Rescue of groundwater level time series
How to visually identify and treat errors”

Retike, Inga; Bikše, Jānis; Kalvāns, Andis; Dēliņa, Aija; Avotniece, Zanita; Zaadnoordijk, Willem Jan; Jemeljanova, Marta; Popovs, Konrāds; Babre, Alise; Zelenkevičs, Artjoms

DOI
10.1016/j.jhydrol.2021.127294

Publication date
2022

Document Version
Final published version

Published in
Journal of Hydrology

Journal of Hydrology 605 (2022) 127294

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Rescue of groundwater level time series: How to visually identify and treat errors

Inga Retike [a,*], Jānis Bikše [a], Andis Kalvāns [a], Aija Dēliņa [a], Zanita Avotniece [a], Willem Jan Zaadnoordijk [b], Marta Jemeljanova [a], Konrāds Popovs [a], Alise Babre [a], Artjoms Zelenkevičs [a], Artūrs Baikovs [a]

[a] University of Latvia, Faculty of Geography and Earth Sciences, 1 Jelgavas Street, LV-1004 Riga, Latvia
[b] TNO Geological Survey of the Netherlands, Princetonlaan 6, Utrecht, Netherlands

ARTICLE INFO

This manuscript was handled by C. Corradini, Editor-in-Chief, with the assistance of Saket Pande, Associate Editor

Keywords:
Groundwater level
Time series
Visual inspection
Monitoring
ARIMA models

ABSTRACT

Groundwater level time series are of great value for a variety of groundwater studies, particularly for those dealing with the impacts of anthropogenic and climate change. Quality control of groundwater level observations is an essential step prior to any further application, e.g., trend analysis. Often the quality control of data is limited to the removal of outliers or elimination of entire time series from a dataset, while such approaches drastically reduce the spatial coverage of initially huge datasets. Frequently studies tend to present already quality-controlled data, but neglect to demonstrate how the data were selected, judged, and modified. We present a data rescue approach developed for correcting the Latvian national groundwater level database, containing 1.68 million groundwater level observations since 1959, including 0.69 million manual measurements. A web-based R-Shiny interface was developed and used for visual identification and manual correction of erroneous measurements in groundwater level time series. All data manipulations were performed programmatically. Reproducibility and traceability were ensured by deploying separate data tables for raw observations, data repair actions and the final dataset. As a result of applied actions, 34.3% of all automatic measurements were either deleted or corrected, while only 6.5% of manual measurements were edited. Commonly found errors in groundwater level time series were grouped into: errors in measurement and data recording; technical problems at the observation site; local anthropogenic impact and other unclassified problems. The improvement from the rescue approach was assessed by comparing the Akaike information criterion derived from fitted ARMA and ARIMA models to both original and repaired time series. The results showed that models fitted using repaired time series were better than those fitted on the original time series for the same time series sections. The presented rescue approach and results can be of great value for all studies using groundwater level time series as an input.

1. Introduction

Groundwater globally ensures water supply, ecosystem functioning and human well-being, and the overall importance is expected to grow as groundwater is more buffered from seasonal and multi-year climate variability than surface water (UNESCO 2015, 2020). Increasing groundwater demand to supply drinking water, agriculture and industry in combination with climate change has highlighted the importance of groundwater protection (EEA, 2018; Naranjo-Fernández et al., 2020; Obergfell et al., 2019; Witte et al., 2019). Timely detection of negative groundwater level trends is crucial to make appropriate decisions and ensure sustainable groundwater management (Bakker and Schaars, 2019; Lehr and Lischeid, 2020), while reliable information on groundwater levels is a prerequisite prior any groundwater resources assessment (Ritzema et al., 2018).

Time series analysis can be of a great value for groundwater studies (Bikše and Retike, 2018; Jarsjö et al., 2020; Marandi et al., 2012; Noorduijn et al., 2019). However, such analysis requires availability of measured heads, sometimes also measured or estimated forcings (e.g., rainfall, evaporation, water pumping) for sufficiently long observation periods. Around the world, groundwater levels are measured in observation wells for a variety of reasons, for instance monitoring of long-

* Corresponding author.
E-mail address: inga.retike@lu.lv (I. Retike).

https://doi.org/10.1016/j.jhydrol.2021.127294
Received 8 June 2021; Received in revised form 16 October 2021; Accepted 27 November 2021
Available online 13 December 2021

term changes, assessment of seasonal variations, or evaluation of response to a particular stress (IGRAC, 2020). Thus, the spatial coverage and density of monitoring networks is uneven (Bakker and Schaars, 2019). In addition, observation periods and frequencies vary, and time series may contain essential gaps (Asgharinia and Petroselli, 2020).

Various data pre-selection criteria have been applied in previous studies depending on the research aim and scale. Zaadnoordijk et al. (2019) proposed to use an 8-year long observation period with a minimum of 84 measurements for adequate time series models reflecting the dynamics of the current groundwater system. Similar results were obtained by Heudorfer et al. (2019) who tested sensitivity of various indexes to changes of observed period location and time series length. In general, a higher sensitivity was observed in indices calculated on weekly rather than daily time series, while a coinciding drop in sensitivity for both daily and weekly indices could be observed when the length of time series reached 8 years. While Haaf and Barthel (2018) applied 10-year observation length criteria with a minimum weekly measurement frequency to capture multi-annual and decadal periodicities in groundwater signal. For the assessment of the impact of groundwater use on groundwater droughts, Wendt et al. (2020) used a dataset of 30-year time series from which they removed all series with more than 6 consecutive months of missing observations. Stoll et al. (2011) used a criterion of at least a 30-year long time series with a monthly temporal resolution to detect groundwater response to climatic variations. And Chen et al. (2004) used a 15 to nearly 40 years long time series to study the historic relationship between groundwater levels and climatic variables such as temperature and precipitation.

Often groundwater levels have been monitored for decades resulting in an extensive number of time series (Berendrecht and Van Geer, 2016). Quality control of the data is an essential first step prior to any further application, e.g., in time series analysis. The presence of various errors, such as outliers, shift and drift require evaluation of each time series (Zaadnoordijk et al., 2019). Post and Von Asmuth (2013) point out that the most common sources of error due to the actual measurement processes are related to the measurement instruments, the conversion from pressure to heads, time lag effects and defects of observation wells. Also, data processing errors (e.g., typing errors, duplicates) generally account for a large proportion of errors in databases (Kandel et al., 2011; Liu et al., 2018; Post and Von Asmuth, 2013). Consequently, groundwater level time series usually contain missing values, including those which are a result of error and outlier removal. There are several methods used to deal with missing data in groundwater level time series (Asgharinia and Petroselli, 2020; Von Asmuth et al., 2002; Wendt et al., 2020; Zaadnoordijk et al., 2019). However, analysis of series with a constant time step between subsequent measurements is easier and computationally less demanding (e.g., Post and Von Asmuth, 2013). For this reason, it may be beneficial to fill in missing values, although filling large gaps remains a challenge (Oikonomou et al., 2018) and using a mixture of measured and modeled values gives additional challenges in the assignment of an accuracy to the values. Therefore, time series with gaps are often removed from further analysis leading to significant reduction of the dataset (Wendt et al., 2020).

Data preprocessing is the most time-consuming and at the same time, the least documented phase in the data analysis pipeline which may strongly affect the quality of study results (Bernard et al., 2019; Kandel et al., 2011; Van den Broeck et al., 2005). Temporal aggregation can level out some random errors, while new errors might be introduced if such data are used in further calculations (Ritzema et al., 2018). The application of fully automatic data quality control procedures is often limited by the uncertainty of errors and the need for an expert judgement to verify the results (Ali et al., 2019; Liu et al., 2018). As concluded by Haaf and Barthel (2018) a visual inspection of groundwater level time series remains a valuable and necessary task in order to understand the data despite some shortcomings that should be taken into account. Visual inspection and manual correction based on the expert judgment might be time consuming, subjective, and hard to replicate (Naranjo-Fernández et al., 2020, Zaadnoordijk et al., 2019), yet it is simple to apply and widely used (Asgharinia and Petroselli, 2020; Avotniece et al., 2017; Haaf and Barthel, 2018) as human eye is very sensitive to spot differences in visual looks (Barthel et al., 2021)

Differentiation between actual quality issues and unusual (however valid) data values requires human interaction (Gschwandtner and Erhart, 2018). For instance, Haaf and Barthel (2018) categorize sudden or continuous changes in groundwater level time series that are hard to explain by natural factors as “irregular” using visual inspection. Suspicious cases can be checked using various accompanying data such as meteorological conditions or known local anthropogenic influences, but the prerequisite for a good assessment is a sufficient understanding of the study area and evaluation results may vary among experts (Ritzema et al., 2018). Lehr and Lischeid (2020) propose a method to identify potential measurement errors and anthropogenic influence using “stable” principal components (PCs) of all groundwater head series to calculate “reference hydrograph” that incorporates general patterns from PCs, but any deviation from actual observations indicates potential errors. However, the method requires observations measured at the same time intervals, thus limiting applicability to often irregularly obtained measurements. In addition, the authors also suggest that visual inspection should be included in the workflow of groundwater level time series assessment. Several interactive data quality control procedures integrating humans into the data treatment process are found to be useful to improve data quality. However, such approaches are task-specific and difficult to apply for other types of data (Liu et al., 2018).

As pointed out by Bernard et al. (2019), there is no single definition of “clean” data and it depends on the application which risks are associated with including wrong or excluding right measurements. Likewise, the assessment is not straightforward whether and how much the correction has improved the data. Models like Autoregressive Moving Average (ARMA) and its integrated variant ARIMA can be used as an approximation to describe the complex fluctuation patterns of groundwater levels using only one variable - the groundwater level itself (Adamowski and Chan, 2011). ARIMA models are frequently used to forecast time series in various disciplines, including hydrogeology (Ahn, 2000; Shirmohammadi et al., 2013; Patle et al., 2015; Mirzavand and Ghazavi, 2015; Gibrilla et al., 2018). The performance of time series models can be evaluated by Akaike’s information criteria (AIC) (Akaike, 1974), which is a relative metric typically used to select the best model created from the same dataset. For the assessment of data quality, there are not models with different structures using the same data, but models with the same structure using the same data (the original and repaired time series). So now, the AIC can be seen as a metric for the data quality and can be used to assess the improvement of the data due to the corrections.

An important purpose of the data cleaning is to improve (or make possible) analysis of long-term structural changes in groundwater level time series. Given the data quality issues, anthropogenic changes are more difficult or even impossible to detect (Barthel et al., 2021). Both anthropogenic changes and data errors deteriorate the performance of the AR(I)MA time series models. The increased performance of the AR(I)MA models indicates that the dataset is adequate, for example, to calibrate a physically based distributed groundwater model for the entire country or a large region with the aim to forecast climate change impacts on groundwater (TACTIC, 2021). In such models it is neither feasible nor relevant to include historic short term groundwater abstractions and other anthropogenic influences on groundwater. Time series modelling can be used to select appropriate long term groundwater level series for such a calibration (Zaadnoordijk and Bakker, 2013).

In this paper we present a data rescue approach and repair results for systematic groundwater level observations collected in the Latvian national database from 1959 till 2019. A visual assessment procedure exploiting a web-based interface was developed for identification and manual correction of erroneous measurements in groundwater level

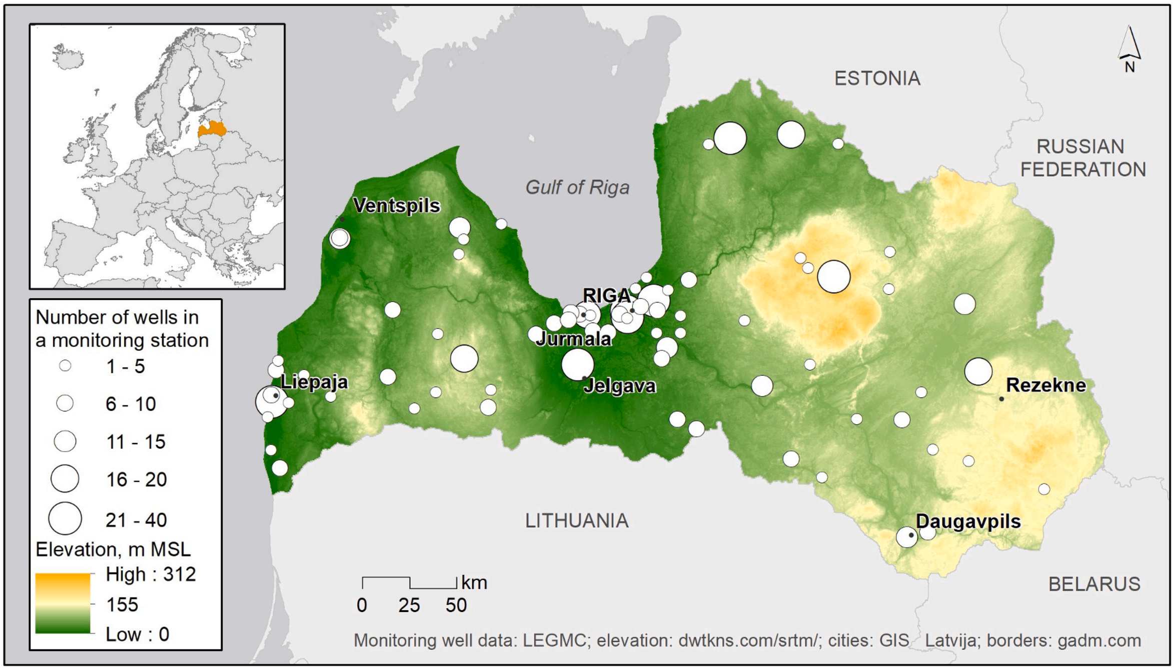

Fig. 1. Groundwater monitoring stations in the Latvian national database, 1959–2019.

time series. All data manipulations were performed programmatically ensuring reproducibility and traceability of the work. We have identified a number of errors commonly present in groundwater level time series and proposed type-specific data rescue actions. Finally, the improvement of time series after applied corrections was quantified by derivation of the Akaike information criterion from fitted ARMA and ARIMA models to both, the original and repaired groundwater level time series. Presented approach and results can be of great value for studies using groundwater level time series as an input.

2. Materials and methods

2.1. Study area

Latvia is located in North Eastern Europe and lies in the central part of the Baltic Artesian Basin. The present topography is shaped mostly by multiple advances and retreats of Pleistocene Ice sheets and the action of Baltic Sea. The elevation varies from few meters below the sea level up to 312 m above the sea level (Kalm and Gorlach, 2014). The thickness of the sedimentary cover varies from 500 m in the northern part increasing up to 2 km in the southwestern part (Lukševičs et al., 2012). Layering of the bedrock sequence is subhorizontal slightly inclining towards southwest direction (Brangulis and Kaņevs, 2002). Therefore, Middle Devonian sandstones, siltstones, dolomites and clays are exposed in the bedrock surface in the northern part of the territory, while in the southern part the bedrock surface exposes carbonate and terrigenous sequences of Upper Devonian and mostly terrigenous Mesozoic deposits (Lukševičs et al., 2012). Overlying Quaternary deposits are composed of interlayers of glacial, glacifluvial and glaciolimnic sediments with the thickness of a few meters in lowland areas increasing up to 200 m in uplands, particularly in central and eastern part of Latvia (Zelčs et al., 2011).

Within Latvia, three hydrodynamically and hydrochemically distinct zones separated by regional aquitards or aquicludes are delineated: stagnation zone (Ediacaran-Cambrian aquifer complex with brines), passive water exchange zone (Lower and Middle Devonian aquifer complex with brackish groundwater) and active water exchange zone (freshwater aquifers) (Jodkazis, 1989; Levins et al., 1998). Lukševičs et al. (2012) have explicitly described geological setting of the study area, while more details on hydrogeological conditions can be found in Babre et al. (2016) and Retike et al. (2016) studies. This study puts an emphasis on the active water exchange zone of aquifers corresponding to the Middle and Upper Devonian as well as Quaternary which are mainly used for water supply in Latvia. Only 6 of the 612 groundwater level time series belong to the passive water exchange zone and none to the stagnation zone.

Climate in Latvia is characterized by its location in the transition zone between continental and maritime conditions – the country lies in the north-western part of the Eurasian continent, but at the same time is strongly affected by maritime climate impacts associated with the proximity to the Atlantic Ocean. Prevailing westerlies and strong cyclonic activity determine a highly variable weather pattern with precipitation dominating over evaporation. Distinct seasonality is characteristic. Air temperature below zero °C and snow accumulation are common in the cold season. Seasonality is also evident in groundwater level patterns in shallow aquifers. Two groundwater level maxima occur, one in spring which is associated with snowmelt water infiltration, and one in autumn – early winter (September-December) induced by increased precipitation and low evapotranspiration (Tolstovs et al., 1986). Usually, groundwater level minima can be observed in late summer and winter (Kalvāns et al., 2020), but a minimum can be absent in mild winters (Lauva et al., 2012).

2.2. Evolution of groundwater level monitoring network in Latvia

The main objective of groundwater monitoring in Latvia is to ensure good quality and sufficient quantity of groundwater resources, which has not changed over the past hundred years. However, specific objectives of the groundwater monitoring have changed over time mainly due to the available funding, existing regulations, and political framework.

The first systematic groundwater observations can be dated back to the end of the 19th century, but observation sites were few and the monitoring initiatives were short-term. The establishment of a systematic national groundwater monitoring network started in 1953 with the first regular observations performed since 1959 (Fig. 1. The initial network in 1959 consisted of 15 observation wells organized into 4 monitoring stations. Wells were mostly installed in unconfined aquifers. The number rapidly expanded to include new well fields around the

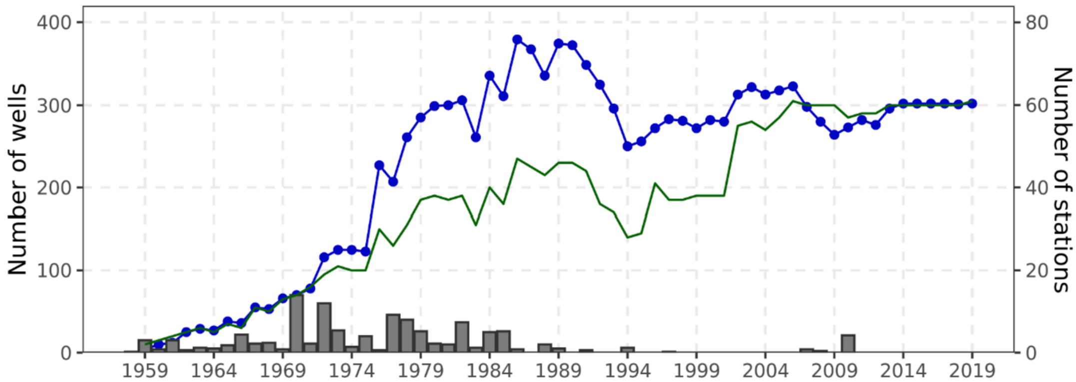

Fig. 2. Total number of active observation wells (blue line with dots, left Y axis) and number of active stations (green line, right Y axis) in corresponding year, and number of newly installed wells (columns, left Y axis) within each year. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

largest cities and to carry out monitoring in vicinities around newly built hydroelectric power plants. In the early 1970s, around 130 new wells were installed to examine waterlogged soil conditions in agricultural lands. In the end of 1975, the groundwater monitoring network had 227 wells grouped into 30 monitoring stations (Jankins et al., 1993; Levina and Levins, 1994).

Since 1976 the national groundwater monitoring network has had two principal branches - regional and local monitoring. Regional networks consisted of transects of monitoring stations each with multiple wells following groundwater flow lines from recharge to discharge areas. Local monitoring networks addressed specific issues at large groundwater abstraction sites; hydroelectric power plants; open pit mains; or heavily contaminated sites. Observation frequency ranged from a few times a year up to 10 times per month. Most groundwater level measurements were made manually (Jankins et al., 1993).

Between 1992 and 1993, after the collapse of the Soviet Union and subsequent decrease in funding, many wells were removed from the groundwater monitoring programs (Jankins et al. 1993). Some monitoring wells were excluded from the monitoring network due to observation well defects (i.e., clogging of the well screen or leaks due to faulty joints). Also, a lack of proper legislation resulted in landowners denying access or even demolishing monitoring wells installed on private lands. Meanwhile, the first digital groundwater database was established and was continuously expanded in the following years by adding observations by the State Geological Survey and its successor - Latvian Environment, Geology and Meteorology Centre (LEGMC).

Since 1999, the monitoring programme has been adapted in line with the EU Water Framework Directive (EU 2000). Most recent and largest establishment of new wells and installation of automatic loggers recording water level twice a day happened from 2010 until 2012. Until now the automatic level measurements are accompanied by 2 to 4 manual observations per year for verification purposes. In 2019, groundwater monitoring was carried out in 301 wells grouped into 60 monitoring stations (see Fig. 2. The database contains observations from altogether 612 wells from 74 stations. It is important to note that observations periods for wells and stations differ, therefore not all wells and stations have been exploited simultaneously.

2.3. The dataset and its repair procedure

A groundwater observation dataset was obtained upon a request from the Latvian Environment, Geology and Meteorology Centre, LEGMC (https://videscentrs.lvgmc.lv/). It included raw groundwater level time series of 612 wells grouped into 74 monitoring stations from 1959 to 2019 as well as coordinates, well depths, screen intervals and represented aquifers. The groundwater level was recorded in meters below soil surface. In case of automatic measurements, barometric pressure and temperature was available.

In total 1.68 million groundwater head records were gathered in the database from which 0.69 million (41%) were manual observations and 0.99 million (59%) were automatic measurements. The frequency of manual observations ranged from several readings per week to few times a year, whereas automatic recordings were performed twice a day. Automatic measurements were accompanied by occasional manual groundwater head observations for verification purposes. As pointed out by Post and Von Asmuth (2013), it is a standard practice to determine possible deviations and provide means to correct errors. According to Kandel et al. (2011) datasets usually contain some proportion of errors and Zaadnoordijk et al. (2019) highlight that assessment of each time series is crucial before being used in further analysis.

The observations in the LEGMC database were recorded without any quality or consistency screening. Thus, a workflow was set up for pre-processing of groundwater level time series (see Fig. 3. Main steps of data processing and repair were defined. First, the data were imported and merged into an SQL database. Then, visual error screening was performed, and any necessary data manipulations were coded into the SQL data table. Lastly, a final repaired dataset was generated automatically using the repair instructions. Corrected groundwater level time series together with the information about applied corrections were returned to the data maintainer LEGMC for further usage in groundwater management.

Potential errors in groundwater level time series were identified by conventional methods, i.e., visually. The four eyes principle (Nihei et al., 2002) was used, thus each time series was reviewed by two experts having distinct roles: Corrector and Controller. The task of the Corrector was visual inspection of the groundwater level time series via a custom-made R-Shiny application in order to identify problems and determine necessary corrections (Fig. 3 – I and II). In case of doubts, the Corrector initiated team discussion in a dedicated online chat to reach an agreement by all members of a group how the case should be treated (Fig. 3 – III). If corrections were necessary, the Corrector introduced repair instructions in the designated SQL table suggesting either to delete an observation or modify it by changing its value using basic mathematical operators. Next, the Controller examined the decisions made by the Corrector (Fig. 3 – IV) and double checked for errors. In a designated shared spreadsheet, the Corrector received comments whether the Controller has approved or declined corrections, or any further action was suggested. Final decision whether to implement or reject the Controller suggestions was made by the Corrector.

As suggested by Rau et al. (2019), the original data (raw groundwater level measurements) were left intact and stored alongside the repaired data, and all applied manipulations were coded in an SQL table and carefully documented in a separate Spreadsheet. It was done to ensure the traceability and repeatability of the work. A chat platform set

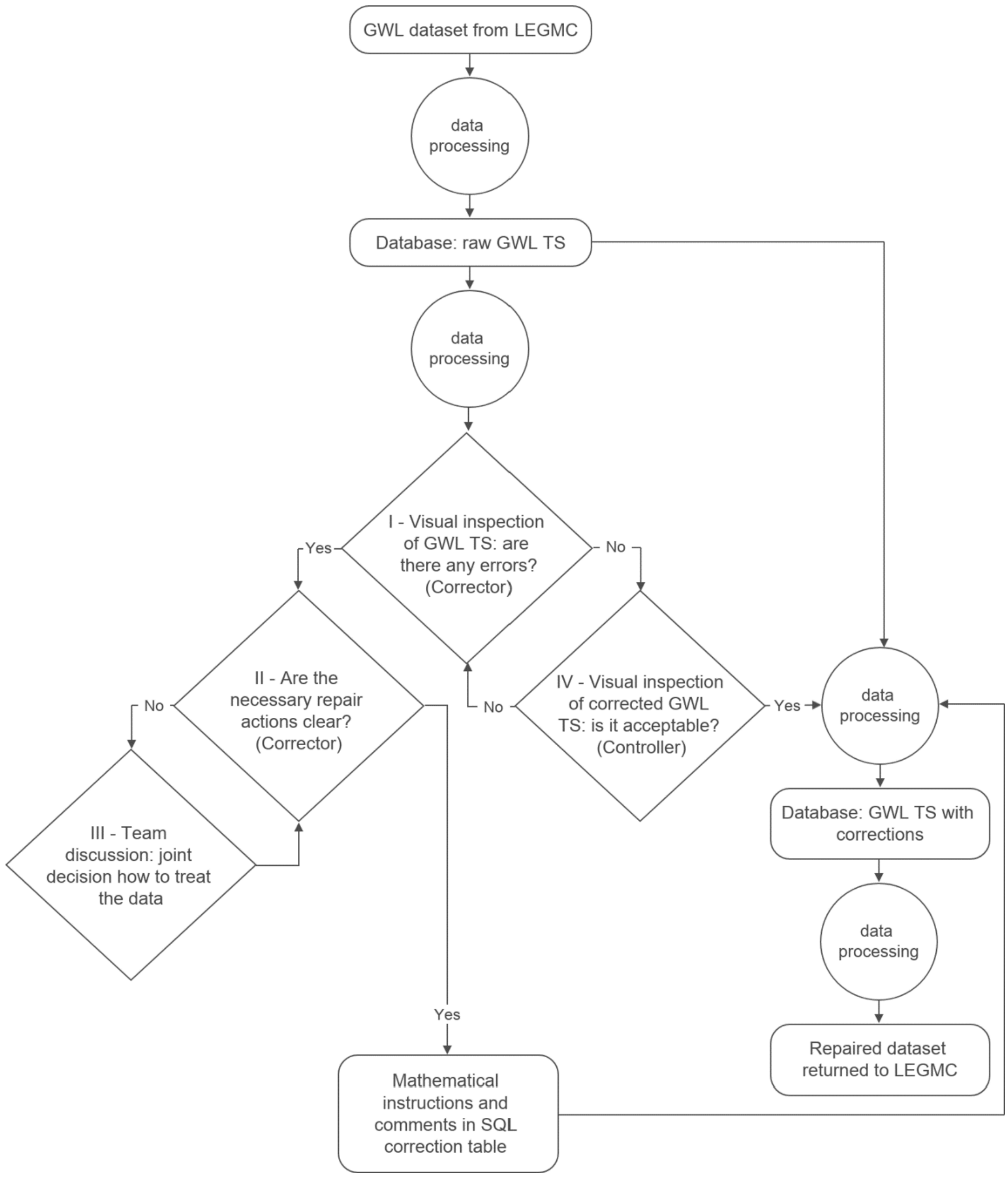

Fig. 3. Workflow for error identification and repair in groundwater level time series (GWL TS): I - visual evaluation of the time series for obvious errors; II - defining necessary corrections or III - team discussion to reach a consensus if and what corrections should be applied; IV - approval of corrections (if any) by the senior team member.

up for team members to share and discuss problematic cases gradually built up and harmonized the collective expertise. Moreover, the accumulated archive was especially useful to train new team members.

2.4. Iterative development of a web-based R-shiny application for visual data assessment

An interactive web-based interface was implemented in R (the R statistical programming language version 3.6.3; R Core Team, 2020) to assist the visual analysis of groundwater level time series. An iterative programming approach was adapted, and new features were added as soon as they were necessary. Open-source Shiny Server version 1.5.12 (Chang et al., 2020) was used to publish the application on a local server. The application incorporated several tools to ease error identification and data repair (Table 1.). Plotly package version 4.9.2 (Sievert, 2020) was used to create interactive figures (Wickham et al., 2019), allowing the user to zoom in for more detailed analysis or to precisely identify an observation of interest using well number and observation timestamp. The code of the developed application has been published on Zenodo (Bikše et al. 2021).

The main window of a web-based application contains inputs to select specific well or change plot parameters and outputs that show general information for the selected well (Fig. 4, a). All tools are described in Table 1 and plotting options can be changed in the upper part of the application (Fig. 4, a). The interactive application is supplemented also by a data table showing original data as a table from database, as well as simple statistics about the number of observations.

The Station wells plot tool will typically be used when the Abrupt change plot tool indicates a sudden change. A sudden rise of the groundwater level can be due to extreme precipitation events (Vidon, 2012) or anthropogenic recharge events, while short extraction events will lower the level (especially of confined groundwater). Such events

Table 1
Descriptions of the tools used in the application for visual data assessment.

Tool	Description
Map	An interactive map created with Leaflet package (Chang et al., 2020) showing all wells and the selected well on the OpenStreetMap. Additional information is shown: well ID, head elevation, depth, screen interval and aquifer (Fig. 4, a).
Repair plot	Simultaneous visualization of both the original and repaired time series in an interactive Plotly plot (see Fig. 4, b). All applied corrections can be represented in the plot as vertical dashed lines.
Station wells plot	An interactive plot that shows the time series of all wells within the selected monitoring station. This plot allows to identify if distinct time series patterns in one well can be observed also in other nearby wells.
Manual vs Automatic scatterplot	A scatterplot of manual versus automatic observations if both data types are present. Mean bias between manual and automatic measurements is shown next to the plot, which is useful to detect a wrong logger reference level.
Abrupt change plot	A plot showing the rate of change between consecutive measurements. Abrupt changes require special attention to determine whether they can be attributed to measurement or data recording errors, or unusual circumstances.

will also affect nearby well in a similar manner (Berendrecht and Van Geer, 2016). The sudden change usually is not found in other wells, in case of a data error. In such a case, there often is a sudden change of the opposite sign later in the time series (e.g., Fig. 4, b).

2.5. The assessment through AR(I)MA models

ARMA and ARIMA models (Box and Jenkins, 1976) were made for the original and repaired time series in order to assess the improvement of applied data rescue actions. The performance of each model was evaluated by Akaike's information criteria (AIC) (Akaike, 1974). The structure of the ARIMA model for each original series and the corresponding repaired time series were the same, so the AIC difference is a proxy for the improvement of the repaired time series.

The Autoregressive (AR) part of the ARIMA explains current value as a linear function of past observation(s) according to order p, while the moving average (MA) part uses white noise (random error) in the past observations (order q) to linearly predict a current value. The integrated (I) part in the order of d removes trends and seasonality to make time series stationary (differencing) (Box and Jenkins, 1976). The integrated part in ARIMA is necessary to deal with non-stationary time series, whereas stationary time series with constant mean level and no trend can be modelled by ARMA models (combined AR and MA parts). A model can be represented as ARIMA(p,d,q) where p,d,q are orders of the AR, I and MA processes.

ARMA(1,1) and ARIMA(1,1,1) models were fitted to both original and repaired time series and AIC values were retrieved from each model/time series combination for evaluation purposes. The Stats package (version 3.6.3) from R (R Core Team, 2020) was used to fit the models and to calculate the AIC calculation. A time step of one day was used and larger time step were filled in with linear interpolation between observations. However, original time series were split in sections when observation gaps longer than 6 months were detected similar to the approach of the Wendt et al. (2020). And at least 8 years long sections as proposed by Zaadnoordijk et al. (2019) were retained for the assessment. Finally, the AIC values retrieved from original and repaired time series models were compared.

The model with the lowest AIC value implies a better model fit to the data. In this study, we compared AIC values derived from the same model type but performed on original and repaired time series. Thus, for a given model type and time series section, two derived AIC values indicated whether a repaired time series resulted in a better fitting model or not. We used a difference between AIC derived from the original and repaired time series and used this ΔAIC value for comparison needs.

As a result, positive ΔAIC values indicate a better model that has been fitted to the repaired time series section and vice versa. We used three categories of ΔAIC values: No change (absolute smaller than the threshold), Improvement (ΔAIC larger than the threshold), and Decrease (ΔAIC smaller than the negative value of the threshold). The 25th percentile of all absolute ΔAIC values was used as the threshold.

3. Results and discussion

3.1. Data treatment according to the cause of errors

The main problems identified in the groundwater level time series were grouped according to their potential cause: errors in measurement and data recording (Table 2; technical problems at the observation site (Table 3; local anthropogenic impact and other unclassified problems (Table 4. The errors are supplemented by illustrative examples from visual analysis (Figs. 5–7. The proposed data treatment for the identified problem categories was based on an extensive summary of the applied repair actions by Correctors. Additionally, confidence levels were added to indicate reliability of expert judgement and applied decisions.

Distinct errors caused by data entry and actual measurements (group 1) were relatively easy to identify (with high confidence) by visual inspection of the time series (Kandel et al., 2011). In case of few errors, the false data points were removed (Fig. 5, a). However, if there were longer time periods with automatic measurements that deviated from previous and following data (and manual measurements if available) by a constant, then these were shifted to fit into the whole time series (Fig. 5, b and d). These deviated time periods were likely bound to the misplaced level loggers after well sampling. Rau et al. (2019) emphasize that frequent removal of loggers (e.g., for data download or water sampling) may cause the wire length to change due to kinks. Also, the logger may not always be returned to the same position. Accompanying manual measurements were mostly assumed to be the correct ones, thus serving as a reference point to shift the mismatching observations (Fig. 5, b and d).

The typical pattern of the automatic measurements was considered when longer periods of manual and automatic measurements did not coincide (Fig. 5, c). Then the automatic measurements were assumed as the correct ones and erroneous manual measurements were deleted. In case of a constant offset between the manual and automatic measurements (Fig. 5, e), the level of the manual measurements was preferred, and the automatic measurements were shifted. When a time series of only manual measurements consisted of seemingly two separate sets of measurements, either one set was selected based on expert judgement or data were left intact for cases when deviations were small (Fig. 5, f). However, uncertainty prevails over the two previously described correction actions. Availability of both automatic and manual measurements in parallel usually accounted for high confidence of the applied corrections as it was clear which records and had to be moved according to Post and von Asmuth (2013). While Rau et al. (2019) suggest performing regular checkups of the performance and adjustment of automatic level loggers at least once in 3 months, the frequency of verification in the Latvian groundwater level database was not regular and ranged from 2 to 4 times a year. The confidence level of decisions decreased when there was no reference indicating which of the observation group is the correct one (Fig. 5, e and f). The site visit with control measurement could resolve such issues if the time series are continued until present.

The second group of errors consists of technical problems with automatic data loggers and piezometers. As stated by Rau et al. (2019), sensor drift is one of the most common errors in automatic level measurements. A pattern of continuous drift of automatic logger data was identified (Fig. 6, a) in a few unrelated wells. We concluded that this

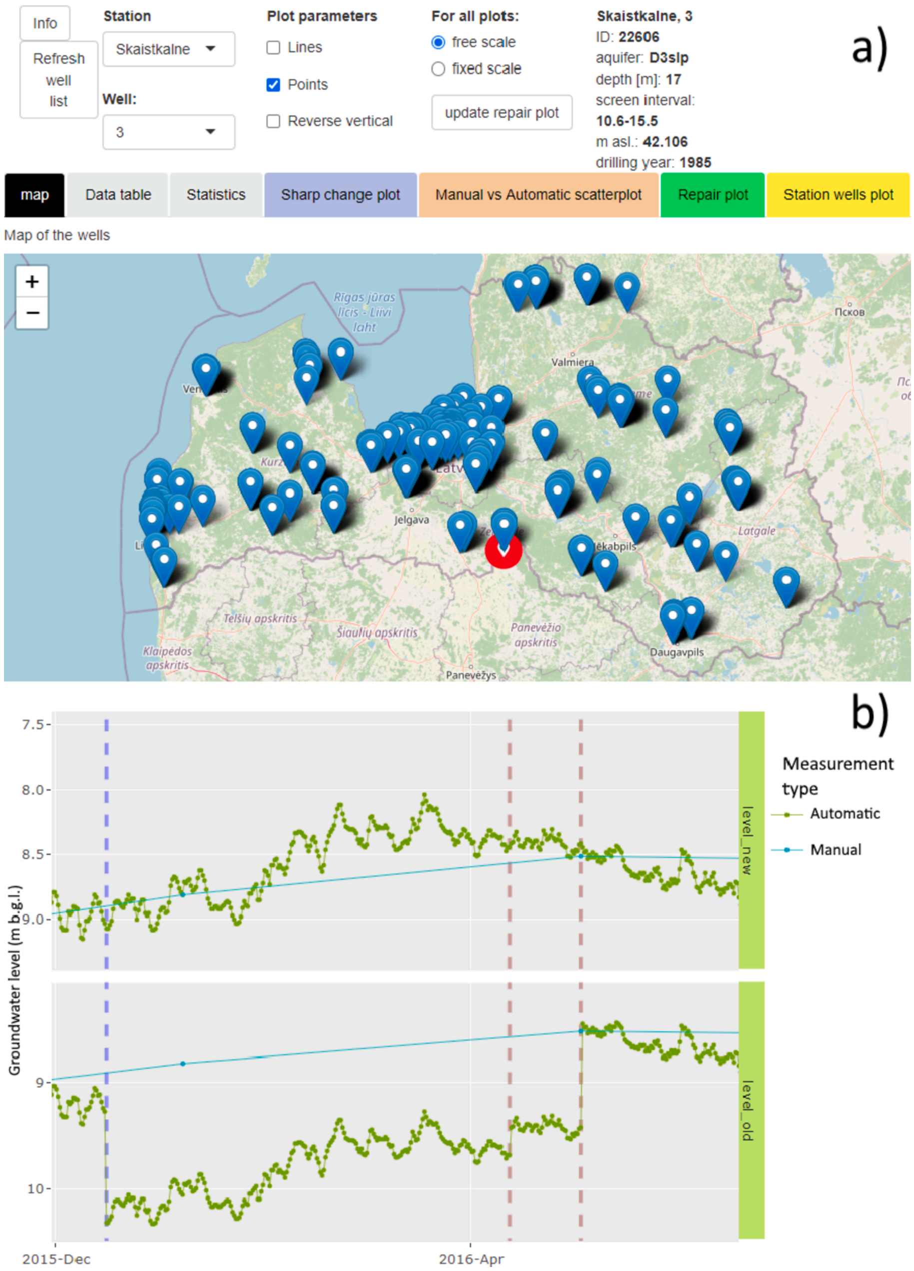

Fig. 4. Interactive visual tool to aid groundwater level correction process comprising (a) interactive map and complementary information, (b) the main repair tool (vertical dashed lines indicate sections where corrections were made).

pattern illustrated malfunction of the automatic level loggers, therefore these drifting measurements were removed.

Errors due to the inadequately installed wells and misplaced measurement equipment also belong to this group. For example, seasonally elevated groundwater levels exceeding the top of the well heads or freezing of the water in the well head could prohibit taking correct measurement by the tape and resulted in a plateau-like pattern of peak groundwater table (Fig. 6, b). Groundwater level can rise above the top of the well head also due to the long-term groundwater recovery from intensive aquifer exploitation, which was the case presented in Fig. 6, c. The well is located near the coastal city Liepāja - an area that historically has been affected by extensive groundwater pumping but currently the levels have recovered (Bikše and Retike, 2018), thus stressing the importance of regular field site maintenance. In such cases automatic level loggers recorded plateau-like measurements that did not indicate true water level (Rau et al., 2019). Similar negative plateau-like groundwater level patterns (Fig. 6, e) can be a result of a shallow well screen. During the dry season, the groundwater level drops below the screen interval, while the water leftovers below the screen is recorded. Usually, the lowest recorded water level coincides with screen bottom. When the water table dropped below the depth of the sensor installed in the well (Fig. 6, d), a plateau-like pattern was identified. Such observation periods were excluded from groundwater level time series. Ha et al. (2021) have observed a similar plateau-like pattern caused by water abstraction induced groundwater level drawdown.

The zoom in functionality of the developed application allowed to spot even minor impacts of freezing and thawing of barometric loggers which later created bias in groundwater level measurements (Fig. 6, f). According to Ritzema et al (2018) the difference between day and night temperatures can account for the level deviations of several centimeters. While Liu and Higgins (2015) emphasized that sensors should be protected from temperatures below freezing point to avoid errors and

Table 2
Identified problems caused by data recordings and measurements and proposed treatment.

Problem	Description of problem and possible cause (representative visual example)	Proposed repair action	Confidence level of identification/ repair process
Distinct errors	One or several data points significantly outside the data range (Fig. 5, a).	Delete the error.	High - medium/ high
Shift in water level	Sudden, sharp level changes for a certain time period due to automatic level logger displacement (Fig. 5, b).	Mathematically adjust (shift) the outstanding data portion to correspond to the adjacent datasets.	high/high
Mismatch between manual and automatic measurements (when both are present)	Manual observation that does not fit into the overall time series (Fig. 5, c). Sharp shift of automatic measurements that does not correspond to the manual observations due to the misplaced level logger (Fig. 5, d).	Delete a single data error. Identify the shift and align the automatic measurements to the manual observations.	high/high
Mismatch between all manual and all automatic measurements	Different reference levels for manual/ control and automatic logger data (Fig. 5, e).	Mathematically adjust (shift) the false automatic measurements to match the manual observations.	high/medium to high
Jagged/toothed level pattern	Levels continuously change from high to low. Possible reason might be two different observers who are inconsistent in the measurement reference or one having an erroneous measuring device (Fig. 5, f).	Delete higher or lower records or ignore the problem in case none of the "tooths" could be assumed as the correct one.	high/low

logger breakdown. Ensure the pressure transducer is protected from temperatures below the freezing point. In general identification of technical problems remained straightforward because of their specific patterns.

The third group consists of measurements which are not representative of the groundwater head for the purpose of monitoring natural groundwater head fluctuations. Influence of direct pumping from the monitoring well (e.g., sampling for water quality) was easy to spot (Fig. 7, a) and to eliminate as the starting timestamp of such an event matched the time of manual control observation. Usually manual control measurements, download of level logger readings and water sampling were made at the same time due to the cost efficiency. Similar groundwater level fluctuations due to well pumping are presented in Ha et al. (2021).

A group of other anthropogenic pressures responsible for misleading records were identified with the help of the developed application and its functionality to show nearby objects on a map. Pumping (Fig. 7, b) or recharge (Fig. 7, c) effects in the nearby wells or vicinity often had a

Table 3
Identified data issues caused by malfunction of loggers or observation well defects and proposed treatment.

Problem	Description of problem and possible cause (representative visual example)	Proposed repair action	Confidence level of identification/ repair process
Malfunction of automatic level logger	Continuous drift of logger data, deviating from control/manual measurements and from previous data (Fig. 6, a).	Identify the start of a drift and delete the subsequent data.	high/high
Well completion problems	Well head is too short, leading to seasonally overflowing wells or freezing of water within the well head (Fig. 6, b). Long-term water level recovery leading to overflowing well (Fig. 6, c). Well is too shallow, leading to seasonal drying up or data logger in the air (Fig. 6, d and e).	Delete the high- or low-level plateau.	high/medium to high
Malfunction of barometric pressure loggers at freezing temperature	Freezing and thawing of barometric loggers in the cold season that creates false atmospheric pressure readings and subsequently causes noise in groundwater level records (Fig. 6, f).	Delete one of the measurements or ignore as the daily average might compensate for the bias.	low/low to medium

particular pattern such as peaks on Mondays or Sundays, or during certain seasons. Chen et al. (2004) observed alike patterns caused by seasonal groundwater pumping in the study area. Similarly, Ha et al. (2021) reported notable groundwater level drops in summer months because of intensive groundwater abstraction for irrigation. While Rau et al. (2019) identified similar sharp groundwater level responses in the observation well caused by nearby pumping activities which are frequent but irregular and stressed that the typical twice a day measurement interval fails to capture such short-term variations.

Visual data assessment tools (Table 1, Fig. 4 eased the identification of anthropogenic influences and facilitated the decision-making by fast and simple evaluation process of supplementary data and level changes in nearby wells. Likewise, Wendt et al. (2020) removed unrealistic observations which after verification with metadata were not explained. Our aim for the data quality control and correction was to rescue groundwater level time series for further modeling of groundwater drought events. However, the decision whether to eliminate such data or apply corrections strongly depends on the future usage - study of near-natural conditions or human induced changes.

The fourth group contains other errors, such as sudden changes in groundwater patterns (Fig. 7, d, e and f) which most likely were associated with data processing errors such as mixed ID numbers between wells or false measurements. However, extreme changes in groundwater level patterns have been also reported by Barthel et al. (2021) as a result of dam construction where the influence lessened in the observation wells further from reservoir. To distinguish between errors or untypical still true groundwater patterns two tools in the developed application for

Table 4
Identified problems caused by anthropogenic or other unclassified impact and proposed treatment.

Problem	Description of problem and possible cause (representative visual example)	Proposed repair action	Confidence level of identification/ repair process
Effects of well pumping/ sampling	Sharp water level dropping due to well pumping, followed by level recovery (Fig. 7, a).	Delete the data associated with the well pumping or ignore.	high/high
Influence of nearby pumping/ recharge	Regular drop of water level, followed by fast recovery (Fig. 7, b) or regular recharge events of anthropogenic origin followed by quick dropdown (Fig. 7, c).	Delete the data associated with pumping/ recharge events or ignore.	high/high
Change in groundwater level pattern	Sudden change in the groundwater level pattern, usually accompanied with the shift in water level (Fig. 7, d). Could be due to mixing well IDs during registration of data.	Delete the data records with differing patterns or shift to the main time series level.	medium to high/ low
Sharp change in water level in all monitoring station wells	The same sharp changes observed in nearby wells likely due to personnel turnover or equipment change (Fig. 7, e).	Shift one part of data records to correspond to the other.	high/low
Diurnal groundwater level fluctuations	Groundwater head fluctuations in shallow wells due to thermal effects associated with recorded barometric pressure or diurnal transpiration cycle (Fig. 7, f).	Ignore as the range of head fluctuations is few centimeters and may represent daily fluctuations.	low/low to medium

visual data assessment were particularly useful. First, the map tool that allowed to see the location of well and nearby objects (such as dams) and second, the station well plots that allowed to see if similar changes can be observed in nearby wells. Identification of such patterns was less straightforward and applied corrections involved more expert judgment.

The decision whether to delete suspicious records or apply repair action depended on the number of suspicious records and the importance of the time series (spatial representativity of monitoring point and length of dataset).

In total the whole data correction process took approximately 325 man-hours including 280 man-hours for the Correctors (data correction) and 45 man-hours for the Controllers (Fig. 3.

3.2. Repair outcomes

In total, 612 groundwater level time series in the Latvian national database were assessed according to the proposed procedure (Fig. 3 and to 536 or 88% of the series corrections or deletions in line with the proposed repair actions (Table 2- 4 were applied. In 196 or 32% of all time series more than ten percent of the initial groundwater level observations were modified. To compare, using conventional approach (exclusion of all time series with at least one identified error), only 76 from initial 612 time series could be retained. Moreover, the retained 76 time series would cover a time period of 9832 months, while our approach retained 605 time series covering 150124 months. For the automatic measurements repair actions modified the groundwater levels from −0.10 to 0.42 m (first and third quartile), while for the manual observations from −0.66 to 0.17 m (first and third quartile). In extreme cases, introduced repair actions modified groundwater level for more than a few meters. In total 5.3% of 0.99 million automatic measurements were deleted while 29% were corrected. And 3.9% of the 0.69 million manual measurements were deleted whereas 2.6% were corrected (Fig. 8.

A larger proportion of the applied corrections and deletions was associated with the most recent observations and especially, with automatic measurements. The frequency of manual observations was much lower, thus distinct outliers or data processing errors accounted for most errors. Automatic level measurements usually did not have isolated errors (outliers). If a problem occurred (e.g., misplaced, or broken logger) the logger continued to record false observations until the next visit by the operator (usually 2–4 times a year). In case of a broken logger, the problem was identified only after data screening at the office, usually once a year.

It is assumed that errors in automatic level measurements observed soon after their installation are the result of misplaced loggers, while in recent years the influence of equipment aging, particularly, malfunctioning of the automatic groundwater level loggers can be observed as a continuous increase of the applied corrections. Von Asmuth (2010) points out that in several groundwater monitoring networks up to half of the sensors should be replaced because of malfunctioning over time, and

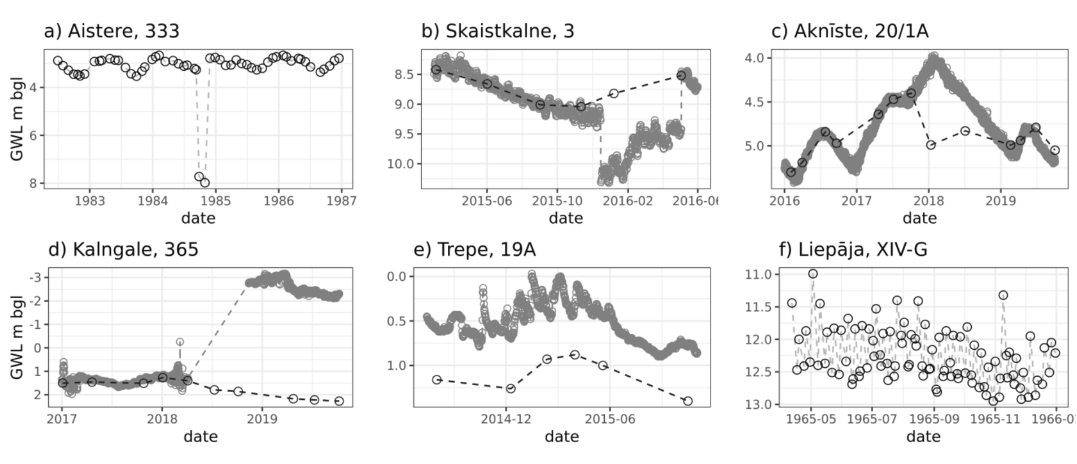

Fig. 5. Representative examples of groundwater level (GWL) time series where problems caused by data recordings and measurements were identified (title represent station name and well number).

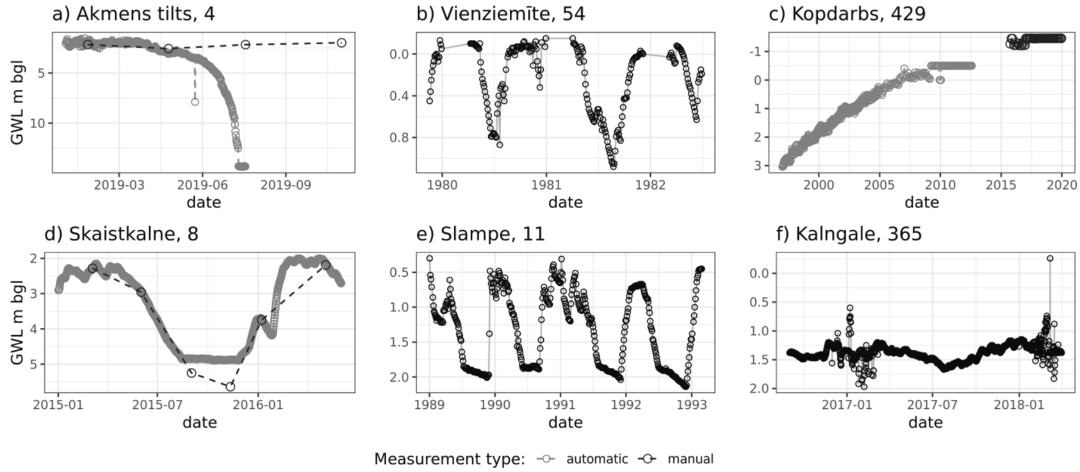

Fig. 6. Representative examples of groundwater level (GWL) time series where data errors were caused by malfunction of automatic level or barometric pressure loggers, or observation well defects (title represent station name and well number).

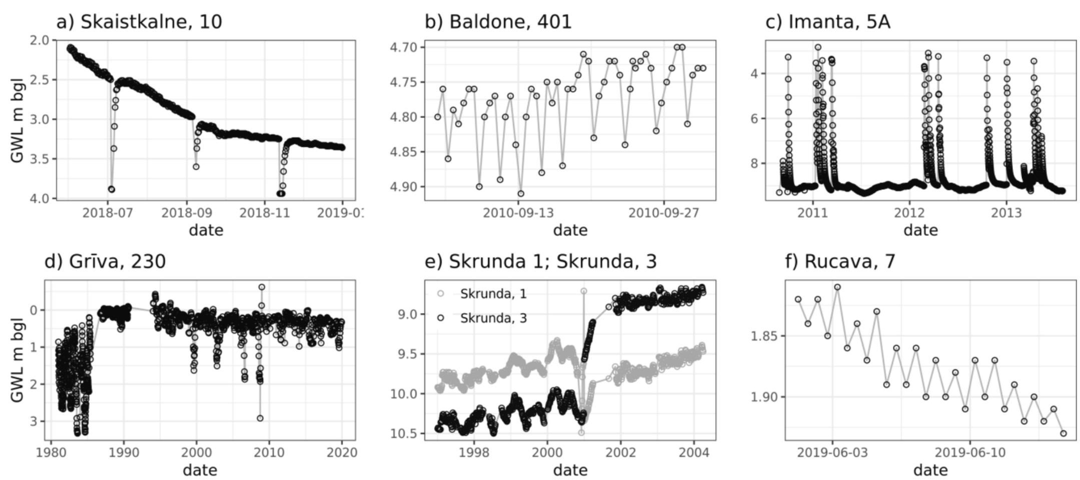

Fig. 7. Representative examples of identified data problems in groundwater level (GWL) time series caused by anthropogenic impact and other unclassified problems (title represent station name and well number).

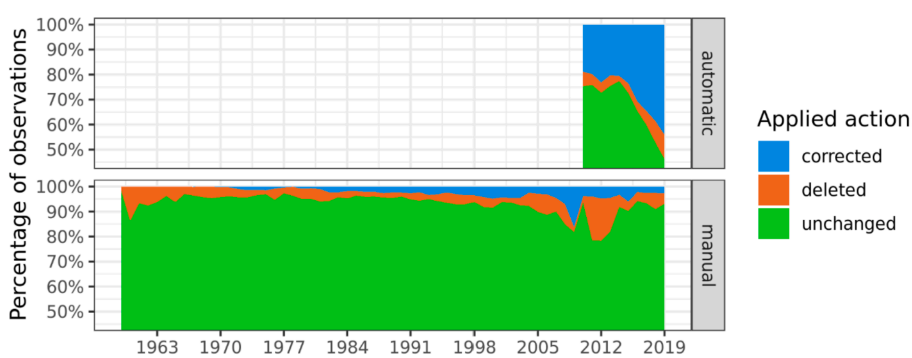

Fig. 8. The fraction of repaired time series for automatic and manual measurements per year.

similarly this study highlights the need for replacement of most currently operational automatic level loggers in the Latvian groundwater monitoring network. In addition, the repair procedure itself contributes to the increase because the historical or earlier measurements were frequently considered to be the correct ones. Thus, the recent observations were adjusted to the historical measurements and resulted in more repairs in the recent observations.

The groundwater monitoring in Latvia was strongly influenced from 2009 until 2013 by reduced funding and negative effects of the global economic crisis. It is suspected that turn-over of employees of the institutions responsible for groundwater monitoring, insufficient funding for observation site maintenance, poor training of new employees, and lack of detailed monitoring guidelines, have resulted in deterioration of data quality, especially for the manual observations.

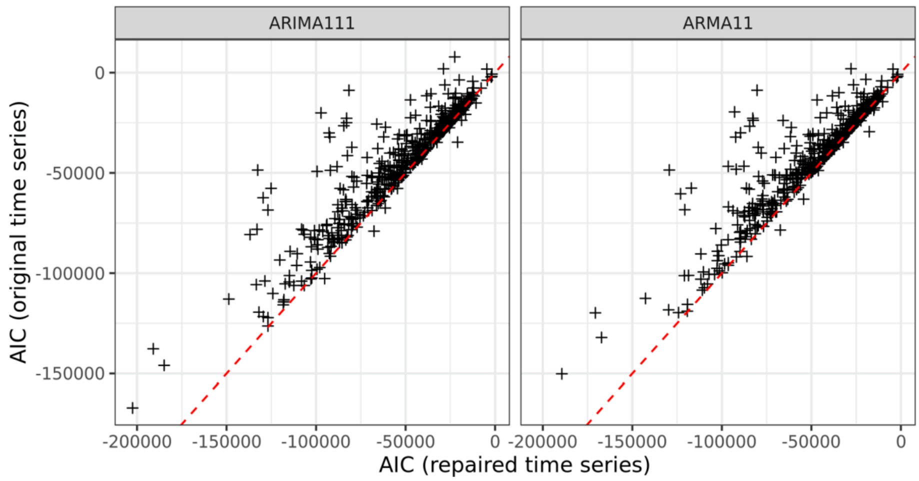

Fig. 9. Comparison of AIC score for original and repaired observation series (only non-discarded sections), ARMA(1,1) and ARIMA(1,1,1) models.

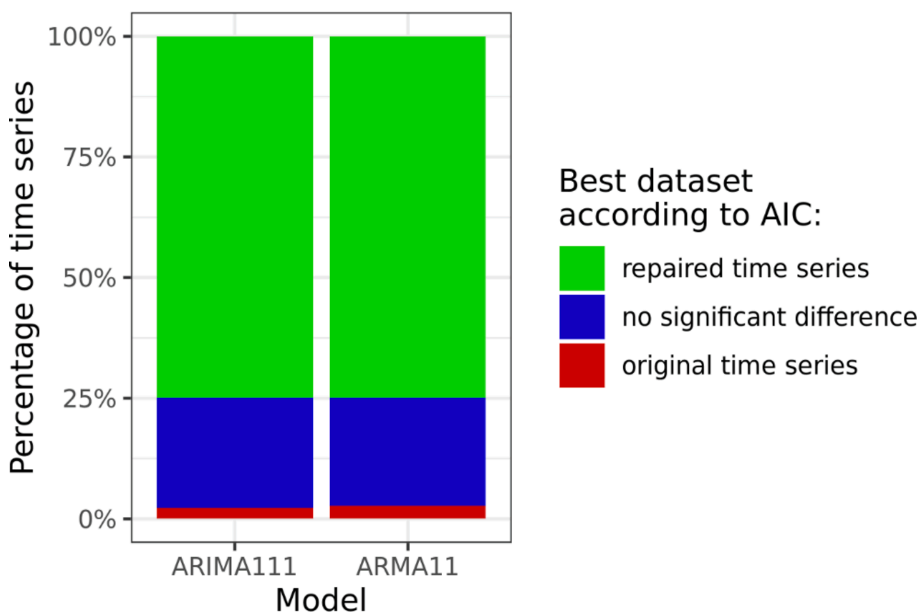

Fig. 10. Evaluation of original versus repaired time series sections (non-discarded only).

3.3. *Quantification of data quality improvement*

Out of the 612 groundwater level time series in the original dataset, 605 time series had at least a single observation left after data was repaired and were used to fit AR(I)MA models. Time series were separated in sections if gaps larger than 6 months were detected. This resulted in 1377 time series sections that were modelled individually. Only 590 individual time series sections from 494 unique wells were longer than 8 years, while the rest of section were discarded from further modelling. For these sections, ARMA(1,1) and ARIMA(1,1,1) models were fitted to both the original and repaired time series. ARIMA(1,1,1) and ARMA(1,1) models could be fitted to 523 and 483 sections respectively. For the rest of the 590 sections the model failed to fit to either the original or the repaired series (or both). Generally, the models fitted using repaired time series were better models than those fitted on the original time series for the same time series sections (Fig. 9.

A majority of assessed time series shows significantly better AIC if repaired time series are used to fit models instead of original ones while 22.8% and 22.4% time series shows insignificant changes in AIC for the ARIMA and ARMA models, respectively (Fig. 10. The AIC was worse after the repair for 2.29% (ARIMA) and 2.69% (ARMA) of the time series.

4. Conclusions

Long and continuous groundwater level time series are of great value, but they usually contain errors, which should be corrected prior to any further application. We propose a data rescue approach which was applied to the Latvian national groundwater level database containing 612 wells comprising 1.68 million groundwater level observations since 1959. We developed web-based interactive tools for visual assessment of time series and manual correction of errors. Implementation of the four-eye principle and documentation of all applied manipulations separately from the raw data ensured traceability and repeatability of the work. Errors were attributed to possible causes and the confidence level of the error repair actions was assigned: high, medium, or low. The identified errors were grouped into: errors caused by data recordings and measurements; technical problems at the observation site; local anthropogenic impact, and other unclassified problems. The Akaike information criterion derived from fitted ARMA and ARIMA models to both original and repaired time series demonstrated substantial improvement of consistency of most time series after applying proposed data rescue approach. The presented approach and results can be of great value for all studies using groundwater level time series as an input.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is funded by the Latvian Council of Science, project "Spatial and temporal prediction of groundwater drought with mixed models for multilayer sedimentary basin under climate change", project No. lzp-2019/1-0165, Ministry of Education and Science, Latvia and University of Latvia grant No. ZD2016/AZ03 "Climate change and sustainable utilization of natural resources", ERDF projects "Groundwater and soil water regime under climate change" (No.1.1.1.2/VIAA/3/19/524) and EU ESF Project No.8.2.2.0/18/A/010 "Academic Staff Renewal and Continuing Proffessional Development at the University of Latvia". The contribution of W.J.Z. was supported by the TACTIC project, part of the GeoERA programme that has received funding by the European Union's Horizon 2020 research and innovation programme under grant agreement number 731166. The authors would like to thank Latvian Environment, Geology and Meteorology Centre (LEGMC) for providing the essential raw groundwater level datasets for this research.

References

Adamowski, J., Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. J. Hydrol. 407 (1–4), 28–40. https://doi.org/10.1016/j.jhydrol.2011.06.013.

Ahn, H., 2000. Modeling of groundwater heads based on second-order difference time series models. J. Hydrol. 234 (1–2), 82–94. https://doi.org/10.1016/S0022-1694(00)00242-0.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19 (6), 716–723. https://doi.org/10.1109/TAC.1974.1100705.

Ali, M., Alqahtani, A., Jones, M.W., Xie, X., 2019. Clustering and Classification for Time Series Data in Visual Analytics: A Survey. IEEE Access 7 (8930535), 181314–181338. https://doi.org/10.1109/ACCESS.2019.2958551.

Asgharinia, S., Petroselli, A., 2020. A comparison of statistical methods for evaluating missing data of monitoring wells in the Kazeroun Plain, Fars Province. Iran. Groundwater for Sustainable Development 10, 100294. https://doi.org/10.1016/j.gsd.2019.100294.

Avotniece, Z., Aniskevich, S., Briede, A., Klavins, M., 2017. Long-term changes in the frequency and intensity of thunderstorms in Latvia. Boreal Environ. Res. 22, 415–430.

Babre, A., Kalvāns, A., Popovs, K., Retiķe, I., Dēliņa, A., Vaikmäe, R., Martma, T., 2016. Pleistocene age paleo-groundwater inferred from water-stable isotope values in the central part of the Baltic Artesian Basin. Isot. Environ. Health Stud. 52 (6), 706–725. https://doi.org/10.1080/10256016.2016.1168411.

Bakker, M., Schaars, F., 2019. Solving Groundwater Flow Problems with Time Series Analysis: You May Not Even Need Another Model. Groundwater 57 (6), 826–833. https://doi.org/10.1111/gwat.12927.

Barthel, R., Haaf, E., Giese, M., Nygren, M., Heudorfer, B., Stahl, K., 2021. Similarity-based approaches in hydrogeology: proposal of a new concept for data-scarce groundwater resource characterization and prediction. Hydrogeol. J. 29, 1693–1709. https://doi.org/10.1007/s10040-021-02358-4.

Berendrecht, W.L., van Geer, F.C., 2016. A dynamic factor modeling framework for analyzing multiple groundwater head series simultaneously. J. Hydrol. 536, 50–60. https://doi.org/10.1016/j.jhydrol.2016.02.028.

Bernard, J., Hutter, M., Reinemuth, H., Pfeifer, H., Bors, C., Kohlhammer, J., 2019. Visual-interactive preprocessing of multivariate time series data. Comput. Graphics Forum 38 (3), 401–412. https://doi.org/10.1111/cgf.13698.

Bikše, J., Kalvāns, A., Popovs, K., 2021. Application for visual inspection of groundwater level time series. Zenodo. https://doi.org/10.5281/zenodo.5573097.

Bikše, J., Retike, I., 2018. An Approach to Delineate Groundwater Bodies at Risk: Seawater Intrusion in Liepāja (Latvia). E3S Web of Conferences 54, 00003. https://doi.org/10.1051/e3sconf/20185400003.

Box, E.P., Jenkins, G.M., 1976. Time Series Analysis: Forecasting and Control. Holden-Day, Oakland, California.

Brangulis A. J., Kaņevs S. (2002) Latvijas tektonika [Tectonics of Latvia]. Valsts ģeoloģijas dienests, Riga, 50 pp. [in Latvian].

Chang W., Cheng J., Allaire J., Xie Y., McPherson J. (2020) Shiny: Web Application Framework for R. R package version 1.5.0. https://CRAN.R-project.org/package=shiny.

Chen, Z., Grasby, S.E., Osadetz, K.G., 2004. Relation between climate variability and groundwater levels in the upper carbonate aquifer, southern Manitoba, Canada 290, 43–62. https://doi.org/10.1016/j.jhydrol.2003.11.029.

EEA (2018) European Waters Assessment of Status and Pressure 2018. EEA Report, 7/2018, vol. 2018, European Environmental Agency, Publications Office of the European Union, Luxembourg. Available: https://www.eea.europa.eu/publications/state-of-water.

EU (2000) Water framework directive, establishing a framework for Community action in the field of water policy, Directive 2000/60/EC.

Gibrilla, A., Anornu, G., Adomako, D., 2018. Trend analysis and ARIMA modelling of recent groundwater levels in the White Volta River basin of Ghana. Groundwater Sustainable Dev. 6, 150–163. https://doi.org/10.1016/j.gsd.2017.12.006.

Gschwandtner T., Erhart O. (2018) Know Your Enemy: Identifying Quality Problems of Time Series Data. IEEE Pacific Visualization Symposium, 2018-April, pp. 205-214. https://doi.org/10.1109/PacificVis.2018.00034.

Ha, K., Lee, E., An, H., Kim, S., Park, C., Kim, G.-B., Ko, K.-S., 2021. Evaluation of Seasonal Groundwater Quality Changes Associated with Groundwater Pumping and Level Fluctuations in an Agricultural Area. Korea. Water 13 (1), 51. https://doi.org/10.3390/w13010051.

Haaf, E., Barthel, R., 2018. An inter-comparison of similarity-based methods for organisation and classification of groundwater hydrographs. J. Hydrol. 559, 222–237. https://doi.org/10.1016/j.jhydrol.2018.02.035.

Heudorfer, B., Haaf, E., Stahl, K., Barthel, R., 2019. Index-Based Characterization and Quantification of Groundwater Dynamics 55 (7), 5575–5592. https://doi.org/10.1029/2018wr024418.

IGRAC (2020) Groundwater monitoring programmes: A global overview of quantitative groundwater monitoring networks.142 pp. Available: https://www.un-igrac.org/sites/default/files/resources/files/National%20groundwater%20monitoring%20programmes%20-%20A%20global%20overview%20of%20quantitative%20groundwater%20monitoring%20networks_0.pdf.

Jankins J., Levina N., Levins I., Prols J., Straume J. (1993) Pazemes ūdeņu monitorings Latvijā (Groundwater monitoring in Latvia, in Latvian), Riga.

Jarsjö, J., Andersson-Sköld, Y., Fröberg, M., Pietroń, J., Borgström, R., Löv, Å., Kleja, D. B., 2020. Projecting impacts of climate change on metal mobilization at contaminated sites: Controls by the groundwater level. Sci. Total Environ. 712, 135560 https://doi.org/10.1016/j.scitotenv.2019.135560.

Jodkazis V. (1989) (Йодказис В.) Региональная гидрогеология Прибалтики (Regional hydrogeology of the Baltic region, in Russian), Vilnius.

Kalm, V., Gorlach, A., 2014. Impact of bedrock surface topography on spatial distribution of Quaternary sediments and on the flow pattern of late Weichselian glaciers on the East European Craton (Russian Plain). Geomorphology 207, 1–9. https://doi.org/10.1016/j.geomorph.2013.10.022.

Kalvāns, A., Dēliņa, A., Babre, A., Popovs, K., 2020. An insight into water stable isotope signatures in temperate catchment. J. Hydrol. 582, 124442 https://doi.org/10.1016/j.jhydrol.2019.124442.

Kandel, S., Heer, J., Plaisant, C., Kennedy, J., Van Ham, F., Riche, N.H., Weaver, C., Lee, B., Brodbeck, D., Buono, P., 2011. Research directions in data wrangling: Visualizations and transformations for usable and credible data. Information Visualization 10 (4), 271–288. https://doi.org/10.1177/1473871611415994.

Lauva, D., Grinfelde, I., Veinbergs, A., Abramenko, K., Vircavs, V., Dimanta, Z., Vitola, I., 2012. The impact of climate change on the annual variation of shallow groundwater levels in Latvia. Environmental and Climate Technologies 8 (1), 41–48. https://doi.org/10.2478/v10145-012-0007-9.

Lehr C., Lischeid G. (2020) Efficient screening of groundwater head monitoring data for anthropogenic effects and measurement errors. Hydrology and Earth System Sciences, 24, pp. 501-513. https://doi.org/10.5194/hess-24-501-2020.

Levina N., Levins I. (1994) 1993.gada pazemes ūdeņu bilance un režīms (Groundwater balance and regime in 1993, in Latvian). Riga, National geological archive No.11097.

Levins, I., Levina, N., Gavena, I., 1998. Latvijas pazemes ūdeņu resursi (Latvian groundwater resources, in Latvian). State Geological Survey, Riga, Latvia.

Liu, S., Andrienko, G., Wu, Y., Cao, N., Jiang, L., Shi, C., Wang, Y.-S., Hong, S., 2018. Steering data quality with visual analytics: The complexity challenge. Visual Informatics 2 (4), 191–197. https://doi.org/10.1016/j.visinf.2018.12.001.

Liu, Z., Higgins, C.W., 2015. Does temperature affect the accuracy of vented pressure transducer in fine-scale water level measurement? Geosci. Instrum. Methods Data Syst. 4 (1), 65–73. https://doi.org/10.5194/gi-4-65-2015.

Lukševičs, E., Stinkulis, Ģ., Mūrnieks, A.K., 2012. Popovs **Geological evolution of the Baltic Artesian Basin** Highlights of groundwater research in the Baltic Artesian. Basin 7–52.

Marandi A., Karro E., Raidla V., Vaikmae R. (2012) Conceptual model of groundwater quality for the monitoring and management of the Voronka groundwater body, Estonia. Estonian Journal of Earth Sciences, 61(4), pp. 328-339. https://doi.org/10.3176/earth.2012.4.11.

Mirzavand M., Ghazavi R. (2015) A Stochastic Modelling Technique for Groundwater Level Forecasting in an Arid Environment Using Time Series Methods. Water Resources Management, 29(4), pp. 1315-1328. 10.1007/s11269-014-0875-9.

Naranjo-Fernandez, N., Guardiola-Albert, C., Aguilera, H., Serrano-Hidalgo, C., Montero-Gonzalez, E., 2020. Clustering Groundwater Level Time Series of the Exploited Almonte-Marismas Aquifer in Southwest Spain. Water 12 (4), 1063. https://doi.org/10.3390/w12041063.

Nihei, Y., Terashima, M., Suzuki, I., Morikawa, S., 2002. Why are four eyes better than two? Effects of collaboration on the detection of errors in proofreading. Jpn. Psychol. Res. 44 (3), 173–179. https://doi.org/10.1111/1468-5884.00020.

Noorduijn, S.L., Cook, P.G., Simmons, C.T., Richardson, S.B., 2019. Protecting groundwater levels and ecosystems with simple management approaches. Hydrogeol. J. 27 (1), 225–237. https://doi.org/10.1007/s10040-018-1849-4.

Obergfell, C., Bakker, M., Maas, K., 2019. Identification and Explanation of a Change in the Groundwater Regime using Time Series Analysis. Groundwater 57 (6), 886–894. https://doi.org/10.1111/gwat.12891.

Oikonomou, P.D., Alzraiee, A.H., Karavitis, C.A., Waskom, R.M., 2018. A novel framework for filling data gaps in groundwater level observations. Adv. Water Resour. 119, 111–124. https://doi.org/10.1016/j.advwatres.2018.06.008.

Patle, G.T., Singh, D.K., Sarangi, A., Rai, A., Khanna, M., Sahoo, R.N., 2015. Time series analysis of groundwater levels and projection of future trend. J. Geol. Soc. India 85 (2), 232–242. https://doi.org/10.1007/s12594-015-0209-4.

Post, V.E.A., Von Asmuth, J.R., 2013. Hydraulic head measurements-New technologies, classic pitfalls. Hydrogeol. J. 21 (4), 737–750. https://doi.org/10.1007/s10040-013-0969-0.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/.

Rau, G.C., Post, V.E.A., Shanafield, M., Krekeler, T., Banks, E.W., Blum, P., 2019. Error in hydraulic head and gradient time-series measurements: A quantitative appraisal. Hydrol. Earth Syst. Sci. 23 (9), 3603–3629. https://doi.org/10.5194/hess-23-3603-2019.

Ritzema, H.P., Heuvelink, G.B.M., Heinen, M., Bogaart, P.W., van der Bolt, F.J.E., Hack-ten Broeke, M.J.D., Hoogland, T., Knotters, M., Massop, H.T.L., Vroon, H.R.J., van den Bosch, H., 2018. Review of the methodologies used to derive groundwater characteristics for a specific area in The Netherlands. Geoderma 14, e00182. https://doi.org/10.1016/j.geodrs.2018.e00182.

Retike, I., Kalvans, A., Popovs, K., Bikse, J., Babre, A., Delina, A., 2016. Geochemical classification of groundwater using Multivariate statistical analysis in Latvia. Hydrol. Res. 47 (4), 799–813. https://doi.org/10.2166/nh.2016.020.

Shirmohammadi, B., Vafakhah, M., Moosavi, V., Moghaddamnia, A., 2013. Application of Several Data-Driven Techniques for Predicting Groundwater Level. Water Resour. Manage. 27 (2), 419–432. https://doi.org/10.1007/s11269-012-0194-y.

Sievert, C., 2020. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC, Florida https://plotly-r.com.

Stoll, S., Hendricks Franssen, H.J., Barthel, R., Kinzelbach, W., 2011. What can we learn from long-term groundwater data to improve climate change impact studies? Hydrol. Earth Syst. Sci. 15 (12), 3861–3875. https://doi.org/10.5194/hess-15-3861-2011.

TACTIC, 2021. Tools for Assessment of ClimaTe change ImpacT on groundwater and adaptation Strategies (TACTIC). GeoERA project. Available. https://geoera.eu/projects/tactic9/.

Tolstovs, A.B., Levina, N.N., Prilukova, E., Dume, M.P., Semenov, I.V., Tracesvskaja, L.P., 1986. Investigations of the regime and balance of the groundwater, exogenous geological processes and the maintenance of the state water cadastre (groundwater) in the Latvian SSR for 1984–1986 (summary report for the period 1976–1986). In Russian, Riga, Latvia.

UNESCO (2015) The United Nations World Water Development Report 2015: Water for a Sustainable World. United Nations World Water Assessment Programme, UNESCO, Paris. Available: https://www.unwater.org/publications/world-water-development-report-2015/.

UNESCO (2020) The United Nations World Water Development Report 2020: Water and Climate Change, United Nations World Water Assessment Programme, UNESCO, Paris. Available: https://www.unwater.org/publications/world-water-development-report-2020/.

Van Den Broeck, J., Cunningham, S.A., Eeckels, R., Herbst, K., 2005. Data cleaning: Detecting, diagnosing, and editing data abnormalities. PLoS Med. 2 (10), e267, 0966–0970. https://doi.org/10.1371/journal.pmed.0020267.

Vidon, P., 2012. Towards a better understanding of riparian zone water table response to precipitation: surface water infiltration, hillslope contribution or pressure wave processes? Hydrol. Process. 26 (21), 3207–3215. https://doi.org/10.1002/hyp.8258.

Von Asmuth J.R. (2010) Over de kwaliteit, frequentie en validatie van druksensorreeksen [On the quality, frequency and validation of pressure sensor series]. Technical report no. KWR 2010.001, KWR Watercycle Research Institute, Nieuwegein, The Netherlands. Available: https://edepot.wur.nl/157876.

Von Asmuth J.R., Maas K., Bakker M., Petersen J. (2002) Transfer function-noise modeling in continuous time using predefined impulse response functions. Water Resources Research, 38(12), pp. 23-1-23-12. 10.1029/2001WR001136.

Wendt, D.E., Van Loon, A.F., Bloomfield, J.P., Hannah, D.M., 2020. Asymmetric impact of groundwater use on groundwater droughts. Hydrol. Earth Syst. Sci. 24 (10), 4853–4868. https://doi.org/10.5194/hess-24-4853-2020.

Wickham, H., Averic, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. Journal of Open Source Software 4 (43), 1686. https://doi.org/10.21105/joss.01686.

Witte, J.-P.M., Zaadnoordijk, W.J., Buyse, J.J., 2019. Forensic Hydrology Reveals Why Groundwater Tables in The Province of Noord Brabant (The Netherlands) Dropped More Than Expected. Water 11 (3), 478. https://doi.org/10.3390/w11030478.

Zaadnoordijk, W.J., Bakker, M., 2013. Application of spatial time-series analysis to determine calibration targets for transient groundwater models, presentation at MODFLOW & More 2013, June 2–5. Golden CO, USA.

Zaadnoordijk, W.J., Bus Stefanie, A.R., Lourens, A., Berendrecht, W.L., 2019. Automated time series modeling for piezometers in the national database of the Netherlands. Groundwater 57 (6), 834–843. https://doi.org/10.1111/gwat.12819.

Zelčs V., Markots A., Nartišs M., Saks, T. (2011) Pleistocene Glaciations in Latvia. In J. Ehlers, P. L. Gibbard, & P. D. Hughes (Eds.), Quaternary Glaciations - Extent and Chronology, pp. 221–229. Elsevier Inc.