

Deep learning shows declining groundwater levels in Germany until 2100 due to climate change

Andreas Wunsch (✉ andreas.wunsch@kit.edu)

Karlsruhe Institute of Technology <https://orcid.org/0000-0002-0585-9549>

Tanja Liesch

Karlsruhe Institute of Technology

Stefan Broda

Federal Institute for Geosciences and Natural Resources

Article

Keywords: climate change, groundwater resources, machine learning, groundwater levels

Posted Date: April 22nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-420056/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Nature Communications on March 9th, 2022. See the published version at <https://doi.org/10.1038/s41467-022-28770-2>.

Abstract

In this study we investigate how climate change will directly influence the groundwater resources in Germany during the 21st century. We apply a machine learning groundwater level prediction framework, based on convolutional neural networks to 118 sites well distributed over Germany to assess the groundwater level development under the RCP8.5 scenario, based on six selected climate projections, which represent 80% of the bandwidth of the possible future climate signal for Germany. We consider only direct meteorological inputs, while highly uncertain anthropogenic factors such as groundwater extractions are excluded. We detected significant declining trends of groundwater levels for most of the sites, revealing a spatial pattern of stronger decreases especially in the northern and eastern part of Germany, emphasizing already existing decreasing trends in these regions. We can further show an increased variability and longer periods of low groundwater levels during the annual cycle towards the end of the century.

Introduction

Climate change is increasingly altering water availability even in generally water-rich areas like Germany, where overall water stress is currently low¹. Nevertheless, hot and dry summers in recent years (especially 2018-2020) led to ongoing exceptional droughts^{2,3} with severe consequences for agriculture and ecology, such as drought damages in forests, reduced crop yields and extreme low flows in rivers. Drought effects accumulated over years, because winter precipitation did not compensate summer deficits. This applies not only, but especially to groundwater resources, which are of major importance since drinking water supply in Germany is strongly dependent on groundwater and springs (almost 70%)⁴. Declining groundwater levels due to generally reduced groundwater recharge and higher water demand in summer, regionally forced water suppliers to exploit their current maximum capacity during dry periods to meet the demand; locally even water supply shortages occurred. During future dry periods strong usage conflicts can be expected in areas of low water availability between water suppliers and industry (process and cooling water), additionally amplified by increasing agricultural irrigation demand, which currently has only minor significance with less than 2% of the total withdrawal volume¹. Knowledge of future groundwater level development, especially in the long-term, is therefore crucial to develop sustainable groundwater management plans to meet future demands, solve usage conflicts and protect ecosystems.

Climate change affects groundwater in several direct and indirect ways⁵. Major direct drivers are changes in precipitation, snowmelt and evapotranspiration⁶. For Germany, climate projections show opposing trends in terms of water availability, with a slight increase in annual precipitation sums, i.e. more water, but at the same time a significant temperature increase of several degrees Celsius by 2100^{34,35}, i.e. less water. The effect on groundwater resources is therefore not directly clear and needs to be analyzed. For Europe in general, higher precipitation is generally expected during winter, which in combination with a generally decreasing amount of snow, thus increasing direct infiltration, leads to higher groundwater recharge during winter and less in spring. Especially for snow dominated regions this might cause changes of seasonality⁶. Weather extremes are expected to intensify, therefore longer droughts and more frequent intense rainfall events will occur⁵. Generally higher temperatures cause higher atmospheric water demand, thus increasing evapotranspiration, leading to less infiltration and therefore less groundwater recharge. Especially unconfined, shallow aquifers are most

likely to be sensitive to direct climate change effects⁷. Indirect climate change influences on groundwater are mostly related to anthropogenic groundwater withdrawals or associated with land-use changes⁵. It is known that the groundwater storage reduction caused by pumping could easily far exceed natural recharge^{6,8}. The impact of these factors will be exacerbated as water demand increases to meet the needs of regionally growing population (mainly due to growing urban areas), industry and agricultural irrigation.

In recent years, artificial neural network (ANN) approaches have proven their usefulness in predicting groundwater levels⁹⁻¹⁴, even using a highly transferable approach with purely climatic input parameters (e.g. ref¹⁵). In a previous study¹⁵ we showed that 1D-Convolution Neural Networks (CNNs) are a good choice for groundwater level simulation, as they can provide high accuracy and furthermore are fast and reliable. Unlike physically-based models, which usually require a very good knowledge of local conditions and need to be time-consumingly built and calibrated, data-driven models such as ANNs are able to predict a target variable using only relevant driving forces. This makes studies on larger areas easier and is therefore the method of choice for this study. To the authors' knowledge, no comprehensive direct evaluation of groundwater level development until 2100 exists for Germany yet. Besides a rather old small-scale study¹⁶ also a regional-scale study for the Danube basin has been conducted to date¹⁷. The latter uses several dynamically-coupled, process-based model components and the authors found strongly declining groundwater levels with declines of up to 10 m close to the Alps in southernmost Germany for their scenario period (2036–2060). Further, several studies investigated future groundwater recharge in different contexts for subregions of Germany using mainly water balance models or process-based models¹⁷⁻²². Furthermore, the application of ANNs to study groundwater level development in the long-term and in the context of climate change for a larger area like Germany has not been performed yet. Related studies with applications of ANNs either used a very small number of wells²³⁻²⁵ and limited time horizons^{23,24} or use ANNs without directly presenting future climate signals to the ANN²⁵. In case of streamflow runoff simulation, however, ANNs have been successfully applied to analyze the future development under climate change influences in several catchments all over California²⁶ as well as two catchments in China^{27,28}.

In this study we use a 1D-CNN approach²⁹ to build 118 site-specific models, well distributed all over Germany in the respective uppermost unconfined aquifer, which are able to predict weekly groundwater levels with high accuracy using only precipitation and temperature as inputs in the past. We visually check the model output plausibility under an artificial extreme climate scenario²⁶ and investigate how the model has learned input-output relationships using an explainable AI approach (SHAP³⁰). We then use the trained CNN models to investigate the future groundwater level development for the selected sites, using precipitation and temperature derived from the RCP8.5 scenario³¹ of bias-corrected and downscaled (5 x 5 km²) climate projection data³² from six climate models as inputs. These six climate models were preselected by the German Meteorological Service to represent 80% of the possible future climate signal ("core-ensemble")³³ for Germany. Table 1 lists these projections, which are part of the EURO-CORDEX Ensemble and assigns them an abbreviation that will be used as a synonym in the remaining part of the paper. As we use purely climatic input parameters we can only project the influence of direct climate change effects, while secondary, most certainly stronger indirect effects, such as increased groundwater pumping, are not included in this study. However, due to high prediction accuracy in the past, the selected sites are unlikely to be under the influence of strong groundwater

withdrawals or comparable effects, and are therefore suitable for predicting that part of the future groundwater level trend that results from direct climatic influences, as long as the basic input-output relationships remain unchanged.

Table 1: Climate projections used in this study and according abbreviations used throughout the text. For more information on the models please visit <https://www.euro-cordex.net/>.

Projection	Abbrev.
CCCma-CanE SM2_rcp85_r1i1p1_CLMcom-CCLM4-8-17	p1
ICHEC-EC-EARTH_rcp85_r1i1p1_KNMI-RACMO22E	p2
MIROC-MIROC5_rcp85_r1i1p1_GERICS-REMO2015	p3
MOHC-HadGEM2-ES_rcp85_r1i1p1_CLMcom-CCLM4-8-17	p4
MPI-M-MPI-ESM-LR_rcp85_r1i1p1_UHOH-WRF361H	p5
MPI-M-MPI-ESM-LR_rcp85_r2i1p1_MPI-CSC-REMO2009	p6

Generally, climate projections show a slight increase in precipitation sums and a significant temperature increase of several degrees Celsius for Germany by 2100^{34,35}, exact values depending on the scenario considered³⁵. Figure 1 shows the change of total annual precipitation (A1-A3) and annual average temperature (B1-B3) for each of the climate projections used in this study in 2100, compared to the start of our investigation in 2014. The change is derived from a linear trend analyses at the 118 sites, that are subject to further investigations in this study. Boxplots (A2, B2) show only significant ($p < 0.05$) changes, according numbers are shown in subplots A3 and B3, further, the order within subplots A1 and B1 does not correspond to the numbering of the projections but to the strength and direction of the trend. We see that many projections of total annual precipitation do not show any significant trend and are therefore marked in grey (especially p2, p4 and p5, s.a. Figure 1-A3). However, for almost all sites we observe significant declines of up to -450 mm per year for p1, but at the same time increases of up to 296 mm per year for p3 and especially p6. Some projections are therefore diverging until 2100 in terms of precipitation sums, which shows that we cover a large range of a possible climate signal under the RCP8.5 scenario. Despite many non-significant trends, a spatial pattern of significant changes with a decreasing tendency in northwestern Germany and less clear increasing tendency in eastern Germany is visible. Strongest decreases are projected to occur in southernmost Germany; however, especially in southern but also in eastern areas two opposing trends usually occur at one site, so the development is not unambiguous. Compared to precipitation sums, the development of the annual average temperature is more consistent for all projections. Overall, temperature increases between 2.4°C and 5.8°C occur. On average, p1 shows the strongest increases, followed by p4. Together with the decreasing precipitation sums, p1 therefore shows the probably most challenging development in terms of water availability compared to the other projections used in this study. Spatially, we observe lighter temperature increases in north-western Germany, which most certainly is linked to a buffer effect near the coast.

Results

Individual projection results

For each of the examined 118 test sites, we simulated the future weekly groundwater level development based on six climate projections (s.a. Table 1). Since these climate projections differ considerably in detail for individual future time periods, we also obtained six different future groundwater level simulations, which should only be interpreted on the basis of longer time periods (at least 30 years)³⁶. Figure 2 depicts the trend

as the relative development in percent of the annual mean for each of the six projections (A) as well as the annual upper extreme (97.5%) quantile (B) and the annual lower extreme (2.5%) quantile (D) for all test sites in 2100, compared to the start of the simulation (2014) and normalized on the individual historic range as explained in the methods section. For each site, all relative developments are shown ordered by the strength of the change, the order does therefore not correspond to the numbering of the projections. The given boxplots in Figure 2C provide more detailed information for the three maps as well as on the development of the 25% and the 75% quantiles, relative and absolute values of the presented changes are given in Table 2. The values of the non-significant trends are not shown in the boxplots, which has to be kept in mind for interpretation, especially for quantiles with many non-significant trends (compare Table 2).

In case of the mean, approximately 54% of all simulations (387 of 708, i.e. six projections for each of the 118 sites) show a significant trend until 2100. At least one of the projected developments is always considered significant ($p < 0.05$) for each site, which, however, also means that there are several sites with mainly non-significant trends (grey). The large majority of the significant trends is negative with a median ranging between -23% in case of p1 and -6.6% in case of p6 (Table 2). In Figure 2C we observe that p1 systematically shows the strongest declines until 2100, being significant for 117 of the 118 wells. The overall maximum decline is -46%, clearly indicating the different character of p1 compared to the other projections. Especially projections p3-p5 show more moderate changes of the mean (median ranges from -8% to -13%), with many non-significant trends (35%-54%). Simulations based on p2 and p6 only find significant trends for around 30% of all sites and additionally are moderate in their significant results. Three projections (p2, p3, but mainly p6, compare Table 2) even show some positive developments until 2100, however overall, such developments are rare and occur at sites, where other projections simultaneously show at least non-significant or even negative trends. In absolute numbers the mentioned median changes are in the order of -0.1 m to -0.4 m, which is highly dependent on the individual groundwater level range at each site. Despite many non-significant and some positive trends, there is a clear tendency of declining mean groundwater levels until 2100. Additionally, we can observe a slight spatial tendency with more and stronger significant negative trends in some areas of northern and eastern Germany, where we also find the strongest overall relative declines. In southern Germany many wells show several non-significant trends and also most positive changes can be found scattered in this region, however, some of the southernmost wells show very strong declines for single simulations, comparably to the strong declines in eastern Germany.

In case of the upper extreme value quantile (97.5%) this spatial pattern is partly confirmed. In Figure 2B we clearly observe many significant declines in eastern Germany, while the large majority (>70%) of the trends in whole Germany is considered to be non-significant. Increasing trends are found comparably often for the 97.5% quantiles, with increases up to 20%. Comparing the projections with each other (Figure 2C), we find a similar behavior as before: p1 shows the strongest significant decreases (down to -47%), p3, p4 and p5 tend to move in the moderate negative range (medians around -12%), while p2 and p6 more often show positive trends (positive medians of the significant trends). We therefore observe partly a contradictory development of the upper extreme values compared to the mean. The absolute numbers of the mentioned changes again are in the order of few tens of centimeters upwards and downwards. The strongest simulated absolute increase (max. of p6) is almost 5 meters, however, in a karstic well in southern Germany, which has a high variability anyway.

The tendency of declining groundwater levels we observed for the mean, gets clearer for the lower extreme values (2.5% quantile) shown in Figure 2D. We still observe 36% non-significant trends, however the remaining 65% show almost exclusively negative changes with a maximum decline of -81% (Table 2). The median change of the 2.5% quantile of all projections ranges between -38% for p1, which again shows the strongest declines, followed by p4 (-21%), as well as p2, p3, p5 and p6 with a median change around -10% each. The latter four, and especially of them p6, contain the majority of non-significant trends, the changes shown in the boxplots therefore tend to be overestimated. There are only few sites where only one result is considered significant. These occur mainly near the Baltic Sea coast, the central and eastern part of northern Germany, and the central area of southern Germany. In the latter, however, there are at the same time quite strong relative decreases, just as we also find them in eastern Germany and in the western part of northern Germany. This pattern is largely consistent with the spatial pattern of the mean mentioned above. Most median decreases (p2-p6) are in the order of -0.1 to -0.4 m, for p1 the median decrease reaches even -0.7 m for the annual lower extreme value quantile. All projections except p6 agree that of all significant changes, at least a decrease of -0.1 m will be observed (max. values for 2.5% quantile in Table 2).

Considering all results, we see a clear tendency toward declining groundwater levels overall, with stronger declines for lower quantiles, i.e. groundwater level lows will occur more frequently and will be more severe in the future. At the same time, mostly no or even increasing trends are found for upper extreme values, which means that the overall variability will increase significantly by the end of the century.

Table 2: Detailed numbers for each projection on relative changes (left), already shown as boxplots (Figure 2C). Right tables show associated absolute changes in meters.

Relative [%]	p1	p2	p3	p4	p5	p6	mean
97.5% quantile	18.9	17.3	23.9	13.7	13.3	20.6	17.9
Upper Quartile	-12.3	10.5	2.2	-8.5	-6.0	13.9	-0.1
Median	-17.8	7.5	-12.0	-12.3	-10.7	10.7	-5.8
Lower Quartile	-23.5	-9.3	-15.6	-16.9	-14.2	-4.7	-14.0
Min	-46.8	-16.3	-30.4	-31.9	-30.6	-16.5	-28.7
No. of sign. samples	45	20	31	34	32	39	201

Absolute [m]	p1	p2	p3	p4	p5	p6	mean
97.5% quantile	2.1	3.1	0.6	1.2	0.2	4.0	1.9
Upper Quartile	-0.2	0.2	0.0	-0.1	-0.1	0.5	0.0
Median	-0.3	0.1	-0.2	-0.2	-0.2	0.2	-0.1
Lower Quartile	-0.6	-0.2	-0.3	-0.4	-0.4	0.0	-0.3
Min	-1.3	-0.4	-0.7	-0.7	-0.8	-0.3	-0.7
No. of sign. samples	45	20	31	34	32	39	201

Relative [%]	p1	p2	p3	p4	p5	p6	mean
75% quantile	18.3	14.0	22.9	12.5	9.4	20.2	16.2
Upper Quartile	-10.6	8.4	-6.7	-8.2	-7.4	12.0	-2.1
Median	-16.3	-7.9	-9.4	-11.1	-8.9	9.0	-7.4
Lower Quartile	-22.2	-12.3	-15.3	-16.6	-13.0	-8.0	-14.6
Min	-44.1	-16.3	-30.6	-33.5	-24.5	-17.7	-27.8
No. of sign. samples	64	25	46	45	47	40	267

Absolute [m]	p1	p2	p3	p4	p5	p6	mean
75% quantile	2.1	2.5	2.0	1.8	0.1	4.8	2.2
Upper Quartile	-0.2	0.1	-0.1	-0.1	-0.1	0.3	0.0
Median	-0.3	-0.1	-0.2	-0.2	-0.2	0.2	-0.1
Lower Quartile	-0.4	-0.2	-0.3	-0.3	-0.3	-0.1	-0.3
Min	-1.6	-0.6	-0.7	-0.7	-0.9	-0.3	-0.8
No. of sign. samples	64	25	46	45	47	40	267

Relative [%]	p1	p2	p3	p4	p5	p6	mean
mean	-6.8	8.5	11.7	-6.5	-4.3	15.0	2.9
Upper Quartile	-17.8	-6.4	-7.8	-9.7	-6.8	7.6	-6.8
Median	-22.9	-8.5	-10.6	-12.7	-8.4	-6.6	-11.6
Lower Quartile	-28.1	-11.9	-12.8	-17.5	-12.1	-9.3	-15.3
Min	-46.0	-18.2	-27.0	-31.4	-22.3	-14.2	-26.5
No. of sign. samples	117	35	66	76	54	39	387

Absolute [m]	p1	p2	p3	p4	p5	p6	mean
mean	-0.2	1.5	1.0	-0.1	-0.1	4.1	1.1
Upper Quartile	-0.3	-0.1	-0.1	-0.2	-0.1	0.2	-0.1
Median	-0.4	-0.1	-0.2	-0.3	-0.2	-0.1	-0.2
Lower Quartile	-0.6	-0.2	-0.3	-0.5	-0.3	-0.1	-0.4
Min	-6.5	-1.1	-3.6	-5.0	-1.1	-0.4	-3.0
No. of sign. samples	117	35	66	76	54	39	387

Relative [%]	p1	p2	p3	p4	p5	p6	mean
25% quantile	-12.2	-5.0	-4.6	-7.3	-5.0	10.0	-4.0
Upper Quartile	-29.9	-8.3	-9.1	-13.4	-7.8	-7.7	-12.7
Median	-34.9	-11.3	-12.3	-17.6	-10.0	-9.0	-15.8
Lower Quartile	-42.2	-14.2	-15.5	-22.8	-14.0	-10.2	-19.8
Min	-67.0	-23.9	-25.1	-41.4	-25.9	-15.5	-33.2
No. of sign. samples	118	53	64	96	41	38	410

Absolute [m]	p1	p2	p3	p4	p5	p6	mean
25% quantile	-0.3	-0.1	-0.1	-0.1	-0.1	1.0	0.1
Upper Quartile	-0.5	-0.2	-0.2	-0.2	-0.1	-0.1	-0.2
Median	-0.6	-0.2	-0.2	-0.3	-0.2	-0.1	-0.3
Lower Quartile	-1.0	-0.3	-0.4	-0.6	-0.3	-0.2	-0.5
Min	-12.6	-3.0	-3.8	-7.6	-1.1	-2.8	-5.1
No. of sign. samples	118	53	64	96	41	38	410

Relative [%]	p1	p2	p3	p4	p5	p6	mean
2.5% quantile	-12.6	-4.2	-5.1	-7.5	-4.0	8.0	-4.2
Upper Quartile	-31.1	-8.8	-8.8	-15.6	-7.8	-7.2	-13.2
Median	-37.7	-10.9	-11.5	-20.8	-9.6	-9.7	-16.7
Lower Quartile	-45.9	-15.0	-14.4	-25.9	-12.9	-11.3	-20.9
Min	-80.8	-27.8	-26.7	-45.1	-25.1	-15.5	-36.8
No. of sign. samples	118	72	66	102	60	37	455

Absolute [m]	p1	p2	p3	p4	p5	p6	mean
2.5% quantile	-0.3	-0.1	-0.1	-0.1	-0.1	1.4	0.1
Upper Quartile	-0.5	-0.2	-0.2	-0.3	-0.1	-0.1	-0.2
Median	-0.7	-0.2	-0.2	-0.4	-0.2	-0.1	-0.3
Lower Quartile	-1.0	-0.3	-0.4	-0.7	-0.3	-0.2	-0.5
Min	-15.0	-4.1	-3.4	-9.9	-2.0	-3.7	-6.3
No. of sign. samples	118	72	66	102	60	37	455

Figure 3 shows the detailed development at four selected sites (black boxes in Figure 2). For each site we plot the six projected groundwater level time series for the far future (2070-2100) (A1-D1), as well as the complete simulations, separately as heatmaps with years as row and weeks as columns (A2-D2). The time series plots show the diverging development of some projections in the far future, however, there is no strict sequence of

projections in terms of absolute groundwater height, the order can change throughout the years. Most heatmaps show the development described above by displaying generally declining groundwater levels (more and darker red, as well as lighter or constant blue shadings towards 2100 in the lower part of the heatmaps). What we additionally see now is that the length of low groundwater levels increases (red shadings get wider) for all sites. The time of higher groundwater levels throughout the year shows two possible developments of either getting shorter (blue shadings get narrower, e.g. B2-p1 or even change to red, e.g. D2-p4) or staying constant in length (width of blue shadings does not change, e.g. A2-p2 and A2-p6), with optionally even increasing peak height (darker blue, e.g. A2-p6). In both plot types we can also recognize sequences of several more extreme years, such as several dry years around 2090 in B1-p4, which also reflects in a dark-red stripe in the corresponding heatmap (B2-p4). Such sequences are especially critical because effects accumulate and dependent ecosystem are not able to recover but are instead particularly vulnerable to further damage in subsequent years due to reduced resilience.

Average projection results

We consolidated the separate projection results for each site into one by calculating the mean of the significant trends shown in Figure 2. Only sites with at least 4 (thus the majority) significant results are included, the rest is depicted as not significant on average. Results are shown in Figure 4. The development of the mean is depicted in the upper left map and we find, that according to the aforementioned definition, 41% of the wells (49 of 118) are considered significant on average and on median show a change of -13%. We do not find any wells with increasing mean trends and observe a similar spatial pattern as before with strongest decreases in eastern Germany. For wells in southern Germany we observe noticeably many non-significant changes. All in all, we simulated significant average decreases between -0.2 m to -2.4 m for about 25 wells, and at least a decrease of -10 cm for all 49 wells in Figure 4A (max. abs. value of the mean in Figure 4D). In case of the upper extreme value quantile (97.5%) we can summarize that the consolidated results show mainly no trends, especially for southern Germany, they will therefore probably remain at their current level. Few sites (5), all of them in northern Germany, are expected to show increased upper extreme values up to a maximum of 15% or 1.5 m, however, we still observe a spatial pattern of decreasing upper extreme values in eastern Germany up to -30% or -0.7 m. Hence, in this area the groundwater levels probably will decrease in every part of the annual cycle and with comparably high certainty (many consistent significant simulations). This applies also to the lower extreme values (2.5% quantile) that show on average significant decreases for more than half of the examined sites all over Germany with median decreases of -19% (equivalent to -0.3 m, comp. Figure 4C, D). On this map, no clear spatial pattern is recognizable any longer.

Annual maximum and minimum timing aspects

Besides the relative and absolute developments of the groundwater height, we also investigated timing aspects of the groundwater dynamics. For a possible shift of the annual minimum (Figure 5) we found significant ($p < 0.05$) results for p1 (41 of 118) and also p4 (33 of 118), with median shifts of 3.4 and 3.1 weeks (positive, i.e. later). A spatial pattern exists, showing significant and stronger shifts with increasing proximity to the coast in the north and no or even negative (i.e. earlier) shifts in the south. However, please note that most results are not significant and the shown pattern may only serve as an indication for further interpretation.

Even fewer significant shifts were found in case of the annual maximum timing (not shown). Especially for snow dominated regions a shift of the peak timing from spring towards the winter is expected in the context of climate change, however, Germany as a whole cannot be considered snow-dominated. This is in accordance with our findings, because we found mainly non-significant shifts (< 10 per projection). Only in case of p4 we detected a slightly larger number of significant shifts (29 of 118). Here, the maximum even occurs on median 4 weeks later during the annual cycle, in contrary to the expected shift for snow-dominated regions.

Model input analysis

From the combined analysis of our groundwater level simulations and the model inputs shown in the introduction, we can conclude, that temperature is mainly the driving factor for declining groundwater levels, rather than precipitation. This applies because mostly no significant or even increasing precipitation is projected, our models, however, still frequently show declining groundwater level tendencies, which therefore most likely are caused by the significantly increased temperature until the end of the century. Therefore, our results are consistent with other studies, which indicate that the reduction in water availability in the future is driven primarily by changes in temperature³⁴.

This reflects also in the model interpretability approach (SHAP values) we used to check the plausibility of our model outputs. The minimum SHAP value for T is mostly lower than the minimum SHAP value observed for P (except for eight sites); i.e. the models have learned that high temperatures can cause stronger decreasing groundwater levels than low precipitation. This is, however, only an interpretation of what was learned, which agrees with our conception. A causality cannot be derived from this.

Discussion

The results of our simulations show a nation-wide decrease in groundwater levels by the end of the century. The absolute changes may seem small, but the fact that we investigated almost exclusively shallow aquifers and sites with very small depths to groundwater, reinforces the importance of the results, predominantly in terms of water availability for vegetation and agriculture. A decline of several tens of centimeters (depending on the projection and the area) can be vital for plants during hot and dry periods, if, as a result, the groundwater is no longer accessible. Furthermore, a related study showed, that for large parts of northern Germany, a decline of the groundwater levels by 10 cm can be considered critical in terms of altered streamflow discharge due to reduced baseflow from groundwater⁸. This has already been visible during the last two years, when simultaneously to low groundwater levels also extremely low water levels in surface waters (even until running dry) have been observed³. Our results show a clearer tendency of declining groundwater levels in the North and East compared to the South (Figure 4A), which emphasizes the already existing trends and patterns. However, in the southernmost part of Germany, for some individual projections, we find also some of the strongest declines (Figure 2). It is very important to note that the assessed results are only long-term averages of a future development. As recent developments showed, the succession of several dry years is much more critical than the overall trend. In such periods, the projected effects accumulate over consecutive years to extremely low groundwater levels, and thus more severe consequences are to be expected. Such longer dry periods are most likely to be averaged out, in a linear trend analysis, as performed in this study, but their existence can be seen in the examples shown in Figure 3. Future research should pay attention to this aspect more intensively. It is also

important to recall that we model simply direct climate effects on groundwater levels, thus the change is based on the development of temperature and precipitation until the end of the century only, and we assume that the basic input-output relationship or system behavior does not change. However, it can be expected that in many cases, the system behavior will be influenced by future changes in groundwater extractions, changes in vegetation and land use, as well as surface sealing and other related factors. Groundwater withdrawals in particular, are expected to increase due to regionally growing population especially in metropolitan areas (drinking water demand) and increasing demand for industry, energy and especially irrigated agriculture. As a result, the groundwater level will inevitably drop further if no active measures, such as limitation of withdrawals, avoidance of irrigated agriculture by changing crop types or even artificial recharge by infiltration, are applied. Despite all these limitations, the results give a good impression of the magnitude of changes to be expected purely due to direct climatic influences.

Methods

Data

We used weekly groundwater level data from 118 different sites, well distributed all over Germany (Figure 6A). All wells are located in the unconfined, uppermost (thus mostly shallow) aquifers, which are most likely to be subject to direct climate change effects. Greater depths to groundwater are predominantly found in fractured and karstic aquifers. For additional details on the sites please refer to our supplementary material. Groundwater level records of all sites show very different lengths (Figure 6B), from 15 to 67 years, with a median length of 36 years. Data gaps were closed using information of several related groundwater level time series with highly correlated dynamics³⁷. Information on interpolated values are included into the dataset (see section data availability).

Input parameters for our models are purely climatic: precipitation (P) and temperature (T). They are widely available and easy to measure in the past and present, and also well evaluated in terms of climate projection output. Precipitation serves as proxy for groundwater recharge, temperature for evapotranspiration. Additionally, temperature usually shows a distinct annual cycle, which also provides the models with valuable information on seasonality. Since we specifically selected wells with high forecast accuracy in the past (see Model Calibration and Evaluation), we can assume that groundwater dynamic at these wells is mainly dominated by climate forcings. As long as no fundamental change of the system relations occurs (e.g. newly installed groundwater pumping or severe changes in land use nearby), we can expect reasonable results for our simulations.

Besides the groundwater level data itself, we based our analysis on several datasets. The models were trained using data from the HYRAS dataset^{38,39}, which is a gridded (5x5 km²) meteorological dataset based on observed data from meteorological stations ranging from 1951 to 2015. To evaluate the influence of climate change we used RCP8.5 scenario data from six selected climate projections that form the so called core-ensemble defined by DWD³³. The core-ensemble is specifically selected for Germany and derived from a larger set of 21 climate projections ('reference-ensemble')³³ to represent 80% of the bandwidth of the possible future climate signal. Further, we received the projection data bias adjusted onto the HYRAS dataset and regionalized on a 5x5 km² grid by ref³².

Convolutional neural networks (CNNs)

Convolutional Neural Networks (CNNs)⁴⁰ are commonly used for image recognition and classification tasks but also work well on sequential data, such as groundwater level time series²⁹. The CNNs used in this study comprise a 1D-Convolutional layer with fixed kernel size (3) and optimized number of filters, followed by a Max-Pooling layer and a Monte-Carlo dropout layer, applying a fixed dropout of 50% to prevent the model from overfitting. A dense layer with optimized size follows, succeeded by a single output neuron. We used the Adam optimizer for a maximum of 100 training epochs with an initial learning rate of 0.001 and applied gradient clipping to prevent exploding gradients. Early stopping algorithm with a patience of 15 epochs was applied as another regularization technique to prevent the model from overfitting the training data. Several model hyperparameters (HP) were optimized using Bayesian optimization⁴¹: training batch size (16 to 256); input sequence length (1 to 52 weeks); number of filters in the 1D-Conv layer (1 to 256); size of the first dense layer (1 to 256). All models were implemented using Python 3.8⁴², the deep-learning framework TensorFlow⁴³ and its Keras⁴⁴ API. Further, the following libraries were used: Numpy⁴⁵, Pandas^{46,47}, Scikit-Learn⁴⁸, BayesOpt⁴¹, Matplotlib⁴⁹, Unumpy⁵⁰ and SHAP³⁰.

Model calibration and evaluation

In this study we used weekly groundwater level time series data of varying length (Figure 6B). To find the best model configuration we split every time series into four parts: training set, validation set, optimization set and test set. The test set uses always the 4-year period from 2012 to 2016 (Figure 7B, s.a. Figure 8A for an example), for few sites where the time series ended slightly earlier, we shifted the test set accordingly. The first 80% of the remaining time series before 2012 were used for training, the following 20% for early stopping (validation set) and for testing during HP optimization (optimization set), using 10% of the remaining time series each (Figure 7B). As acquisition function during HP optimization we chose the sum of Nash-Sutcliffe efficiency (NSE) and squared Pearson r (R^2) (compare ref¹⁵), because in this study we used mainly these two criteria to judge the accuracy of the final optimized model in the test section. For each model we used a maximum optimization step number of 150 or stopped after 15 steps without improvement once a minimum of 60 steps was reached. Generally, we scaled the data to [-1,1] and used an ensemble of 10 pseudo-randomly initialized models to reduce the dependency towards the random number generator seed. For each of the ten ensemble members, we applied Monte-Carlo dropout during simulation to estimate the model uncertainty from 100 realizations each. We derived the 95% confidence interval from these 100 realizations by using 1.96 times the standard deviation of the resulting distribution for each time step. Each uncertainty was propagated while calculating the overall ensemble median value for final evaluation in the test set (2012-2016). We calculated several metrics to judge the simulation accuracy: Nash-Sutcliffe efficiency (NSE), squared Pearson r (R^2), absolute and relative root mean squared error (RMSE/rRMSE), as well as absolute and relative Bias (Bias/rBias). Note that we calculate NSE with a long term mean GWL before the test set. Please see ref²⁹ for more details on calculation as the same approach was used. We use only wells, at which the models showed a very high forecast accuracy in the test-set (mostly NSE and R^2 larger than 0.8, compare Figure 7A). Some models were included with slightly lower accuracy (at least NSE and R^2 larger than 0.7) to improve the spatial coverage resulting in a set of 118 wells from all over Germany. For additional details on the error measures and hyperparameters for all sites please refer to our supplementary material.

Model plausibility and interpretability

To perform groundwater level simulations until 2100 we retrained all models using the defined hyperparameters and all data until 2014. Hence, we split the time series only in two parts: 80% for training and 20% for early stopping (Figure 7B). Afterwards, we assessed the model stability and the plausibility of the output values in the extrapolated regime accordingly to ref²⁶ by evaluating the model output using artificially altered input data based on historical observed climatology with quadruple precipitation and systematically 5°C higher temperature (Figure 8B). As long as the model output does not “blow up” or produce meaningless outputs²⁶, we hereby improve confidence in the model output when investigating the RCP8.5 climate change scenario. Models showing implausible behavior were not considered for this study. We additionally applied an explainable AI approach to check, whether the models have learned correctly in terms of our hydrogeological understanding. We calculated SHAP³⁰ values that explain the influence (sign and strength) of every input feature value on the model output (Figure 8C). Generally, our models showed that the relationship between input and output was captured plausibly. For example, high precipitation inputs (red) produce high SHAP values and therefore have a strongly positive influence on the model output, which corresponds to our basic understanding of the influence of recharge, leading to increasing groundwater levels. Low or no precipitation (blue) has a comparably very slight negative influence on GWL, whereas high temperature inputs (red) have a strong negative influence on the model output. Again, this corresponds with our basic understanding of the governing processes, where high temperature usually means high evapotranspiration, which causes less recharge or even direct groundwater evaporation in some cases. This sounds trivial, however, during preliminary work for this study we found that not all models capture these relations correctly, which also partly caused erroneous values in the extrapolated regime. Figure 8 exemplarily summarizes the model evaluation (A) and plausibility checks (B, C) for one well. Check the supplement for the respective figures of all other sites.

Evaluation of the groundwater results

For our simulation results until 2100, we examined the relative development of the mean and the following quantiles over time: 2.5% (lower extreme quantile), 25% (lower quartile), 75% (upper quartile), and 97.5% (upper extreme quantile). All were site-specifically calculated on a yearly basis for each individual projection, followed by a linear trend analysis. In doing so, we are able to capture both the range and the individual development of all considered future climate projections. To make comparisons between different sites possible, results are normalized on the individual range of each historic groundwater level time series between the years 2000 and 2014 (start of simulation). Even though all climate projections are bias-adjusted on the HYRAS training dataset, they still do not depict the real climate development for individual years (also historically), which can cause a bias between the end of historic data records and the start of our simulations. We therefore investigated the trend of the aforementioned quantities between the start of the simulation and the end in 2100 and did not directly consider the end of the historic records. We examined each quantity development using Mann-Kendall linear trend test⁵¹ and derived the relative development in percent from a linear fit using Theil-Sen slope. We considered a trend significant for $p < 0.05$.

Declarations

Data availability

All groundwater level data are available free of charge from the respective websites of the local authorities. We used data interpolated based on previous knowledge and therefore publish the used data with the kind permission of the local authorities under:

<https://doi.org/10.5281/zenodo.4683879>

All climate data are available on request and free of charge for non-commercial purposes from the German Meteorological Service.

Code availability

The code necessary to reproduce our results is available on GitHub under:

<https://github.com/AndreasWunsch/Long-Term-GWL-Simulations>

Author contributions

All authors contributed to conceptualization of this study. AW and TL contributed to the methodology, AW further contributed to writing the software code, validation, formal analysis, investigation, visualization and wrote the original draft. All authors contributed to reviewing and editing the draft. TL and SB both supervised the work and were involved in project administration.

Funding

Open Access funding enabled and organized by Project DEAL.

Competing interests

The authors declare no competing interests.

References

1. UBA. Trockenheit in Deutschland – Fragen und Antworten. *Umweltbundesamt* <https://www.umweltbundesamt.de/themen/trockenheit-in-deutschland-fragen-antworten> (2020).
2. UFZ. UFZ Dürremonitor Deutschland. <https://www.ufz.de/index.php?de=37937> (2021).
3. Gunter Wriedt. *Grundwasserbericht Niedersachsen: Sonderausgabe zur Grundwasserstandssituation in den Trockenjahren 2018 und 2019*. (2020).
4. Destatis. Wassergewinnung: Bundesländer, Jahre, Wasserarten. *GENESIS-Online* <https://www-genesis.destatis.de/genesis//online?operation=table&code=32211-0002&bypass=true&levelindex=0&levelid=1611589342283#abreadcrumb> (2016).
5. Taylor, R. G. *et al.* Ground water and climate change. *Nature Climate Change* **3**, 322 (2012).
6. Wu, W.-Y. *et al.* Divergent effects of climate change on future groundwater availability in key mid-latitude aquifers. *Nat Commun* **11**, 3710 (2020).
7. Kløve, B. *et al.* Climate change impacts on groundwater and dependent ecosystems. *Journal of Hydrology* **518**, 250–266 (2014).

8. de Graaf, I. E. M., Gleeson, T., (Rens) van Beek, L. P. H., Sutanudjaja, E. H. & Bierkens, M. F. P. Environmental flow limits to global groundwater pumping. *Nature* **574**, 90–94 (2019).
9. Afzaal, H., Farooque, A. A., Abbas, F., Acharya, B. & Esau, T. Groundwater Estimation from Major Physical Hydrology Components Using Artificial Neural Networks and Deep Learning. *Water* **12**, 5 (2020).
10. Bowes, B. D., Sadler, J. M., Morsy, M. M., Behl, M. & Goodall, J. L. Forecasting Groundwater Table in a Flood Prone Coastal City with Long Short-term Memory and Recurrent Neural Networks. *Water* **11**, 1098 (2019).
11. Guzman, S. M., Paz, J. O. & Tagert, M. L. M. The Use of NARX Neural Networks to Forecast Daily Groundwater Levels. *Water Resources Management* **31**, 1591–1603 (2017).
12. Jeong, J. & Park, E. Comparative applications of data-driven models representing water table fluctuations. *Journal of Hydrology* **572**, 261–273 (2019).
13. Müller, J. *et al.* Surrogate optimization of deep neural networks for groundwater predictions. *J Glob Optim* (2020) doi:10.1007/s10898-020-00912-0.
14. Zhang, A., Winterle, J. & Yang, C. Performance comparison of physical process-based and data-driven models: a case study on the Edwards Aquifer, USA. *Hydrogeol J* (2020) doi:10.1007/s10040-020-02169-z.
15. Wunsch, A., Liesch, T. & Broda, S. Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrology and Earth System Sciences* **25**, 1671–1687 (2021).
16. Eckhardt, K. & Ulbrich, U. Potential impacts of climate change on groundwater recharge and streamflow in a central European low mountain range. *Journal of Hydrology* **284**, 244–252 (2003).
17. Barthel, R. *et al.* Integrated Modeling of Global Change Impacts on Agriculture and Groundwater Resources. *Water Resour Manage* **26**, 1929–1951 (2012).
18. Wegehenkel, M. & Kersebaum, K.-C. An assessment of the impact of climate change on evapotranspiration, groundwater recharge, and low-flow conditions in a mesoscale catchment in Northeast Germany. *Journal of Plant Nutrition and Soil Science* **172**, 737–744 (2009).
19. Neukum, C. & Azzam, R. Impact of climate change on groundwater recharge in a small catchment in the Black Forest, Germany. *Hydrogeol J* **20**, 547–560 (2012).
20. Kersebaum, K. C. & Nendel, C. Site-specific impacts of climate change on wheat production across regions of Germany using different CO₂ response functions. *European Journal of Agronomy* **52**, 22–32 (2014).
21. Kreins, P., Henseler, M., Anter, J., Herrmann, F. & Wendland, F. Quantification of Climate Change Impact on Regional Agricultural Irrigation and Groundwater Demand. *Water Resour Manage* **29**, 3585–3600 (2015).
22. Herrmann, F., Kunkel, R., Ostermann, U., Vereecken, H. & Wendland, F. Projected impact of climate change on irrigation needs and groundwater resources in the metropolitan area of Hamburg (Germany). *Environ Earth Sci* **75**, 1104 (2016).
23. Ghazi, B., Jeihouni, E. & Kalantari, Z. Predicting groundwater level fluctuations under climate change scenarios for Tasuj plain, Iran. *Arab J Geosci* **14**, 115 (2021).
24. Jeihouni, E., Eslamian, S., Mohammadi, M. & Zareian, M. J. Simulation of groundwater level fluctuations in response to main climate parameters using a wavelet–ANN hybrid technique for the Shabestar Plain, Iran. *Environ Earth Sci* **78**, 293 (2019).

25. Idrizovic, D. *et al.* Impact of climate change on water resource availability in a mountainous catchment: A case study of the Toplica River catchment, Serbia. *Journal of Hydrology* **587**, 124992 (2020).
26. Duan, S., Ullrich, P. & Shu, L. Using Convolutional Neural Networks for Streamflow Projection in California. *Front. Water* **2**, 28 (2020).
27. Lee, D., Lee, G., Kim, S. & Jung, S. Future Runoff Analysis in the Mekong River Basin under a Climate Change Scenario Using Deep Learning. *Water* **12**, 1556 (2020).
28. Gao, C. *et al.* Projected streamflow in the Huaihe River Basin (2010–2100) using artificial neural network. *Stoch Environ Res Risk Assess* **24**, 685–697 (2010).
29. Wunsch, A., Liesch, T. & Broda, S. Groundwater Level Forecasting with Artificial Neural Networks: A Comparison of LSTM, CNN and NARX. *Hydrology and Earth System Sciences Discussions* **2020**, 1–23 (2020).
30. Lundberg, S. M. & Lee, S.-I. A Unified Approach to Interpreting Model Predictions. in *Advances in Neural Information Processing Systems* (eds. Guyon, I. *et al.*) vol. 30 4765–4774 (Curran Associates, Inc., 2017).
31. Moss, R. *et al.* *Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies*. 132 http://ipcc-data.org/docs/ar5scenarios/IPCC_Final_Draft_Meeting_Report_3May08.pdf (2008).
32. Brienen, S. *et al.* Klimawandelbedingte Änderungen in Atmosphäre und Hydrosphäre: Schlussbericht des Schwerpunktthemas Szenarienbildung (SP-101) im Themenfeld 1 des BMVI-Expertenetzwerks. (2020) doi:10.5675/expnbs2020.2020.02.
33. DWD. Kern-Ensemble v2018. https://www.dwd.de/DE/klimaumwelt/klimaforschung/klimaprojektionen/fuer_deutschland/fuer_dtld_rcp-datensatz_node.html (2018).
34. Thober, S., Marx, A. & Boeing, F. *Auswirkungen der globalen Erwärmung auf hydrologische und agrarische Dürren und Hochwasser in Deutschland*. 20 (2018).
35. Marx, A. *et al.* Climate change alters low flows in Europe under a 1.5, 2, and 3 degree global warming. 24 (2017).
36. Kreienkamp, F., Huebener, H., Linke, C. & Spekat, A. Good practice for the usage of climate model simulation results - a discussion paper. *Environ Syst Res* **1**, 9 (2012).
37. Wunsch, A. & Liesch, T. *Entwicklung und Anwendung von Algorithmen zur Berechnung von Grundwasserständen an Referenzmessstellen auf Basis der Methode Künstlicher Neuronaler Netze*. 191 https://www.bgr.bund.de/DE/Themen/Wasser/Projekte/laufend/F+E/Mentor/mentor-abschlussbericht-1.pdf?__blob=publicationFile&v=2 (2020).
38. Frick, C. *et al.* Central European high-resolution gridded daily data sets (HYRAS): Mean temperature and relative humidity. *Meteorologische Zeitschrift* **23**, 15–32 (2014).
39. Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A. & Gratzki, A. A Central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorol. Z.* **22** (2013) doi:10.1127/0941-2948/2013/0436.
40. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
41. Nogueira, F. *Bayesian Optimization: Open source constrained global optimization tool for Python*. (2014).

42. van Rossum, G. Python tutorial. (1995).
43. Abadi, M. *et al.* TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. 19 (2015).
44. Chollet, F. *Keras*. (Keras, 2015).
45. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.* **13**, 22–30 (2011).
46. McKinney, W. Data Structures for Statistical Computing in Python. in 56–61 (2010). doi:10.25080/majora-92bf1922-00a.
47. Reback, J. *et al.* *pandas-dev/pandas: Pandas 1.0.3*. (Zenodo, 2020). doi:10.5281/ZENODO.3509134.
48. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).
49. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* **9**, 90–95 (2007).
50. Lebigot, E. O. *Uncertainties: a Python package for calculations with uncertainties*. (2010).
51. Md. Manjurul Hussain Shourov, Ishtiak Mahmud & Niemeyer, K. *mmhs013/pyMannKendall: v1.4.1*. (Zenodo, 2020). doi:10.5281/ZENODO.3876036.

Figures

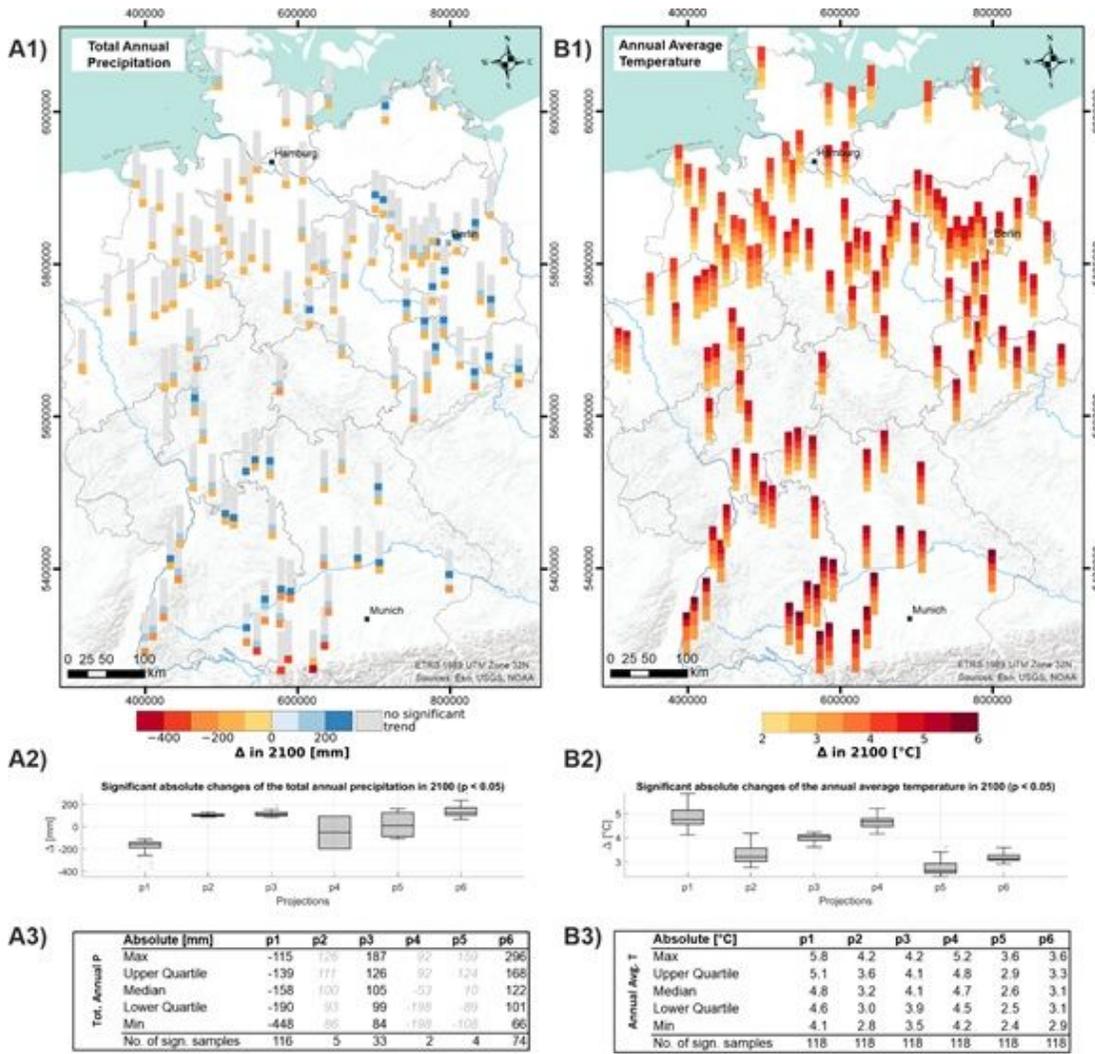


Figure 1

Absolute changes of total annual precipitation sums (A) and annual average temperature (B) projected by climate models for the relevant sites used in this study. Single squares depict results of a single projection, ordered by the strength and sign of the change. A2 and B2 summarize all significant ($p < 0.05$) results from A1 and B1, Tables (A3 and B3) give detailed numbers on the boxplots.

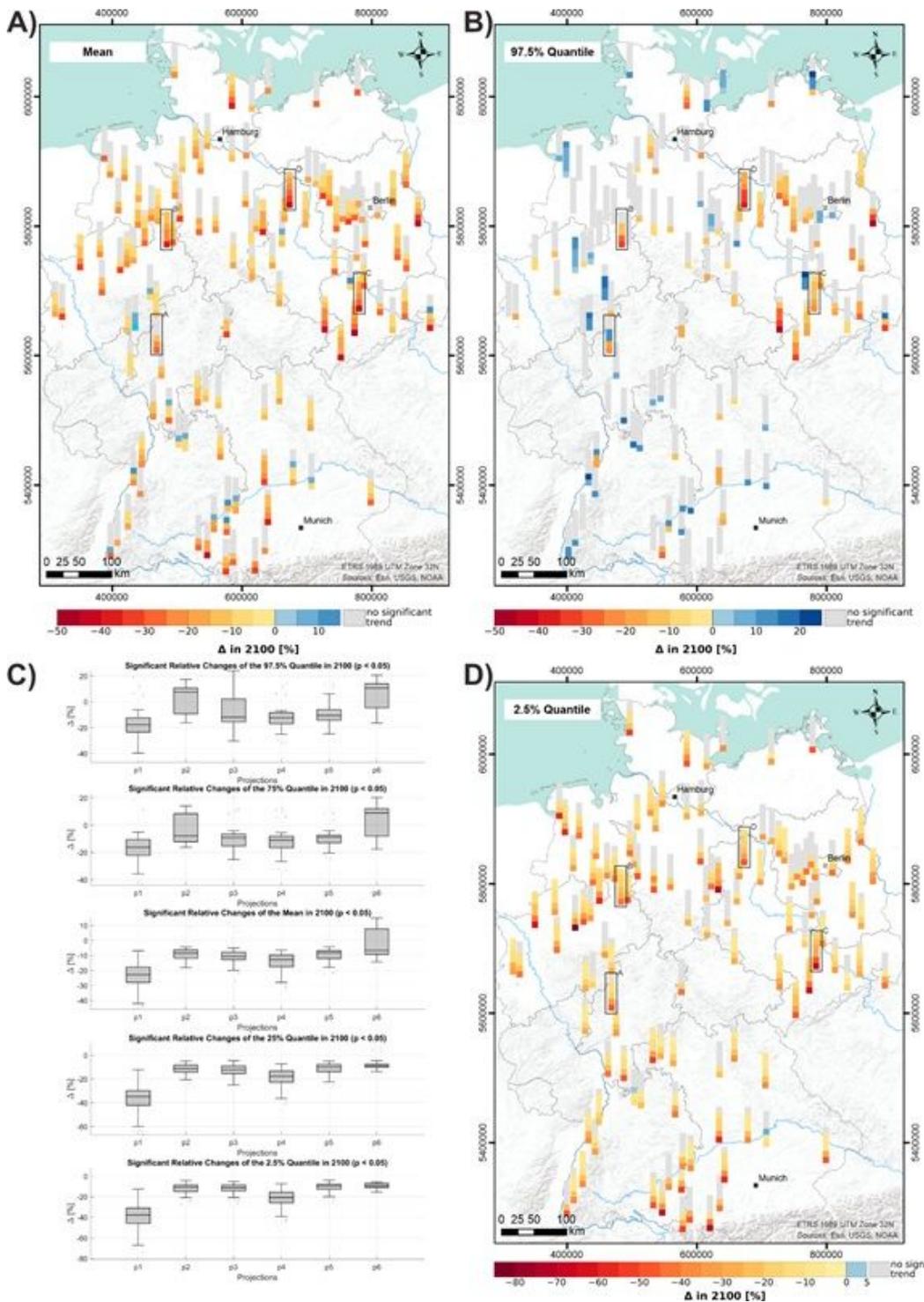


Figure 2

Change of groundwater level [%] in 2100 relative to 2014 (start of sim.) for each site and each climate projection, based on a linear trend analysis: A) mean, B) 97.5% quantile, D) 2.5% quantile; the order corresponds to the strength and sign of the change. C) Boxplots showing the significant changes for A, B, D as well as the 25% and 75% quantiles. Black boxes mark four sites (A-D) shown in detail in Figure 3.

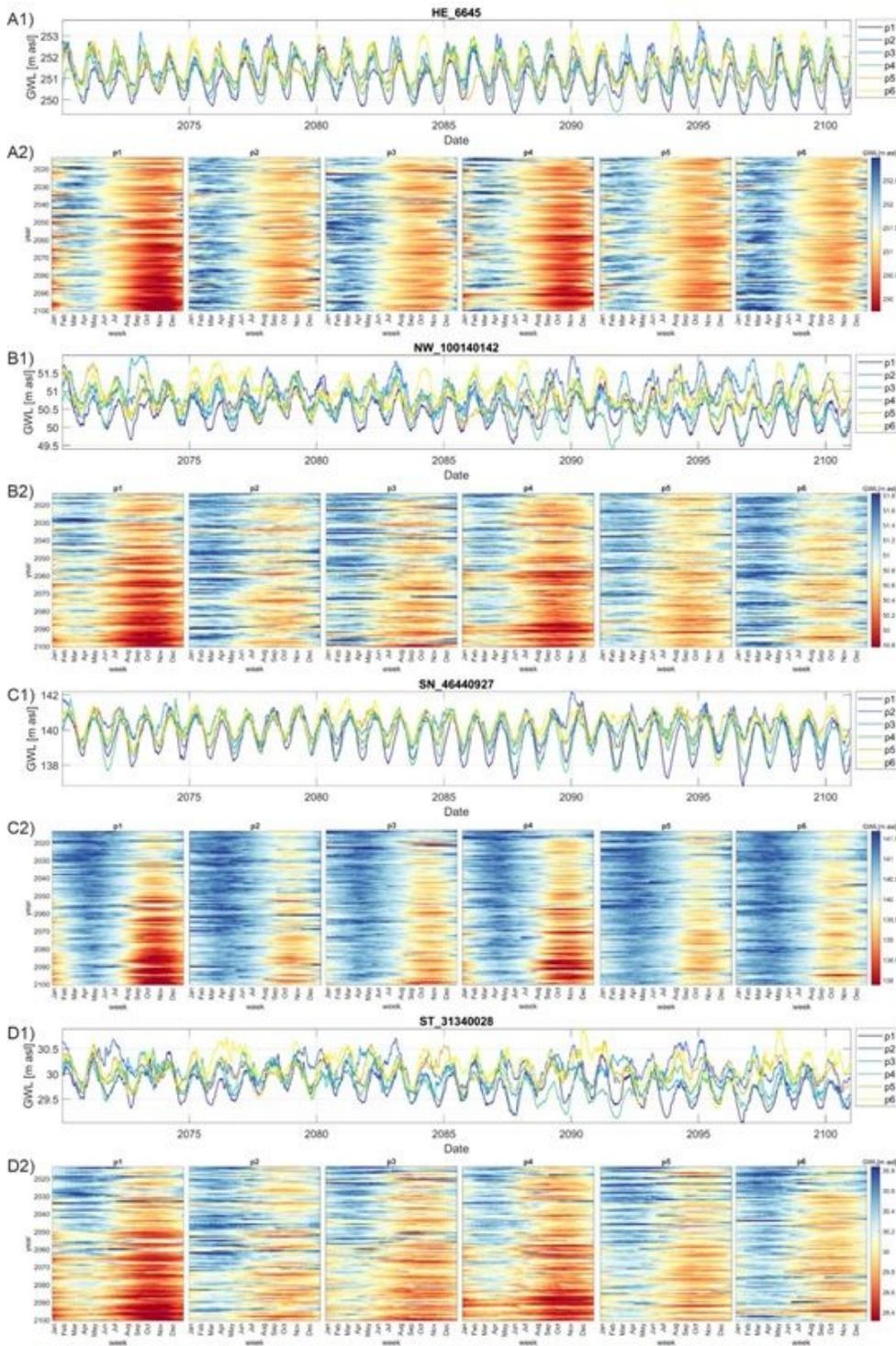


Figure 3

Detailed results on four sites (marked by black boxes in Figure 2): Time series plots of the far future (2070-2100) simulation results (A1-D1); Heatmap plots (A2-D2) of the whole simulation for each of the projections with columns as weeks during the year and rows as the year (up: 2104 – down: 2100).

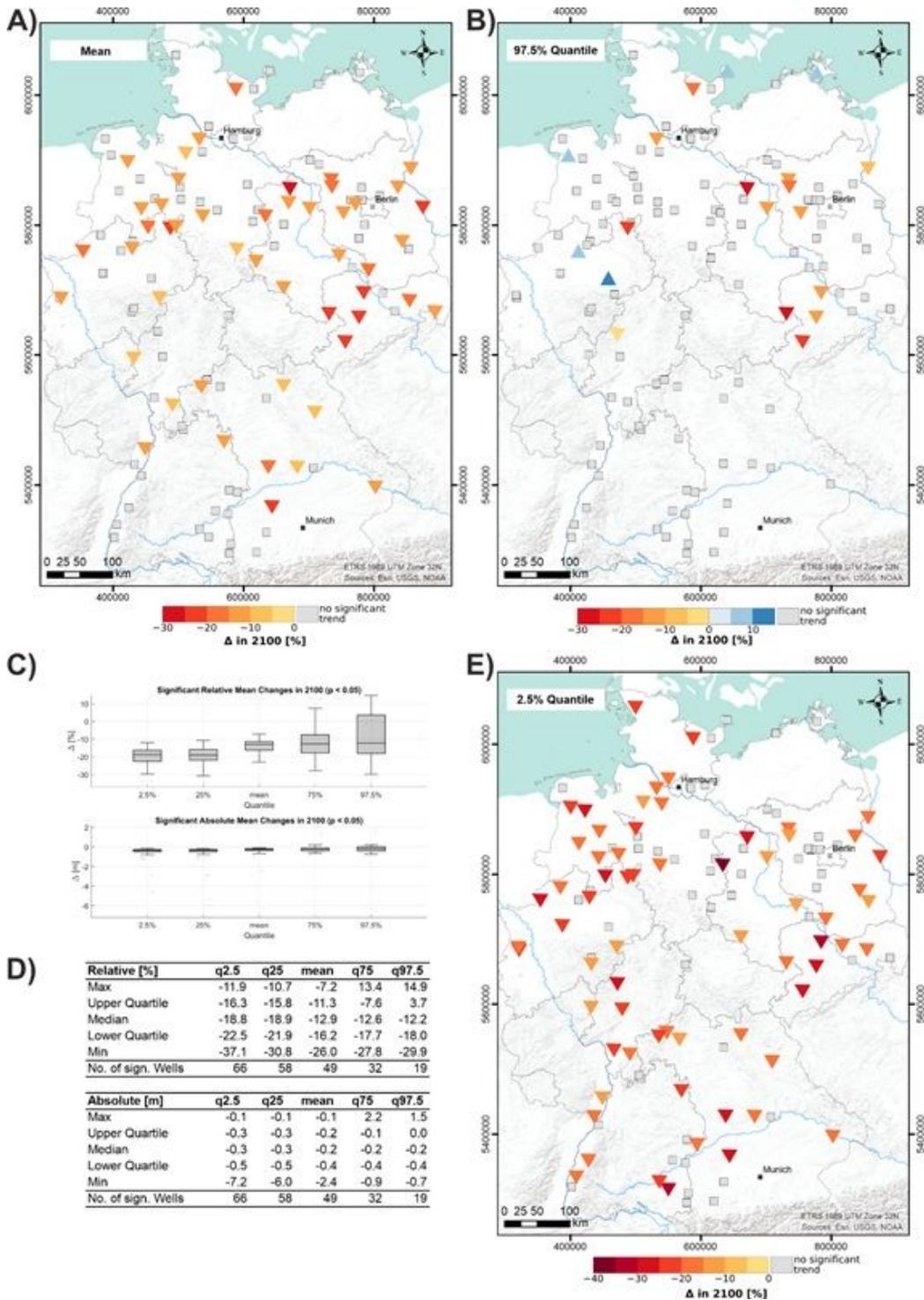


Figure 4

Means of the significant trends of the mean (A), the 97.5% (B) and the 2.5% (E) quantiles shown also in Figure 2. Subplot C shows the associated boxplots (also for 25% and 75% quantiles) and the corresponding absolute changes (lower boxplots). Tables in D show detailed numbers describing the boxplots.

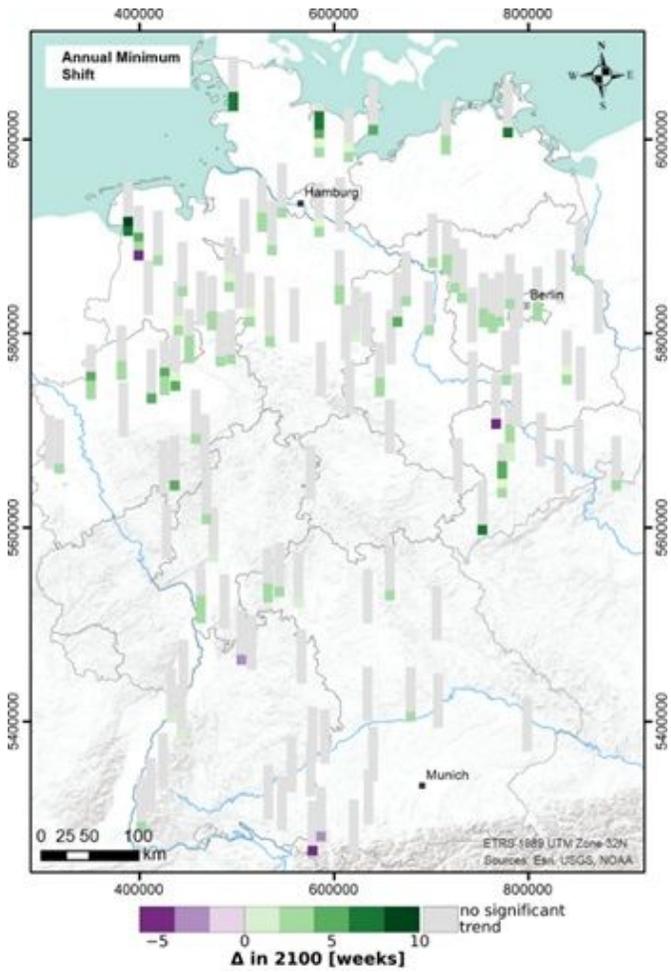


Figure 5

Shift of the annual minimum in weeks until 2100 compared to the start of the simulation (2014). Negative means earlier, positive later during the annual cycle.

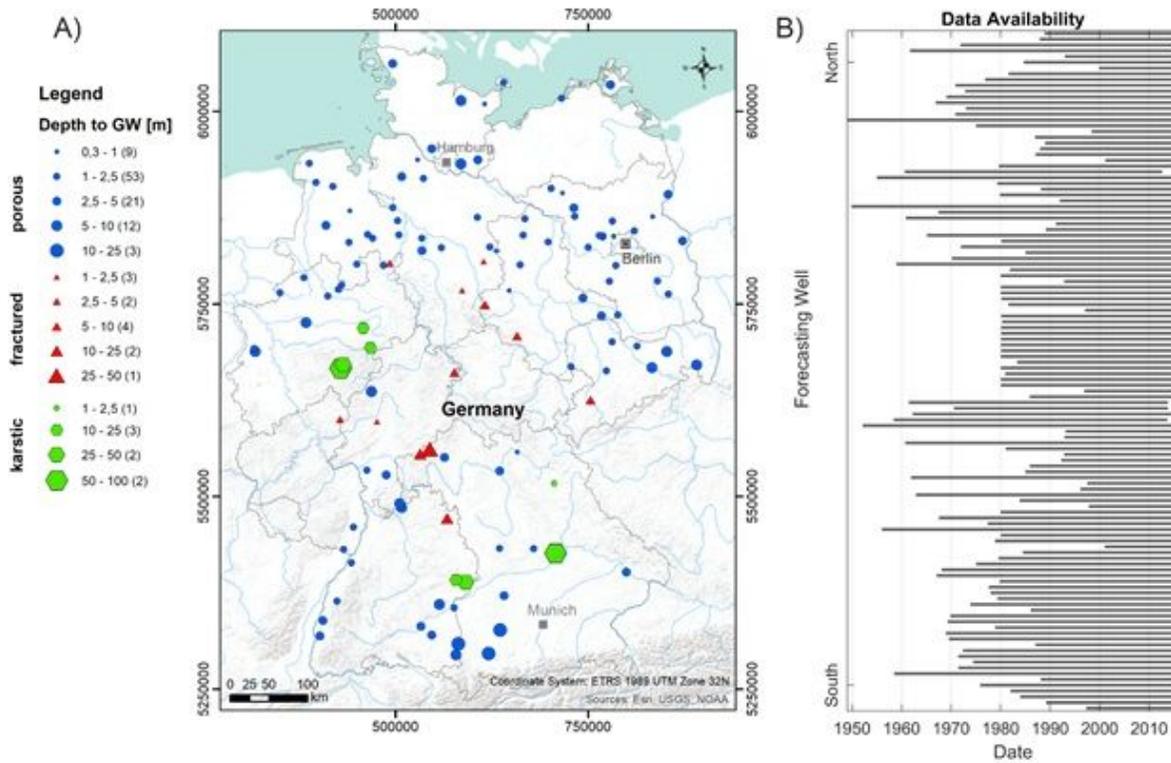


Figure 6

Position, type of aquifer and depth to groundwater for each study site, B: time series length of all study sites North-South ordered.

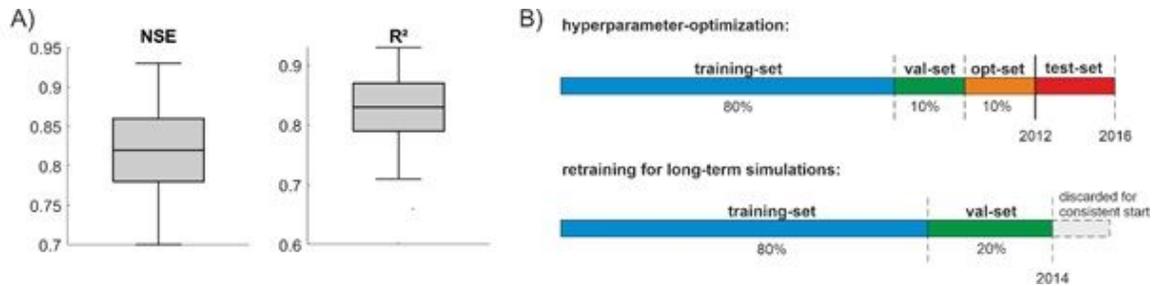


Figure 7

A) Model performance of all models for the test-set (2012-2016), B) time series splitting scheme for optimization (upper) and retraining (lower).

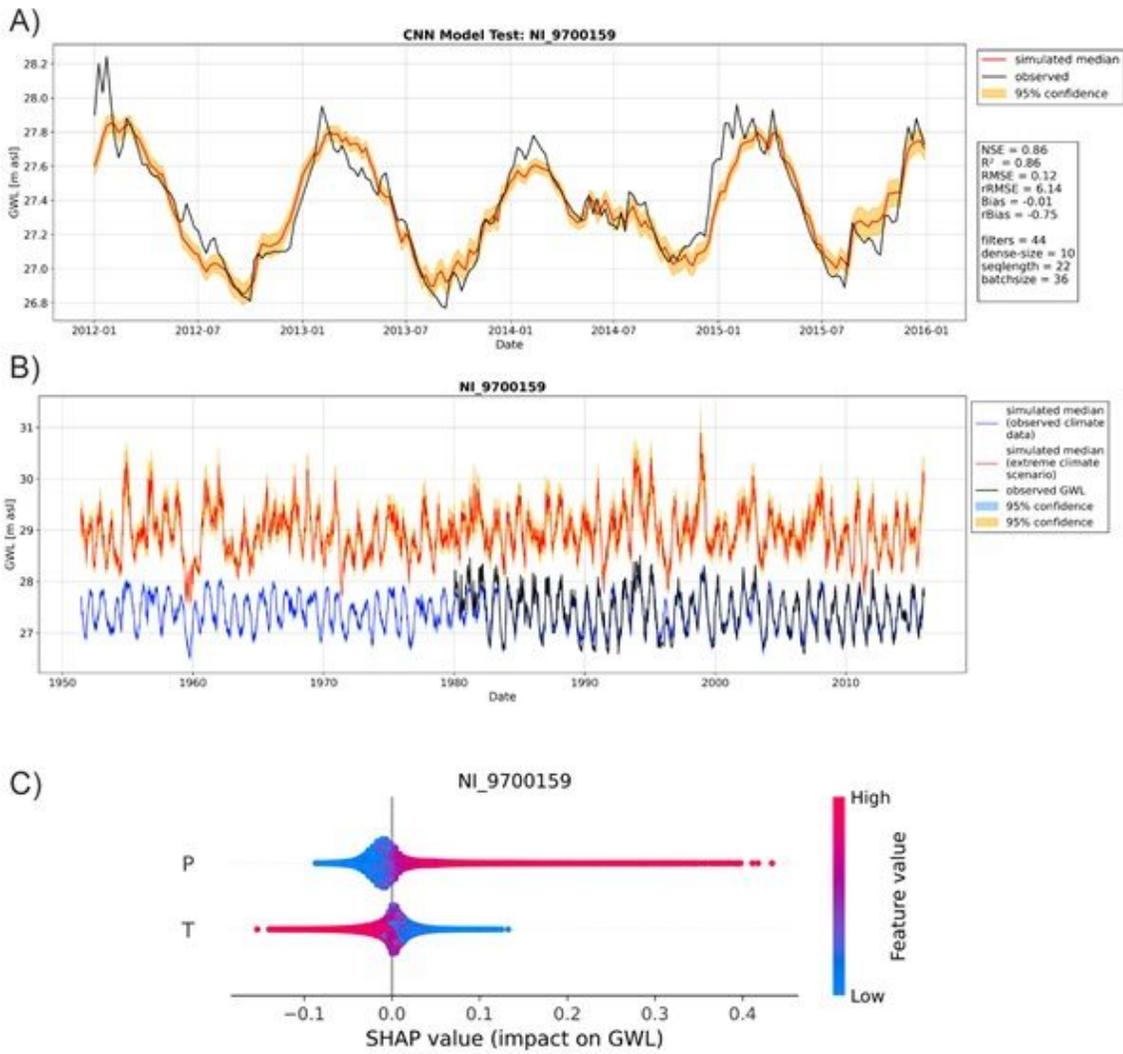


Figure 8

A) Optimized model evaluation in the past for the test set (2012-2016), B) Model output under an artificial extreme climate scenario in the past, C) SHAP Summary plot

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupportingInformation100MB.pdf](#)