Groundwater level projections for aquifers affected by annual to decadal hydroclimate variations

Sivarama Krishna Reddy Chidepudi, Nicolas Massei, Abderrahim Jardani, Abel Henriot, Matthieu Fournier, Bastien Dieppois

Groundwater level projections for aquifers affected by annual to decadal hydroclimate variations

Sivarama Krishna Reddy Chidepudi [1, 2], Nicolas Massei[1], Abderrahim Jardani[1], Abel Henriot[2], Matthieu Fournier[1], Bastien Dieppois[3]

[1]Univ Rouen Normandie, UNICAEN, CNRS, M2C UMR 6143, F-76000 Rouen, France

[2]BRGM, 3 av. C. Guillemin, 45060 Orleans Cedex 02, France

[3]Centre for Agroecology, Water and Resilience, Coventry University, Coventry, UK

Corresponding author: Sivarama Krishna Reddy Chidepudi (sivaramakrishnareddy.chidepudi@univrouen.fr)

ORCID : https://orcid.org/0000-0001-9394-7970

Key Points:

- Future changes indicated predominantly decreasing trends in groundwater levels and variability, intensifying from SSP2-4.5 to SSP5-8.5.
- Variability of annual-type aquifers increased for all emission scenarios.
- Paradoxically, groundwater levels seemed higher in the future compared to the historical period for all scenarios.

Abstract

In a context where anticipating future trends and long-term variations in water resources is crucial, improving our knowledge about most types of aquifer responses to climate variability and change is necessary. Aquifers with variability dominated by seasonal (marked annual cycle) or low-frequency variations (interannual to decadal variations driven by large-scale climate dynamics) may encounter different sensitivities to climate change. We investigated this hypothesis by generating groundwater level projections using deep learning models for annual, inertial (low-frequency dominated) or mixed annual/low-frequency aquifer types in northern France from 16 CMIP6 climate model inputs in an ensemble approach. Generated projections were then analysed for trends and changes in variability. Generally, groundwater levels tended to decrease for all types and scenarios across the 2030-2100. The variability of projections showed slightly increasing variability for annual types for all scenarios but decreasing variability for mixed and inertial types. As the severity of the scenario increased, more mixed and inertial-type stations appeared to be affected by decreasing variability. Focusing on low-frequency confirmed this observation: while a significant amount of stations showed increasing variability for the less severe SSP 2-4.5 scenario, low-frequency variability eventually showed slight yet statistically significant decreasing trends as the severity of the scenario increased. For the most severe scenario, almost all stations were affected by decreasing low-frequency variability. Finally, groundwater levels seemed, in most instances, slightly higher in the future than in the historical period, without any significant differences between emission scenarios.

Plain Language Summary

The groundwater levels in aquifers can vary seasonally but also over much longer time scales, displaying longer-term variations than just annual. According to such behaviours, their future evolution regarding climate variability and change may vary significantly. Using different greenhouse gas emissions and socioeconomic trajectories, various climate models that eventually released different scenarios of precipitation and temperature, and deep learning models to simulate groundwater level changes, we investigated how the different types of groundwater level behaviours would respond to climate change, taking the example of northern France where all these groundwater level types can be found. Future trends would show

decreasing levels between 2030 and 2100. Yet, the amplitude of groundwater level variations would increase for aquifers with dominating seasonal variability, while it decreased for aquifers varying more slowly through time. Paradoxically, in many, if not most cases, the expected groundwater levels - in particular for slow-varying types - would be higher in the near or even far future as compared to the historical 1970-2022 period, even for the most severe climate change scenario. Although in line with other studies in the same or nearby regions, understanding this phenomenon will deserve deeper investigation.

1 Introduction

Groundwater is a vital freshwater resource for sustaining domestic, agricultural, and industrial activities. However, the sustainability of groundwater resources is under threat due to the impacts of climate change and human interventions (Jasechko et al., 2023; Mishra et al., 2024; Scanlon et al., 2023; Taylor et al., 2013). Assessing the potential effects of climate change on groundwater levels is of paramount importance, particularly in regions that heavily rely on groundwater resources and are vulnerable to changes in climate, such as northern France (Habets et al., 2013; Vergnes et al., 2023). Robust projections of future groundwater levels under different climate change scenarios are essential for informed water resource management and the development of effective adaptation strategies. The Coupled Model Intercomparison Project Phase 6 (CMIP6) provides state-of-the-art information on plausible global to regional climate changes in the past, present, and future (Eyring et al., 2016). However, climate models show large uncertainty due to model physics differences, emission scenarios sensitivity, and internal climate variability (Atawneh et al., 2021; Hawkins & Sutton, 2009; Lehner et al., 2023). In addition, when analysing climate change impacts on hydrological variations and trends, additional uncertainties arise from the hydrological model structures and parameterisations (Clark et al., 2016; Melsen et al., 2018; Wu et al., 2024; Yuan et al., 2023). Nevertheless, only 20% of groundwater impact studies considered the influence on climate model uncertainties (Atawneh et al., 2021).

Characterising such uncertainties is crucial for enhancing the reliability of climate change impact scenarios for groundwater resources.

Most of the time, physically-based hydrological models have been employed to assess the impacts of climate change on groundwater resources (Costantini et al., 2023; Halloran et al., 2023; Mishra et al., 2024; Vergnes et al., 2023). However, due to the substantial computational time and specific data requirements associated with these models, data-driven methods have become increasingly popular complements or sometimes alternatives (Bhasme et al., 2022; Chidepudi et al., 2023a; Hauswirth et al., 2021; Rehana & Rajesh, 2023; Wunsch et al., 2022). In recent years, some studies have employed artificial intelligence (AI), machine learning (ML), and Deep Learning (DL) models in conjunction with CMIP5 and CMIP6 climate projections to assess the impacts of climate change on groundwater levels (Chakraborty et al., 2021; Kayhomayoon et al., 2023; Nozari et al., 2022; Roshani & Hamidi, 2022; Xiong et al., 2022; Wunsch et al., 2022; Nourani et al., 2023; Secci et al., 2023). Most of these approaches used neural networks and at least one deep neural network architecture (i.e., DL). Many different architectures exist that can be mostly suited for specific tasks in handling time series data. Recently, Secci et al. (2023) compared different types of DL models (NARX, LSTM, and CNN) and found that long short-term memory neural networks (LSTM) outperformed the others due to their ability to handle long-range dependencies. Long et al. (2024) reached similar conclusions and demonstrated the ability to capture complex spatiotemporal patterns and nonlinear relationships between climate variables and hydrological processes.

Most of the studies dealing with groundwater level simulation using DL approaches either mainly considered aquifers dominated by seasonal variability or develop forecasts on quite short forecasting horizons, up to a few days or weeks (Boo et al., 2024; Rajaee et al., 2019; Tao et al., 2022; Uc-Castillo et al., 2023). However, it is now well-recognised that interannual to decadal variability affecting groundwater level originates from large-scale climate variability (Baulon et al., 2022; El Janyani et al., 2012; Hanson et al., 2006; Holman et al.,2011; Liesch & Wunsch, 2019; Massei et al.,2010; Neves et al., 2019; Rust et al., 2019), and can significantly impact decadal trends at the regional scale in climate change projections. Such interannual to decadal variability

is represented differently by different climate models and individual simulations from the same climate model (i.e., different ensemble members; Deser et al., 2012; Deser & Phillips, 2023). Emission scenarios can also modify low-frequency variability as climate change impacts large-scale modes of variability and teleconnections (Klavans et al., 2022; Mahmood et al., 2022; Terray, 2012). Therefore, It is also crucial to consider aquifers that behave on more low-frequency dynamics, develop and apply DL tools that can appropriately describe such variability in groundwater systems, and assess how systems subject to low-frequency variability would behave under climate change compared to those dominated by seasonal variations, which seemed to have received more attention so far.

The study presented here uses DL techniques and CMIP6 climate change scenarios to provide an overview of the potential impacts of climate change on groundwater levels for different types of aquifers dominated by seasonal variability or low-frequency variability (or a mix of the two). These three examples of contrasted behaviours are observed in northern France, which is here used as a case study. The research assesses potential alterations in groundwater level trends and variability due to future climate conditions. Employing a multi-station deep learning approach, we generated groundwater level projections for the region, incorporating three different climate change scenarios and socioeconomic pathways. This approach aims to capture the spatial patterns and temporal evolution of projected groundwater level changes across northern France according to the hydrological systems' characteristic behaviour (annual, mixed, inertial) to assess their sensitivity to different climate change scenarios. Furthermore, the study evaluates the performance, uncertainties, and limitations of the deep learning methodology and the climate model projections utilised in the groundwater level projection analysis.

The rest of the paper is structured as follows: Section 2 presents the data and study area. Section 3 presents the methodology of models and the trend and variability assessment of projections. Section 4 presents the dispersion of GWL projections under different scenarios. Section 5

discusses the time evolution of GWL projections. Section 6 presents our comparison with other relevant studies and conclusions.

2 Study Area and Data

The study focuses on the northern region of France, primarily encompassing the Paris Basin and its surrounding areas. This region was selected due to long-term groundwater level (GWL) data availability, which is crucial for accurate projections. A significant advantage of this area is the presence of three distinct GWL behaviours despite its relatively limited spatial coverage: 1- reactive aquifers dominated by seasonal variability ("annual" type), 2- aquifers with marked seasonal variations along with significant low-frequency variability (“mixed" type), and 3- aquifers dominated by low-frequency variability ("inertial" type). Example time series of these three types are provided in Figures 3-5 (observational data, left panel). This diversity in groundwater level patterns provides a valuable opportunity to assess the performance of the deep learning models in capturing various hydrogeological conditions and responses to climate variability.

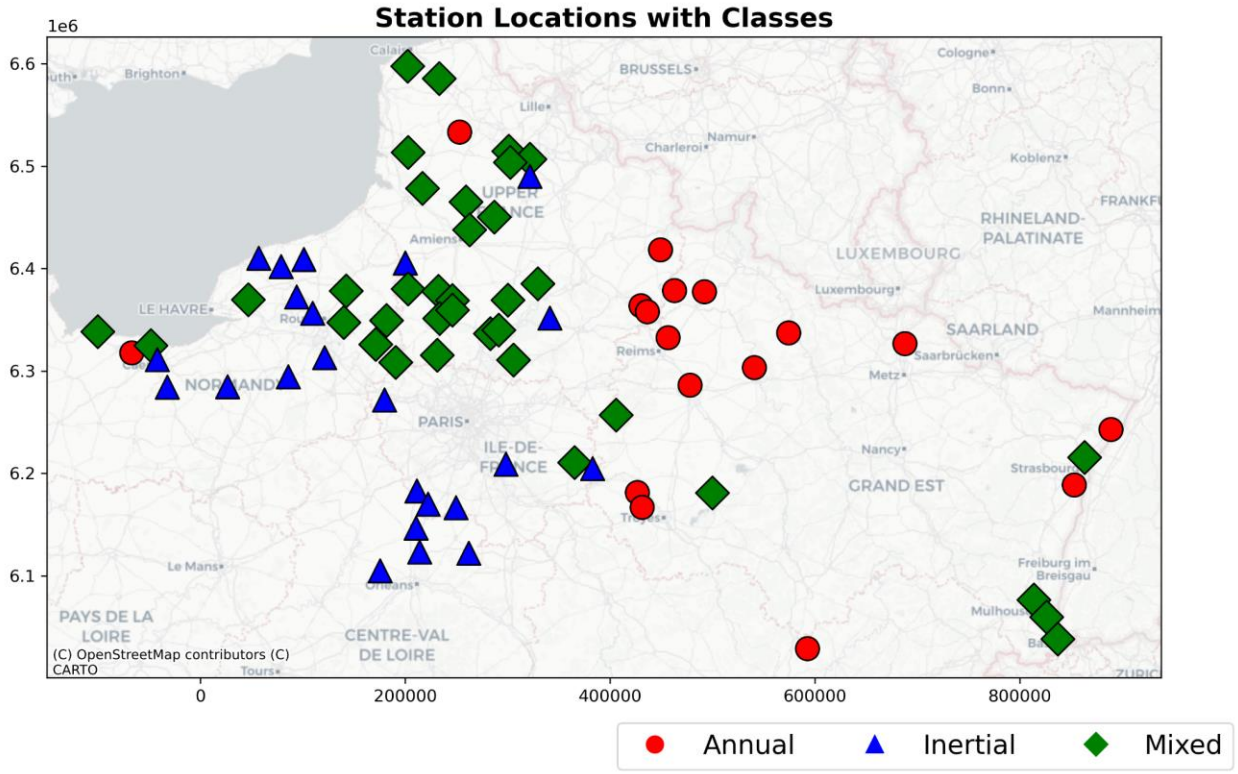


Figure 1: Study area and location of stations with details of GWL types: Annual (red dots), Inertial (blue triangles), Mixed (green lozenges)

The dynamic climate variables (precipitation and temperature) were obtained from the ERA5 reanalysis dataset (Hersbach et al., 2020), which provides forcing data at a high spatial resolution of 0.25 degrees. The selection of these two variables was made for several reasons. First, precipitation and temperature are available across all 16 selected climate models and three scenarios, ensuring consistency and reliability in the data used for analysis. Second, using these variables maintains coherence with other studies within the same research framework, allowing for better comparability and integration of results. Third, precipitation and temperature are fundamental drivers of hydrological processes, making them the most relevant basic variables for projections. Finally, focusing on these two key variables keeps the approach parsimonious regarding data availability and processing requirements, enhancing the efficiency and reproducibility of the analysis while still capturing the essential climate dynamics for groundwater projections. The groundwater level data were sourced from the ADES (Accès aux Données sur les Eaux Souterraines) database (https://ades.eaufrance.fr/; Winckel et al., 2022), specifically focusing on climate-sensitive wells minimally influenced by human activities and exhibiting strong sensitivity to climate variability (Baulon et al., 2022).

To generate future GWL projections, climate data from 16 CMIP6 models were used as inputs in deep learning models trained on data from ERA5. Three Shared Socioeconomic Pathway (SSP) scenarios were considered: SSP2-4.5 (moderate emissions), SSP3-7.0 (severe emissions), and SSP5-8.5 (extreme emissions). These scenarios represent different future pathways of greenhouse gas emissions and socioeconomic factors, allowing for a comprehensive assessment of potential climate change impacts on groundwater resources. We chose to utilise the bias-corrected datasets from the NEX-GDDP-CMIP6 dataset, which has data for only one variant for each CMIP6 model. Hence, the uncertainty related to internal climate variability is not considered (Deser et al., 2012).

3. Methodology

3.1 DL models and neural network architectures

This study employs a deep learning approach to project future groundwater levels in northern France using climate projections from CMIP6 models under different emission scenarios. The methodology builds upon Chidepudi et al.'s (2024b) findings, which demonstrated the effectiveness of models trained with a multi-station approach on clustered data combined with wavelet preprocessing.

While the efficiency of LSTMs has long been demonstrated, other recurrent-based architectures have emerged more recently, like the Gated Recurrent Unit (GRU), as Cho et al. (2014) described. GRUs have a

simpler architecture than LSTMs, which can lead to faster training. Additionally, GRUs have shown comparable or sometimes superior performance to LSTMs in various sequence modelling tasks, making them an attractive alternative for groundwater level modelling (Chidepudi et al., 2023a). The core of our approach then utilises GRU neural networks according to their ability to capture long-term dependencies in sequential data and their computational efficiency.

Our neural network models are trained with inputs from ERA5 data preprocessed using the Boundary Corrected Maximal Overlap Discrete Wavelet Transform (BC-MODWT), implemented with a 'la8' (least asymmetric) wavelet and four decomposition levels, as in Chidepudi et al. (2023a). This technique effectively separates input signals into different frequency bands while preserving time information and mitigating boundary effects (Quilty & Adamowski, 2018; Chidepudi et al., 2023a). It is important to note that for each input time series or feature, BC-MODWT produces five new time series: four wavelet detail levels corresponding to the decomposition levels and one smooth approximation. For more details about discrete wavelet transform and MODWT, the reader can refer to the very rich literature on the application of wavelet methodology to hydrology, such as Labat et al. (2000), Percival & Walden (2000) and Massei et al. (2017). This preprocessing step is applied to climate reanalysis data, which serves as input for the GRU models during training and validation. Specifically, if the original input consisted of N features, the BC-MODWT preprocessing would result in a 5N time series as input to the model, each maintaining the original temporal resolution but capturing different frequency components of the original signals. The observed groundwater levels are used as the target variable.

This study adopts a multi-station approach, where GRU models are trained within each GWL cluster (annual, mixed, or inertial) using aggregated data from multiple stations. Most recent studies have focused on single-station approaches, which consist of training DL models based on a single target time series (i.e. the DL model learns from known values of the one-time series to be eventually simulated). For the last couple of years, recent research has suggested that models trained with more diverse data can result in more reliable hydrological projections (Wi & Steinschneider, 2022), as this would enhance the capabilities of the models for generalizability and transferability. The approach consists of training DL models for time series regression on multiple time series (in our case, GWL time series) simultaneously, hence leveraging a wider range of relevant values or hydrological events than available in a single station time series. The approach has already been called the “global models” or “multi-well” approach (Heudorfer et al., 2024) or “multi-station” (Chidepudi et al., 2023b & 2024b). This key finding highlights the potential advantages of multi-station approaches over single-station methods. Typically, the multi-

station approach used herein leverages collective information from stations with at least 42 years of data (1970-2022), enhancing the model's ability to capture spatial patterns and increasing the generalizability of the projections. Only dynamic variables, such as precipitation and temperature, are used as input features.

Data preprocessing includes normalising each input variable individually to be in the 0-1 range and reshaping it into a 3D format suitable for GRU models. A sequence length of 48 months is used to capture long-term patterns in the data. To enhance robustness and mitigate the effects of random weight initialisation, multiple GRU models (10) are trained with different initialisations, creating an ensemble approach. As described in Chidepudi et al. (2024b), hyperparameter tuning was performed using Bayesian optimisation.

The models are trained using early stopping and model checkpointing techniques to prevent overfitting and save the best-performing model. A 20% validation split is used to monitor performance during training. Importantly, models are trained with collective data from all stations in each cluster with at least 42 years of data (1970-2022), ensuring a comprehensive historical context.

3.2 Assessing trends and variability in projected GWL under three SSP scenarios

The trained ensemble GRU models are applied to the wavelet-transformed CMIP6 climate projection data to generate projections. As a reminder, the total number of features would equal 10 (two variables, i.e. precipitation and temperature, each decomposed into five wavelet components). The median of the ensemble predictions is calculated for each CMIP6 model, as shown in Figure 2. Then, the median across all models is computed to produce a robust projection for each scenario.

For each GWL station, up to 160 projected time series were generated, corresponding to 16 climate models and 10 initialisations of the DL model. For each climate scenario, the results are presented either: i) for each climate model individually (i.e., for each climate scenario, 10 GWL projected time series corresponding to 10 DM model initialisations; cf. Fig. 2); ii) 2- for all climate models together (i.e., 160 GWL projected time series corresponding to 16 climate models with ten initialisations of the DL model). A median time series of the 10 or 160 projections is derived in these two cases. It is then used to assess the ensemble trend and explore a possible change in the amplitude of the variability of GWL. In other words, the ensemble median time series (ensemble of either only 10 or 160 projected time series) is tested for stationarity in the weak sense (i.e. change in mean and variance) throughout the period 2030-2100 to assess whether a change in GWL is to be expected on average, and if the overall variability would

also tend to change. Figure 3 illustrates such changes using one randomly selected GWL projection and is described later in the text. Ultimately, it comes to exploring whether water levels and their amplitude of variation (difference between high and low levels) will be expected to increase, decrease, or remain unchanged.

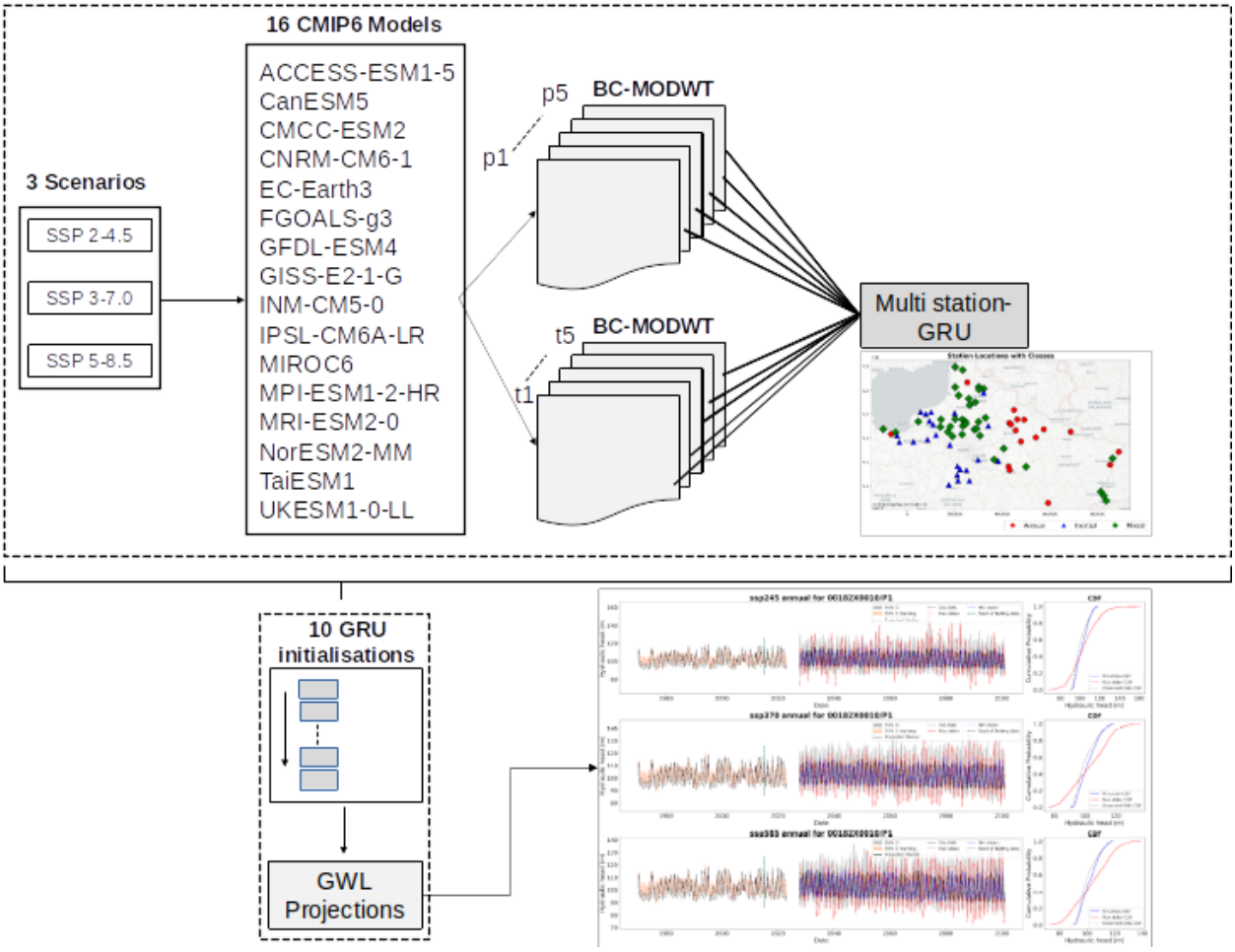


Figure 2: Structural workflow of the methodological approach for projections.

Trend analysis is performed using the Correlated Seasonal Mann-Kendall (CSMK) test, an extension of the classical Mann-Kendall trend test proposed by Hipel & McLeod (1994). The CSMK test is a non-parametric statistical method that detects monotonic trends in seasonal time series data with serial correlation. It is particularly suitable for hydrological and climate variables (Hussain et al., 2019). This test accounts for both seasonality and the correlation between observations in consecutive months or seasons, addressing the limitations of the classical MK test. The CSMK test does not require the data to follow a specific

distribution and can handle missing values and outliers. It operates by separating the time series into seasonal groups, calculating the MK test statistic for each season, adjusting for serial correlation, and then combining the results to obtain an overall trend assessment. A standardised test statistic is computed and compared to a critical value (using a P-value of 0.05) to determine if a statistically significant trend exists. The CSMK test is robust against non-normality and censored data, making it particularly valuable for analysing trends in groundwater levels, precipitation, temperature, and other variables relevant to groundwater projections using deep learning models, especially when these variables exhibit strong seasonal patterns and serial correlation. Sen's slope is also computed as part of the CSMK test.

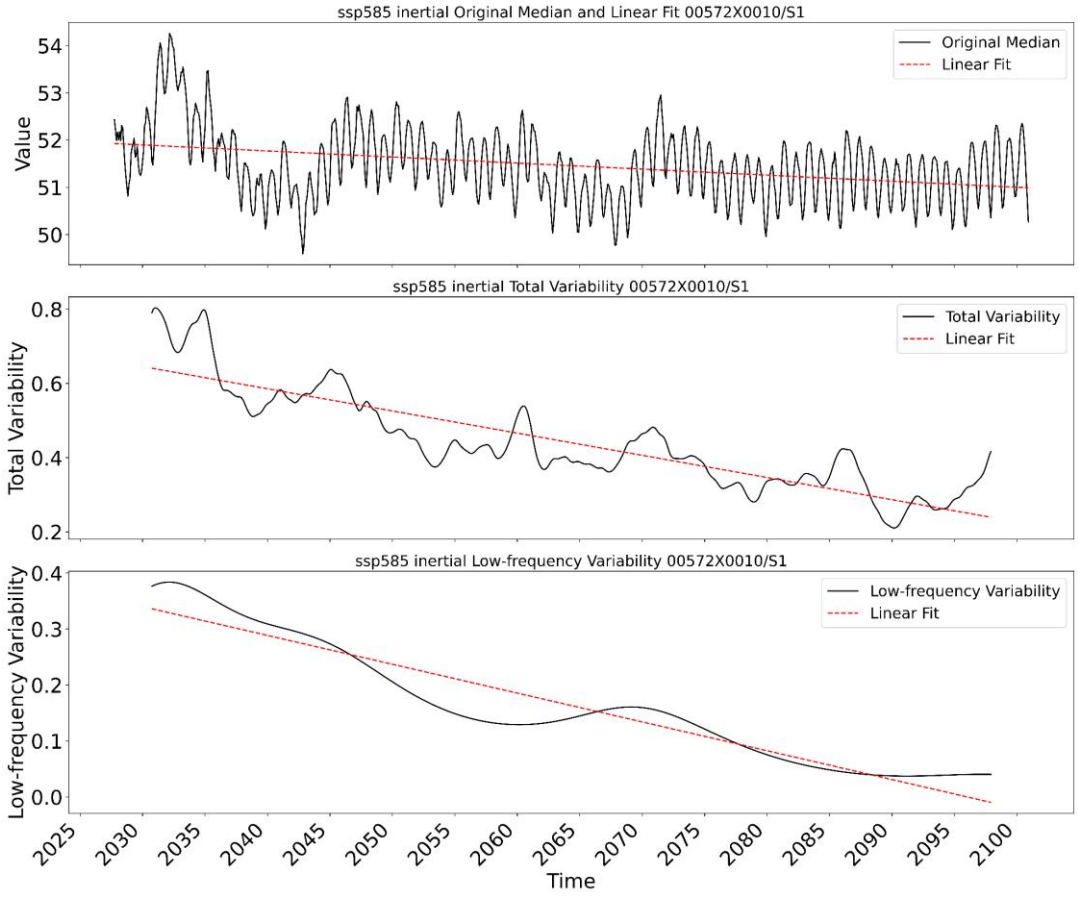
This analysis is conducted for each groundwater station, climate scenario, and GWL type. For assessing trends in GWL, the time series used for the MK trend test is the median of the ensemble. The ensemble would consist of either 10 or 160 projected time series:

- 10 projected time series when climate models are taken separately, i.e. for each climate model, ten projections corresponding to the ten different initialisations of the DL model are obtained;

- 160 projected time series when all climate models are grouped, leading to 10 DL initialisations * 16 climate models.

For assessing a change in the variability at each station, the time series used for the MK trend test is the estimated variance across time of the median of several projections. The times series' variance is estimated using so-called scale-averaged wavelet power, following the procedure described in Torrence and Compo (1998). At each time point, the scale-averaged wavelet power measures the variance in GWL for either all-time scales or a range of time scales. In the first case (all time scales), it measures the total variance of the time series across time (i.e. for all time steps). In contrast, the second case (selected range of time scales) measures the variance associated with one particular frequency band across time. The reader is referred to Torrence and Compo (1998) for a detailed and comprehensive explanation of wavelet scale-averaging's calculation and main interest. Figure 3 illustrates this for one randomly selected GWL projection, representing the variance through time obtained for one GWL projection for a time scale greater than five years (i.e., low-frequency fluctuations of GWL at this station). In this example, scale-averaged wavelet power shows that total variability (Fig. 3b) and >5-yr low-frequency variability (Fig. 3c) tend to decrease through time. In order to prevent the results from being too much affected by edge effects, we removed the first and last 36 months from these scale-averaged power time series (Fig. 3b and c), which corresponded to removing as many wavelet coefficients as possible falling into the so-called

cone of influence (cross-hatched area in figure 3d). Such coefficients can be identified on the continuous wavelet spectrum of the time series and are located before the first. After the second vertical dashed line (Fig. 3d). The continuous wavelet spectrum also clearly shows that low-frequency variability (i.e. variance for periods higher than ~5 years / ~60 months on the plot) tends to decrease through time.



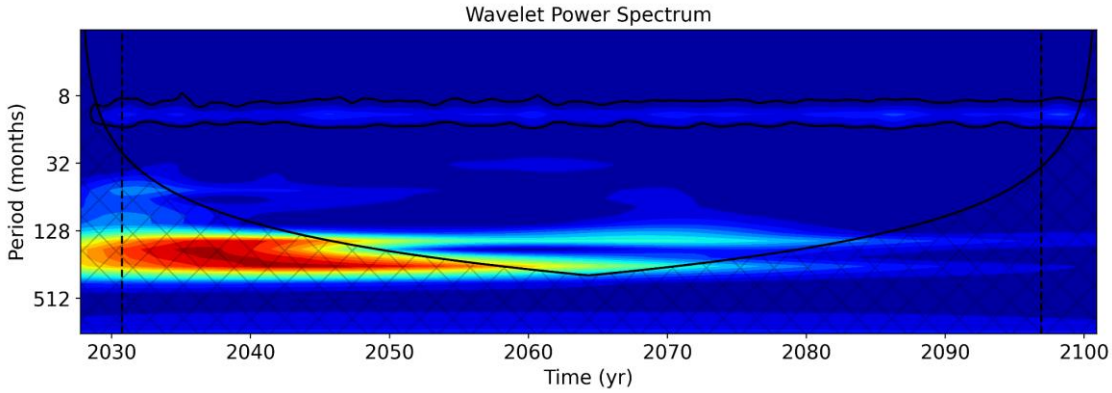


Figure 3: a) median of projections and fit line b) Scale averaged wavelet power for total variability c) Scale averaged wavelet power for low-frequency variability d) Continuous wavelet spectrum (scalogram) of GWL time series: red indicates high variance, and blue indicates low variance.

4. Dispersion of climate change impact projections on various GWL types under contrasting emission scenarios

Here, all 16 downscaled GCM were used as input to the DL models with ten different parameter initialisations (i.e. initialisation of the neural network weights) to create an ensemble of 160 projections per SSP scenario at each time point. While projections are usually presented and quantified using the ensemble mean or median, relying on these metrics alone could suppress the information on the variations and uncertainty. Here, we chose to represent our results in percentile rather than box plots, as recently suggested by Müller & Döll (2024), who found such a representation more suited to support participatory climate change adaptation processes and uncertain local climate hazards.

Figures 4-6 show the groundwater projections of three types (annual, mixed and inertial) for three SSPs. The left panel shows the training and test results of the historical period. The central panel shows the projected median groundwater level and the 95% confidence interval. We highlighted the projections with the highest and lowest variability (from now on referred to as HV and LV, resp. represented as red and blue lines in Figures 4-6), as these naturally correspond to either more pronounced or more dampened extremes resp, for HV and LV projections. In addition, comparisons between GWL types, i.e. with different amplitudes of low-frequency variability, provide information on how various representations of climate variability may impact the projections released.

The results for the training and testing stage are always quite satisfactory, with small confidence intervals (Figures 4a-6a). On the other hand, big differences may exist between the different projections for any given GWL type (annual, mixed and inertial), hence translating the uncertainty linked to the various climate models (Figures 4b-6b). While LV projections are relatively close to the median, HV ones would deviate substantially from the median and LV time series. The median time series always displayed a much more pronounced seasonal variability and a much lower low-frequency variability (the median time series appears somewhat “shrunk”) compared to historical observations - noticeable for all GWL types - and even more particularly for the inertial. This increased seasonal variability is due to the stochastic nature of low-frequency variability (stochastic noise eventually cancels out on average). In contrast, the annual

cyclicity is, by definition, almost entirely deterministic: median computation from all 160 projections at each time point ultimately results in a low amplitude of any other variability than the annual cycle, which, in contrast, is over-expressed.

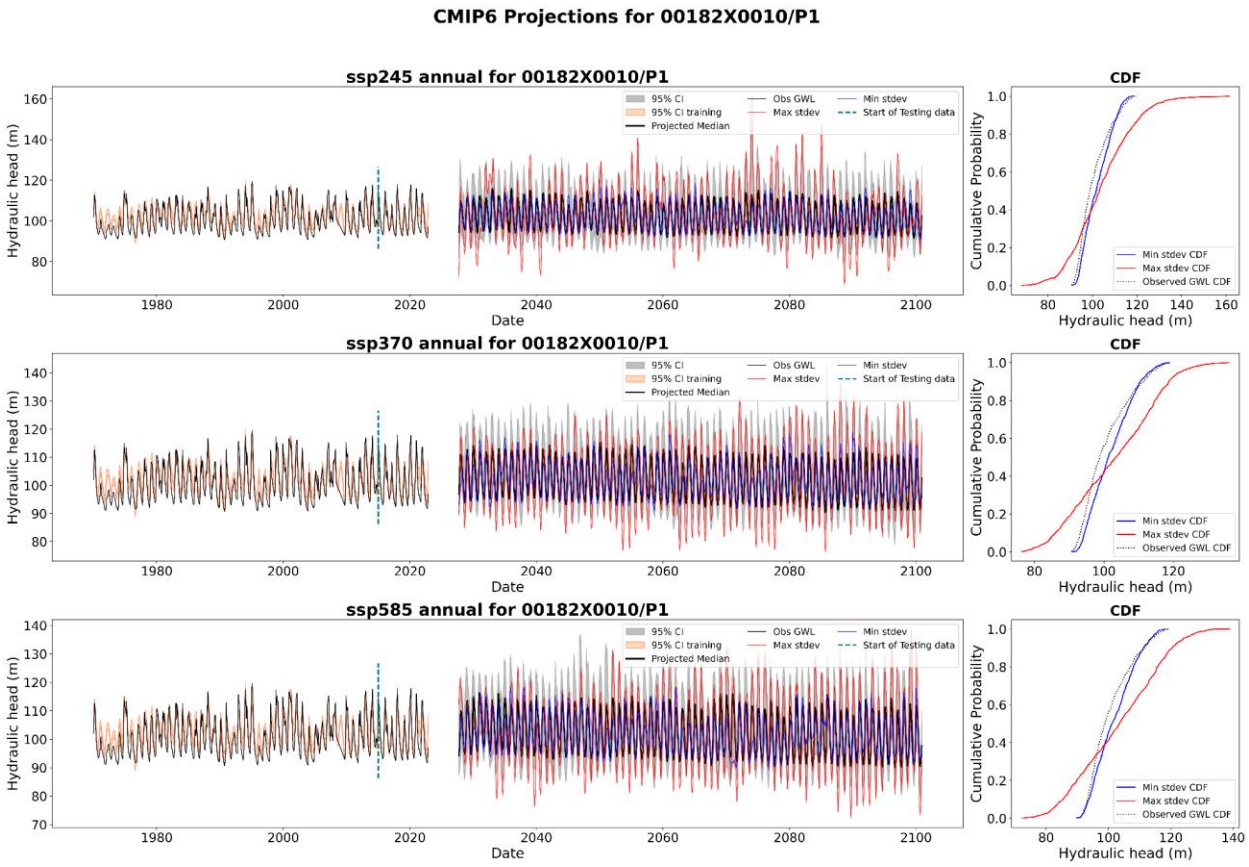


Figure 4: Annual type groundwater projections for three SSP scenarios (Top-bottom): a)(left panel)Training and testing results with confidence intervals, b)(middle panel) Projections for SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios show median (black), high variance (HV, red), low variance (LV, blue) projections, and confidence intervals (grey). c) (right panel) Cumulative Distribution Functions (CDFs) comparing historical observations (dotted black) with HV and LV projections(resp. as red and blue CDFs).

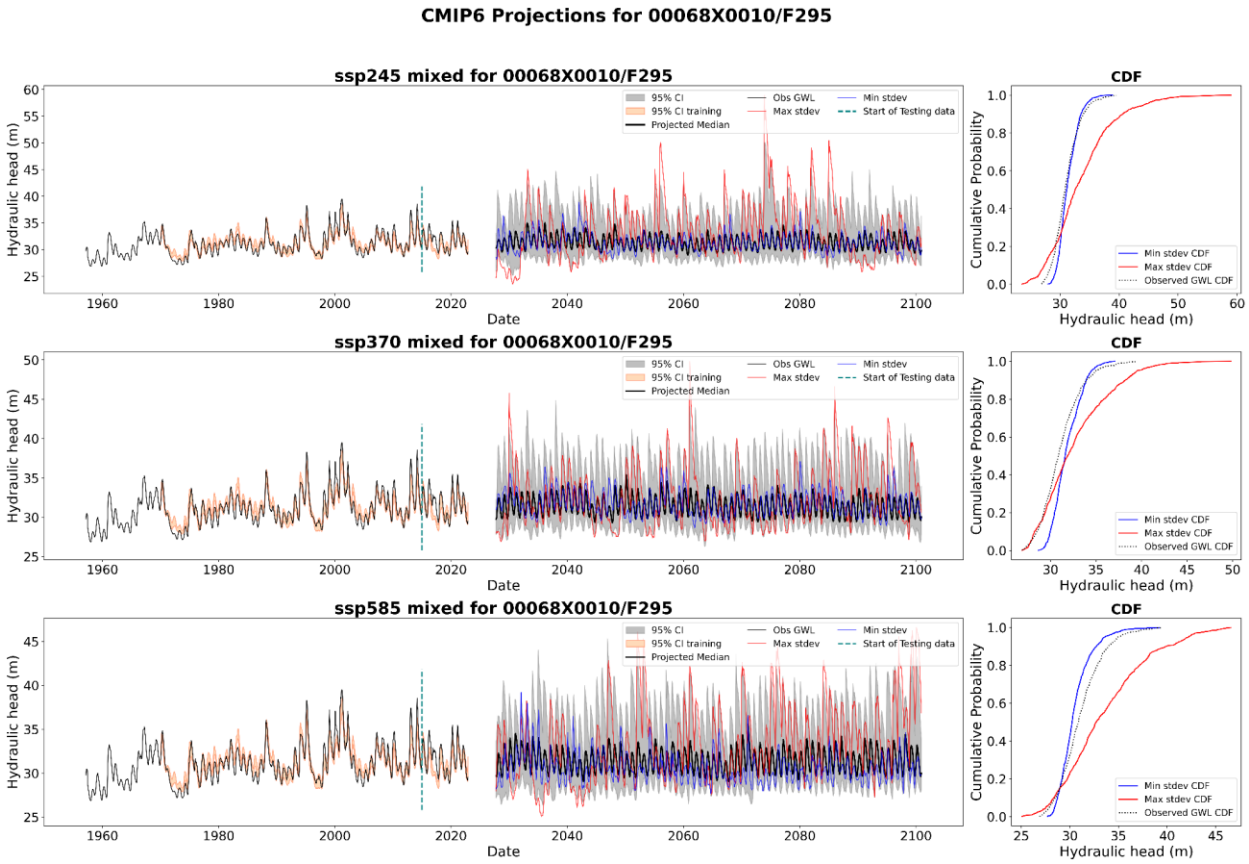


Figure 5: Mixed-type groundwater projections for three SSP scenarios (Top-bottom): a)(left panel)Training and testing results with confidence intervals, b)(middle panel) Projections for SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios show median (black), high variance (HV, red), low variance (LV, blue) projections, and confidence intervals (grey). c) (right panel) Cumulative Distribution Functions (CDFs) comparing historical observations (dotted black) with HV and LV projections(resp. as red and blue CDFs).

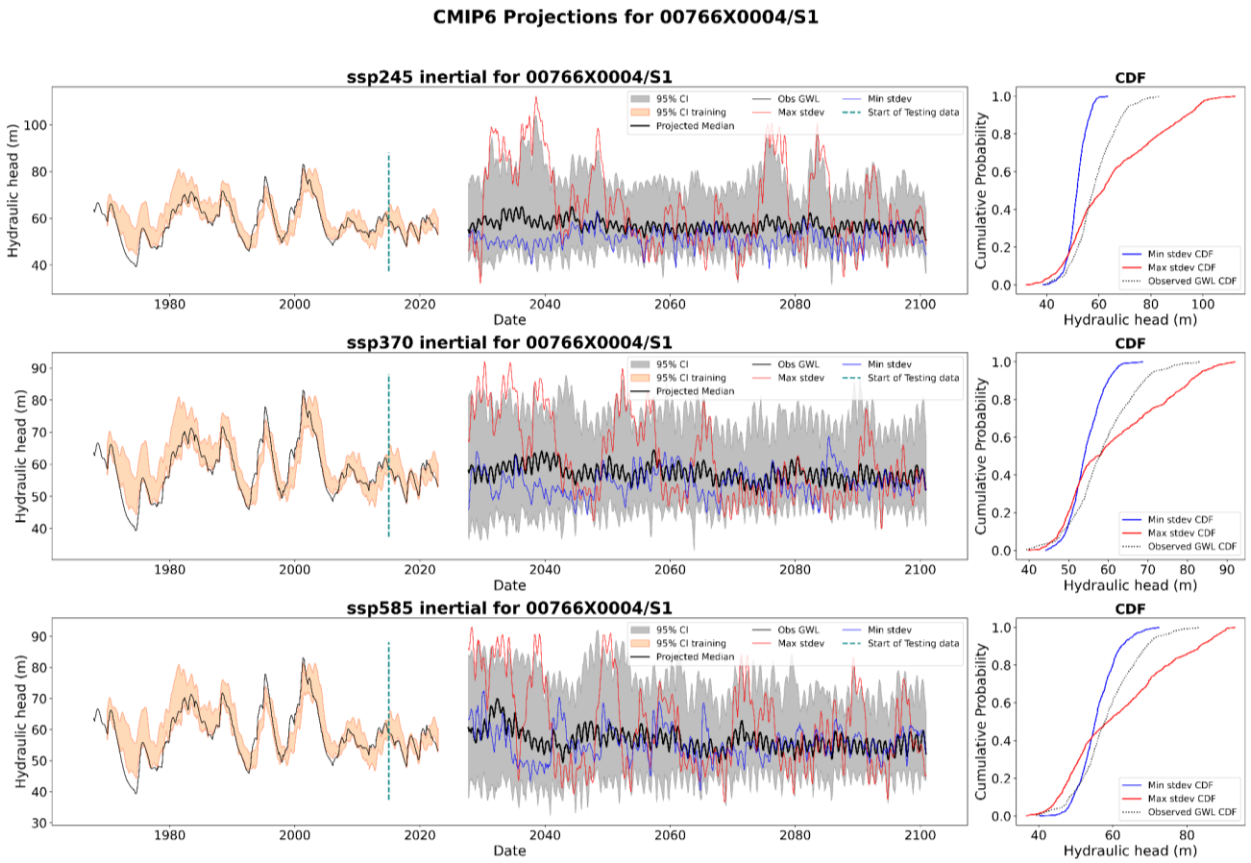


Figure 6: Inertial type groundwater projections for three SSP scenarios (Top-bottom): a)(left panel)Training and testing results with confidence intervals, b)(middle panel) Projections for SSP1-2.6, SSP2-4.5, and SSP5-8.5 scenarios show median (black), high variance (HV, red), low variance (LV, blue) projections, and confidence intervals (grey). c) (right panel) Cumulative Distribution Functions (CDFs) comparing historical observations (dotted black) with HV and LV projections(resp. as red and blue CDFs).

On the other hand, comparisons of the CDFs (Figures 4c-6c) of the LV and HV projections with the historical observed time series show that for the annual and mixed types, LV projections seem to approximately fit the historical observations for all 3 SSP scenarios (Figures 4-5). which is not the case for the inertial type (Figure 6), for which the variability of the LV projection is always lower than for the historical period. To check whether this observation drawn from the three randomly selected annual, mixed and inertial GWL stations presented in these figures can be generalised, we compared the variability of HV and LV projections to that of the historical GWL time series at each station. To this aim, we simply computed the standard deviation of HV, LV and observed GWL time series at each station, and represented the ratios of HV standard deviation to observed standard deviation and of HV standard deviation to observed standard deviation (Figures 7-9) for all 3 GWL types and all emission scenarios. The

results confirmed the conclusions drawn from Figures 4 to 6 presented above. For all stations of the annual type and all emission scenarios (Fig.7), the LV to observation ratio is always close to 1, meaning that the overall variability of LV and observed historical GWL are always close.

In contrast, the HV projection always has a much higher standard deviation than the observation (~1.5 to 2 times), which is also true for the mixed-type GWL, albeit to a - slightly - lesser extent but does not hold for GWL projections of inertial types (Fig.9). Finally, as already mentioned above, one can notice here that except for the annual GWL type, the variability of the median projection time series is always lower than the LV projection for any given GWL station; the same explanation as provided above can be given here. No obvious differences among the different emission scenarios could be observed for any GWL type (Fig. 7 to 9).

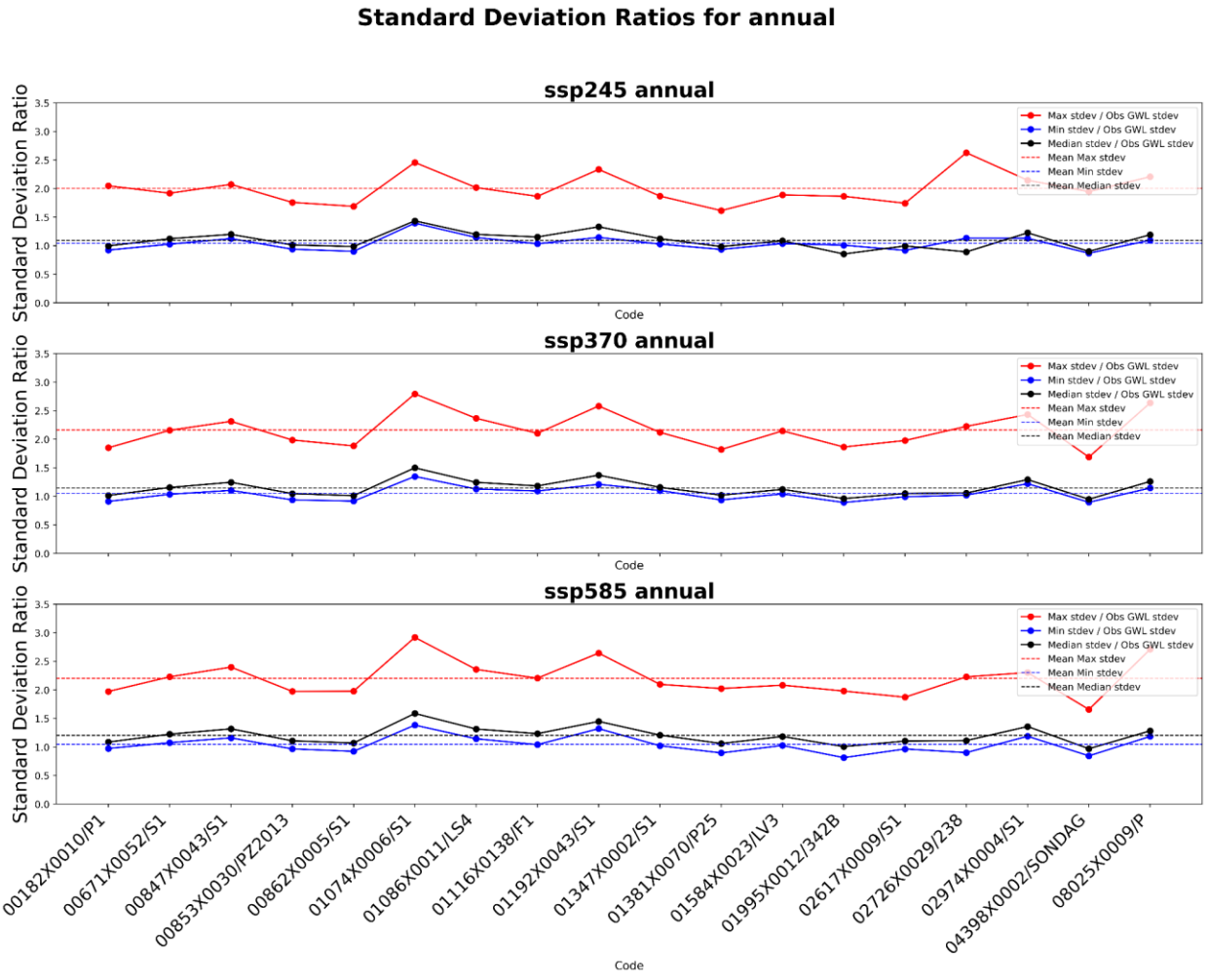


Figure 7: Variability comparison of all stations in annual-type GWL: Ratios of Median(black), HV (red) and LV (blue) projection standard deviations to observed standard deviations across all stations(X-axis) and emission scenarios (top-bottom).

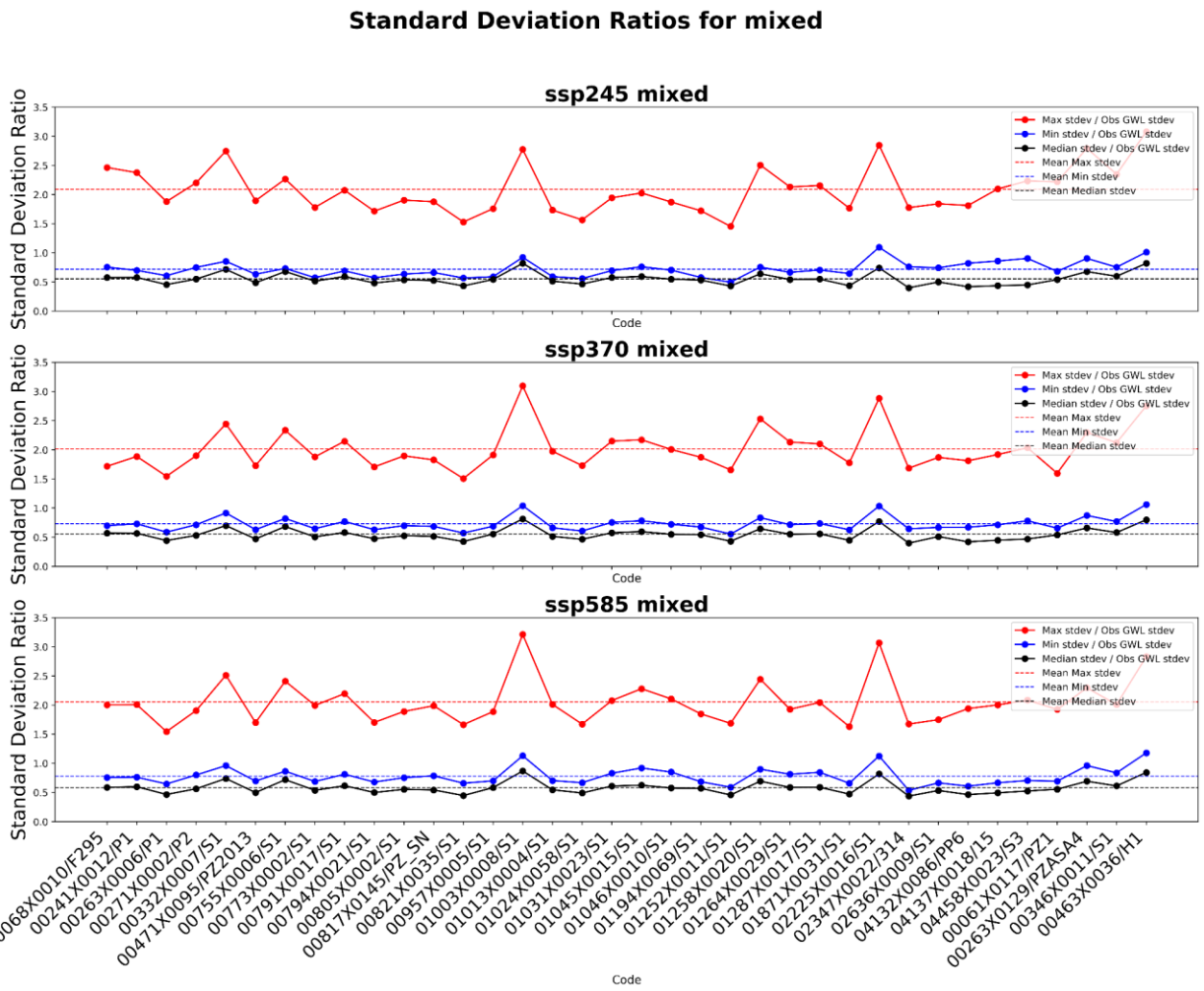


Figure 8: Variability comparison of all stations in mixed-type GWL: Ratios of Median(black), HV (red) and LV (blue) projection standard deviations to observed standard deviations across all stations(X-axis) and emission scenarios (top-bottom).

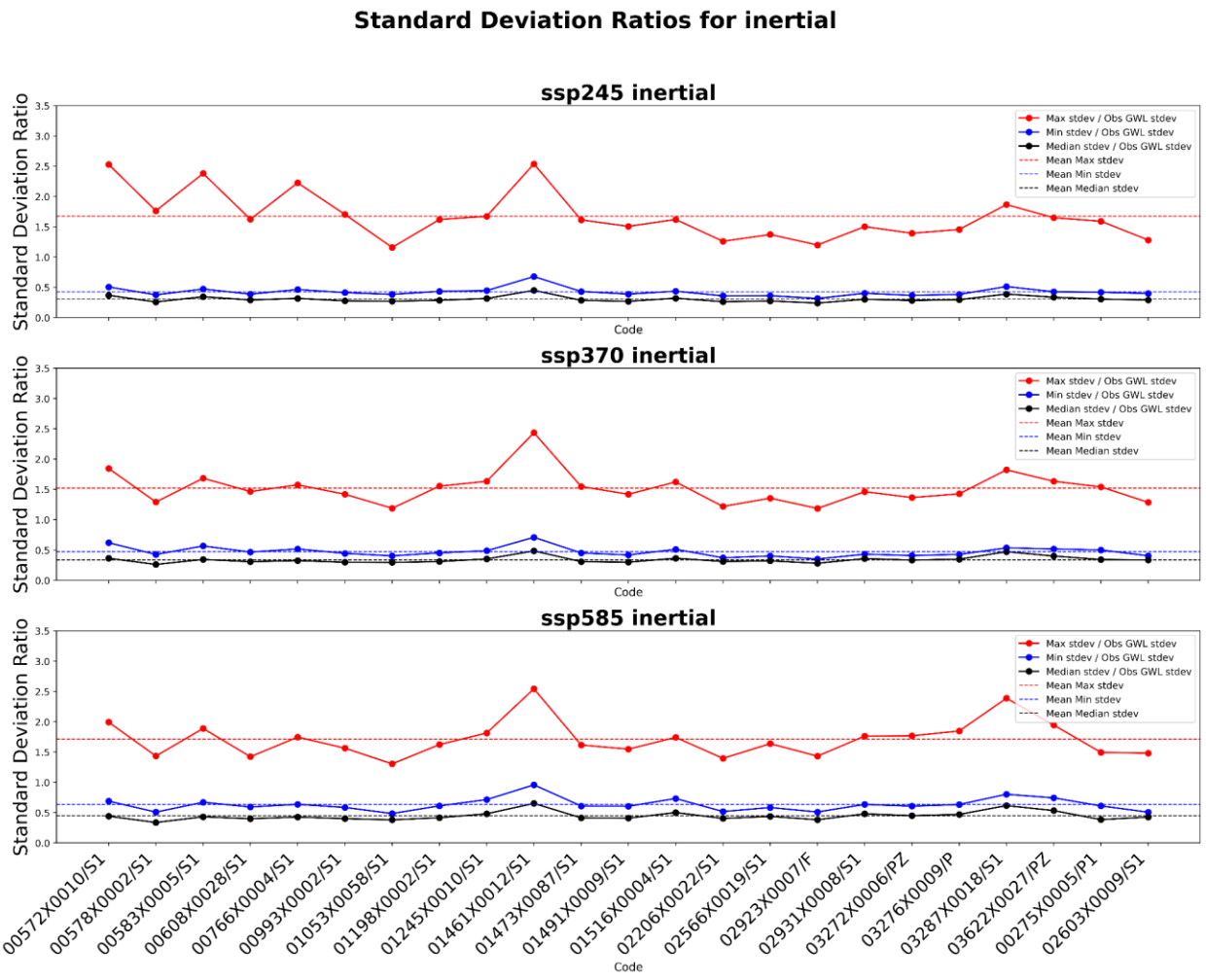


Figure 9: Variability comparison of all stations in inertial-type GWL: Ratios of Median(black), HV (red) and LV (blue) projection standard deviations to observed standard deviations across all stations(X-axis) and emission scenarios (top-bottom).

5. Time evolution of GWL: future trends and variability for annual, mixed and inertial types

5.1 Trends based on the ensemble approach

In the previous section, we described the different projections that could be obtained owing to the use of 16 different models for different GWL types and three emission scenarios. Here, we focus the analysis on the time evolution of GWL. For clarity, we evaluated such changes using the median time series only. Although we showed that the total amplitude of median projections is artificially lower than that of the observed historical time series, the aim here is not to assess the change between the historical and the future periods but only the change during the future period. We use the median as in most other works (Martel et al., 2022; Wunsch et al., 2022;). In particular, we examined three different aspects of GWL time

evolution: 1- we first assessed the potential changes in water levels using the MK trend test performed on the median GWL time series, 2- we tested for changes in the variability of GWL using the MK trend test on scale-averaged power of the median GWL series, as described in the methodology section, 3- we repeated step 2- with a specific focus on the low-frequency variability. Here, low-frequency variability corresponds to time scales above five years, which have been recognised as responsible for several extreme events in groundwater levels in northwestern Europe (Massei et al., 2010; Baulon et al., 2022; Rust et al., 2019; Neves et al., 2019; Liesch & Wunsch, 2019).

It is clear from Figure 10 that almost all GWL would decrease, regardless of the emission scenario and the GWL type. On the other hand, the overall variability of GWL through time showed more contrasted results (Figure 11). In all three SSP scenarios, the total variance of annual-type GWL is expected to increase while decreasing for almost all inertial and mixed types. No clear difference shows up between the scenarios. GWL types strongly influence the trends in total variance. Specifically, Annual type aquifers predominantly show increasing variance (blue circles). Inertial and mixed-type aquifers mostly exhibit decreasing variance (red triangles and diamonds). The increasing variance in annual type aquifers suggests these areas may face more extreme fluctuations in groundwater levels, i.e., be more prone to extremely high and low groundwater levels. The spatial patterns of increasing and decreasing variance trends persist across all three SSP scenarios (SSP2-4.5, SSP3-7.0, SSP5-8.5), suggesting that the overall pattern of change is relatively robust to different emission scenarios.

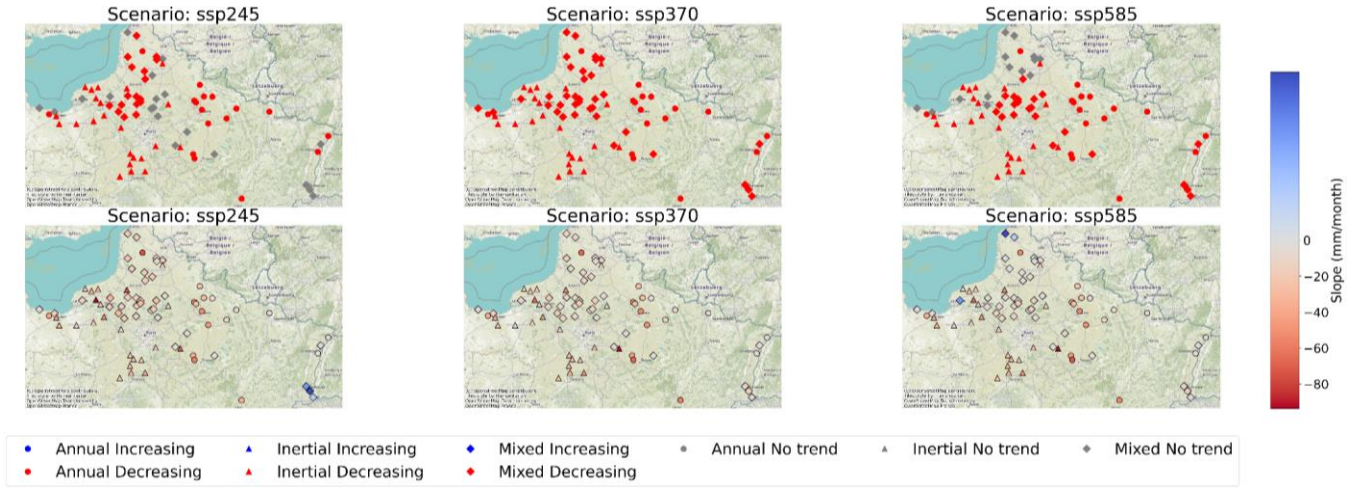
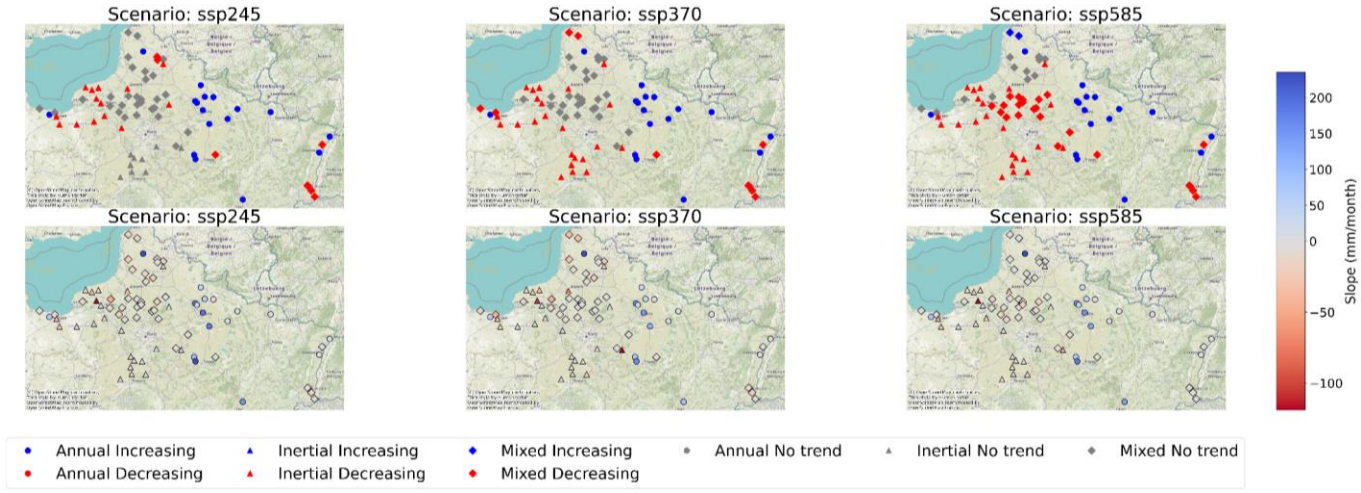


Figure 10: GWL trend on multi-model ensemble median projections from 16 CMIP6 models and 10 DL models for each scenario: Blue (Increasing) and Red (Decreasing)

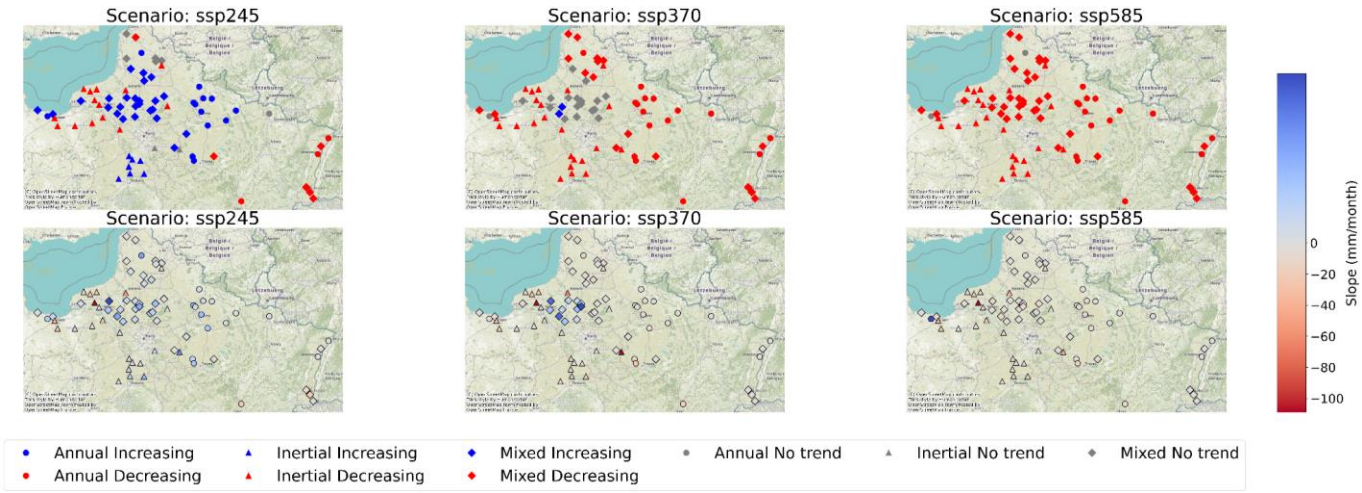
397 The trend in total variability (Figure 11) of annual type is always increasing. In contrast, for mixed and
398 inertial types, it either decreases or shows no significant trends (central and South-central parts of the
399 area).



400

401 *Figure 11: Trend in the total variability of projected GWL*

402 We then focused on low-frequency, i.e. interannual-to-decadal variability only, defined as fluctuations
403 over time scales exceeding five years. Such time scales are usually employed to compare hydro-climatic
404 variability and large-scale climate variations (often along with their moving average) as depicted by
405 climate indices and teleconnections mentioned in the introduction. Figure 12 shows the trends and slopes
406 of low-frequency variance obtained from CWT. Changes in low-frequency variance could impact the
407 occurrence and intensity of multi-year drought or wet periods. Our analysis (from Figure 12) revealed
408 distinct trends across different emission scenarios. Under the SSP 2-4.5 scenario, many stations showed
409 increasing trends. In contrast, the number of stations with increasing trends significantly decreased to
410 very few for the SSP 3-7.0 scenario, while all stations displayed decreasing trends under the SSP 5-8.5
411 scenario.



412

413 *Figure 12: Trends and slopes in low-frequency variability from CWT*

414 To summarise (Figure 11-12), as emission scenarios worsen, the overall variability increases for annual-
415 type stations, whereas interannual to decadal variability decreases for all stations. Generally, more
416 pessimistic emission scenarios correlate with a reduction in long-term groundwater level variability. While
417 the magnitude of variability increase is typically low, it appears more pronounced in less pessimistic
418 scenarios. These findings raise an important question: Could this be an effect of climate change reducing
419 the amplitude of natural climate variability? The results suggest that low-frequency natural climate
420 variability may diminish under more pessimistic emission scenarios, directly impacting interannual to
421 decadal water level variations. This observation warrants further investigation into the complex
422 interactions between anthropogenic climate change and natural climate variability that we did not explore
423 in the current study.

424 ## 5.2 Differences in GWL projections according to different climate models

425 In the previous section, we explored trends in GWL and temporal changes in GWL variability using the
426 median time series computed from all projections generated with different climate models (16 models)
427 and different initialisations (10 initialisations). It is well known that estimating the differences in
428 hydrological projections from climate change models and scenarios is crucial for understanding the
429 sources of uncertainty and communicating contrasted but equally plausible outcomes to stakeholders.
430 This is why we extracted and studied the highest and lowest variability of all projections in section 4
431 (Figures 4 to 6). The results of section 4 showed that large differences between projections would most
432 likely be due to the use of 16 climate model inputs rather than from DL models. Although we did not
433 specifically assess their respective parts in the total uncertainty, comparing the range of uncertainty

related to DL model initialisation on observed data (Figures 4a to 6a) and that obtained for projections (Figures 4b to 6b) showed quite clearly that climate model inputs would result in a much higher difference in GWL projections. In this section, we then explored the differences in trends and changes in variability of projections owing to the 16 climate models taken separately. All the figures associated with this section (Fig. S1 to S3) are provided in the supplementary information.

Similarly to what was done for the ensemble approach in section 5.1, we examined the direction (increasing or decreasing) and slope magnitudes of trends (level and variability) at all stations. As explained in the methodology section, in this case, the time series used for calculating GWL trends corresponded to the median of the 10 DL model initialisations for each climate model and each scenario. For assessing trends (i.e. potential change) in GWL variability (either total or considering low-frequency variations only), the time series used was the scale-average wavelet spectrum of the same median of the 10 DL model initialisations.

Figures S1a and S1b show, respectively, the trend direction and Sen's slopes of GWL projections from 16 CMIP6 models (as lines) and three scenarios each (as columns). Although the results seem rather contrasted, three main outcomes could be distinguished. First, for all SSP scenarios, decreasing or non-significant trends (with very slight decreasing or increasing slopes less than 10 mm/month) appeared to dominate. Amongst all models, most of the statistically significant increasing trends were observed for only SSP 2-4.5 with 4 out of 16 models (CMCC, GFDL, MRI and, to a lower extent, FGOALS). These models showed increasing trends with magnitudes up to more than 25 mm/month. Second, for SSP 3-7.0 and SSP 585 (2nd and 3rd column), all models showed mostly decreasing GWLs and some non-significant trends, except the CanESM5 and MPI-ESM1 models. The magnitude of negative slopes for SSP 3-7.0 and SSP 5-8.5 showed decreasing GWLs at rates as low as, or lower than, -50 mm/month. Although slightly increasing slopes could be identified in many models for those 2 SSP scenarios (e.g., EC-Earth3 SSP 3-7.0, CNRM SSP 3-7.0 and SSP 5-8.5), they were never statistically significant, except for the CanESM5 and MPI-ESM1 models. Third, it was noticeable that for the worse scenario, SSP 5-8.5, all stations located in the most eastern part of the area were affected by decreasing trends (or with no significant trends) for all 16 climate models; these stations consisted mainly of annual-type GWLs.

As done with the ensemble approach, we then investigated how the different climate models led to potential changes in both total GWL variability and low–frequency variability. Figures S2a and b illustrate the results obtained for the total variability of the projected (median) time series. Unlike the results about changes in GWL, it seemed less easy to distinguish any general tendency in terms of changes in time series

total variability according to the emission scenario for any model. As a first step, for the sake of simplicity, examination of trend direction only (Fig. S2) showed the same consistent pattern as previously noticed with the ensemble approach, albeit less clearly, where an obvious distinction between the eastern part (mainly annual-type aquifers) and the western part of the area was quite readily visible (Fig. 10). The magnitude of the slopes confirm this finding: for most models, one could observe either rather strong variability trends separating the eastern (increasing variability, blue-labeled stations) and western regions (decreasing variability, red-labeled stations), or only low-magnitude trends (most of the time not statistically significant). In brief, although differences between climate models exist, the inertial and mixed types (i.e. those with strong low-frequency variability) were mostly affected by decreasing variability through time over the period 2030 to 2100, whereas the annual-type aquifers would be characterised by increasing variability over the same period.

It then seemed that the amplitude low-frequency variability would decrease over time during the 2030 to 2100 period. The results of the ensemble approach showed that apart from SSP 2-4.5, the more pessimistic the emission scenario, the higher the number of stations (and the larger the region) affected (Fig. 12): there is a clear evolution from SSP 2-4.5 to SSP 5-8.5 for low-frequency variability. However, considered individually (Fig. S3), the 16 different models could show noticeable discrepancies in some cases: for instance, INM-CM5, NorESM2 or CanESM5 shows increasing low-frequency variability in terms of trend direction, contrary to many other models for SSP 3-7.0. Despite such disagreements, it seems rather clear that many models still display significantly increasing low-frequency variability for all station types (inertial, mixed or annual) with Sen's slopes of rather high magnitude for SSP 2-4.5, and that conversely, many models would display decreasing variability with strong Sen's slope magnitudes. Decreasing variability is also even more pronounced for SSP 5-8.5 than SSP 3-7.0. These results are consistent with the conclusions of the ensemble approach (Fig. 12). However, they also confirm that using the median time series in the ensemble approach seems well suited to identify clear general trends properly.

6. Discussion and Conclusion

In this study, we aimed to develop projections of GWL under three different climate change scenarios, focusing on three different GWL types: annual, mixed, and inertial. While some aquifers can be rather reactive to seasonal changes, dominated mainly by annual cyclicity, others are sometimes largely

dominated by interannual to decadal variations that are driven by large-scale climate (Hanson et al.,2006; El Janyani et al., 2012; Rust et al.,2018; Baulon et al., 2022). However, most studies dealing with groundwater level simulation or forecasting mainly considered time series represented by seasonal variations with a strong annual cycle governing water level variability. It was then desirable to study more complex aquifer dynamics mainly controlled by internal climate oscillations and assess how these different types of aquifers might be affected by climate change. In this framework, potential changes in variabilities under different SSP scenarios for such aquifers were explored by using deep learning GRU with wavelet pre-processing and CMIP6 bias-corrected precipitation and temperature data as input from 16 climate models. We analysed trends in groundwater levels and changes in variability (i.e. the amplitude of GWL variations) over time across the 2030-2100 period. Ten different initialisations of DL models and 16 climate model inputs resulted in an ensemble of 160 projections for each of the 3 SSP scenarios (SSP 2-4.5, 3-7.0, 5-8.5).

The analysis of the ensemble of 160 projections revealed that the lowest-variability (LV) GWL projections closely aligned with the range of variation of historical observations for annual and mixed types. In other words, only projections with the lowest possible variability would allow for maintaining the same range of variability encountered during the last ~60 decades. This was not the case for inertial-type GWL, for which the LV projection still has a significantly lower variability than previously observed in the historical period. The highest-variability (HV) projected GWL time series exhibited a much greater variability than all observed time series, with standard deviations approximately 1.5 to 2 times higher than for observed GWL. The annual-type GWL showed increasing total variability, while mixed and inertial types indicated decreasing across all scenarios. Additionally, distinct patterns for total variability showed up separating the eastern and western parts of the area (resp. increasing and decreasing variability): it seems like this is because the eastern part comprises almost all annual-type GWL, which are the ones exhibiting increasing variability. On the other hand, low-frequency variability seemed to be decreasing for almost all stations except for the less severe SSP 2-4.5 scenario; as well, more stations tended to exhibit decreasing low-frequency variability as emission scenarios became more severe (with all stations in this case for the most severe scenario).

Our CSMK test results on the median time series of the 160 projection ensemble indicated decreasing trends in groundwater levels for all scenarios and GWL types in northern France. This finding aligns with Habets et al. (2013), who projected a significant decrease in water resources for rivers and aquifers in two northern French basins despite model differences and uncertainties.

Conversely, Vergnes et al. (2023) projected higher groundwater levels on average over France in the future, including northern France, with mean annual GWL increasing by up to approximately 2 m. Vergnes et al. (2023) noted that their results did not match those of Wunsch et al. (2022) and were in contradiction with previous studies from Caballero et al. (2007) or Dayon et al. (2018). However, Wunsch et al. (2022) computed trends and relative changes essentially for the future period (2014-2100) from the projected levels alone, whereas Vergnes et al. (2023) compared the future period (2070-2099) to the historical period (1976-2005). In Germany, Wunsch et al. (2022) projected a median relative decrease in groundwater levels between 2014 and 2100, which seems to agree with our findings on trends in projected levels. It is also interesting to note that these authors showed increased variability in the annual cycle towards 2100, while our results highlighted an increase in the total variability of annual-type GWL in the eastern part of northern France (i.e. closest to Germany).

To facilitate comparison with these studies, we computed the mean difference (expressed as the relative change in %) between future periods and two historical reference periods: 1. Near future (2030-2050), 2. Middle future (2051-2070), 3. Far future (2071-2100), 4. Whole future period (2030-2100). We used 1970-2022 as our primary reference historical period (Figure. 13) and 1976-2005 as a secondary reference period (Figure. 14) to compare with Vergnes et al. (2023) directly. Figures 13 and 14 are used to support this analysis, from which several main points could be underlined. Typically, changes are generally small for all periods and scenarios, ranging from +8% to -6% (1970-2022 reference) and +8% to -10% (1976-2005 reference). The magnitude of changes is similar to Wunsch et al. (2022), although slightly less pronounced. There is a distinct pattern between the western and eastern parts, with future levels higher in the west and lower in the east compared to the historical reference as noticed for projected change in total variability (decreasing variability to the West, increasing to the East). No significant differences were observed between scenarios. Yet, a noticeable difference exists between the periods. Relative change is higher in mid and far-future periods than in the near future. Moreover, in the near future period (2030-2050), more stations indicated positive changes in mean, followed by stabilisation in later periods. These increasing mean changes in some areas align with results from Vergnes et al. (2023).

Our results reconcile those from Vergnes et al. (2023) and Wunsch et al. (2022), although according to Vergnes et al. (2023), they were apparently contradictory. Indeed, while projected GWL is expected to decrease over time (our study and Wunsch et al., 2022), changes in the annual mean projected levels remained slightly higher than historical levels (our study and Vergnes et al., 2023). This intriguing result warrants further investigation and could be an interesting research question for future studies. It is worth

noting that the relative positive change is mainly observed for inertial and mixed GWL types and much lower for annual types. Additionally, the differences between periods seemed consistent with the expected decreasing trends: for the far future, the difference in mean compared with the historical reference period seemed a little less than that of the near future, although the changes are so slight they are barely visible on the maps in figures 13 and 14.

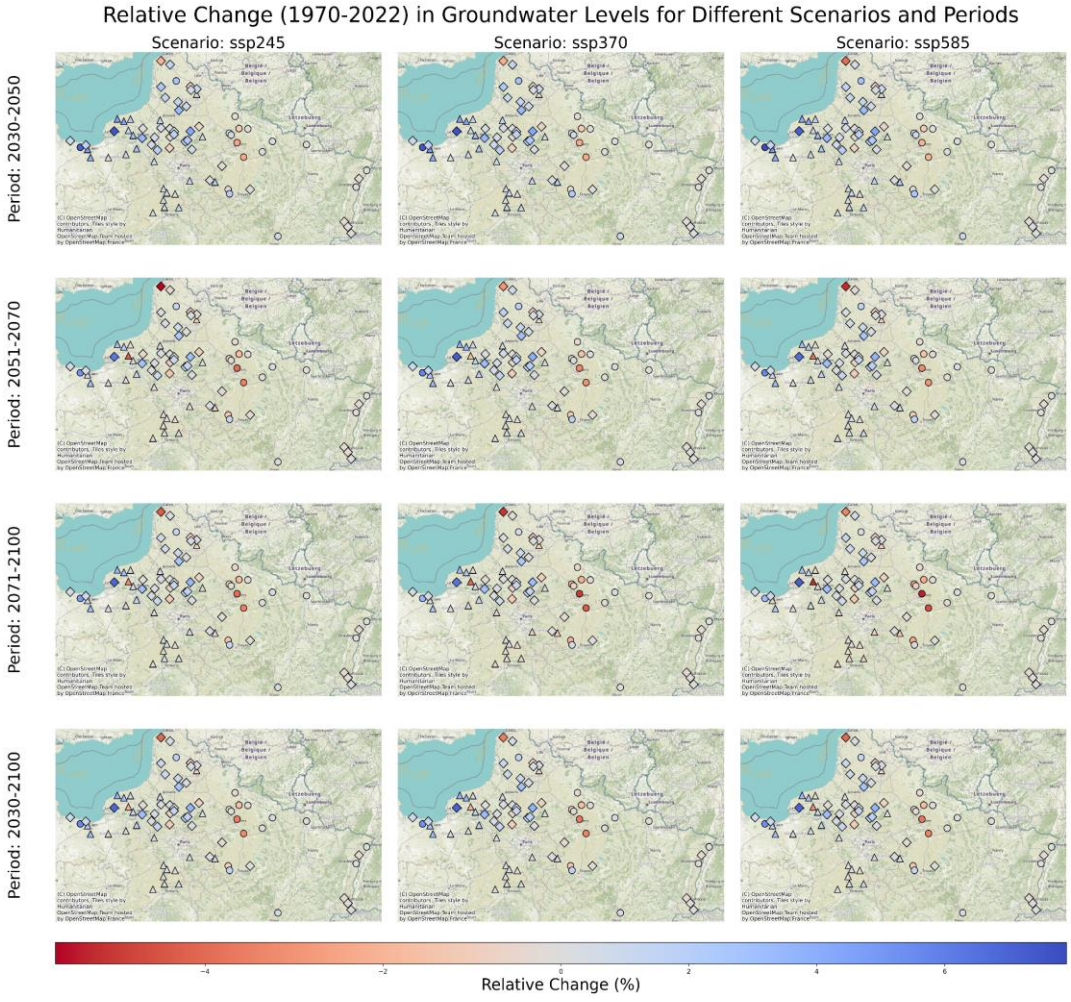


Figure 13: Relative change for future periods with respect to historical mean from 1970-2022

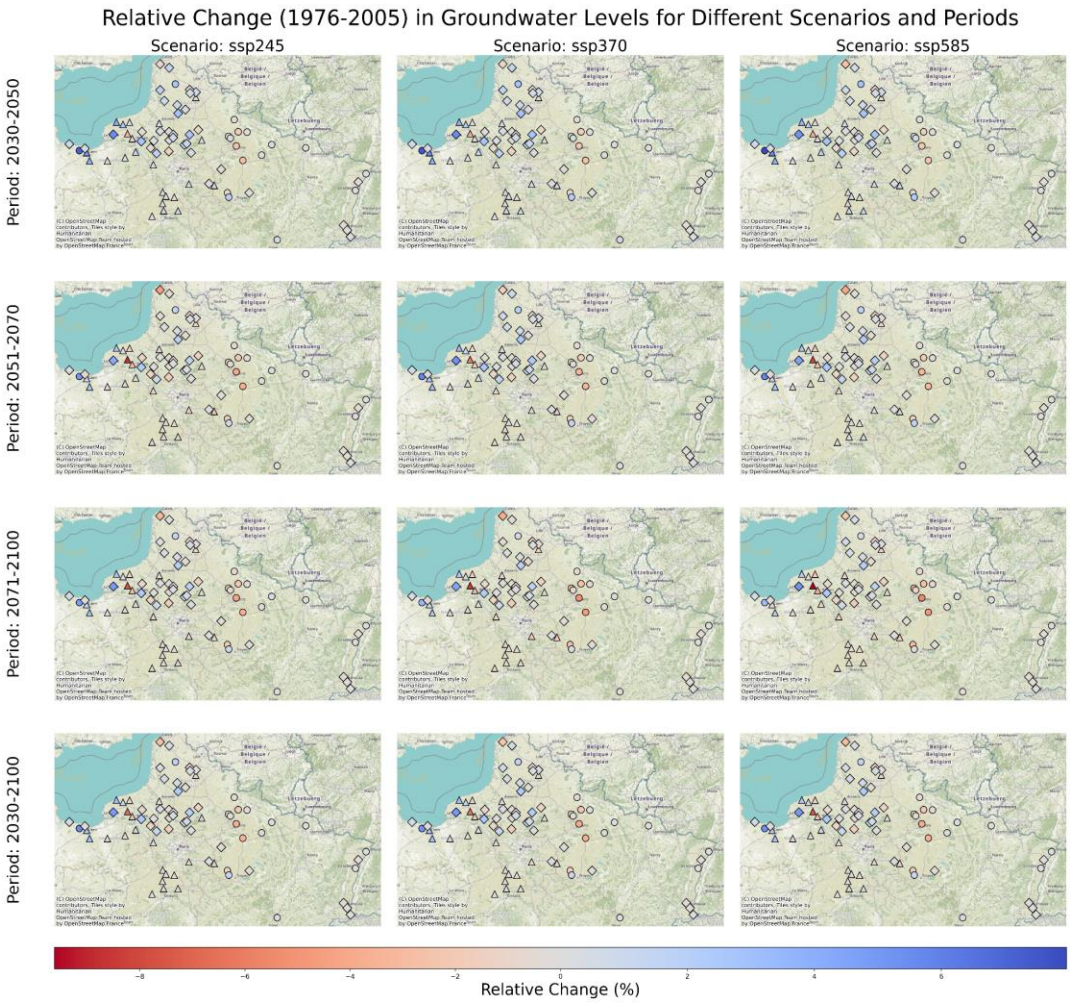


Figure 14: Relative change for future periods with respect to historical mean from 1976-2005 (similar to Vergnes et al. 2023)

Finally, it is also important to note that all these studies used different types of projected climate inputs (CMIP5 RCP or CMIP6 SSP scenarios, CMIP5 Euro-CORDEX regionalised climate projections for France, different ensembles or number of members), which makes direct comparison challenging. This highlights the need for a comprehensive, community-wide benchmarking experiment to understand better and reconcile these differences in future groundwater level projections.

While our study employed advanced techniques like utilising multiple CMIP6 climate models and scenarios, as well as training different initialisations for the GRU deep learning models, there are inherent strengths and limitations associated with these approaches that contribute to the total uncertainty affecting the results. The uncertainty associated with hydrological impact projections arises from multiple origins: it is partly linked to the climate model used (physics, initialisation), to the internal or "natural"

variability, to the different emission scenarios, and ultimately to the hydrological model used. In particular, as mentioned in the introduction, low-frequency/long-time scale, natural variability can play a crucial role in modulating the effects of climate change by amplifying or attenuating (masking) hydroclimatic trends and associated extremes, e.g. as shown by Boé and Habets (2014) and emphasised by Massei et al. (2020). On interannual to decadal scales, climate oscillations and teleconnections such as the North Atlantic Oscillation (NAO) or the El Nino-Southern Oscillation (ENSO) were identified as a significant forcing factor of groundwater resources variations (Holman et al., 2011; Liesch & Wunsch, 2019; Massei et al., 2007, 2010; Rust et al., 2019). In many instances, such fluctuations may correspond to significant hydrological events, as shown in Baulon et al. (2022) and describe multi-year periods of successive dry or wet years, to which human activities are particularly vulnerable. This led Blösch et al. (2019) to classify the understanding of these phenomena as one of the "23 unsolved hydrological problems". The 16 model variants used in the work presented herein still represented the stochastic low-frequency climate variability differently, which allowed us to appreciate its potential impacts on GWL projections. However, using only one variant for precipitation and temperature of each of the 16 climate models prevented us from analysing the contribution of natural variability to the total uncertainty in the GWL projections released. A larger-ensemble approach would be needed to properly assess the contribution of natural climate variability to the uncertainty of GWL projections.

Finally, from a more technical standpoint, the DL models used herein have shown strong performance in identifying complex patterns in large datasets and capturing GWL variability, as also emphasised in our previous studies (Chidepudi et al., 2023a, 2024a), making them efficient for long-term simulations. They have proven to be a relevant alternative or complement to more complex modelling frameworks, like physically-based models, that are often more difficult to develop and implement. Their efficiency could be leveraged to facilitate benchmarking of hydrological and hydrogeological projections under climate change by conducting more similar studies globally.

Acknowledgements:

Climate scenarios were from the NEX-GDDP-CMIP6 dataset, prepared by the Climate Analytics Group and NASA Ames Research Center using the NASA Earth Exchange and distributed by the NASA Center for Climate Simulation (NCCS). SKR Chidepudi acknowledges the funding from BRGM and Normandie Region.

Data Availability Statement

All data used in this study are publicly available. The ERA5 reanalysis dataset can be accessed from Hersbach et al. (2020). Groundwater level data were obtained from the ADES (Accès aux Données sur les Eaux Souterraines) database (Winckel et al., 2022). CMIP6 projections were sourced from the NEX-GDDP-CMIP6 dataset (Thrasher et al., 2021).

References

Atawneh, D. A., Cartwright, N., & Bertone, E. (2021). Climate change and its impact on the projected values of groundwater recharge: A review. Journal of Hydrology, 601, 126602. https://doi.org/10.1016/j.jhydrol.2021.126602

Baulon, L., Allier, D., Massei, N., Bessiere, H., Fournier, M., & Bault, V. (2022). Influence of low-frequency variability on groundwater level trends. Journal of Hydrology, 606, 127436. https://doi.org/10.1016/j.jhydrol.2022.127436

Bhasme, P., Vagadiya, J., & Bhatia, U. (2022). Enhancing predictive skills in physically-consistent way: Physics Informed Machine Learning for hydrological processes. Journal of Hydrology, 615, 128618. https://doi.org/10.1016/j.jhydrol.2022.128618

Boé, J. and Habets, F. (2014). Multi-decadal river flow variations in France, Hydrol. Earth Syst. Sci., 18, 691–708, https://doi.org/10.5194/hess-18-691-2014

Boo, K. B. W., El-Shafie, A., Othman, F., Khan, M. M. H., Birima, A. H., & Ahmed, A. N. (2024). Groundwater level forecasting with machine learning models: A review. Water Research, 252, 121249. https://doi.org/10.1016/j.watres.2024.121249

Caballero, Y., Voirin-Morel, S., Habets, F., Noilhan, J., LeMoigne, P., Lehenaff, A., & Boone, A. (2007). Hydrological sensitivity of the Adour-Garonne river basin to climate change. Water Resources Research, 43(7). https://doi.org/10.1029/2005WR004192

Chakraborty, D., Başağaoğlu, H., Gutierrez, L., & Mirchi, A. (2021). Explainable AI reveals new hydroclimatic insights for ecosystem-centric groundwater management. Environmental Research Letters, 16(11), 114024. https://doi.org/10.1088/1748-9326/ac2fde

Cho, K., van Merriënboer, B., Bahdanau, D., Bengio, Y., 2014. On the properties of neural machine translation. Encoder-Decod. Approaches 103–111. https://aclanthology.org/W14-4012

Chidepudi, S. K. R., Massei, N., Jardani, A., Henriot, A., Allier, D., & Baulon, L. (2023a). A wavelet-assisted deep learning approach for simulating groundwater levels affected by low-frequency variability. Science of The Total Environment, 865, 161035. https://doi.org/10.1016/j.scitotenv.2022.161035

Chidepudi, S. K. R., Massei, N., Henriot, A., Jardani, A., and Allier, D. (2023b). Local vs regionalised deep learning models for groundwater level simulations in the Seine basin., EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-3535, https://doi.org/10.5194/egusphere-egu23-3535

Chidepudi, S. K. R., Massei, N., Jardani, A., & Henriot, A. (2024a). Groundwater level reconstruction using long-term climate reanalysis data and deep neural networks. Journal of Hydrology: Regional Studies, 51, 101632. https://doi.org/10.1016/j.ejrh.2023.101632

Chidepudi, S. K. R., Massei, N., Jardani, A., Dieppois, B., Henriot, A., and Fournier, M. (2024b). Training deep learning models with a multi-station approach and static aquifer attributes for groundwater level simulation: what's the best way to leverage regionalised information?, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-794,

Clark, M.P., Wilby, R.L., Gutmann, E.D. et al. Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. Curr Clim Change Rep 2, 55–64 (2016). https://doi.org/10.1007/s40641-016-0034-x

Costantini, M., Colin, J., & Decharme, B. (2023). Projected Climate-Driven Changes of Water Table Depth in the World's Major Groundwater Basins. Earth's Future, 11(3), e2022EF003068. https://doi.org/10.1029/2022EF003068

Dayon, G., Boé, J., Martin, É., & Gailhard, J. (2018). Impacts of climate change on the hydrological cycle over France and associated uncertainties. Comptes Rendus Geoscience, 350(4), 141-153. https://doi.org/10.1016/j.crte.2018.03.001

661 Deser, C., Phillips, A., Bourdette, V. et al. Uncertainty in climate change projections: the role of
662 internal variability. Clim Dyn 38, 527–546 (2012). https://doi.org/10.1007/s00382-010-0977-x

663 Deser, C. and Phillips, A. S. (2023). A range of outcomes: the combined effects of internal
664 variability and anthropogenic forcing on regional climate trends over Europe, Nonlin. Processes
665 Geophys., 30, 63–84, https://doi.org/10.5194/npg-30-63-2023

666 El Janyani, S., Massei, N., Dupont, J., Fournier, M., & Dörfliger, N. (2012). Hydrological responses
667 of the chalk aquifer to the regional climatic signal. Journal of Hydrology, 464-465, 485-493.
668 https://doi.org/10.1016/j.jhydrol.2012.07.040

669 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016).
670 Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design
671 and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016

672 Habets, F., Boé, J., Déqué, M. et al. Impact of climate change on the hydrogeology of two basins
673 in northern France. Climatic Change 121, 771–785 (2013). https://doi.org/10.1007/s10584-013-
674 0934-x

675 Halloran, L. J., Millwater, J., Hunkeler, D., & Arnoux, M. (2023). Climate change impacts on
676 groundwater discharge-dependent streamflow in an alpine headwater catchment. Science of The
677 Total Environment, 902, 166009. https://doi.org/10.1016/j.scitotenv.2023.166009

678 Hanson, R.T., Dettinger, M.D. & Newhouse, M.W. Relations between climatic variability and
679 hydrologic time series from four alluvial basins across the southwestern United States. Hydrogeol
680 J 14, 1122–1146 (2006). https://doi.org/10.1007/s10040-006-0067-7

681 Hawkins, E., & Sutton, R. (2009). The Potential to Narrow Uncertainty in Regional Climate
682 Predictions. Bulletin of the American Meteorological Society, 90(8), 1095-1108.
683 https://doi.org/10.1175/2009BAMS2607.1

684 Hauswirth, S. M., Bierkens, M. F., Beijk, V., & Wanders, N. (2021). The potential of data driven
685 approaches for quantifying hydrological extremes. Advances in Water Resources, 155, 104017.
686 https://doi.org/10.1016/j.advwatres.2021.104017

687 Heudorfer, B., Liesch, T., and Broda, S. (2024). On the challenges of global entity-aware deep
688 learning models for groundwater level prediction, Hydrol. Earth Syst. Sci., 28, 525–543,
689 https://doi.org/10.5194/hess-28-525-2024

690 Holman, I.P., Rivas-Casado, M., Bloomfield, J.P. et al. Identifying non-stationary groundwater
691 level response to North Atlantic ocean-atmosphere teleconnection patterns using wavelet
692 coherence. Hydrogeol J 19, 1269–1278 (2011). https://doi.org/10.1007/s10040-011-0755-9

693 Hussain et al., (2019). pyMannKendall: a python package for non parametric Mann Kendall family
694 of trend tests. Journal of Open Source Software, 4(39), 1556,
695 https://doi.org/10.21105/joss.01556

696 Jasechko, S., Seybold, H., Perrone, D., Fan, Y., Shamsudduha, M., Taylor, R. G., Fallatah, O., &
697 Kirchner, J. W. (2023). Rapid groundwater decline and some cases of recovery in aquifers globally.
698 Nature, 625(7996), 715-721. https://doi.org/10.1038/s41586-023-06879-8

699 Kayhomayoon, Z., Jamnani, M. R., Rashidi, S., Ghordoyee Milan, S., Arya Azar, N., & Berndtsson,
700 R. (2023). Soft computing assessment of current and future groundwater resources under CMIP6
701 scenarios in northwestern Iran. Agricultural Water Management, 285, 108369.
702 https://doi.org/10.1016/j.agwat.2023.108369

703 Klavans, J. M., Clement, A. C., Cane, M. A., & Murphy, L. N. (2022). The Evolving Role of External
704 Forcing in North Atlantic SST Variability over the Last Millennium. Journal of Climate, 35(9), 2741-
705 2754. https://doi.org/10.1175/JCLI-D-21-0338.1

706 Labat, D., Ababou, R., & Mangin, A. (2000). Rainfall–runoff relations for karstic springs. Part II:
707 Continuous wavelet and discrete orthogonal multiresolution analyses. Journal of Hydrology,
708 238(3-4), 149-178. https://doi.org/10.1016/S0022-1694(00)00322-X

709 Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., Knutti, R., and Hawkins,
710 E. (2020). Partitioning climate projection uncertainty with multiple large ensembles and
711 CMIP5/6, Earth Syst. Dynam., 11, 491–508, https://doi.org/10.5194/esd-11-491-2020

712 Liesch, T., & Wunsch, A. (2019). Aquifer responses to long-term climatic periodicities. Journal of
713 Hydrology, 572, 226-242. https://doi.org/10.1016/j.jhydrol.2019.02.060

Long, J., Wang, L., Chen, D., Li, N., Zhou, J., Li, X., Guo, X., Liu, H., Chai, C., & Fan, X. (2024). Hydrological Projections in the Third Pole Using Artificial Intelligence and an Observation-Constrained Cryosphere-Hydrology Model. Earth's Future, 12(4), e2023EF004222. https://doi.org/10.1029/2023EF004222

Mahmood, R., Donat, M. G., Ortega, P., Doblas-Reyes, F. J., Delgado-Torres, C., Samsó, M., and Bretonnière, P.-A. (2022). Constraining low-frequency variability in climate projections to predict climate on decadal to multi-decadal timescales – a poor man's initialized prediction system, Earth Syst. Dynam., 13, 1437–1450, https://doi.org/10.5194/esd-13-1437-2022

Martel, J.-L., Brissette, F., Troin, M., Arsenault, R., Chen, J., Su, T., & Lucas-Picher, P. (2022). CMIP5 and CMIP6 model projection comparison for hydrological impacts over North America. Geophysical Research Letters, 49, e2022GL098364. https://doi.org/10.1029/2022GL098364

Massei, N., Kingston, D. G., Hannah, D. M., Vidal, J.-P., Dieppois, B., Fossa, M., Hartmann, A., Lavers, D. A., and Laignel, B. (2020). Understanding and predicting large-scale hydrological variability in a changing environment, Proc. IAHS, 383, 141–149, https://doi.org/10.5194/piahs-383-141-2020

Massei, N., Dieppois, B., Hannah, D.M., Lavers, D.A., Fossa, M., Laignel, B., Debret, M., (2017). Multi-time-scale hydroclimate dynamics of a regional watershed and links to large-scale atmospheric circulation: application to the Seine river catchment, France. J. Hydrol. 546, 262–275. https://doi.org/10.1016/j.jhydrol.2017.01.008

Massei, N., Laignel, B., Deloffre, J., Mesquita, J., Motelay, A., Lafite, R., & Durand, A. (2010). Long-term hydrological changes of the Seine River flow (France) and their relation to the North Atlantic Oscillation over the period 1950–2008. International Journal of Climatology, 30(14), 2146-2154. https://doi.org/10.1002/joc.2022

Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J. J. F., Clark, M. P., Uijlenhoet, R., and Teuling, A. J. (2018). Mapping (dis)agreement in hydrologic projections, Hydrol. Earth Syst. Sci., 22, 1775–1791, https://doi.org/10.5194/hess-22-1775-2018

Mishra, V., Dangar, S., Tiwari, V. M., Lall, U., & Wada, Y. (2024). Summer monsoon drying accelerates India's groundwater depletion under climate change. Earth's Future, 12, e2024EF004516. https://doi.org/10.1029/2024EF004516

Müller, L. and Döll, P. (2024). Quantifying and communicating uncertain climate change hazards in participatory climate change adaptation processes, Geosci. Commun., 7, 121–144, https://doi.org/10.5194/gc-7-121-2024

Neves, M. C., Jerez, S., & Trigo, R. M. (2018). The response of piezometric levels in Portugal to NAO, EA, and SCAND climate patterns. Journal of Hydrology, 568, 1105-1117. https://doi.org/10.1016/j.jhydrol.2018.11.054

Nourani, V., Ghareh Tapeh, A. H., Khodkar, K., & Huang, J. J. (2023). Assessing long-term climate change impact on spatiotemporal changes of groundwater level using autoregressive-based and ensemble machine learning models. Journal of Environmental Management, 336, 117653. https://doi.org/10.1016/j.jenvman.2023.117653

Nozari, S., Bailey, R. T., Haacker, E. M., Zambreski, Z. T., Xiang, Z., & Lin, X. (2022). Employing machine learning to quantify long-term climatological and regulatory impacts on groundwater availability in intensively irrigated regions. Journal of Hydrology, 614, 128511. https://doi.org/10.1016/j.jhydrol.2022.128511

Percival, D.B., Walden, A.T., (2000). Wavelet Methods for Time Series Analysis. Cambridge University Press https://doi.org/10.1017/CBO9780511841040

Quilty, J., & Adamowski, J. (2018). Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework. Journal of Hydrology, 563, 336-353. https://doi.org/10.1016/j.jhydrol.2018.05.003

Rajaee, T., Ebrahimi, H., & Nourani, V. (2019). A review of the artificial intelligence methods in groundwater level modeling. Journal of Hydrology, 572, 336-351. https://doi.org/10.1016/j.jhydrol.2018.12.037

766 Rehana, S., & Rajesh, M. (2023). Assessment of impacts of climate change on Indian riverine
767 thermal regimes using hybrid deep learning methods. Water Resources Research, 59,
768 e2021WR031347. https://doi.org/10.1029/2021WR031347

769 Roshani, A., Hamidi, M. (2022). Groundwater Level Fluctuations in Coastal Aquifer: Using Artificial
770 Neural Networks to Predict the Impacts of Climatical CMIP6 Scenarios. Water Resour Manage 36,
771 3981–4001 https://doi.org/10.1007/s11269-022-03204-2

772 Rust, W., Holman, I., Bloomfield, J., Cuthbert, M., and Corstanje, R. (2019). Understanding the
773 potential of climate teleconnections to project future groundwater drought, Hydrol. Earth Syst.
774 Sci., 23, 3233–3245, https://doi.org/10.5194/hess-23-3233-2019

775 Scanlon, B. R., Fakhreddine, S., Rateb, A., De Graaf, I., Famiglietti, J., Gleeson, T., Grafton, R. Q.,
776 Jobbagy, E., Kebede, S., Kolusu, S. R., Konikow, L. F., Long, D., Mekonnen, M., Schmied, H. M.,
777 Mukherjee, A., MacDonald, A., Reedy, R. C., Shamsudduha, M., Simmons, C. T., . . . Zheng, C.
778 (2023). Global water resources and the role of groundwater in a resilient water future. Nature
779 Reviews Earth & Environment, 4(2), 87-101. https://doi.org/10.1038/s43017-022-00378-6

780 Secci, D., Giovanna Tanda, M., D'Oria, M., & Todaro, V. (2023). Artificial intelligence models to
781 evaluate the impact of climate change on groundwater resources. Journal of Hydrology, 627,
782 130359. https://doi.org/10.1016/j.jhydrol.2023.130359

783 Tao, H., Hameed, M. M., Marhoon, H. A., Zounemat-Kermani, M., Heddam, S., Kim, S., Sulaiman,
784 S. O., Tan, M. L., Sa'adi, Z., Mehr, A. D., Allawi, M. F., Abba, S., Zain, J. M., Falah, M. W., Jamei, M.,
785 Bokde, N. D., Bayatvarkeshi, M., Al-Mukhtar, M., Bhagat, S. K., . . . Yaseen, Z. M. (2022).
786 Groundwater level prediction using machine learning models: A comprehensive review.
787 Neurocomputing, 489, 271-308. https://doi.org/10.1016/j.neucom.2022.03.014

788 Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., Longuevergne, L., Leblanc,
789 M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M.
790 F., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., . . . Treidel, H. (2013). Ground water and
791 climate change. Nature Climate Change, 3(4), 322-329. https://doi.org/10.1038/nclimate1744

Terray, L. (2012). Evidence for multiple drivers of North Atlantic multi-decadal climate variability. Geophysical Research Letters, 39(19). https://doi.org/10.1029/2012GL053046

Thrasher, B., Wang, W., Michaelis, A. Nemani, R. (2021). NEX-GDDP-CMIP6. NASA Center for Climate Simulation. https://doi.org/10.7917/OFSG3345

Torrence, C., & Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological Society, 79(1), 61–78. https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Uc-Castillo, J. L., Marín-Celestino, A. E., Martínez-Cruz, D. A., Tuxpan-Vargas, J., & Ramos-Leal, J. A. (2023). A systematic review and meta-analysis of groundwater level forecasting with machine learning techniques: Current status and future directions. Environmental Modelling & Software, 168, 105788. https://doi.org/10.1016/j.envsoft.2023.105788

Vergnes, J. P., Caballero, Y., & Lanini, S. (2023). Assessing climate change impact on French groundwater resources using a spatially distributed hydrogeological model. Hydrological Sciences Journal, 68(2), 209–227. https://doi.org/10.1080/02626667.2022.2150553

Wi, S., & Steinschneider, S. (2022). Assessing the physical realism of deep learning hydrologic model projections under climate change. Water Resources Research, 58, e2022WR032123. https://doi.org/10.1029/2022WR032123

Winckel, A., Ollagnier, S., Gabillard, S., (2022). Managing groundwater resources using a national reference database: the French ADES concept. SN Appl. Sci. 4 (8), 1–12. https://doi.org/10.1007/S42452-022-05082-0

Wu, Y., Miao, C., Slater, L., Fan, X., Chai, Y., & Sorooshian, S. (2024). Hydrological Projections under CMIP5 and CMIP6: Sources and Magnitudes of Uncertainty. Bulletin of the American Meteorological Society, 105(1), E59-E74. https://doi.org/10.1175/BAMS-D-23-0104.1

Wunsch, A., Liesch, T., & Broda, S. (2022). Deep learning shows declining groundwater levels in Germany until 2100 due to climate change. Nature Communications, 13(1), 1-13. https://doi.org/10.1038/s41467-022-28770-2

818 Xiong, J., Guo, S., & Kinouchi, T. (2022). Leveraging machine learning methods to quantify 50
819 years of dwindling groundwater in India. Science of The Total Environment, 835, 155474.
820 https://doi.org/10.1016/j.scitotenv.2022.155474

821 Yuan, Q., Thorarinsdottir, T. L., Beldring, S., Wong, W. K., & Xu, C. (2023). Assessing uncertainty
822 in hydrological projections arising from local-scale internal variability of climate. Journal of
823 Hydrology, 620, 129415. https://doi.org/10.1016/j.jhydrol.2023.129415