

Article

Time-Series Analysis of Monitoring Data from Springs to Assess the Hydrodynamic Characteristics of a Coastal Discharge Zone: Example of Jurjevska Žrnovnica Springs in Croatia

Andrej Stroj ¹ , Jasmina Lukač Reberski ^{1,*} , Louise D. Maurice ² and Ben P. Marchant ²¹ Croatian Geological Survey, Milana Sachsa 2, 10000 Zagreb, Croatia; astroj@hgi-cgs.hr² British Geological Survey, Nicker Hill Road, Keyworth, Nottingham NG12 5GG, UK; loma@bgs.ac.uk (L.D.M.); benmarch@bgs.ac.uk (B.P.M.)

* Correspondence: jlukac@hgi-cgs.hr

Abstract: This study assesses the functioning of the karst aquifer system located on the Croatian coast of the Adriatic Sea, where saltwater intrusion often presents a major problem for freshwater supply. We use two years of sensor data collected from two coastal springs to conduct a range of time-invariant and time-variant statistical analyses over various timescales. We perform separate analyses of the within-day and longer-term variation in the data as well as the interactions between the spring levels, salinity, rainfall, and sea levels. Such comprehensive analyses provide a greater understanding into the inner functioning of the intricate, heavily karstified aquifers. Time-invariant time-series analyses of the hourly data indicate that the spring levels and salinity are strongly controlled by sea levels. Furthermore, time-variant wavelet analyses demonstrate that the variation in spring levels in both springs has two modes defined by flow regime. Increases in the delay of the spring response to sea level indicate that aquifer diffusivity decreases in low flow conditions. Analyses facilitated the development of a conceptual model of the karst subsurface in the discharge zone. Using daily data, we constructed a linear mixed model of the spring levels. This model identified long-term sea level changes, rainfall from previous weeks, and seasonal recharge patterns as the primary factors influencing longer-term spring dynamics.

Keywords: karst; coastal spring; salinity; time-series analyses; wavelet; linear mixed models



Academic Editor: Yong Xiao

Received: 10 April 2025

Revised: 25 April 2025

Accepted: 8 May 2025

Published: 13 May 2025

Citation: Stroj, A.; Lukač Reberski, J.; Maurice, L.D.; Marchant, B.P.

Time-Series Analysis of Monitoring Data from Springs to Assess the Hydrodynamic Characteristics of a Coastal Discharge Zone: Example of Jurjevska Žrnovnica Springs in Croatia. *Hydrology* **2025**, *12*, 118.

<https://doi.org/10.3390/hydrology12050118>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Autonomous sensors and data loggers are increasingly used to monitor hydrogeological processes and properties, such as the fluctuation of groundwater levels and various physico-chemical properties. These sensors can provide detailed high-frequency information regarding the recent variation in the hydrogeological system. They support assessments of the short-term (e.g., within a day) variations and oscillations of these time series. They also support an assessment of the timescales over which aquifer or spring levels interact with, or are driven by, other temporally varying parameters or variables such as, in the case of coastal aquifers, rainfall or sea levels. Sensor data can be used to determine which hydrogeological processes tend to lead or drive short-term variation in other properties and the magnitude of the time lags in these relationships, providing insight into the hydrogeological properties of the aquifer. However, suitable data loggers have only been available in recent decades, and they require regular maintenance to record reliable data, so the duration of the observed series of measurements is often limited (e.g., to the duration of a research project). Therefore, they are often less suitable for assessing the longer-term (e.g.,

between years) variation in hydrogeological properties or for understanding multi-year trends and interactions between different properties. The study of coastal karst aquifers is a good example of using high-frequency logger data, which are necessary to characterise the typically rapid response of karst springs to precipitation and sea tides.

Karst aquifers are widespread in many countries across the Mediterranean coast. These aquifers are valuable freshwater sources [1]. However, seawater intrusion into the aquifers presents a major problem for freshwater supply in the area. Seawater intrusion is density-driven as seawater has a higher density than freshwater. Consequently, seawater tends to create a wedge beneath freshwater in the aquifer [2]. Highly conductive karst conduits positioned below sea level create routes for concentrated seawater intrusion into the aquifer [3–5]. Even coastal springs situated above sea level can discharge brackish water if they receive seawater from conduits positioned below sea level. Generally, seawater intrusion into the karst springs is controlled by two main factors: the dynamic difference between sea level and groundwater head in the aquifer and characteristics (especially depths) of the conduits that connect the spring to the sea. Submarine springs (vruljas) that are common along the eastern Adriatic Sea coast, together with brackish coastal springs, are indicators of aquifers karstified deep below the present sea level. Vruljas often function as submarine springs during the wet periods of the year and as openings for sea intrusion into the aquifer during dry periods. Karstification of many coastal aquifers deep below the present sea level occurred during periods of geological history when global sea levels were considerably lower than today (e.g., up to 120 m lower sea level during the last glacial maximum approximately 20,000 years ago) [6].

We explore the functioning of the coastal aquifer system from hourly spring level and salinity measurements over almost two years. The investigated springs are part of the abundant Jurjevska Žrnovnica discharge zone, which is located at the northern part of the Croatian Adriatic Sea coast within the broader Dinaric Karst region. Dinaric Karst, the largest continuous karst area of Europe, spreads across several countries (from Italy to Slovenia, Croatia, Bosnia and Herzegovina, Montenegro, and northern Albania) parallel to the eastern Adriatic Sea coast (Figure 1).

Karst rocks, mostly limestone and limestone breccia of Mesozoic and Cenozoic age, underlie the entire catchment area of the Jurjevska Žrnovnica. Both limestone and limestone breccia are compact and well-cemented, with negligible intergranular porosity. However, groundwater tracer tests in the broader area have demonstrated that the limestone subsurface is extensively karstified. Groundwater flow is primarily concentrated in karst conduits, surrounded by a generally low-permeability rock mass. Therefore, the geometry of the conduit network is the dominant control of flow directions and velocities [7]. Local conduit geometry also influences the characteristics and dynamics of the freshwater–seawater interface in the subsurface at the coastal discharge zones.

The tracer tests have also demonstrated that the studied springs historically received water from an extensive catchment, which includes the coastal mountains in their immediate hinterland (North Velebit mountain, up to 1700 m elevation) and the sinkholes of the sinking rivers that flow across two extensive karst poljes in the mountain hinterland (Lika and Gacka rivers, more than 20 km away from the springs). The tracer tests revealed a very complex network of groundwater paths from the sinkholes through the karstic massif to an approximately 70 km wide discharge area along the Adriatic Sea coast [7–9]. It should be noted that today the Lika and Gacka rivers are used for energy production within the hydroelectric power plant system and no longer sink into natural sinkholes. Consequently, the Northern Velebit mountain now represents the main recharge area for the Jurjevska Žrnovnica springs.

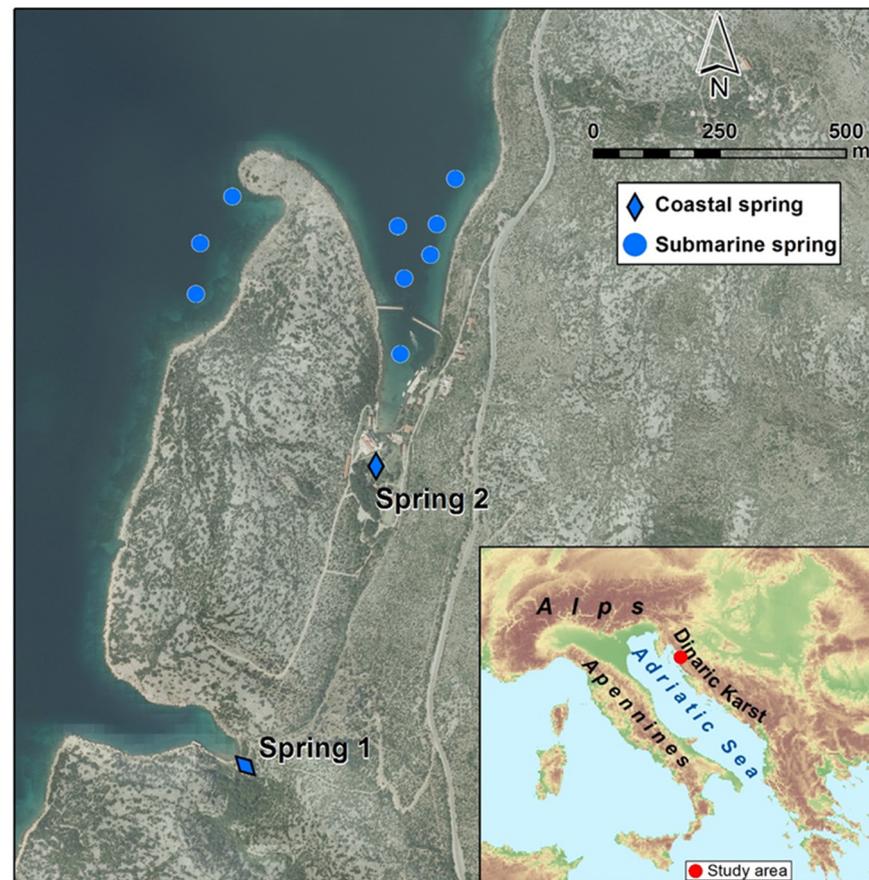


Figure 1. Studied springs (Spring 1 and Spring 2) are located within one of the most abundant coastal groundwater discharge zones on the Adriatic coast of Croatia, called Jurjevska Žrnovnica, which consists of a few coastal and several submarine springs (vruljas). (Orthophoto copyright: Croatian State Geodetic Administration (<http://www.geoportal.dgu.hr>, accessed on 12 February 2025); European elevation map copyright: European Environment Agency (<http://www.eea.europa.eu>, accessed on 12 February 2025).

The two studied springs are the most abundant coastal discharge locations within the Jurjevska Žrnovnica discharge zone, which also includes many perennial and intermittent submarine springs (i.e., vruljas, Figure 1). However, it is estimated that submarine groundwater discharge largely surpasses the discharge of the coastal springs [10]. To date, this has not been determined in detail due to the inability to precisely measure and monitor submarine discharges. Therefore, we analysed only the dynamics of the coastal springs in relation to sea tides and rainfall dynamics. However, due to the high degree of karstification in this aquifer the springs within the zone are presumably hydraulically connected below the surface, meaning that the dynamics of the coastal springs likely reflect the dynamics of the entire discharge zone. Spring 1 is located only 30 m from the sea coast at an elevation of slightly less than 1 m above mean sea level (AMSL). Discharge in the spring ranges between 0 and 150 L/s. Spring 2 is approximately 90 m away from the sea at 1.5 m AMSL, and it is part of a group of several springs situated very close to each other. Their total discharge ranges from 0 to almost 1000 L/s. However, total discharge of all coastal and submarine springs in the area of Jurjevska Žrnovnica is estimated to be an order of magnitude larger, ranging from several hundred to several thousand L/s. The distance between Springs 1 and 2 is 600 m (Figure 1), and they both discharge fresh water for most of the year but become brackish during low-flow periods. Both springs discharge from spring pools, which are formed by man-made overflows. Generally, water from the springs

is not utilised and flows naturally, but there is occasional low-intensity pumping for local irrigation from Spring 1.

While the exact positions of the karst conduits in the subsurface cannot be directly determined, spring and sea level dynamics can be directly measured and monitored. Tidal changes produce fluctuations in spring flow and salinity, as well as in groundwater level in the coastal aquifers. Comparison and analysis of groundwater and seawater dynamics can give valuable information about aquifer structure in the discharge area, i.e., conduit depth, volume, and openness to the sea. For instance, occasional increases in salinisation of Spring 2 imply, according to the Ghyben–Herzberg ratio for hydrostatic conditions [11,12], that conduits drained at the spring must be at least 60 m below sea level (in hydrodynamic conditions they must be even more to account for a loss of pressure along the conduits). Such information is crucial for evaluating measures to prevent seawater intrusion at the spring.

The primary objective of this study is to apply various statistical methods to analyse time-series data on the dynamics of the springs. This analysis aims to provide a detailed understanding of the karst underground structure and to develop a conceptual model of the Jurjevska Žrnovnica discharge area. Previous research primarily focused on hydrogeological mapping, tracer tests, and measuring the outflow and salinity of the spring water [13] without conducting an in-depth analysis of the spring dynamics. We consider different statistical techniques that can be used to interpret the spring level data over different temporal scales. We perform separate analyses of the within-day and longer-term variation in the data and the interactions between the spring levels, salinity, rainfall, and sea levels. Our analyses focus on understanding the temporal scales over which these parameters interact, the nature of their relationships, and the temporal lags between perturbations in one property leading to responses in others.

Other authors have previously used various similar techniques to characterise karstic systems. For instance, Padilla and Pulido-Bosch [14] and Larocque et al. [15] used time-invariant time-series and spectral methodologies to characterise French and Spanish karstic aquifers. These time-invariant approaches treat the relationships between variables as if they remain fixed across the study period. Jukić and Denić-Jukić [16] used partial spectral analysis and higher-order partial correlation functions [17] to solve ambiguities in the correlation functions related to various space–time-variant processes in the study of Jadro karst spring in Croatia. Charlier et al. [18] used time-variant wavelet analyses to observe the time-varying responses caused by pumping from a Mediterranean karstic system and Massei et al. [19] also used wavelet analyses to identify strong relationships between turbidity and rainfall that were not apparent from time-invariant analyses of a French karstic system. More recently, Razaee and Saatsaz [20], Wu et al. [21], Zhang et al. [22], and Çeliker et al. [23] employed wavelet analysis to identify multiscale impacts of climatic or anthropogenic factors on karstic watersheds. Meanwhile, Lovrinović et al. [24], Trevino et al. [25], and Yang and McCoy [26] used wavelet analysis to identify major factors influencing groundwater and groundwater-related surface water in porous coastal aquifers. To the best of our knowledge, no previous studies have employed wavelet analysis specifically to investigate karstic coastal systems.

In the within-day analyses, we initially assume that, after the removal of long term trends, behaviour is time-invariant with the same oscillations and interactions being evident for the entire duration of the sensor data. Rainfall data are only available on a daily time scale, and, therefore, we only consider interactions between spring levels, salinity, and sea levels. We use time-domain and frequency-domain time-series techniques to quantify the period or frequency of the oscillations within these measured properties and to determine

the lags between oscillations occurring in the time series. We then consider whether further insights into this variation can be obtained from time-variant wavelet analyses.

Analyses of daily data are designed to quantify more persistent trends in spring levels and how they can be explained by parameters such as sea levels and rainfall. Exploratory temporal-domain analyses are used to understand the timescales over which the spring levels are related to rainfall and sea levels. The results of these analyses inform the development of a linear mixed model for the variation in the levels of each spring. We highlight the information that can be gleaned from different time-series analysis methods and develop a conceptual model based on the hydrogeological implications of the time-series results.

2. Methods

2.1. Data Collection

The continuous monitoring of two coastal springs of Jurjevska Žrnovnica (Figure 1) started in August 2006 and ended in July 2008 (17,448 hourly measurements in total). Monitoring was discontinued after that period and has not resumed to date. However, the discharge zone remains predominantly in its natural condition and the hydrodynamic characteristics of the springs have remained unchanged since the monitoring period.

The monitoring of spring water levels was achieved by using automatic measuring devices with internal memory, i.e., data loggers. The loggers were placed within spring pools formed by weirs placed immediately downstream from the locations of water emergence. Onset HOBO water level loggers were used to measure the water level and temperature with a temporal frequency of 1 h. The specified accuracy of the loggers is ± 1 cm pressure of H₂O. The water level measurement data were subsequently compensated for changes in atmospheric pressure. Atmospheric pressure was measured by a separate device of the same type as the water level.

To determine the stage–discharge consumption curve, i.e., the functional dependence of discharge on changes in the pool water level, several flow measurements were performed under different hydrological conditions. The flow was measured with a mechanical current meter in the stream channel downstream of the spring pools. The calibrated stage–discharge curve was used to convert the water level to the rate of discharge from the spring. However, during portions of the monitoring period, the water level was below the weir of the spring pool, resulting in an absence of discharge despite variations in the water level still occurring. Such periods were more frequent on Spring 2.

During the second year of monitoring, specific electrical conductivity (EC) of Spring 2 was also measured with 1 h frequency using an Eijkelkamp CTD-Diver (accuracy 1% of the measured value).

The nearest sea-tide gauges with available hourly sea level data are located in Bakar (maintained by the Faculty of Science in Zagreb) and Zadar (maintained by the Hydrographic Institute of the Republic of Croatia). Sea level data for the spring location were obtained by a weighted average of data from these two stations according to their distance from the study area (the Bakar gauge is located approximately 50 km to the NW, and the Zadar gauge 90 km to the SSE). It should be noted that the tidal dynamics of these two stations was generally in phase with slightly different amplitudes (correlation coefficient between the two data series is 0.96).

Daily rainfall data were measured at the Zavižan measuring station (maintained by the Croatian Meteorological and Hydrological Service) located on the ridge of the Velebit Mountain, in the near hinterland approximately 10 km away from the springs and within their watershed.

2.2. Analyses of Complete Hourly Time Series

Since rainfall data were only available at a daily sampling frequency, hourly analyses were limited to spring and sea level data (and EC data for Spring 2 during the second year of monitoring). Due to the continuous character of the water level data series, in contrast to the discharge data which are interrupted by no-flow periods, the level was the primary variable considered in the analyses. Application of standard temporal- and frequency-domain time-series methodologies require the assumption that the time series are time-invariant or have time-invariant spectra. The long term trend was removed from the spring and sea level data by subtracting the daily mean of these quantities from the hourly data. The daily mean was assumed equal to a moving average of the hourly data with a window length of 24.

The de-trended spring and sea level data were then visually inspected for any evidence of time-variance. Note that long term trends represent one type of non-stationarity. The de-trended time series might also be time-variant in the sense that their degree of variability or the relationships between different series varies with time.

Auto-correlation functions (ACFs) were calculated for each de-trended time series, and then the cross-correlation function (CCF) between the sea level and each spring level were calculated. From plots of these ACFs and CCFs, the dominant timescales of oscillations in these de-trended time series and the temporal lag between the spring and sea level data were determined.

Analogous analyses were then performed in the frequency domain using Fourier techniques. Spectra of each individual de-trended time series were calculated along with the coherence and cross-spectra between the sea level and each spring level. The coherence plot was used to identify frequencies where both series have substantial power, and the cross-spectra were used to identify the phase lag between the series.

Wavelet analyses were then applied to the raw (i.e., not de-trended) sea and spring level time series to explore the impacts of time-variance upon the results of the temporal- and frequency-domain analyses. These analyses were based on the analytic Morlet wavelet. Continuous wavelet transforms of each spring and sea level time series were calculated along with wavelet coherences between sea level and each spring level series. Where coherence between the two signals was evident, the (potentially temporally varying) phase lag was determined from the wavelet cross spectrum.

Where the wavelet analyses had identified time-variant behaviour (beyond the long term trends) in the time series, the series were split into portions that were approximately time-invariant. The temporal and frequency domain analyses were then applied to these seemingly time-invariant de-trended series to test whether these methods identified the same oscillatory behaviour and lags between signals as the wavelet analyses.

2.3. Analysis of Daily Data

The ACF and partial auto-correlation function (PACF) of each time series of daily spring level data were calculated to investigate the degree of temporal auto-correlation amongst these response variables and the type of model that might be used to represent it. The CCF between the potential drivers (sea level and rainfall) and the daily spring levels were calculated to investigate the strength of any relationship between these variables and the relevant timescales of the driving variables.

The results of these descriptive analyses were used to inform the development of linear mixed models of the variation in the daily level of each spring. The choice of auto-covariance function for the random effects of this model were informed by the ACF and PACF, and the different CCF informed the driving variables that were included in the fixed effects and whether it was necessary to adjust the timescale over which these

driving variables were accumulated. The driving variables were added to each linear mixed model iteratively, starting from a model with constant fixed effects, and the Akaike information criterion (AIC) [27] was used to determine whether each driving variable led to an improved model. The final models underwent leave-one-out cross-validation and the bias and mean squared prediction errors were calculated. The proportion of variance explained by the model (i.e., one minus the variance of the errors divided by the variance of the response variable) and the proportion of variance explained by the fixed effects term in isolation were also calculated. Finally, the hydrogeological implications of the terms included in the model were discussed.

A concise theoretical overview of all applied statistical methods is given in Appendix A.

3. Results

3.1. Data Summary

The water level of Spring 1 (Figure 2, top) generally ranges between 0.8 and 0.9 m AMSL, with two significant decreases down to 0.4 m AMSL. One occurs for an extended period in the second half of 2006 and the other is shorter and occurs at the start of August 2007. The level of Spring 2 (Figure 2, second row) varies between 1.2 and 2.0 m AMSL. Several sharp dips in the time series are evident, but these do not dominate the pattern of variation to the degree seen for Spring 1. However, sharp decreases in Spring 2 are more frequent and are distributed across the monitoring period. The levels of both springs are more variable during these periods of decreased levels. Sharp decreases in the observed level reflect periods of no-flow at the spring (i.e., when the spring level is below the weir of the spring-pool). Generally, days where flow is partly absent are more frequent in Spring 2 (Figure 3), although, in contrast to Spring 1, there are no days with permanent absence of flow in that spring.

The Spring 2 EC was only measured in the second year of the monitoring (Figure 2, third row). There are substantial periods where the EC levels are low and stable, characterised by oscillations within a narrow range of 0.4–0.5 mS/cm. These correspond to periods of mostly continuous flow from the spring (Figure 3) when the spring discharges fresh water only. In periods of discontinuous flow EC variation is much more pronounced with daily peaks of over ten mS/cm, indicating seawater intrusion into karst conduits where it mixes with fresh water before emerging. Within-day oscillations are evident in the sea level data (Figure 2, fourth row) in addition to some longer-term trends such as dips in sea levels at the start of 2007 and in February 2008. Longer-term trends of the sea level can be caused by atmospheric pressure changes and/or strong winds, such as a strong offshore directed wind called Bora, which is typical for the area.

The relationships between the different measured quantities can be better visualised by focussing on data from week-long periods (Figures 4–7). During the August 2007 period of low flow from Spring 1 (Figure 4), the relationship between sea level and the Spring 1 level is not completely clear. The spring dynamics are dominated by a few high-intensity drops of an order of magnitude greater than the usual amplitude of the spring level oscillations. These drops appear to occur 5–6 h after drops in sea levels. In contrast, the Spring 2 levels have a within-day pattern of variation that is consistent and approximately in phase with the sea level variation. The pattern of sea level variation appears to have a dominant oscillatory component of approximately 24 h period and a weaker component of 12 h period. Flow from Spring 1 becomes continuous after rainfall occurs on 2 August.

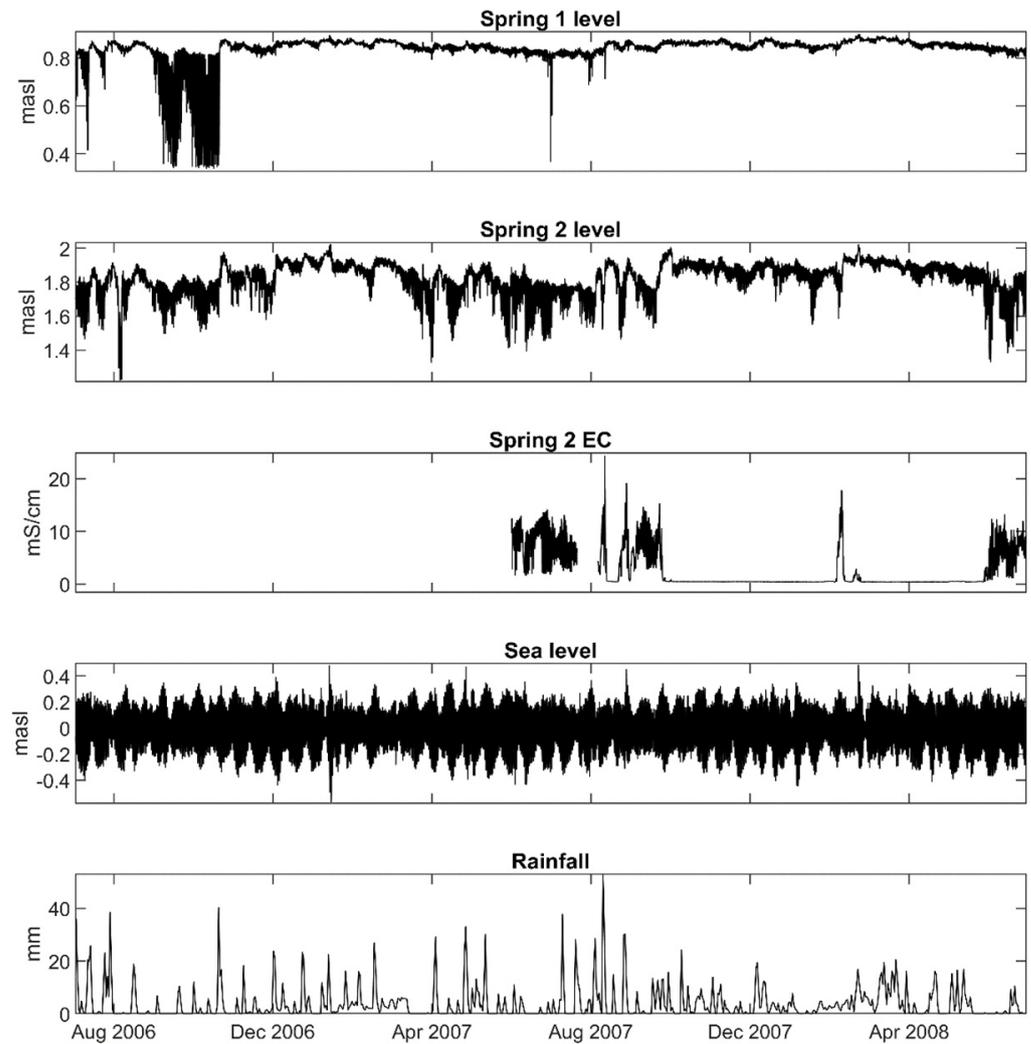


Figure 2. Observed hourly time series for the water level of Spring 1 (**top row**), level of Spring 2 (**second row**), electrical conductivity (EC) of spring 2 (**third row**) and sea level (**fourth row**). Daily time series for observed rainfall (**fifth row**). Notice that EC is very close to zero on the graph during periods with low EC (values of 0.4–0.5 mS/cm which indicates freshwater flow), while there are no EC measurements before June 2007.

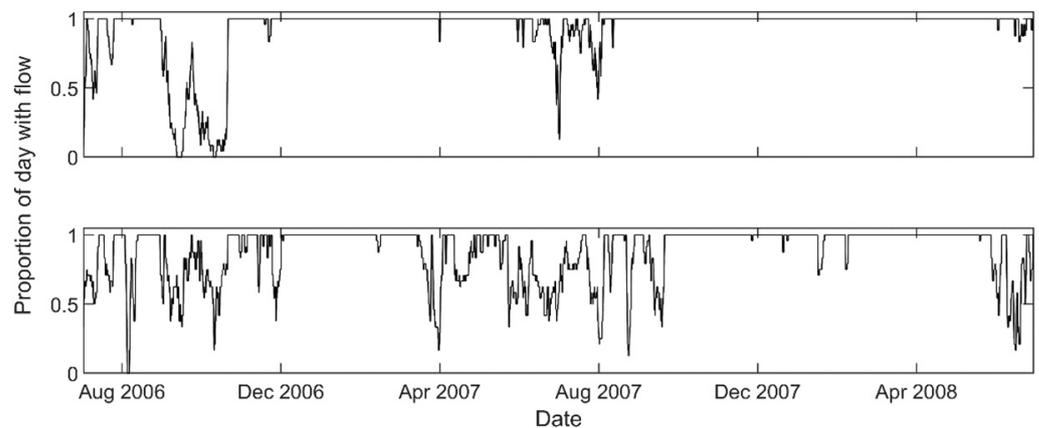


Figure 3. Proportion of each day where flow is recorded for Spring 1 (**top**) and Spring 2 (**bottom**).

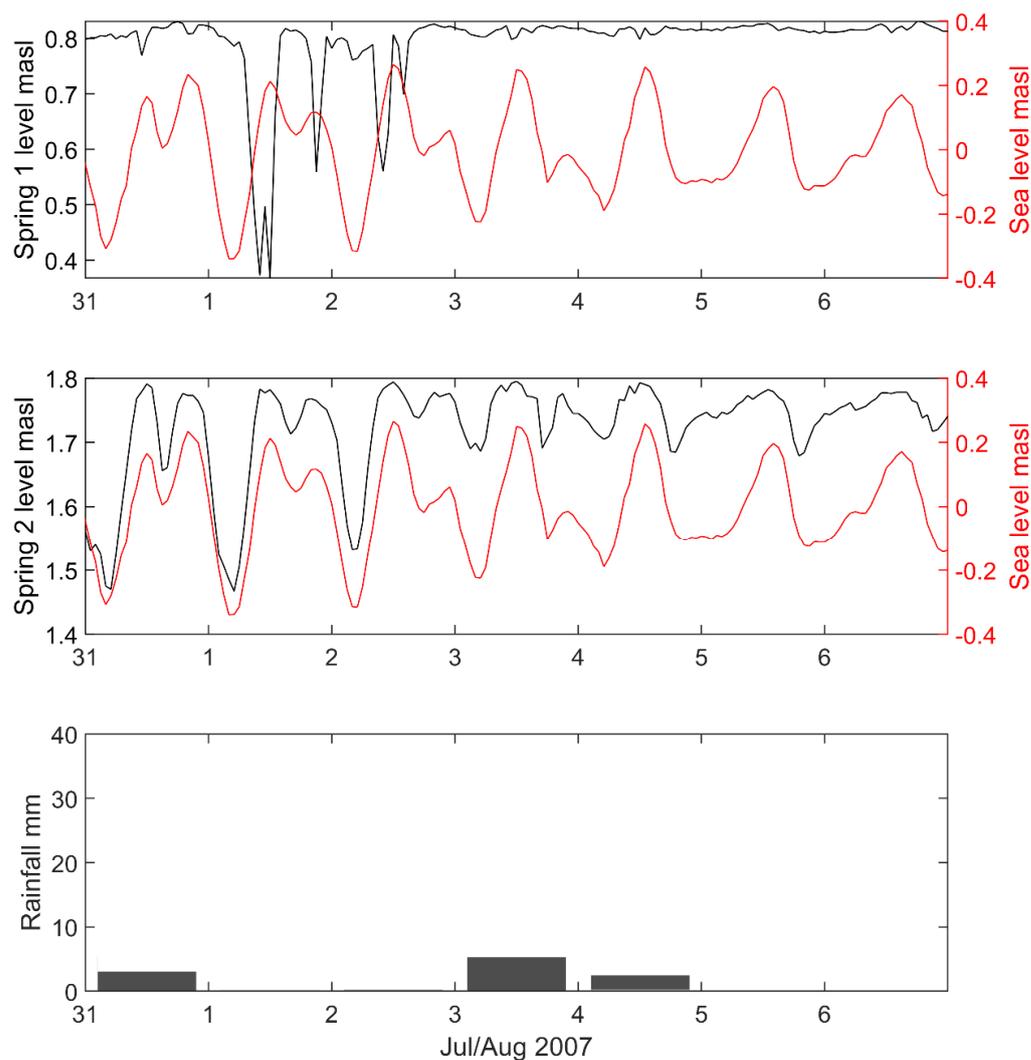


Figure 4. One-week time series of Spring 1 level (**top row**) and Spring 2 level (**second row**) measured with hourly frequency during a period with only occasional flow from Spring 1. Daily rainfall totals over the same week-long period (**third row**). X-axis ticks occur at midnight at start of day.

From 8 to 12 August 2007 (Figure 5), the Spring 1 level has the same broad pattern of within-day variation as the sea level (note the small scale of the spring level variation). However, this pattern is distorted by short-lived and sharp dips in the Spring 1 level, probably caused by irregular pumping from the spring pool. Local people occasionally pump water from the spring pool for irrigation, unfortunately, without making notes on pumping periods. The pumping mostly happens during dry periods and it only affects the spring level considerably during periods of discontinuous or very-low flow from the spring. From 13 August onwards, the within-day Spring 1 level variation is considerably more aligned to the sea level variation. This follows a large amount of rainfall on 10 August, after which pumping probably stopped. The within-day variation in Spring 2 remains approximately aligned to the sea levels through the 8–14 August period.

In December 2007 (Figure 6), there is almost continuous flow from each spring. Both spring levels are broadly aligned with the within-day sea level pattern, but the Spring 1 levels are less smooth than those of Spring 2 and include high-frequency fluctuations, again probably influenced by irregular pumping.

For the two periods with discontinuous flow, the EC at Spring 2 is very much aligned with the sea level, whereas for the period with continuous flow from each spring the EC is low and almost constant (Figure 7).

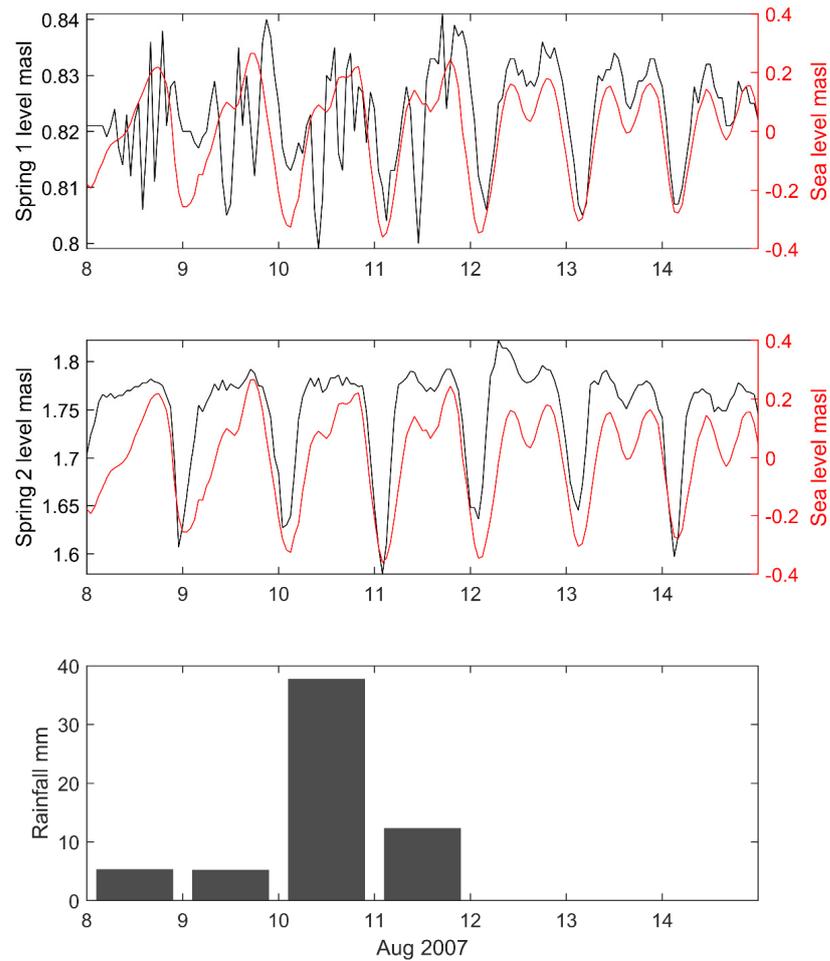


Figure 5. One-week time series of Spring 1 level (**top row**) and Spring 2 level (**second row**) measured with hourly frequency during a period of discontinuous flow from each spring. Daily rainfall totals over the same week-long period (**third row**). X-axis ticks occur at midnight at start of day.

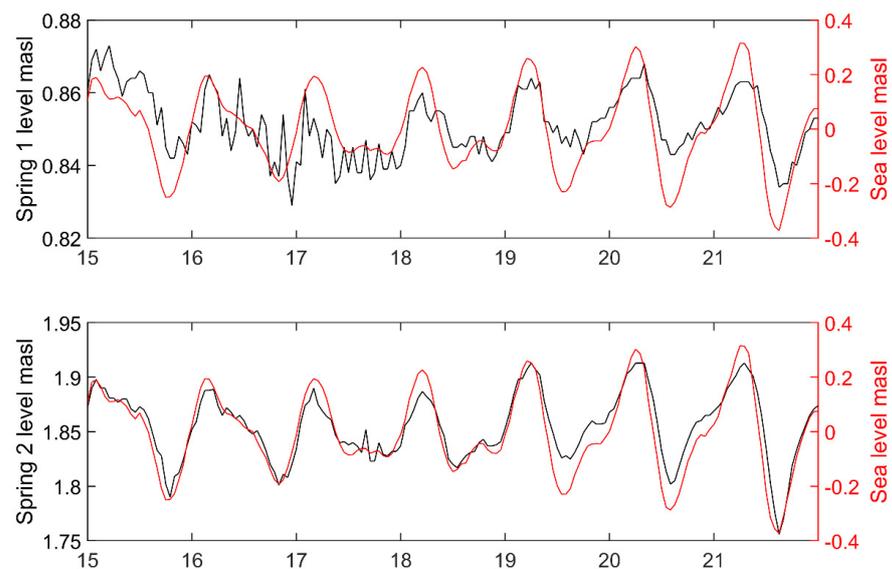


Figure 6. One-week time series of Spring 1 level (**top row**) and Spring 2 level (**second row**) measured with hourly frequency during a period of continuous flow from each spring. X-axis ticks occur at midnight at start of day.

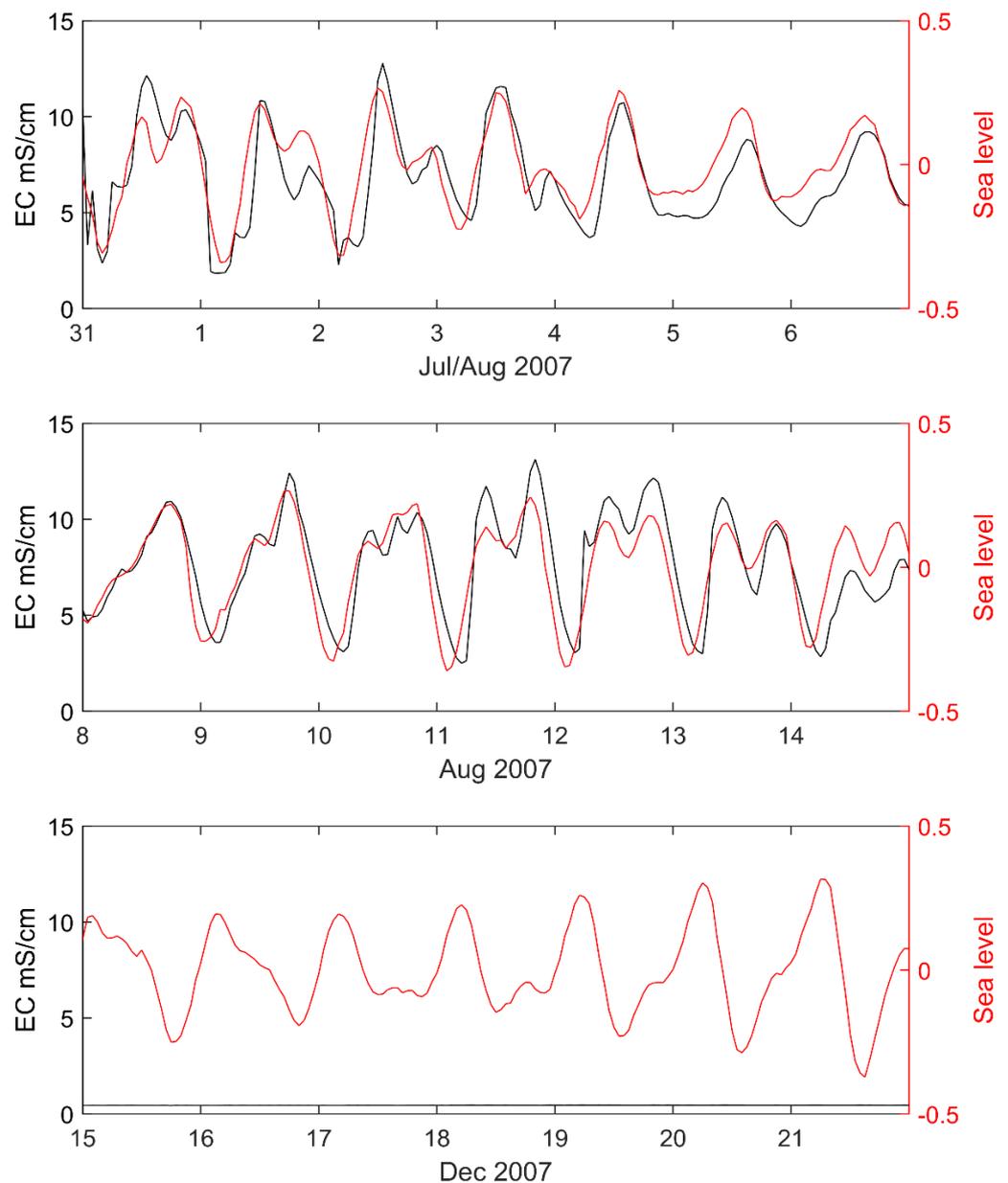


Figure 7. Week-long time series of electrical conductivity (EC) of Spring 2 water for the time periods shown in Figures 4–6. Notice that EC is very close to zero (0.4–0.5 mS/cm) in the graph in the third row, indicating exclusively freshwater flow.

3.2. Analysis of Within-Day Spring and Sea Level Variation

The ACF, spectrum, and wavelet spectrum of the Spring 1 levels (de-trended for the ACF and spectrum) all emphasise the strong oscillations of periods of approximately 12 and 24 h (Figure 8). Note that oscillations of period 24 h have a frequency of 0.042 cycles per hour. Spectral power is also evident in the Fourier and wavelet spectrums for approximately 12 h. In the ACF, there is strong negative auto-correlation at a lag of 12 h. In addition to the two horizontal bands corresponding to oscillations of periods 12 and 24 h, the wavelet spectrum for periods of no or discontinuous flow includes substantial power across all scales at times (e.g., in October 2006 and June 2007), reflecting the rapid changes in the spring levels (see Figure 8, right).

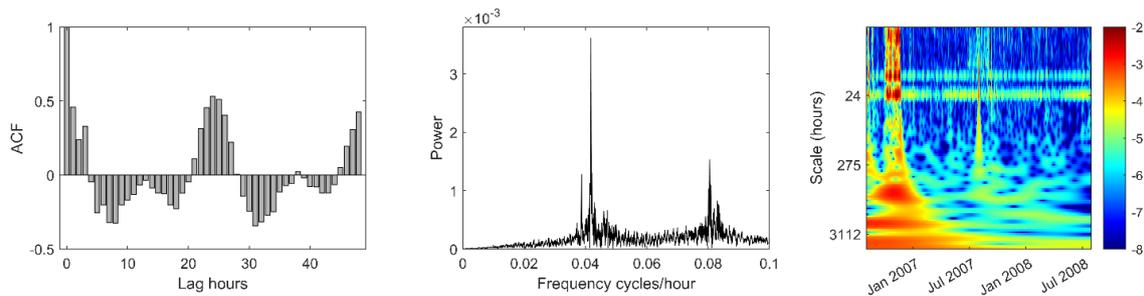


Figure 8. Auto-correlation of hourly and de-trended Spring 1 level (**left**); Fourier spectrum of hourly and de-trended Spring 1 level (**centre**); continuous wavelet transform of Spring 1 level (**right**).

The ACF, Fourier spectrum, and wavelet spectrum of the sea level data (Figure 9) are very similar to the corresponding plots for the Spring 1 level data, although the power across all scales in times of no or little flow is not evident in the wavelet spectrum. The Fourier coherence plot (Figure 10, right) between these two series confirms shared power at the frequency corresponding to 24 h. The phase of this component of the Fourier spectrum indicates that the sea level leads the spring level by 3.4 h. The peak in the corresponding CCF (Figure 10, left) indicates that the sea levels lead the spring levels by four hours. The wavelet power spectrum indicates that this phase lag is not constant for the duration of the time series (Figure 11, right). At times when there is flow for fewer than 12 h of the day, the phase lag averages 5.6 h whereas for time with more than 12 h of flow a day, the average phase lag is 1.1 h. Wavelet coherence analysis enables determination of relationship and phase lag between the sea and Spring 1 in low-water periods despite significant disturbances of the spring level due to irregular pumping (Figures 4 and 5).

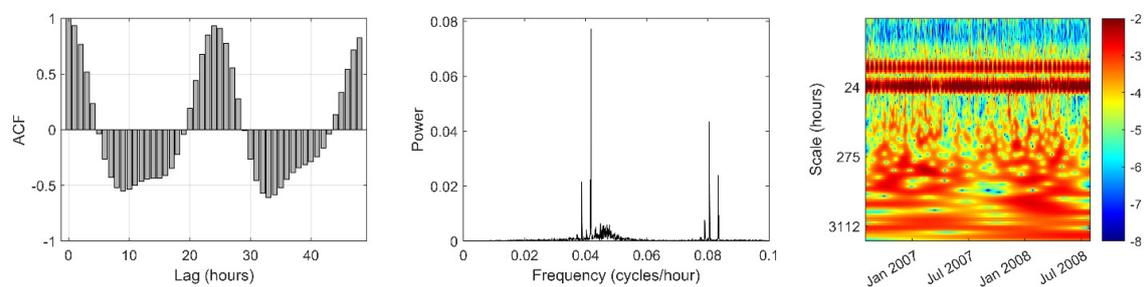


Figure 9. Auto-correlation of hourly and de-trended sea level (**left**); Fourier spectrum of sea level (**centre**); logarithm of continuous wavelet transform spectrum of sea level (**right**).

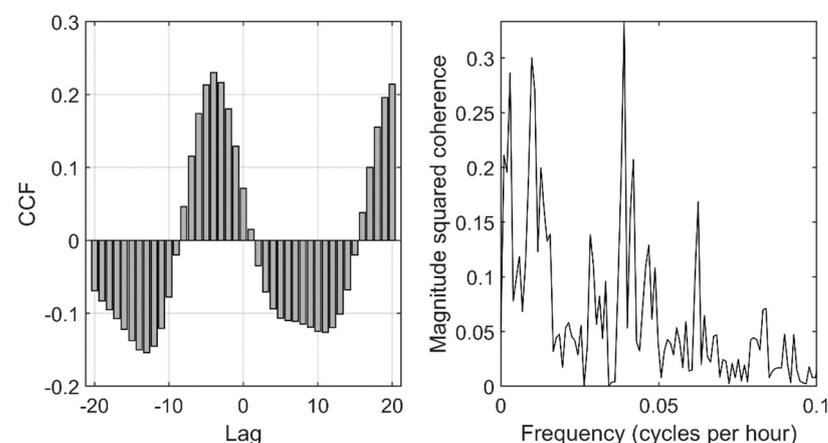


Figure 10. Cross correlation function between sea level and Spring 1 (**left**); fast Fourier transform derived coherence between sea level and Spring 1 as a function of frequency (**right**).

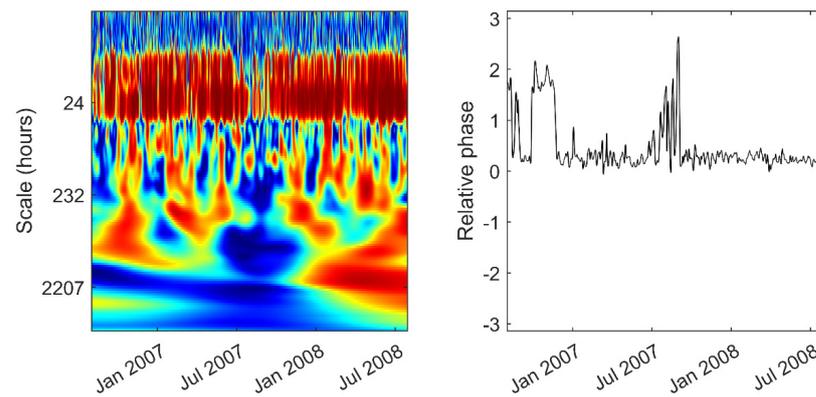


Figure 11. Wavelet coherence for sea level and Spring 1 level (**left**) and relative phase for 24 h scale wavelet spectrogram (**right**).

These phase lags are consistent with the results of temporal- and frequency-domain analyses limited to two seemingly time-invariant portions of the Spring 1 level time series (Figure 12). During one of these periods there is almost constant flow from the spring, whereas during substantial portions of the other period there is no flow. The CCFs indicate that the largest correlation between sea level and Spring 1 level occur for a lag of -6 h in the low-flow period, whereas the corresponding maximum correlation occurs for a lag of -1 h during the period with almost constant flow. The phases of the corresponding Fourier cross-spectrums suggest that in the low-flow period the sea level leads the Spring 1 level by 7.0 h, whereas in the period with flow this lag is reduced to 1.1 h.

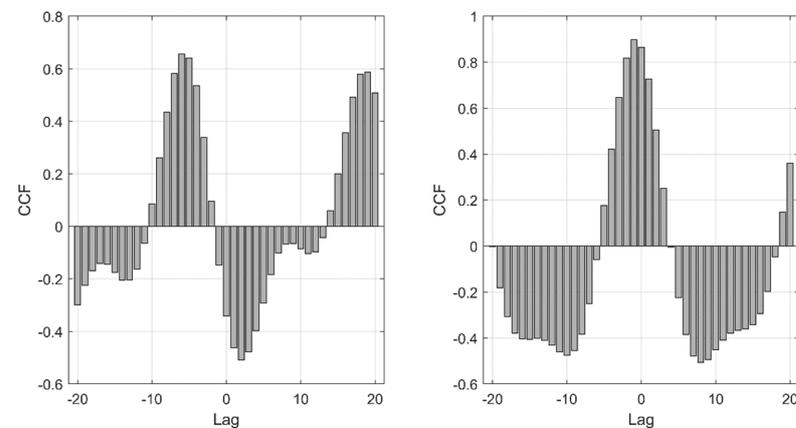


Figure 12. Cross-correlation functions of discontinuous flow period (**left**) and continuous flow period (**right**) of hourly sea and de-trended Spring 1 level time series.

There are similarities in the results of wavelet analyses for Spring 2 using hourly data (Figure 13). The dominant oscillations within the time series have periods of 24 and 12 h and continue for the duration of the time series. These frequencies also dominate the wavelet coherence plot relating sea levels to Spring 2 levels, and generally, the phase in the sea level/Spring 2 wavelet 24 h period component is close to zero. However, the phase does become negative during discontinuous periods of the time series. On average, during continuous flow periods the spring level lags behind the sea level by 0.2 h, whereas for periods with discontinuous flow the lag is -0.9 h. These periods of negative phase are counter-intuitive since they indicate that the spring level is leading the sea level. In Figure 4, the daily rising limb of Spring 2 level does occur slightly ahead of the corresponding increase in sea level, while falling limbs are mostly aligned. It, therefore, appears that the negative phases are a feature of the data and not artefacts of the wavelet analysis. The process causing the negative phases is unclear, although it is likely to reflect the complex

hydraulics with several coastal springs connected by fractures and conduits within the subsurface and could be connected to transition from a flow to a non-flow regime on Spring 2.

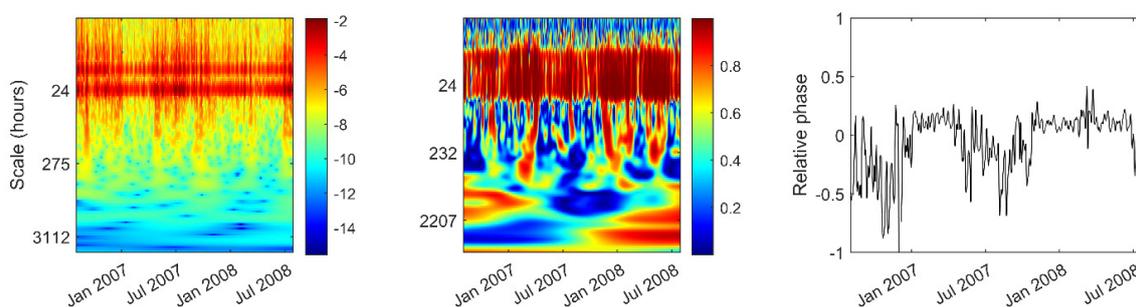


Figure 13. Logarithm of continuous wavelet transform spectrum of Spring 2 level (**left**); wavelet coherence between sea level and Spring 2 level (**centre**); relative phase between sea level and Spring 2 level for continuous wavelet transform (CWT) [28] component with period 24 h (**right**).

The EC for Spring 2 is low and relatively constant (within 0.4–0.5 mS/cm range) for periods with continuous flow. When there is discontinuous flow, and hence seawater intrusion, the EC increases, becomes more variable, and exhibits substantial with-day variation. The wavelet coherence between sea level and EC exhibits substantial power at 12 and 24 h for the discontinuous flow periods (Figure 14). The phase for the period 24 h component during these periods is noisy, perhaps reflecting the transition between flow and no-flow states. It is generally small and positive, indicating that, on average, the fluctuations in EC lag behind sea level fluctuations by 1.5 h. EC lag reflects the actual delay of seawater reaching the spring in relation to the change in tides.

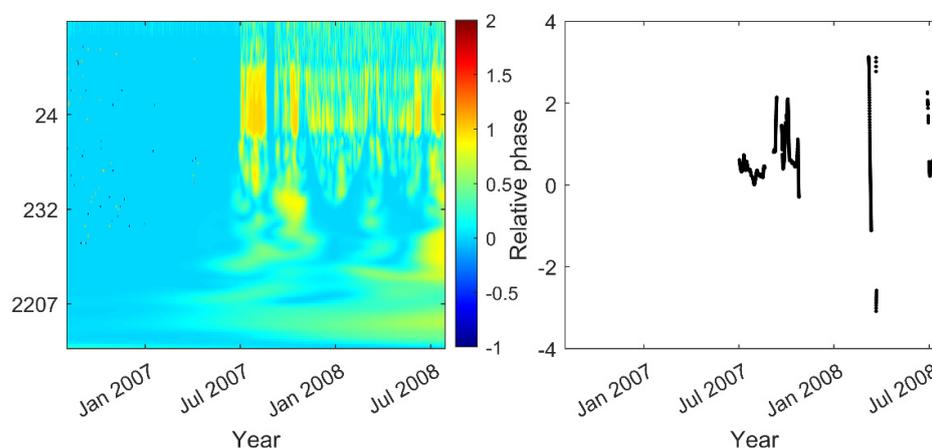


Figure 14. Wavelet coherence between sea level and Spring 2 EC (**left**); relative phase between sea level and Spring 2 EC for CWT component with period 24 h (**right**). The relative phase is only shown for periods where EC > 2 mS/cm.

3.3. Auto-Correlation and Cross-Correlation Functions of Daily Data

The moving averaged daily spring-level data are shown in Figure 15. As with the hourly data, the sharp dip in Spring 1 levels in the second half of 2006 is evident, although the smoothing of the time series has largely removed the shorter dip in 2007. Our analyses of these daily data are limited to the periods marked in black on Figure 15 which correspond to periods of continuous flow from the springs. The ACFs for the smoothed data from both springs indicate that observations are significantly correlated beyond a lag of 30 days (Figure 16). However, the PACF for Spring 1 is dominated by a positive and significant value at lag 1. Negative auto-correlations at lag 2, which just exceed the 95% confident

interval, are also evident, but the dominance of the lag 1 value indicates that an autoregressive (AR) model of order 1 might be sufficient to represent the auto-correlation in this time series. The daily data from Spring 2 has a more complex auto-correlation structure and highly significant PACF values are evident for a number of positive lags.

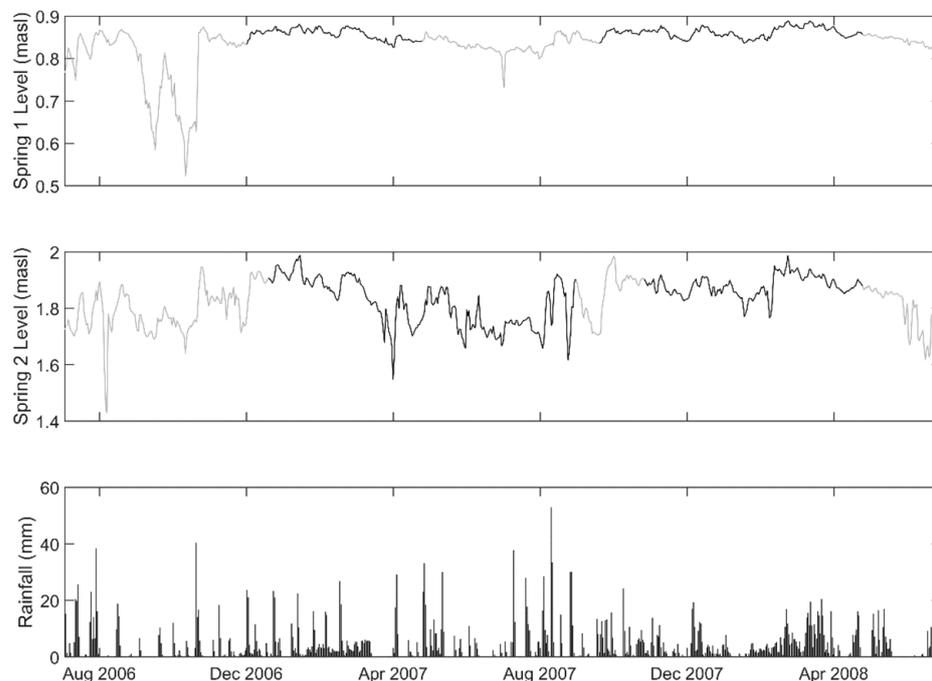


Figure 15. Time series of daily Spring 1 levels (**top row**); time series of daily Spring 2 levels (**second row**); daily rainfall (**third row**). Portions of time series used in linear mixed models are marked in black.

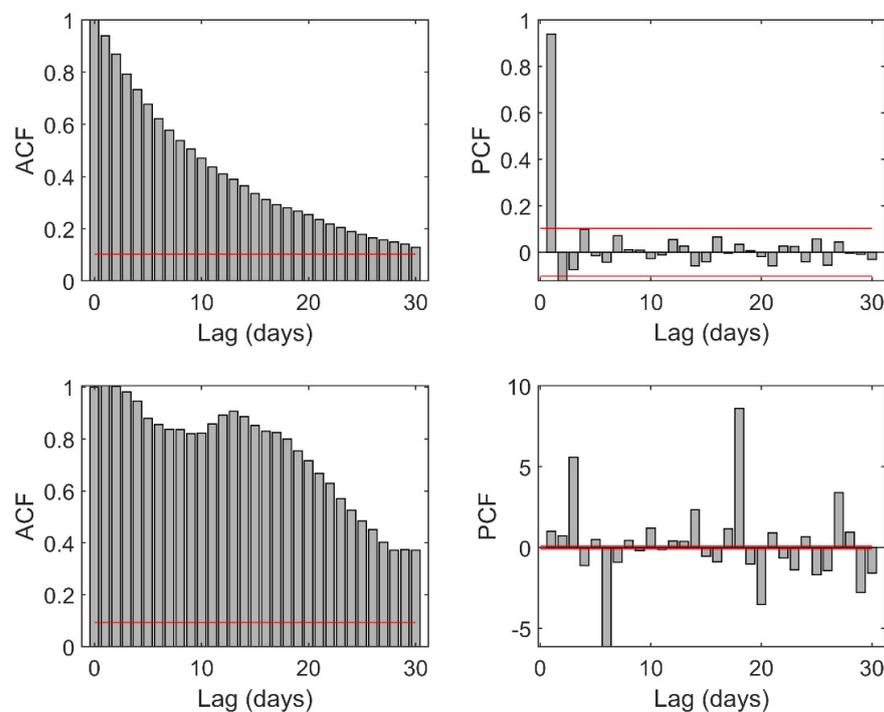


Figure 16. Auto-correlation (**top left**) and partial auto-correlation functions (**top right**) for daily observed Spring 1 level. Auto-correlation (**lower left**) and partial auto-correlation functions (**lower right**) for daily observed Spring 2 level. The 95% confidence limits for time series with no auto-correlation are marked in red.

The CCF between the sea level and daily Spring 1 level data indicate a significant correlation between lags of -8 and 3 days (Figure 17). The correlation is largest for lag -1 day. Similarly, the corresponding CCF for Spring 2 is significant between -3 and 2 days, with the largest correlation again occurring for a lag of -1 days. There is a second significant peak in this CCF at 15 days. This is likely caused by the peak in the Spring 2 ACF at around 12 days. The CCFs with rainfall are much less symmetrical and indicate that past rainfall over a number of days appear to influence spring levels. For Spring 1, the cross-correlations remain positive for lags beyond -20 days, although none of these cross-correlations are significant. For Spring 2, the cross-correlations are positive for lags of up to -11 days and are significant for lags between -2 and -5 days.

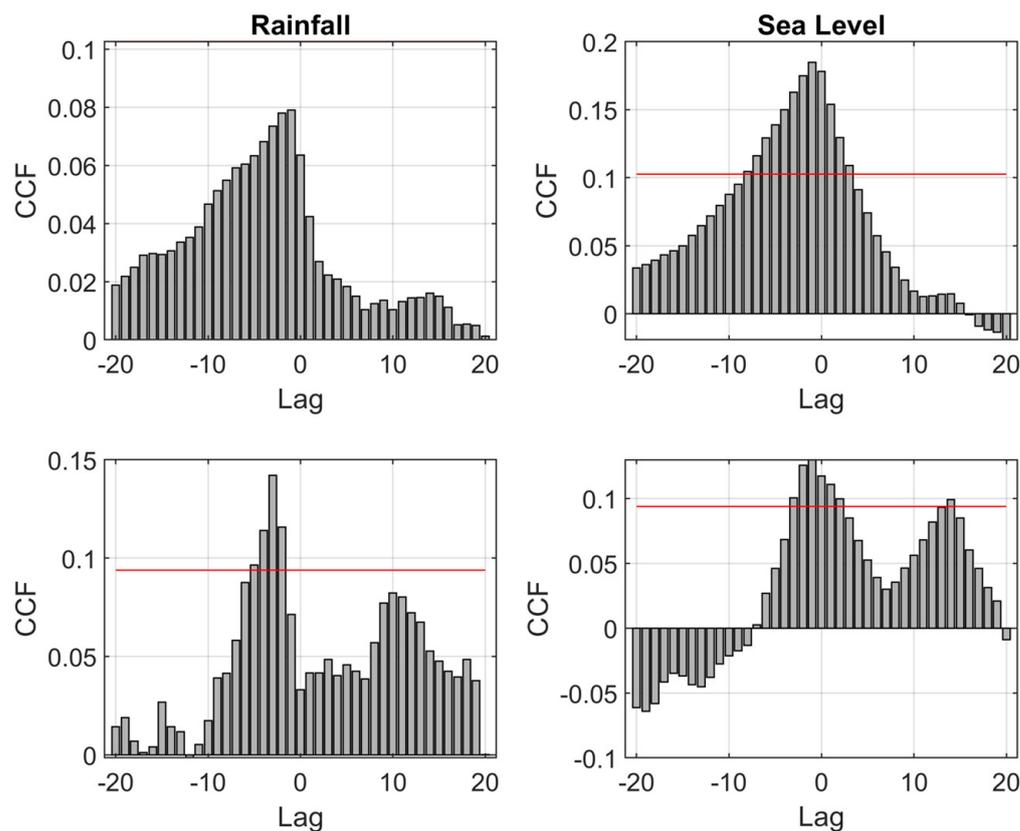


Figure 17. Cross-correlation functions between daily rainfall and Spring 1 level (**top left**); daily observed sea level and Spring 1 level (**top right**); daily rainfall and Spring 2 level (**lower left**); daily observed sea level and Spring 2 level (**lower right**). The upper limit of 95% confidence interval for uncorrelated time series is marked in red.

3.4. Linear Mixed Model for Daily Variation in Levels

The correlation functions for Spring 1 are consistent with variation being driven by concurrent sea levels and rainfall over the previous two-or-three weeks and with the auto-correlation in spring levels being derived from low-order auto-regressive terms (Figures 17 and 18). For Spring 2, the auto-correlation structure appears to be more complex. The validation statistics for linear mixed models, including these drivers as fixed effects and an exponential (equivalent to auto-regressive order 1) covariance function, are shown in Table 1. The rainfall fixed effect utilised an impulse response function (IRF) with contributions up to 30 days. These models were estimated upon the daily data marked in black in Figure 15 and correspond to days when there was flow from the springs.

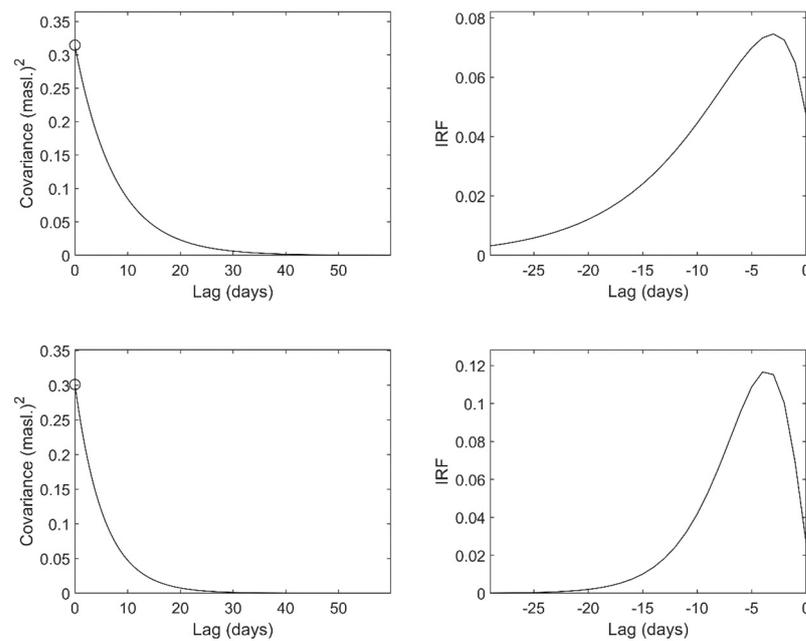


Figure 18. Estimated auto-correlation function for residuals from Model 3 of Spring 1 levels (**top left**); estimated impulse response to rainfall for Model 3 of Spring 1 (**top right**); estimated auto-correlation function for residuals from Model 3 of Spring 2 levels (**lower left**); estimated impulse response to rainfall for Model 3 of Spring 2 (**lower right**).

Table 1. Validation statistics of the estimated linear mixed models for each spring. Fixed effects are constant (C), rainfall IRF (R), sea level (Sea), and seasonal term (Seas).

Spring	Fixed Effects	AIC	Bias m	Variance Explained %	Mean SSE	Variance Explained by Fixed Effects %
1	C	218.54	0.00	95	0.86	0
1	C+R	157.06	0.00	96	0.96	49
1	C+R+Sea	110.71	0.00	96	1.00	59
1	C+R+Sea+Seas	104.52	0.00	96	1.00	70
2	C	292.05	0.00	96	0.75	0
2	C+R	257.72	0.00	96	0.81	17
2	C+R+Sea	243.82	0.00	96	0.81	17
2	C+R+Sea+Seas	219.51	0.00	96	0.78	69

All of the fixed effect terms lead to an improvement (i.e., decrease) in the AIC, justifying their inclusion in the models for both springs. In particular, the rainfall IRF led to a significant increase in the likelihood, despite the rainfall at all individual lags being non-significant predictors of the level of Spring 1. Upon leave-one-out cross-validation, all of the models are unbiased. The mean squared standardised error is 0.86 for the Spring 1 model with constant fixed effects but increases to the required value of 1.0 when other terms are included in the fixed effects. In each case, the linear mixed model explains 95–96% of the variation in the level of Spring 1. It should be noted that in this validation exercise the predictors use observations of spring levels from the day before and day after the prediction date and that the auto-correlation within the random effects is sufficient to produce accurate predictions. When fixed effects are added to the model they explain variation, which over these small lags was previously explained by the random effects. The variation explained by the fixed effects is more useful when there are larger gaps in the data which cannot be infilled by the random effects. The inclusion of the rainfall IRF term led to 49% of variation being explained by the fixed effects and this increased to 59% when the

sea level term was included. Visual inspection of the residuals from the model including constant rainfall and sea level in the fixed effects indicated that they contained a strong seasonal component. Therefore, a seasonal term (a sinusoid with a period of 365 days) was also added to the model. The resultant validation statistics indicated that this model was accurate and quantified uncertainty appropriately and that the fixed effects explained 70% of variation. The estimated covariance function and rainfall IRF for this model are shown in Figure 18.

When a similar model was estimated for Spring 2, each potential fixed effect term again improved the AIC (Table 1) and upon cross-validation all of the models were unbiased. The proportion of variation explained by the linear mixed model was almost identical to that of Spring 1 (96%). However, the mean squared standardised errors were between 0.75 and 0.81 (rather than 1.0), indicating that the uncertainty in model predictions was quantified less accurately. This possibly reflects the complex auto-correlation structure observed in Figure 17. In isolation, the rainfall fixed effects explained 17% of variation. Although the AIC improved, the addition of the sea level term led to a marginal decrease in the proportion of variation explained by the fixed effects. This illustrates how, for these relatively short time series, it can be challenging to discriminate between the effectiveness of models with explanatory fixed effects and those where variation is explained by the random effects. Again, addition of the seasonal term led to the largest increase in the variance explained by the fixed effects as it reached 69%. In agreement with the corresponding CCFs, the rainfall IRF for Spring 2 averaged rainfall over fewer days than that for Spring 1.

4. Discussion

4.1. Statistical Methodologies to Interpret Spring Level Time Series

Several time-series analysis techniques were employed in this paper to interpret the variation in spring levels. Each method highlighted different features of this variation. We now summarise the type of helpful information which could be inferred from each methodology.

Visual Inspection (Figures 2–7): The broad patterns of variation in the spring levels were apparent from visual inspection of the time series. These patterns included the absence of long term trends but clear differences in the patterns of variation during high- and low-flow conditions. By focussing on relatively short duration portions of the time series, it was possible to identify the short-term interactions between sea level and spring level time series and the time lag separating changes in these two series. However, it was not immediately apparent to what extent these short-term behaviours generalised across the entire time series.

Auto-correlation functions and Fourier spectra (Figures 8 and 9): These methodologies highlighted the dominant 12 and 24 h periods of oscillation in the spring and sea level time series.

Wavelet spectra (Figures 13 and 14): The wavelet spectra confirmed that the dominant periods of oscillation in the time series persisted for the entire observation period.

Fourier cross-spectra (Figure 10, right): These emphasised that the spring and sea level time series shared power in the oscillations of periods 12 and 24 h. The average lag between these dominant oscillatory components in each time series could be determined, although it was not apparent whether this lag varied in time. It was possible to localise these analyses to specific sections of the time series, but then it was not immediately apparent whether the observed behaviours generalised.

Cross-correlation functions (Figure 10, left): These also emphasised the average time lag between variation in the spring level and sea level time series, which differed for each

spring. When examining the relationship between daily rainfall and spring level, it was possible to see the temporal scale over which rainfall influenced spring level.

Wavelet coherence analyses (Figure 12): These also highlighted the shared power of the 12 and 24 h oscillations in the sea level and spring level time series. It was possible to see that the relationship persisted throughout the observation period, although the delay between these two signals varied depending on hydrological conditions, differing at each spring.

Linear mixed models (Figure 18): The linear mixed models were used to summarise the pertinent behaviour and drivers of spring level variation on a daily timescale. The relationship between daily rainfall over weeks and months and the spring levels could be expressed by simple impulse response functions. The model framework permitted tests of whether different covariates or processes were influencing spring level dynamics. However, some subjective decisions were required to decide which processes should be included in these tests.

4.2. Hydrogeological Implications of the Time-Series Analysis Results

The results confirm that sea level variations are the main driver of the spring level dynamics on a within-day timescale. However, the temporal dynamics of the spring level data clearly differ during periods of continuous and discontinuous flow, with much more variability occurring during the discontinuous flow periods. During continuous flow conditions, fluctuations in spring flow reflect groundwater head fluctuations in the draining aquifer. When flow on the higher spring (i.e., Spring 2) decreases, flow at the nearby lower (non-monitored) springs increases, keeping the total flow of the coastal spring zone unchanged (excluding multi-day trends). In this way, groundwater head fluctuations, induced by the sea level fluctuations, cause fluctuations in the flow ratio at individual springs within the zone. However, when the water level drops below the overflow (weir) height of Spring 2 its outflow stops, while the water level within the spring pool continues to fluctuate, directly reflecting aquifer head fluctuations. In contrast, when Spring 2 water level rises above the overflow, oscillations of the hydraulic head in the aquifer produce flow oscillations, while level variability is greatly reduced. In such conditions, overflow height dominantly controls the spring water level, with just a small-scale oscillation above the overflow reflecting flow oscillations. This results in substantially different intensities of water level fluctuation in flow and no-flow conditions. Differences in dynamics, particularly the different phase lags between the 24 h oscillations in sea level and spring level, are only apparent when a time-variant wavelet analysis is applied to the data (Figures 11 and 12). Differing phase lags are not apparent from time- or frequency-domain analyses which assume time-invariant variation.

The magnitude of the delays between the spring responses to sea level dynamics are important indicators of aquifer characteristics. Tidal fluctuations in the coastal aquifers can be used for calculating aquifer diffusivity under the assumption of homogenous and isotropic aquifer with horizontal flow [29,30]. Aquifer diffusivity is defined as the ratio of aquifer transmissivity and storativity so that both higher transmissivity and lower storativity result in higher diffusivity. Diffusivity of a coastal aquifer is negatively correlated to time delay between sea level and aquifer variations, and positively correlated to tidal efficiency, i.e., the ratio of aquifer to sea level variations. Although this rule does not strictly apply in extremely heterogeneous karst aquifers (both laterally and vertically), it can still be used to indicate the lumped system properties. In previous studies, karst aquifer diffusivity was mostly calculated by direct observations of delay between sea level and borehole hydrographs [31] or based on time-invariant time-series analysis [32]. Here, time-variant wavelet analysis allowed identification of diffusivity changes related to the water level in

the aquifer. Diffusivity variations connected to water level variations in a karst aquifer are likely to be caused by heterogeneity in the vertical distribution of high-permeability zones (e.g., karst conduits and intensely fractured zones) within the system.

Generally, the delay of Spring 1 level response to sea level variations is significantly longer than that of Spring 2, where response is almost instant and even seems to slightly precede sea levels during discontinuous flow periods. This is contrary to expectation as Spring 1 is situated much closer to the coast than Spring 2 and illustrates the extreme heterogeneity of the studied karst aquifer. In the case of a homogenous aquifer, delays in the spring response to sea level variations increase with distance from the sea. However, the absence of delay at Spring 2 indicates a direct connection with nearby submarine springs (Figure 1) through highly permeable conduits situated below the water table, while the connection between sea level and Spring 1 is less direct. According to the Ghyben–Herzberg principle of fresh to saltwater interface position in the subsurface of coastal aquifer [33,34] and considering the vertical position of Spring 2 (approx. 1.5 m AMSL), occasional salinisation of that spring demonstrates the existence of conduits at least 60 m below the sea level, and they are probably even deeper due to the loss of pressure along conduits in hydrodynamic conditions.

The counter-intuitive negative delay, i.e., the slightly earlier increase in Spring 2 level compared to the increase in sea level during discontinuous flow conditions, probably reflects the complex hydraulics during the transition from a flow to a non-flow regime. Specifically, Spring 2 is the highest among a group of several coastal springs located close to each other at slightly different elevations and connected by fractures and conduits within the subsurface. The Spring 2 water level was monitored in the spring pool, which is separated from the lower downstream springs. During the transition from high to low sea level stands, flow gradually moves from higher to lower discharge locations and vice versa. The negative shift at Spring 2, which is mostly observed on the rising limbs of the water level curve, could be connected to hydraulic functioning during the transition from a non-flow to a flowing regime in relation to flow changes at the lower surrounding springs in the surroundings. However, a more reliable determination of the mechanism causing the negative shift requires more detailed monitoring data of the individual discharge locations. During much of the monitoring period, when Spring 2 was overflowing for most of the day, the delay was slightly positive (approx. 0.2 h).

Occasional and irregular pumping from the Spring 1 pool during dry periods causes a significant distortion of its signal. This results in substantial power across all scales in the wavelet spectrum for such periods (Figures 11 and 12), while during the rest of the time periods approximately 12 and 24 h are dominant. Despite these occasional distortions during low-flow periods, the applied analysis enables determination of the time-variant relation between sea level variations and Spring 1 dynamics. Namely, the Spring 1 level delay compared to the sea level dynamics increases from approximately 1 h in continuous flow periods up to 5–6 h in discontinuous flow conditions (Figure 11). This implies a significant reduction in system diffusivity, i.e., decreased conduit transmissivity and/or increased storativity when there are low water levels in the aquifer. This, in turn, implies that Spring 1 is connected with the sea indirectly, probably via the central group of springs (Spring 2) by relatively shallow-positioned conduits, which are partly above groundwater level in low-flow periods. Loss of function of the uppermost conduits decreases transmissivity, while periodic filling and emptying of partly submerged voids increases system storativity, both of which can result in increased system inertia. In any case, it can be concluded that there is no direct connection between Spring 1 and the sea, while Spring 2 is directly connected to the sea through a highly transmissive conduit system which spreads at least 60 m below sea level. This is further supported by the general absence

of significant submarine springs in the vicinity of Spring 1 (Figure 1). Occasional slight salinisation of Spring 1 in dry periods (according to field observations) could originate through a connection with the Spring 2 conduit system rather than directly from the sea. Also, according to the available measurements, salinization of Spring 1 is less intensive than that of Spring 2. All this identifies Spring 1 as less vulnerable to salinization increases due to water pumping, i.e., it is more suitable for water supply despite its lower flow quantity and lower elevation compared to Spring 2. Figure 19 shows the conceptual model of the Jurjevka Žrnovnica discharge zone based on all the obtained findings.

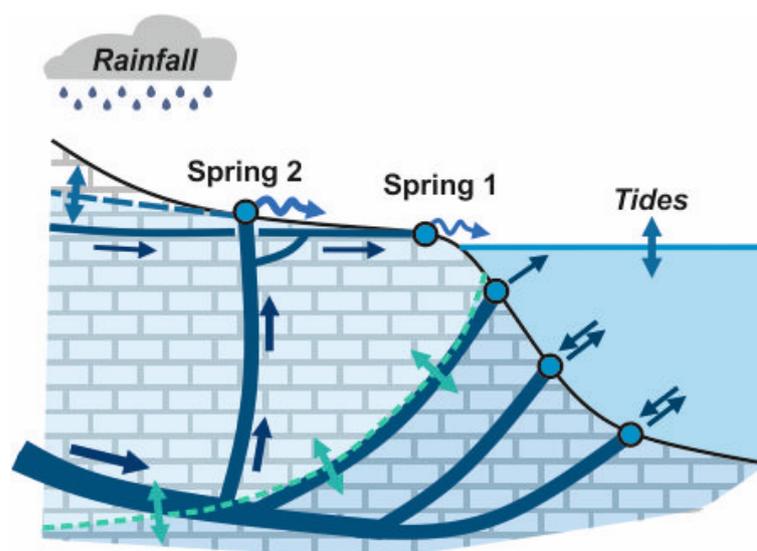


Figure 19. Schematic conceptual model of the karst subsurface at the Jurjevka Žrnovnica discharge zone (not to scale; detailed explanation provided in the text). Arrows show water flow direction.

Linear mixed models enable the identification of the main drivers that influence the flow of the springs on a multi-day timescale, as well as the significance and characteristics of their influence. On the daily timescale, sea level variations influence spring level instantly, while rainfall over the preceding period of 20–30 days affects it cumulatively. The models of the daily variation in spring levels were limited to periods of continuous flow at Spring 1 and largely continuous flow at Spring 2. Sea level variations and rainfall explained a substantially larger proportion of variation in Spring 1 levels relative to Spring 2. This is likely to have been because it was not possible to separate substantial periods of completely continuous flow for Spring 2 and so some non-flow periods were included in the model. However, the daily model of variation should be treated with some caution since it has been estimated using data covering less than two years. More time-variant behaviour could potentially be apparent in time series of longer duration.

The analysis also identifies a third important driver, which is the seasonality of the flow on a yearly timescale (Table 1). Effective infiltration, i.e., recharge of the karst aquifer, is mediated by surficial soil and the epikarst reservoir [35–37], which is filled by precipitation and drained both by evapotranspiration and percolation to the deeper parts of the hydrological system. Seasonality of the flow is a direct reflection of the strong seasonality of potential evapotranspiration, which is controlled by air temperature and vegetation cycles during the year. During the warmer part of the year, emptying by evapotranspiration is much more efficient, resulting in a usually emptier soil/epikarst reservoir prior to rainfall. Rainfall water firstly replenishes this reservoir, and only after its replenishment does effective recharge of the aquifer occur. This implies lower effective infiltration to the karst aquifer and consequently attenuation or absence of spring response to rainfall during the warm season. This explains why the inclusion of a seasonal component, in

addition to rainfall and sea level variations, largely increases model efficiency for both monitored springs.

5. Conclusions

A range of time-series methods were used to gain insight into the structure and functioning of the karst system of the Jurjevska Žrnovnica coastal and submarine discharge zone. Time-series data of the two coastal springs within the discharge zone were analysed. The analysis focused on variations in spring levels in relation to external drivers on different timescales (within-day and multi-day). Time-invariant analysis (e.g., auto-correlation, cross-correlation, Fourier spectra) provided average characteristics of spring dynamics throughout the monitoring period, while time-variant analysis (e.g., wavelets) revealed differences in dynamics related to changing hydrological conditions. On a multi-day time scale, the linear mixed models enabled the identification of the main drivers of the long-term spring level variations. The joint interpretation of these results led to the creation of a conceptual model for the discharge zone.

Sea level changes were identified as the primary driver of spring level variations on a within-day timescale. However, in variable hydrological conditions, non-stationarity in-spring level dynamics were observed, particularly in Spring 1 where phase lags between sea and spring levels were attributed to the vertical heterogeneity of the aquifer. Spring 2, identified as a central coastal spring, is directly connected to the sea via a submarine spring, while Spring 1 is connected to the sea indirectly via the conduits feeding the central spring. Occasional salinization of Spring 2 suggests deep-seated karst conduits linking it to the sea, estimated to be at least 60 m below sea level. In contrast, Spring 1 appears to be connected with the main system by shallower conduits, making it less vulnerable to large-scale salinization and more suitable for freshwater supply. Deep and highly permeable conduits connecting Spring 2 to the sea reduce the possibility of mitigating sea intrusion through engineering interventions.

On daily and longer timescales, flows from both springs are primarily governed by cumulative rainfall over the preceding 20–30 days, in conjunction with sea level fluctuations and the seasonality of potential evapotranspiration, which significantly influences effective rainfall infiltration. While the coastal springs exhibit dynamics typical of other karst systems, they are further modified by both short- and long-term sea level variations. Salinization is prevented only by maintaining a sufficient groundwater gradient toward the sea, underscoring the system's vulnerability to reduced recharge under future climate change scenarios. Our results demonstrated the importance and usability of time-series analysis in the study of complex coastal karst systems. The analysis of hourly data on a within-day timescale has proven to be particularly useful for determining the local characteristics of the system. In contrast, the analysis of daily data is more suitable for recognising and characterising the main drivers of system dynamics on a multi-day timescale.

Author Contributions: Conceptualization, A.S. and J.L.R.; methodology, A.S., B.P.M. and J.L.R.; formal analysis, B.P.M.; data curation, A.S.; writing—original draft preparation, A.S. and B.P.M.; writing—J.L.R. and L.D.M.; visualisation, A.S. and B.P.M.; supervision, B.P.M. and L.D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This paper is partly the result of collaboration, training, and education conducted through the GeoTwinn project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 809943. BGS authors publish with the permission of the BGS-UKRI director. The field investigations are funded and supported by the Croatian Geological Survey, Department of Hydrogeology and Engineering Geology.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to internal agreement.

Acknowledgments: The authors express their gratitude to the Croatian Meteorological and Hydrological Service for the provided data.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AMSL	above mean sea level
EC	electrical conductivity
CCF	cross-correlation function
ACFs	auto-correlation functions
PACF	partial auto-correlation function
CWT	continuous wavelet transform
AIC	Akaike information criterion
IRFs	impulse response functions
AR	auto-regressive
XWT	cross-wavelet transform
BLUP	best linear unbiased predictor

Appendix A. Theory

Appendix A.1. Time-Series Descriptive Statistics

Two critical aims of time-series analyses are to understand and interpret the relationships between measurements of a temporally varying response variable and (i) measurements of the same variable made at different times and (ii) measurements of other variables which potentially drive variation in the response. An understanding of how measurements of the response at different times are related provides insights into the timescale over which a perturbation of the response might continue to be evident. The relationships with other variables provide insights into the processes driving variation in the response and the timescales over which these processes are relevant.

For a time-invariant time series the degree of correlation between measurements of a temporally varying response and other temporally lagged measurements of the same variable can be quantified by the ACF. For a time-series of n measurements, y_t , made at equally spaced times $t = 1, \dots, n$, then the auto-covariance for lag k is estimated by the following [38]:

$$c_k = \frac{1}{n} \sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y}) \quad (\text{A1})$$

where \bar{y} is the mean of the y_t and the ACF for lag k is estimated by the following:

$$r_k = \frac{c_k}{c_0} = \text{Corr}(y_t, y_{t+k}) \quad (\text{A2})$$

The ACF is only calculated for non-negative k since $r_k = r_{-k}$. The upper and lower bounds of the 95% confidence interval of the ACF for a white noise time series (i.e., a sequence of mutually independent random variables with zero mean and finite variance) are as follows [38]:

$$\pm \frac{1.96}{\sqrt{n}} \quad (\text{A3})$$

Therefore, time series are considered to have significant auto-correlation for lags where r_k lies outside these bounds. If some data are missing from the time series then the

summation in Equation (A1) should be limited to cases where both y_t and y_{t+k} are available and the denominator n should be adjusted.

If there is significant auto-correlation between a response variable and lagged copies of itself up to a lag of m times, the time-step between measurements, the value of m provides an indication of the timescale of fluctuations in the response. However, the process causing the auto-correlation could act over a much shorter timescale. For example, if there is a direct relationship between successive measurements of the response this could lead to an indirect relationship and significant auto-correlation over lags of two, three, or more time-steps. In contrast to the ACF, PACF quantifies the correlation between time-series and replicates lagged by k time-steps once the effects of auto-correlation over shorter lags have been removed.

The PACF can be calculated from the ACF using the Durbin–Levinson algorithm [39]. If the PACF is only significantly different to zero for $k = 1$ then this indicates that an AR model of order 1 (i.e., a model where the value of the time series at time t depends linearly on the value at time $t - 1$ and a noise term) is sufficient to approximate the temporal correlation within the time series.

The CCF quantifies the relationship between time-series measurements of a property and lagged versions of another property. The cross-covariance between properties x and y at lag k is estimated by the following [38]:

$$g_k^{xy} = \frac{1}{n} \sum_{t=1}^{n-k} (y_t - \bar{y})(x_{t+k} - \bar{x}) \quad (\text{A4})$$

and the corresponding CCF is as follows:

$$r_k^{xy} = \frac{g_k^{xy}}{\sigma_x \sigma_y} \quad (\text{A5})$$

where σ_x and σ_y are the standard deviations of x and y . Similarly to the ACF, properties x and y are thought to be significantly related when x is displaced by lag k if the r_k^{xy} lies outside the 95% confidence interval (Equation (A3)). The CCF is plotted for both positive and negative k since, in general, $r_k^{xy} \neq r_{-k}^{xy}$. If r_k^{xy} indicates significant correlation over a negative lag k then y is responding to past values of driving variable x .

Appendix A.2. Time-Series Analyses in Frequency Domain

The methods described in the previous section can be used to analyse time series in the temporal domain. They interpret temporal variation by looking at the relationships between measurements made at different times. Time series can also be analysed in the frequency domain. Here, the time series is decomposed into the sum of sinusoids of a range of frequencies and the variation is interpreted in terms of the weight or power that is associated with each frequency.

Fourier's theorem [40] states that any reasonably continuous and periodic time series can be approximately decomposed into a finite sum of sinusoids, and it is as follows:

$$y_t = a_0 + \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \{a_m \cos(2\pi f_m t) + b_m \sin(2\pi f_m t)\} \quad (\text{A6})$$

where the $\lfloor x \rfloor$ indicates the integer part of x , the a_m and b_m are coefficients or weights, and the $f_m = \frac{m}{n}$ are frequencies that take values between 0 and 0.5. For a given length n time series, y_t , the coefficients a_m and b_m can be determined by the fast Fourier transform (FFT) [41]. Thus, the signal can be decomposed into a sum of sinusoids of frequency f_m

and amplitude $c_m = \sqrt{a_m^2 + b_m^2}$. It can be shown that the variance of component m is equal to $c_m^2/2$. A plot of this variance against f_m is referred to as the spectrum of the time series. Peaks in the spectrum are indicative of the frequencies of dominant oscillations within the series. When the spectrum is calculated using the FFT, resolution in the spectral peaks can be lost because of effective discontinuities (jumps to zero) beyond the start and end of the time series. These artefacts can be minimised by multiplying the signal by a length n window function that decays to zero at its start and end.

Analogous to the cross-covariance and cross-correlation functions, relationships between two time series can also be studied in the frequency domain. The cross-spectrum is defined as the Fourier transform of the cross-covariance function. The squared magnitude of the cross-spectrum describes the product of the power of the two time series for each frequency. The phase of the cross-spectrum can indicate whether oscillations in one of the time series tend to lead or follow oscillations in the other time series. The phase is generally quoted as being between $-\pi$ and π radians and indicates the degree to which the oscillations are aligned, with a phase of zero indicating alignment.

A large cross-spectrum power for a particular frequency does not necessarily indicate that the two time series are related. Such a cross-spectral peak could occur if just one of the time series has substantial oscillations at this frequency. The coherence function scales the cross-spectrum by the product of the square root of each of the spectra of each series. Thus, peaks in the coherence function indicate that both time series have substantial power at the corresponding frequencies.

Appendix A.3. Wavelet Analyses for Time-Variant Time Series

Time series measured in the environment often have time-variant spectra. For example, the mean of the time series might vary in time or according to a temporally varying covariate such as the amount of rainfall that has fallen at a location. Also, the oscillations within the time series or the delay between changes to a covariate leading to a response in the time series can vary according to the state of the system. The oscillations of the sinusoids in a Fourier series continue for the entire length of the time series and therefore the FFT cannot be used to identify oscillations that are localised in time. In contrast, wavelet functions are localised in time (i.e., they are only non-zero with a specified time window) and can be used to explore variation in the oscillations contained in a time series.

When the CWT is applied, the width or scale of the wavelet is varied and the wavelet is shifted across the signal before its convolution (i.e., similarity) with the time series is calculated. This leads to an indication of how the timescale of oscillations within the signal vary in time. The CWT of signal y is written as follows:

$$W_t^y(s) = \sqrt{\frac{\delta t}{s}} \sum_{t'=1}^n y_{t'} \psi_0 \left[(t' - t) \frac{\delta t}{s} \right] \quad (A7)$$

where s is the scale, δt is the time-step, and ψ_0 is the mother wavelet function. The quantities in square brackets after the mother wavelet describe the shifting and scaling of this function. The power of the CWT is defined as $|W_t^y(s)|^2$. Many different wavelet functions exist. They differ in terms of their shape, which in turn influences the temporal and scale resolution of the resultant spectra. Generally, a wavelet achieves good time resolution at the expense of scale resolution and vice versa. The Morlet wavelet provides a good compromise in this regard [42].

The cross-wavelet transform (XWT) between two signals x_t and y_t is defined as $W^{xy} = W^x * W^{y*}$ where the $*$ denotes complex conjugation. The power of this transform is defined as the magnitude of the CWT and the complex argument of W^{xy} is the local relative

phase between the two signals. If, for a particular time and scale, this complex argument is zero then the oscillations are exactly aligned. If this argument is π then the oscillations at this time and scale are exactly out of phase, with the peaks in one signal corresponding to the troughs in the other.

Analogously to a covariance function, the power of XWT increases according to the power in each time series. The XWT can be rescaled in line with a correlation function to become a wavelet coherence, such as the following:

$$R_t^2(s) = \frac{|S(s^{-1}W_t^{xy}(s))|^2}{S(s^{-1}|W_t^x(s)|^2) \cdot S(s^{-1}|W_t^y(s)|^2)} \quad (\text{A8})$$

where S is a smoothing function in time and scale [18]. The smoothing functions reflects that each of the wavelet transform terms in Equation (A8) corresponds to the expected value of that transform rather than the observed value. Without the window, all values of $R_t^2(s)$ would equal 1, but the smoothing function leads to the expected values being estimated by a weighted average of the observed values in the vicinity of t and s .

Appendix A.4. Time-Series Modelling

Mathematical models can combine relationships between driving variables and the response variable (e.g., spring levels), and temporal auto-correlation amongst the model residuals to predict the value of the response variable at times when it was not measured. Models of regularly spaced time series often approximate the auto-correlation by an AR or moving average process [43]. However, in this paper we utilise linear mixed models which include auto-correlated residuals [44]. These linear mixed models are more frequently used to model spatial variation and can accommodate missing data. The linear mixed model is written as follows:

$$\mathbf{y} = \mathbf{M}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (\text{A9})$$

where \mathbf{y} is a vector of n (possibly transformed) measurements of the response variable y_t ; \mathbf{M} is an $n \times q$ design matrix containing q temporally driving variables recorded at each of the n observation times; $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_q)^T$ is a vector of q regression coefficients; and $\boldsymbol{\varepsilon}$ is a vector containing n residuals. The $\mathbf{M}\boldsymbol{\beta}$ term, which is referred to as the fixed effects, relates the driving variables to the response variable. The $\boldsymbol{\varepsilon}$ term contains the residuals or difference between the measured y_t and the fixed effects are referred to as the random effects. The random effects are assumed to have been realised from a normal distribution and have zero mean and covariance matrix \mathbf{C} . Auto-correlation amongst the random effects is accommodated by permitting non-diagonal elements of \mathbf{C} to be non-zero. These non-zero values are determined by a parametric authorised function which relates the auto-correlation to the time lag separating two measurements. One commonly applied authorised function [45] is the nested nugget and exponential model, which is as follows:

$$C(\tau) = \begin{cases} c_0 + c_1 & \text{if } \tau = 0 \\ c_1 \exp(-\frac{\tau}{a}) & \text{for } \tau > 0 \end{cases} \quad (\text{A10})$$

where c_0 is the nugget variance, c_1 the partial sill variance, and a a temporal parameter which indicates the time-period over which the random effects are auto-correlated. The exponential function is equivalent to an order 1 AR auto-correlation model.

The parameters of the model (i.e., the elements of $\boldsymbol{\beta}$, c_0 , c_1 and a) must be calibrated to ensure consistency between the model predictions and the observed data. This calibration can be achieved using the maximum likelihood estimator [46] which uses a numerical

optimising procedure to find the parameter values which leads to the largest likelihood that the data would have arisen from the parameterised model. The maximised likelihood value can also be used to compare the effectiveness of different driving variables within the fixed effects and to decide whether specific driving variables should be included in the model. The set of fixed effects that lead to the lowest value of the AIC are considered to be the best balance between model fit and complexity. The AIC is written:

$$\text{AIC} = 2p - 2L \quad (\text{A11})$$

where p is the number of model parameters to be estimated and L the log-likelihood.

A response variable could potentially be controlled by the variation in a driving variable over an extended period of time. For example, the level of a spring could be the result of rainfall over the previous few weeks or months. Therefore, rather than one column of \mathbf{M} containing the observations of the driving variable made at the same times as the response, the column contains a weighted average of the driving variable over the relevant time-period. Marchant et al. [44] demonstrated how IRFs could be incorporated into the linear mixed model to automatically determine the optimal weighting of such an averaged term. They wrote the averaged IRF term at time t as follows:

$$\sum_{\tau=0}^{n_w-1} w(\tau)x_{t-\tau} \quad (\text{A12})$$

where $w(\tau)$ is the weight associated with the driving variable τ time-steps prior to the observation, $p(m_i)$ is the average value of that driving variable for month m_i from the start of the study period, and n_w is the number of time-steps included in the average. The weights are selected according to the following parameterised IRF:

$$w(\tau) = \frac{\alpha a^s \tau^{s-1} \exp(-a\tau)}{\Gamma(s)}, \quad (\text{A13})$$

where α , a , and s are the IRF parameters and $\Gamma(s)$ is the gamma function of order s . The IRF parameters are optimised as part of the maximum likelihood model estimation procedure.

Once a linear model has been estimated it can be used to predict values of the response variable at times when they were not measured and to determine the variance (i.e., uncertainty) of these predictions via the best linear unbiased predictor (BLUP) [46]. Validation of the model can be performed by predicting the value of the response variable at times when it was observed and by comparing the prediction with the observed value. For example, leave-one-out cross-validation removes a single observation and predicts the response variable at the corresponding time. If this process is repeated for all times then the average difference between the predicted and observed values and the average squared standardised errors can be calculated and used to assess the effectiveness of the model. The bias (i.e., mean difference between observed and predicted values) should be close to zero to indicate that on average the response is correctly predicted. The squared standardised error is the squared difference between the observed and predicted values divided by the prediction variance. If the mean squared standardised prediction error is close to 1 then this is an indication that the uncertainty of the predictions is accurately quantified.

References

1. Bakalowicz, M. Karst and karst groundwater resources in the Mediterranean. *Environ. Earth Sci.* **2015**, *74*, 5–14. [[CrossRef](#)]
2. Bear, J.; Cheng, A.H.D.; Sorek, S.; Ouazar, D.; Herrera, I. (Eds.) *Seawater Intrusion in Coastal Aquifers: Concepts, Methods and Practices*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1999; Volume 14.

3. Gjurašin, K. *Prilog Hidrografiji Primorskog Krša (Contribution to the Littoral Karst Hydrography)*; Tehnicki Vjesnik 59/4-6: Zagreb, Croatia, 1942; pp. 1–6. (In Croatian)
4. Kuscer, I.; Kuscer, D. *Observations on Brackish Karst Sources and Sea Swallow-Holes on the Yugoslav Coast*; Mémoires A.I.H.; Réunion d’Athènes: Athènes, Greece, 1964; pp. 344–353.
5. Fleury, P.; Bakalowicz, M.; de Marsily, G. Submarine springs and coastal karst aquifers: A review. *J. Hydrol.* **2007**, *339*, 79–92. [[CrossRef](#)]
6. Chappell, J.; Shackleton, N. Oxygen isotopes and sea level. *Nature* **1986**, *324*, 137–140. [[CrossRef](#)]
7. Stroj, A.; Paar, D. Water and air dynamics within a deep vadose zone of a karst massif: Observations from the Lukina jama–Trojama cave system (–1,431 m) in Dinaric karst (Croatia). *Hydrol. Process.* **2019**, *33*, 551–561. [[CrossRef](#)]
8. Paar, D.; Mance, D.; Stroj, A.; Pavić, M. Northern Velebit (Croatia) karst hydrological system: Results of a preliminary 2H and 18O stable isotope study. *Geol. Croat.* **2019**, *72*, 205–213. [[CrossRef](#)]
9. Bakšić, D.; Paar, D.; Stroj, A.; Lacković, D. Northern Velebit deep caves. In Proceedings of the 16th International Congress of Speleology, Brno, Czech Republic, 21–28 July 2013; Filippi, M., Bosak, P., Eds.; Czech Speleological Society and UIS: Brno, Czech Republic, 2013; Volume 2, pp. 24–29, ISBN 978-80-87857-08-3.
10. Stroj, A. *Underground Water Flows in the Hinterland of the Velebit Channel Coastal Karst Springs*. Ph.D. Thesis, University of Zagreb, Zagreb, Republic of Croatia, 2010.
11. Bonacci, O.; Roje-Bonacci, T. Sea water intrusion in coastal karst springs: Example of the Blaž Spring (Croatia). *Hydrol. Sci. J.* **1997**, *42*, 89–100. [[CrossRef](#)]
12. Ploessel, M.R. Ghyben-Herzberg ratio. In *Beaches and Coastal Geology*; Encyclopedia of Earth Sciences Series; Springer: New York, NY, USA, 1982. [[CrossRef](#)]
13. Biondić, R. Selected Coastal Karst Aquifers (Croatia): Jurjevska Žrnovnica. In *The Main Karstic Aquifers of Southern Europe*; Calaforra, J.M., Ed.; European Commission, Directorate-General for Research, EUR 20911 (Cost Action 621), EC: Brussels, Belgium, 2004; 123p.
14. Padilla, A.; Pulido-Bosch, A. Study of hydrographs of karstic aquifers by means of correlation and cross-spectral analysis. *J. Hydrol.* **1995**, *168*, 73–89. [[CrossRef](#)]
15. Larocque, M.; Mangin, A.; Razack, R.; Banton, O. Contribution of correlation and spectral analyses to the regional study of a large karst aquifer (Charente, France). *J. Hydrol.* **1998**, *205*, 217–231. [[CrossRef](#)]
16. Jukić, D.; Denić-Jukić, V. Partial spectral analysis of hydrological time series. *J. Hydrol.* **2011**, *400*, 223–233. [[CrossRef](#)]
17. Jukić, D.; Denić-Jukić, V. Investigating relationships between rainfall and karst-spring discharge by higher-order partial correlation functions. *J. Hydrol.* **2015**, *530*, 24–36. [[CrossRef](#)]
18. Charlier, J.-B.; Ladouche, B.; Maréchal, J.-C. Identifying the impact of climate and anthropic pressures on karst aquifers using wavelet analysis. *J. Hydrol.* **2015**, *523*, 610–623. [[CrossRef](#)]
19. Massei, N.; Dupont, J.P.; Mahler, B.J.; Laignel, B.; Fournier, M.; Valdes, D.; Ogier, S. Investigating transport properties and turbidity dynamics of a karst aquifer using correlation, spectral and wavelet analyses. *J. Hydrol.* **2006**, *329*, 244–257. [[CrossRef](#)]
20. Rezaei, A.; Saatsaz, M. Large-scale climate indices teleconnections with hydrochemical and isotopic characteristics of a karst spring using wavelet analysis. *Environ. Earth Sci.* **2021**, *80*, 335. [[CrossRef](#)]
21. Wu, L.; Wang, S.; Bai, X.; Chen, F.; Li, C.; Ran, C.; Zhang, S. Identifying the Multi-Scale Influences of Climate Factors on Runoff Changes in a Typical Karst Watershed Using Wavelet Analysis. *Land* **2022**, *11*, 1284. [[CrossRef](#)]
22. Zhang, J.; Zhu, Z.; Hao, H. The Effects of Climate Variation and Anthropogenic Activity on Karst Spring Discharge Based on the Wavelet Coherence Analysis and the Multivariate Statistical. *Sustainability* **2023**, *15*, 8798. [[CrossRef](#)]
23. Çeliker, M.; Uzun, S.; Yıldırım, G. Periodic variations of karstic spring discharge and precipitation from the perspective of wavelet analysis techniques: A case study of tacin spring (Kayseri, Türkiye). *Environ. Earth Sci.* **2025**, *84*, 46. [[CrossRef](#)]
24. Lovrinović, I.; Srzić, V.; Matić, I.; Brkić, M. Combined Multilevel Monitoring and Wavelet Transform Analysis Approach for the Inspection of Ground and Surface Water Dynamics in Shallow Coastal Aquifer. *Water* **2022**, *14*, 656. [[CrossRef](#)]
25. Treviño, J.; Rodríguez-Rodríguez, M.; Montes-Vega, M.J.; Aguilera, H.; Fernández-Ayuso, A.; Fernández-Naranjo, N. Wavelet Analysis on Groundwater, Surface-Water Levels and Water Temperature in Doñana National Park (Coastal Aquifer in Southwestern Spain). *Water* **2023**, *15*, 796. [[CrossRef](#)]
26. Yang, G.; McCoy, K. Modeling groundwater-level responses to multiple stresses using transfer-function models and wavelet analysis in a coastal aquifer system. *J. Hydrol.* **2023**, *627*, 130426. [[CrossRef](#)]
27. Akaike, H. Information theory and the maximum likelihood principle in 2nd International Symposium on Information Theory. In Proceedings of the 2nd International Symposium on Information Theory; Petrov, B.N., Csäki, F., Eds.; Akademiai Kiado: Budapest, Hungary, 1973.
28. Mallat, J. *A Wavelet Tour of Signal Processing*; Academic Press: New York, NY, USA, 1999.
29. Ferris, J.; Branch, G.S.G.W. *Cyclic Fluctuations of Water Level as a Basis for Determining Aquifer Transmissibility*; US Dept. of the Interior, Geological Survey, Water Resources Division, Ground Water Branch: Washington, DC, USA, 1952. [[CrossRef](#)]

30. Merritt, M.L. *Estimating Hydraulic Properties of the Floridan Aquifer System by Analysis of Earth-Tide, Ocean-Tide, and Barometric Effects, Collier and Hendry Counties, Florida*; US Department of the Interior, US Geological Survey: Reston, VA, USA, 2004; No. 3. [[CrossRef](#)]
31. Perriquet, M.; Leonardi, V.; Henry, T.; Jourde, H. Saltwater wedge variation in a non-anthropogenic coastal karst aquifer influenced by a strong tidal range (Burren, Ireland). *J. Hydrol.* **2014**, *519*, 2350–2365. [[CrossRef](#)]
32. Zhang, H. Characterization of a multi-layer karst aquifer through analysis of tidal fluctuation. *J. Hydrol.* **2021**, *601*, 126677. [[CrossRef](#)]
33. Ghyben, B.W. Nota in verband met de voorgenomen put boring nabij, Amsterdam Kononkl. *Inst. Ing. Tijdschr.* **1889**, 1888–1889, 8–22.
34. Herzberg, B. Die Wasserversorgung einiger Nordseebäder. *J. Gasbeleucht. Wasserversorg.* **1901**, *44*, 815–844.
35. Williams, P.W. The role of the epikarst in karst and cave hydrogeology: A review. *Int. J. Speleol.* **2008**, *37*, 1–10. [[CrossRef](#)]
36. Jones, W.K. Physical structure of the epikarst. *Acta Carsologica* **2013**, *42*, 311–314. [[CrossRef](#)]
37. Stroj, A.; Briški, M.; Oštrić, M. Study of Groundwater Flow Properties in a Karst System by Coupled Analysis of Diverse Environmental Tracers and Discharge Dynamics. *Water* **2020**, *12*, 2442. [[CrossRef](#)]
38. Diggle, P.J. *Time Series: A Biostatistical Introduction*; Oxford Science Publications: Oxford, UK, 1990. [[CrossRef](#)]
39. Durbin, J. The fitting of time series models. *ISI* **1960**, *28*, 233–244. [[CrossRef](#)]
40. Bracewell, R.N. *The Fourier Transform and Its Applications*; McGraw-Hill: New York, NY, USA, 1986. [[CrossRef](#)]
41. Cooley, J.W.; Tukey, J.W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **1965**, *19*, 297–301. [[CrossRef](#)]
42. Torrence, C.; Compo, G.P. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* **1998**, *79*, 61–78. [[CrossRef](#)]
43. Hipel, K.W.; McLeod, A.I. *Time Series Modelling of Water Resources and Environmental Systems*; Elsevier: Amsterdam, The Netherlands, 1994.
44. Marchant, B.; Mackay, J.; Bloomfield, J. Quantifying uncertainty in predictions of groundwater levels using formal likelihood methods. *J. Hydrol.* **2016**, *540*, 699–711. [[CrossRef](#)]
45. Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*; John Wiley & Sons: Chichester, UK, 2007. [[CrossRef](#)]
46. Lark, R.M.; Cullis, B.R.; Welham, S.J. On spatial prediction of soil properties in the presence of a spatial trend: The empirical best linear unbiased predictor (E-BLUP) with REML. *Eur. J. Soil Sci.* **2006**, *57*, 787–799. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.