best-performing common trends (Table A1) to external driving phenomena. The SGW wells primarily related to DF1 were mostly located in the close vicinity of the open water surfaces, such as the Lake “Neusiedlersee” and the small brackish ponds (Lacken) of the area (Figure 3A), which proved to be driven by pET \( r = -0.41, p = 0.0597 \). Here, the depth of the water table is generally lower (Figure A3B), which results in higher evapotranspiration. In these SGW wells the amplitude of the change in SGW levels and pET varied (Figure 2A).

This is most probably caused by the intensive irrigation of the croplands in the area [75], which results in water outtake from the SGW primarily in the months when pET shows a secular increase [76]. Moreover, the varying water requirements of the different cultivated species in the area with the highest water demand (e.g., flowers, field vegetables, leaf vegetables/glasshouse, turnips, potato, intensive fruit (e.g., strawberries), sweet corn, etc.) [22] also change water outtake on a yearly basis. The SGW wells related to DF2, were primarily located in the NE part of the area, overlapping with the Seewinkel, and their water level fluctuations were determined by Prec. \( r(192) = 0.76, p = 0.028 \). DF2 was also correlated with the common trend of the Prec. time series obtained by DFA. The correlation coefficient was close to the value calculated with the areal average Prec. However, in cases where the regional distribution of precipitation is uneven (e.g., large differences in relief conditions) common trends for precipitation are derived and correlated with the SGW dynamic factors rather than the areal precipitation averages. It may occur that not only one but multiple common trends of the SGW time series have a strong connection with precipitation, which can be distinguished by deriving the previously mentioned common trends of precipitation. These in turn can provide a spatial picture on the strength of the relationship between precipitation and SGW levels even considering their lagged correlation structure [50].

The determination of the spatial patterns of the common hydrographs trends is useful in (i) exploring the areal characteristics of recharge and discharge, (ii) determining the importance of the obtained driving factors (iii) and selecting candidates for representative index wells (or sets of wells) for long-term monitoring [41] and water resource management purposes.

The SGW monitoring wells were grouped on the basis of their DF loadings, and two groups were obtained (LDA classification 98.01%); this was in contrast to other studies [41–43], in which the SGW wells were grouped arbitrarily using the factor loadings’ 2D or 3D plots. In the present case, however, the procedure was carried out on an objective reproducible basis (See Section: Estimation of common trends and driving factors of the SGW levels and their spatial distribution), that is, based on the DF scores of the wells. While the loadings of DF2 were higher in the SGW wells in the least elevated parts (GR2) of the Seewinkel area, in relation to DF1 no such distinction was found. This implies that
there is a clear sub-regional difference in the importance of precipitation (related to the SGW wells in GR2), which cannot be observed in the case of evapotranspiration [77].

The seasonal pattern of SGW level fluctuations also displayed a distinctive spatial pattern. While the SGW level fluctuations in GR1 is moderate, GR2 can be characterized by a range over twice the size (GR1: 0.50 m GR2: 1.11 m, Figure 5). A similar pattern has been observed in the North China Plain, where smaller seasonal fluctuations were observed in the more highly elevated areas of the region than the lower plains [42]. In flat sedimentary landscapes (e.g., the Seewinkel) the SGW table is typically shallow [4,78] and local climatic change (the ratio of precipitation vs. evapotranspiration) plays a significant role in raising/lowering the water table level, as also in other lake watersheds in the U.S. [41]. To tackle the phenomenon of high evapotranspiration and/or low precipitation leading to negative groundwater recharge in the area, the introduction of further external water supplies to the Seewinkel area should be considered as a management measure. In parallel, a rethink of agriculture in the direction of the cultivation of plants that use less water should also be considered. Both courses of action would be in accordance with various evaluations indicating the sensitive nature of the Seewinkel area.

![Figure 5](Image)

**Figure 5.** Mean seasonal cycles of SGW level fluctuation on the basis of the two groups obtained from the cluster analysis of DFA loadings. For all the time series, a centered 12-month simple moving average was calculated and subtracted from the original time series to eliminate the trend component.

The annual period present in most of the SGW wells may be explained by the infiltration processes characteristic of spring and fall. In certain areas, the replenishment of the subsurface water resource has already begun in October, while in other places it starts a couple of months later, and this is highly dependent on the SGW table depth [32].

Furthermore, the depletion of groundwater supplies, conflicts between groundwater users and surface water users, and the potential for groundwater contamination are concerns that will become increasingly important as further aquifer development takes place here, or indeed in any basin [3].

On the basis of the DF models, the water level time series of the wells proved to be efficiently estimated in most cases. The best results were obtained in the Seewinkel area, and the worst along the Wulka stream, and this was presumably because of the
numerous inflows bringing water from outside of the region with different characteristics. Similar applications on SGW along the banks of the River Danube (HU) [50,63], or even in agricultural areas adjacent to the Everglades National Park (USA) [40,47] highlight that, with DFA, an enhanced understanding of the reasons for changes in the fluctuation of groundwater can be achieved which could assist in the development of sustainable water management and irrigation strategies and preparation for varying potential climate change scenarios. Such results cannot be obtained with common multivariate techniques which search for common trends (e.g., PCA, FA) due to their lack of consideration of the lagged correlation structure of the data at hand.

5. Conclusions and Outlook

In the present study the assessment of common trends of 101 SGW monitoring wells in the Lake “Neusiedlersee” area recorded between 1997–2012 allowed the determination of their primary driving factors, and the spatial distribution of the weights of these factors. The most important drivers were precipitation and evapotranspiration. Instead of conveniently choosing only two factors to assess, multiple models were put to trial with different parameter settings (a combination of a different number of obtained common trends and the structure of variance-covariance matrix) to achieve a best fit. The accuracy of these models was compared on the basis of the AICc values and the factor loadings of the best fit model facilitated the objective grouping of the wells through the use of cluster analysis, relating the groups primarily to the elevation and the geology of the area. The dynamic factor models:

(i) provided a detailed insight into the most important drivers of the SGW table in the area
(ii) yielded an accurate estimation of the SGW table fluctuations
(iii) facilitated the spatial grouping of the wells

These were achieved by taking into account the lagged correlation structure of the time series, something that had not previously been done in any model of the region. With the model presented here and using historical data, sub-regionalized forecasts could be made with reasonable accuracy, thereby facilitating location-based water resource management.

The combined application of the used data analysis methods (DFA, HCA, LDA) to the SGW data and additional explanatory environmental parameters have proven to be sufficient to provide an extensive overview of their relationship with the regional meteorological conditions and topography as determining their primary characteristics; while the different behavior of the clusters of the SGW wells provides the foundation to delineate the different flow regimes (recharge and discharge areas) [17].

Although this study has focused on the thorough assessment of SGW hydrographs in the watershed of the Lake “Neusiedlersee”, the principles applied here could be extended to other, similar study areas. The present study provides a preliminary first-order approach to the understanding of the spatiotemporal controls of the SGW table in this critical region, the watershed of Lake “Neusiedlersee”. A knowledge of these factors is crucial to a better understanding of the hydrogeological processes that characterize the water table, something which is crucial to the development of an appropriate water resource management strategy for the area. The excess knowledge provided by the present study could serve as a benchmark for further preparation for the severe negative effects likely to come about as a result of climate change in the area [17,79,80].

Author Contributions: N.M., J.K., M.A. and I.G.H. designed the experiments and N.M. and B.T. carried them out. N.M. and M.A. developed the model code and performed the analyses. N.M., J.K., I.G.H. and A.P.B. prepared the manuscript and its revision with contributions from all co-authors.” The authors applied the SDC approach for the sequence of authors. See https://doi.org/10.1371/journal.pbio.0050018 for further details. All authors have read and agreed to the published version of the manuscript.
**Funding:** This work was completed in the ELTE Institutional Excellence Program (grant number: TKP2020-IKA-05) financed by the Hungarian Ministry of Human Capacities.

**Data Availability Statement:** The data was obtained from the Hungarian North-Trans Danubian Water Directorate, and the Austrian Federal Ministry of Agriculture, Regions and Tourism, thus is available from the authors after the permission of the above authorities is given.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A**

**Table A1.** Model selection results in increasing order of Akaike Information Criterion (\(\Delta AICc\)). 
\(AICc = -178.5\) is determined as the benchmark; all other \(AICc\) are relative to this value given as \(\Delta AICc\).

| R Matrix                                      | Number of Factors | \(\Delta AICc\) |
|-----------------------------------------------|-------------------|-----------------|
| different variances & covariances             | 3                 | 0               |
|                                               | 2                 | 36.8            |
|                                               | 1                 | 191.4           |
| different variances & no covariance           | 3                 | 19,326.2        |
|                                               | 2                 | 24,461.2        |
| same variances & same covariance              | 3                 | 24,674.1        |
| same variances & no covariance                | 2                 | 25,846.9        |
| same variances & same covariance              | 2                 | 28,591.4        |
| same variances & no covariance                | 3                 | 30,351.3        |
| same variances & same covariance              | 1                 | 32,570.6        |
| different variances & no covariance           | 1                 | 34,279.2        |
| same variances & no covariance                | 1                 | 37,725.7        |

**Figure A1.** Dendrogram of DFA loadings, where the different colors mark the two groups of SGW wells obtained. The dashed horizontal line marks the intersection of the dendrogram at 60% of the maximum linkage distance.
Figure A2. Boxplot of the mean squared errors (MSE) values by the obtained groups. The boxes show the interquartile range and the black lines within the boxes are the medians. The horizontal lines end at the minimum/maximum value within 1.5 times the interquartile range.

Figure A3. The spatial distribution of the factor loadings of dynamic factor (DF)1 (A) and the depth of water table (B).
29. 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]

30. Hatvani, I.G.; Magyar, N.; Zessner, M.; Kovács, J.; Blaschke, A.P. The Water Framework Directive: Can more information be extracted from groundwater data? A case study of Seewinkel, Burgenland, eastern Austria. Hydrogeol. J. 2014, 22, 779–794. [CrossRef]

31. Magyar, N.; Trásky, B.; Kutrucz, G.; Dinka, M. Delineating water bodies on the Hungarian side of Lake Fertő/Neusiedler See. In *Theories and Applications in Geomathematics: Selected Studies of the 2012 Croatian-Hungarian Geomathematical Convent*; GeoLiftera: Opatija, Croatia, 2013.

32. Blaschke, A.; Gschöpf, C. *Groundwater Flow Model Seewinkel*; Burgenländische Landesregierung: Eisenstadt, Austria, 2011; (In German). Available online: https://wasser.bgld.gv.at/fileadmin/user_upload/news/Kurzfassung_Bericht_GWM.pdf (accessed on 10 July 2020).

33. Karner, K.; Mitter, H.; Schmid, E. The economic value of stochastic climate information for agricultural adaptation in a semi-arid region in Austria. *J. Environ. Manag.* 2019, 249, 109431. [CrossRef]

34. Kottek, M.; Grieser, J.; Beck, C.; Rudolf, B.; Rubel, F. World map of the Köppen-Geiger climate classification updated. *Geotherm. Acta* 2006, 15, 259–263. [CrossRef]

35. Appelo, C.; Postma, D. Geochemistry, *Groundwater and Pollution*; Balkema: Rotterdam, The Netherlands, 2005.

36. Wang, Y.; Ma, T.; Luo, Z. Geostatistical and geochemical analysis of surface water leakage into groundwater on a regional scale: A case study in the Liulin karst system, northwestern China. *J. Hydrol.* 2001, 246, 223–234. [CrossRef]

37. Anim-Gyampo, M.; Anornu, G.K.; Agodzo, S.K.; Appiah-Adjei, E.K. Groundwater Risk Assessment of Shallow Aquifers within the Atankwidi Basin of Northeastern Ghana. *Earth Syst. Environ.* 2019, 3, 59–72. [CrossRef]

38. Blöschl, G.; Blaschke, A.P.; Haslinger, K.; Holstätter, M.; Parajka, J.; Salinas, J.; Schön, W. Impact of climate change on Austria’s water sector—An updated status report. *Österreichische Wasser- Und Abfallwirtschaft* 2019, 70, 462–473. [CrossRef]

39. Huang, J.; Ji, M.; Xie, Y.; Wang, S.; He, Y.; Ran, J. Global semi-arid climate change over last 60 years. *Clim. Dyn.* 2016, 46, 1131–1150. [CrossRef]

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

41. Winter, T.C.; Mallory, S.E.; Allen, T.R.; Rosenberry, D.O. The Use of Principal Component Analysis for Interpreting Ground Water Hydrographs. *Groundwater* 2000, 38, 234–246. [CrossRef]

42. Zhang, R.G. Groundwater Hydrograph Patterns in North China Plain during 1982–1986 Interpreted Using Principal Component Analysis. *Adv. Mater. Res.* 2012, 356–360, 2320–2324. [CrossRef]

43. Seferli, S.; Modis, K.; Adam, K. Interpretation of groundwater hydrographs in the West Thessaly basin, Greece, using principal component analysis. *Environ. Earth Sci.* 2019, 78, 257. [CrossRef]

44. Moon, S.-K.; Woo, N.C.; Lee, K.S. Statistical analysis of hydrographs and water-table fluctuation to estimate groundwater recharge. *J. Hydrol.* 2004, 292, 198–209. [CrossRef]

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

46. Hatvani, I.G.; Kovács, J.; Márkus, L.; Clement, A.; Hoffmann, R.; Korponai, J. Assessing the relationship of background factors governing the water quality of an agricultural watershed with changes in catchment property (W-Hungary). *J. Hydrol.* 2015, 521, 460–469. [CrossRef]

47. 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]

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

49. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L. The shuttle radar topography mission. *Rev. Geophys.* 2007, 45, RG2004. [CrossRef]

50. Kovács, J.; Márkus, L.; Halupka, G. Dynamic factor analysis for quantifying aquifer vulnerability. *Acta Geol. Hung.* 2004, 47, 1–17. [CrossRef]

51. Kovács, J.; Márkus, L.; Szalai, J.; Barcza, M.; Bernáth, G.; Székely, I.K.; Halupka, G. Exploring Potentially Hazardous Areas for Water Quality Using Dynamic Factor Analysis. In *Water Quality Monitoring and Assessment*; InTech: Rijeka, Croatia, 2012; pp. 227–256.

52. 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]

53. Kisekka, I.; Migliaccio, K.W.; Muñoz-Carpena, R.; Schaffer, B.; Li, Y.C. 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]

54. Kroiss, H.; Zessner, M.; Schilling, C.; Kavka, G.; Farnleitner, A.; Mach, R.; Blaschke, A.; Kirnbauer, R.; Tentschert, E.; Hassler, C. Effect of seepage and trickling of wastewater mechanically and biologically treated by small sewage treatment plants in decentralized locations. In *Endbericht. Im Auftrag des Bundesministeriums für Land-und Forstwirtschaft und Umwelt; Bundesministerium für Landwirtschaft, Regionen und Tourismus: Stubenring, Austria, 2006; pp. 1–249.* (In German)
55. Kroiss, H.; Matsche, N.; Vogel, B.; Zessner, M.; Kavka, G.; Farnleitner, A.; Mach, R.; Gutnecht, D.; Blaschke, A.; Heinecke, U. Effects of the infiltration of biologically treated wastewater on the groundwater. In Report for BuMi Wirtschaft u. Arbeit, BuMi Bildung Wissenschaft u. Kultur, BuMi Land-Forstwirtschaft, Umwelt und Wasserwirtschaft, Amt der Burgenländischen Landesregierung Abteilung; Amt der burgenländischen Landesregierung: Eisenstadt, Austria, 2002; Volume 9. (In German) [CrossRef]

56. Kersebaum, K.; Steidl, J.; Bauer, O.; Pierr, H.-P. Modelling scenarios to assess the effects of different agricultural management and land use options to reduce diffuse nitrogen pollution into the river Elbe. Phys. Chem. Earth Parts A/B/C 2003, 28, 537–545. [CrossRef]

57. Allen, R. An update for the calculation of reference evapotranspiration. ICID Bull. 1994, 43, 35–92.

58. Ben-Gal, I. Outlier Detection. In Data Mining and Knowledge Discovery Handbook; Maimon, O., Rokach, L., Eds.; Springer: Boston, MA, USA, 2005; pp. 131–146. [CrossRef]

59. Bánkóvági, G.; Ziermann, M. Questions of dynamic forecasts of economic relations. Közgazdasági Szle. 1973, 11, 1269–1286. (In Hungarian)

60. Geweke, J. The dynamic factor analysis of economic time series. In Latent Variables in Socio-Economic Models; Elsevier: Amsterdam, The Netherlands, 1977.

61. Zuur, A.; Pierce, G.J. Common trends in northeast Atlantic squid time series. J. Sea Res. 2004, 52, 57–72. [CrossRef]

62. Mendelssohn, R.; Schwing, F. Common and uncommon trends in SST and wind stress in the California and Peru–Chile current systems. Prog. Oceanogr. 2002, 53, 141–162. [CrossRef]

63. Trasy, B.; Magyar, N.; Havril, T.; Kovács, J.; Garamhégyi, T. The Role of Environmental Background Processes in Determining Groundwater Level Variability—An Investigation of a Record Flood Event Using Dynamic Factor Analysis. Water 2020, 12, 2336. [CrossRef]

64. Tabachnick, B.G.; Fidell, L.S.; Ullman, J.B. Using Multivariate Statistics; Pearson: Boston, MA, USA, 2007; Volume 5.

65. Fisher, R.A. An update for the calculation of reference evapotranspiration. Ann. Eugen. 1936, 7, 179–188. [CrossRef]

66. R Core Team. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2013.

67. Holmes, E.; Ward, E.; Kellie Wills, N.; Federal, M.E.H.-N. Package ‘MARSS’. 2018. Available online: https://cran.r-project.org/web/packages/MARSS/MARSS.pdf (accessed on 14 April 2020).

68. Blaschke, A.; Merz, R.; Parajka, J.; Salinas, J.; Blöschl, G. Effects of climate change on the water supply of ground and surface water. Österreichische Wasser-und Abfallwirtschaft 2011, 63, 31–41. (In German) [CrossRef]

69. Blöschl, G.; Schöner, W.; Kroiß, H.; Blaschke, A.; Böhm, R.; Haslinger, K.; Kreuzinger, N.; Merz, R.; Parajka, J.; Salinas, J. Adaptation strategies to climate change for Austria’s water management—Goals and conclusions of the study for federal and state governments. Österreichische Wasser-und Abfallwirtschaft 2011, 63, 1–10. (In German) [CrossRef]

70. Çimani, B.; Heinrich, G.; Hofstätter, M.; Kerschbaumer, M.; Kienberger, S.; Leuprecht, A.; Lexer, A.; Pesenteiner, S.; Poetsch, M.; Salzmann, M. ÖKS15 climate scenarios for Austria. Daten Methoden und Klimaanalyse Report Vienna 2016. (In German). Available online: https://data.ccca.ac.at/dataset/endbericht-oks15-klimaszenarien-fur-osterreich-daten-methoden-klimaanalyse-v01/resource/06edd0c9-6b1b-4198-9f4f-8d550309f35b (accessed on 7 August 2020).

71. Gobiet, A.; Kotlarski, S.; Beniston, M.; Heinrich, G.; Rajczak, J.; Stoffel, M. 21st century climate change in the European Alps—A review. Sci. Total Environ. 2014, 493, 1138–1151. [CrossRef]

72. Green, T.R.; Taniguchi, M.; Kooi, H.; Gurdak, J.J.; Allen, D.M.; Hiscock, K.M.; Treidel, H.; Aureli, A. Beneath the surface of global change: Impacts of climate change on groundwater. J. Hydrol. 2011, 405, 532–560. [CrossRef]

73. Scipione, D.; Trutvett, H.; Parajka, J.; Blaschke, A.P.; Hepp, G.; Kirchner, M.; Mitter, H.; Schmid, E.; Strein, B.; Zessner, M. Modelled impacts of policies and climate change on land use and water quality in Austria. Land Use Policy 2018, 76, 500–514. [CrossRef]

74. Bond, N.A.; Bumbaco, K.A. Summertime Potential Evapotranspiration in Eastern Washington State. J. Appl. Meteorol. Climatol. 2015, 54, 1090–1101. [CrossRef]

75. Allen, R. An update for the calculation of reference evapotranspiration. ICID Bull. 1994, 43, 35–92.

76. Ben-Gal, I. Outlier Detection. In Data Mining and Knowledge Discovery Handbook; Maimon, O., Rokach, L., Eds.; Springer: Boston, MA, USA, 2005; pp. 131–146. [CrossRef]

77. Blöschl, G.; Schöner, W.; Kroiß, H.; Blaschke, A.; Böhm, R.; Haslinger, K.; Kreuzinger, N.; Merz, R.; Parajka, J.; Salinas, J. Effects of climate change on the water supply of ground and surface water. Österreichische Wasser-und Abfallwirtschaft 2011, 63, 31–41. (In German) [CrossRef]

78. Bánkóvági, G.; Ziermann, M. Questions of dynamic forecasts of economic relations. Közgazdasági Szle. 1973, 11, 1269–1286. (In Hungarian)

79. Geweke, J. The dynamic factor analysis of economic time series. In Latent Variables in Socio-Economic Models; Elsevier: Amsterdam, The Netherlands, 1977.

80. Zuur, A.; Pierce, G.J. Common trends in northeast Atlantic squid time series. J. Sea Res. 2004, 52, 57–72. [CrossRef]