Impact of climate change on water sources and river-floodplain mixing in the natural wetland floodplain of Biebrza River

Tomasz Berezowski[1] and Daniel Partington[2]

[1]Gdańsk University of Technology
[2]Flinders University

March 16, 2023

Abstract

The origins of river and floodplain waters (groundwater, rainfall, and snowmelt) and their extent during overbank flow events strongly impact ecological processes such as denitrification and vegetation development. However, the long-term sensitivity of floodplain water signatures to climate change remains elusive. We examined how the integrated hydrological model HydroGeoSphere and the Hydraulic Mixing-Cell method could help us understand the long-term impact of climate change on water signatures and their spatial distribution in the protected Biebrza River Catchment in northeastern Poland. Our model relied on 20th century Reanalysis Data from 1881 to 2015 and an ensemble of EURO-CORDEX simulations for RCP 2.6, 4.5, and 8.5 from 2006 to 2099. The historical component of the simulations was subjected to extensive multiple-variable validation from 1881 to 2019. The results show that the extents of water sources were rather stable in the floodplain in the 1881-2015 period. The projected future impacts were variable with each analyzed RCP, but in all cases, different significant trends were present for the spatial distribution of water sources and for the river-floodplain mixing. However, the total volume of water from different sources was less sensitive to climate change than the dominant sources and spatial distribution of water. The simulation results highlight the impact of climate change on the extent of water sources in temperate zone wetlands with significant implications for ecological processes and management. These results also underscore the urgent need to leverage such modeling studies to inform protective and preservation strategies of floodplain wetlands.

Impact of climate change on water sources and river-floodplain mixing in the natural wetland floodplain of Biebrza River

Tomasz Berezowski [1]*, and Daniel Partington [2]

[1]Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Gabriela Narutowicza 11/12, 80-233 Gdansk, Poland
[2]National Centre for Groundwater Research & Training (NCGRT), College of Science & Engineering, Flinders University, Adelaide, SA, Australia

Key Points:

- The extent of water sources in inundation will strongly vary in the future climate.
- The volume of water sources will vary considerably less than the extent in the floodplain.
- The shifted extents of water may have implications for floodplain management and ecology.

*This research was financed by grant: 2017/26/D/ST10/00665 funded by the National Science Centre, Poland

Corresponding author: Tomasz Berezowski, `tomberez@eti.pg.edu.pl`

Abstract

The origins of river and floodplain waters (groundwater, rainfall, and snowmelt) and their extent during overbank flow events strongly impact ecological processes such as denitrification and vegetation development. However, the long-term sensitivity of floodplain water signatures to climate change remains elusive. We examined how the integrated hydrological model HydroGeoSphere and the Hydraulic Mixing-Cell method could help us understand the long-term impact of climate change on water signatures and their spatial distribution in the protected Biebrza River Catchment in northeastern Poland. Our model relied on 20th century Reanalysis Data from 1881 to 2015 and an ensemble of EURO-CORDEX simulations for RCP 2.6, 4.5, and 8.5 from 2006 to 2099. The historical component of the simulations was subjected to extensive multiple-variable validation from 1881 to 2019. The results show that the extents of water sources were rather stable in the floodplain in the 1881-2015 period. The projected future impacts were variable with each analyzed RCP, but in all cases, different significant trends were present for the spatial distribution of water sources and for the river-floodplain mixing. However, the total volume of water from different sources was less sensitive to climate change than the dominant sources and spatial distribution of water. The simulation results highlight the impact of climate change on the extent of water sources in temperate zone wetlands with significant implications for ecological processes and management. These results also underscore the urgent need to leverage such modeling studies to inform protective and preservation strategies of floodplain wetlands.

Plain Language Summary

In this study, we used a hydrological model that was capable to simulate volumes of water from rain, snowmelt, groundwater discharge, and river flooding to investigate how these volumes will vary with the climatic conditions. For the study site, we selected the Biebrza River wetland floodplain, where former research highlighted the presence of these water sources in inundation during flooding. It was also known that the water sources have different chemical (e.g. nutrients) and physical (e.g. sediments) compositions and they correlate with the vegetation in the wetland. Hence, any change in the extent of these water sources (driven e.g. by climate change) may affect vegetation. Our research indicated that indeed the spatial extent of water sources will strongly vary with the future climate projection while the less detailed floodplain-wise volume of the water sources will not vary that much. We also showed that the direction of change in the water sources' extent will be different given the analyzed climate scenario. These results should be taken into account especially by the natural conservation managers to prepare for the changes.

1 Introduction

Mixing of river and floodplain water during floods, also known as perirheic mixing (Mertes, 1997), has great significance for ecological and hydrochemical processes. This significance in floodplain ecology is reflected by the floodplain vegetation zonation, which is related to the differences in the chemical or sediment composition of water from river and groundwater, rain and snowmelt inundation in the floodplain (Chormański et al., 2011; Keizer et al., 2014). Similar relations are present in the Amazon floodplain, where the mixing of sediment-rich and sediment-poor water near the confluences is related to vegetation (Park & Latrubesse, 2015), and avifauna (Laranjeiras et al., 2021). Also, in the Amazon floodplain, the river-floodplain water frontier is controlling the crevasse splays occurrence (Aalto et al., 2003). The hydrochemical significance of water mixing is mainly due to nitrate removal by denitrification. This process occurs in the flow-through wetlands, where nitrate- and oxygen-rich water from a river mixes with the oxygen-poor floodplain water. Although this effect was reported in several floodplains, including Atchafalaya (Jones et al., 2014; Scott et al., 2014), Po (Racchetti et al., 2011), and Wisconsin (Forshay & Stanley, 2005), to achieve considerable nitrate removal a significant floodplain area has be connected to the river (Natho et al., 2020). As we have shown previously for a natural temperate zone wetland floodplain - Biebrza River, the river-floodplain water mixing, or the active perirheic zone, is very dynamic in space and time (Berezowski et al., 2019). In that study, we used state-of-the-art modeling tools for a single flood event study, hence we were not able to assess the active perirheic zone's long-term variability and the role of the changing climate.

Hydrological impact models of climate change predict a shift of the highest and lowest discharges at the end of the twenty-first century for several regions of the world (Prudhomme et al., 2013; Giuntoli et al., 2015; Arnell & Gosling, 2016). These regions include the major floodplain and wetlands, where the shift in flooding pattern may influence ecological processes such as vegetation development (Murray-Hudson et al., 2006; Garris et al., 2014; Zulkafli et al., 2016; Thompson et al., 2016). The hydrological shifts in the future will also lead to changes in floodplain connectivity in unregulated floodplains. This may result in increased nitrate removal by denitrification, as simulated for the Lower Missouri River (Jacobson et al., 2022). Nitrate removal varies in floodplain habitats with different contact with river water (Scaroni et al., 2011). Since, the zonation of water sources within the flooding extent is relevant for vegetation development and denitrification, more precise quantification of these ecological processes in the scope of climate change could be achieved by analyzing water sources' zonation. This remains a gap in the literature.

Modeling of climate change impact on floodplain inundation is usually done using either 1D or 2D hydrodynamic models. Such models require a precise definition of boundary conditions for which coupling with catchment-based hydrological models is often used (Thompson et al., 2008; Karim et al., 2015; Zhang et al., 2019). Another approach is to drive a hydrodynamic model using boundary conditions, such as surface runoff, from hydrological components of general circulation models (GCM), or climate reanalysis (Mohanty & Simonovic, 2021). In either case, the surface water in the floodplain lacks or has limited, feedback with parts of the catchment that are not represented by the hydrodynamic model, which includes groundwater, tributary inflow, or surface runoff. These feedbacks are important in the proper modeling of floodplain inundation, as those minor water sources produce the inundation in remote parts of the floodplain and determine the river-floodplain water frontier (Berezowski et al., 2019) and groundwater mixing zone (Nogueira et al., 2022). Therefore, to achieve full feedback between all water sources integrated hydrological models (IHMs) are required (Sebben et al., 2013). The computational complexity of these models often requires some simplifications or limiting the simulation area (Barthel & Banzhaf, 2015) to achieve feasible run times. Also, the application of IHMs to climate change impact research is limited in scenarios and analysis periods lengths (e.g. Ferguson and Maxwell (2010); Sulis et al. (2011); Erler et al. (2019)), while using a GCM ensemble reduces uncertainty related to future climate projections impact on hydrology Z. Kundzewicz et al. (2018). Currently, this research area remains relatively unexplored, as only a few studies run such models with long-term forcing data from GCMs ensembles, such as the Intergovernmental Panel on Climate Change (IPCC) emission scenarios (Goderniaux et al., 2009; Sulis et al., 2012; Perra et al., 2018; Boko et al., 2020; Ramteke et al., 2020; Yuan et al., 2021) and no such models have analyzed the extent of water from different sources.

Except for the GCM ensemble, the credibility of the modeling results in climate change studies is achieved by comprehensive model validation and using multiple impact models. The latter is especially important in large-scale (regional and continental) cases, where some parts of the study area are ungauged (Krysanova et al., 2018). Further, it seems that using IHMs ensembles may not be crucial, since contrary to many conceptual models used in climate change impact studies, they are almost entirely physically based and perform similarly (Kollet et al., 2017). On the other hand, a comparison of conceptual, physically based, and fully integrated hydrological models in a climate change impact study revealed that the models showed the same direction of change for most of the indicators, however, a fully coupled IHM indicated an the opposite trend in mean annual evapotranspiration when compared to remaining models (Perra et al., 2018). Either way, using an IHM ensemble would be beneficial for the results by increasing credibility, although this comes with an associated computational burden.

Comprehensive validation of an IHM, should, therefore, be of greatest concern in climate change impact studies. Most often the observations of river discharge are used for the validation of impact models, while IHMs, due to the simulation of surface water hydrodynamics can be further validated against water levels. The spatial aspect of validation can be achieved by using multiple gauges, however, flooding water extent can serve this purpose as well. The latter is often achieved using multi-temporal remote sensing data providing spatiotemporal insight into model performance (e.g. Paiva et al. (2012)), however, in vegetated areas such validation can be problematic, due to obscuring by vegetation canopy. IHMs are usually also validated against groundwater levels, which gives further insight into processes relevant to catchment functioning that are not depicted by surface water. Also, if transport or water mixing is simulated, IHMs can be validated against hydrochemical parameters. This list of validation variables for IHMs does not include all the possibilities. Instead, it indicates that, contrary to conceptual models, the physically based IHMs can be validated comprehensively to minimize uncertainty related to the simulated complex interactions, such as mixing of water from different sources.

To examine the impact of climate change on spatiotemporal water signatures during flooding in a natural temperate zone wetlands, this research aims to employ a robust IHM for the Bierbza catchment to investigate the long-term variability of the extent and mixing of water from different sources during flooding. The model for the Biebrza will be run for a historical period using 20th Century Reanalysis data and a GCM ensemble for representative concentration pathways (RCP) 2.6, 4.5, and 8.5 scenarios for the future. With this model, the aims of the research are:

- To determine if the past climate and future climates under RCPs 2.6, 4.5 and 8.5 will drive any significant changes in the spatial distribution and dominance of water sources in the Biebrza floodplain.
- To determine if the volume of water in the floodplain will significantly change under past climate and possible future climates with RCPs 2.6, 4.5 and 8.5.
- To highlight the implications for ecological processes, modeling, and management strategies under climate change.

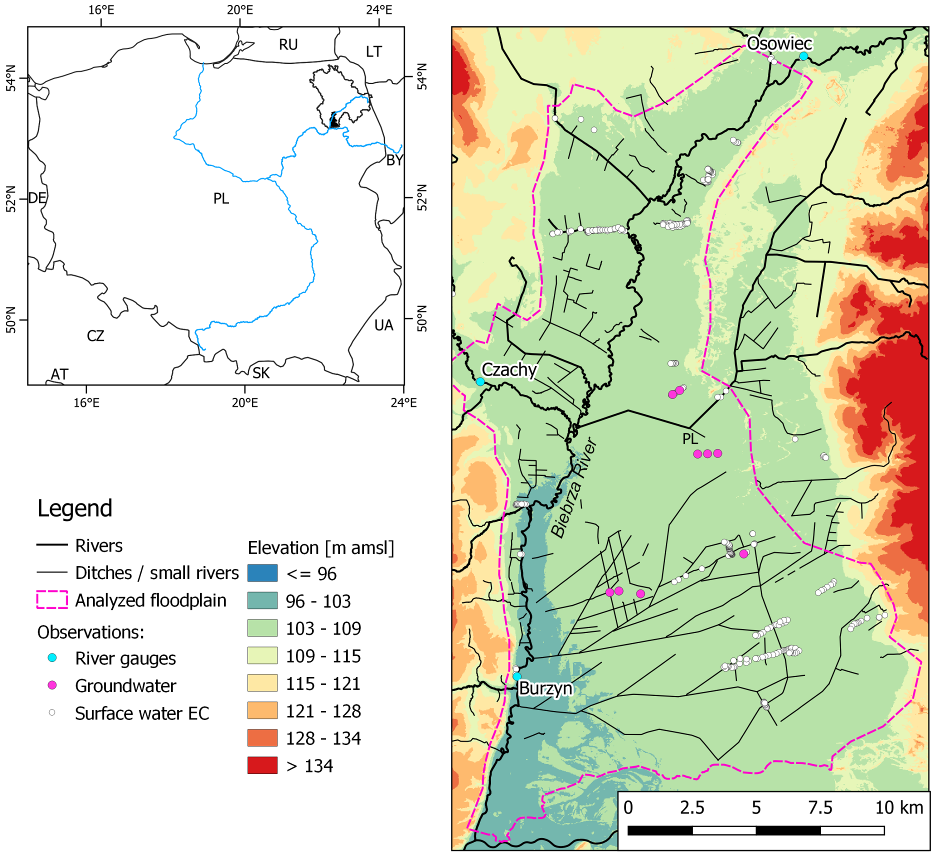

Figure 1. The floodplain area and the measurements points (right panel). Location of the study area in Poland (left panel) with the major rivers (blue lines), Biebrza river catchment (black outline), and the floodplain (black patch). The legend concern only the right panel.

2 Methods

2.1 Study area

The Biebrza catchment (22.7° E, 53.7°N) is of medium size, 7091 km ^{2} and the lower Biebrza valley (hereinafter referred to as floodplain), where we focus our analysis comprises 297 km^2 (Figure 1).

We chose the Biebrza valley as the study area because of its natural character and ecological significance. The major river engineering work was conducted in the area in the first half of the 19th century to establish a waterway between Biebrza and Neman Rivers. Next, in the middle of the 19th century parts of the Wetlands located in the lower and middle parts of the valley were meliorated. In the 20th century, only minor melioration work was conducted except in the middle part of the Biebrza valley (Banaszuk, 2004). Currently, the anthropogenic pressure is low, as the population density in the region where the Biebrza River catchment is located is the lowest in Poland (58 people per km^2) (Statistics Poland, 2021). The future population projections for this region predict a 32% decline between 2020 and 2100 (Eurostat, 2019). The Biebrza valley was grazed and mowed in the past and aquatic vegetation in the river was occasionally removed (Berezowski et al., 2018). Since the establishment of the Biebrza National Park in 1993 mowing and grazing is continued as an active protection measure (Kotowski et al., 2013). Currently, the Biebrza National Park is one of the largest active protection areas in Europe (59223 ha), with the Biebrza Wetlands listed as Ramsar and Natura 2000 sites.

Long-term average discharge in Biebrza River has been 38.1 m^3s^{-1} (1970-2005), with a minimum of 4.33 m^3s^{-1} and maximum of 517 m^3s^{-1}. The river flooding area reaches up to 52.5 km^2 and inundation can last on average between 121 to 193 days depending on location (Grygoruk et al., 2021). The average annual precipitation over the period 1970-2005 in the catchment has been 672 mm, of which 88 mm was snow, whereas the yearly potential evapotranspiration (PET) was 621 mm.

Wetland vegetation in the floodplain exhibits zonation related to flooding (Pałczyński, 1984). The Phragmition belt is located around the river up to about 500-900 m, further away up to 2500 m from the river, Magnocaricion vegetation is present, and further again, Fen vegetation, such as Calamagrostion neglectae, Caricion diandrae, or Caricion demissae is located up to the valley margin.

The Quaternary deposits are 130-212 m deep and the majority consist of glacial till with minor sand layers deposited during the Riss glaciation. Middle and lower parts of the Biebrza valley have a sand layer deposited during the Weichselian glaciation on top of which the Holocene sand and peat layers are present (Banaszuk, 2004).

Given undisturbed vegetation, unregulated river, natural hydrology, and low contamination in relation to European standards the Biebrza wetlands may be considered as a reference site for similar fen wetlands (Wassen et al., 2006).

2.2 Forcing data

Hydrological simulations for over two hundred years period required several sources of forcing data (Table S1). The criteria for selecting a data source were daily (or higher) temporal resolution and availability of the required forcing variables (precipitation, snow cover dynamics, air temperature, and PET).

For the historical 1880-2015 period we used the 20th century climate reanalysis (20CR) data (Slivinski et al., 2019). Out of this data-set, we used ensemble mean of water equivalent of accumulated snow depth (WEASD) [kg m^{-2}], daily mean of 3-hour accumulated precipitation amount (APCP) [kg m^{-2}], air temperature at 2m (air2m) [K], and potential evaporation rate (PEVPR) [W m^{-2}]. We used the following preprocessing steps before bias correction. The difference of WEASD between subsequent days was calculated. Then, the negative values were multiplied by -1 and used as uncorrected daily snowmelt, $\dot{s}$, [mm] and the positive values were used as uncorrected daily snowfall ($\dot{p}_s$) [mm]. For PET, the PEPVR values were multiplied by 0.01152 to change units to mm.

For the future period, we used the EURO-CORDEX data (Jacob et al., 2014) from ten simulations using different GCMs (Table S1). Each simulation used the SMHI-RCA4 regional climate model (RCM). We selected all available simulations from the EURO-CORDEX archive that had the required forcing data for the hydrological model. Only four out of ten simulations had the required forcing data for RCP 2.6. To investigate the effect of greenhouse gases emission scenarios on water sources mixing in the floodplain we used the following RCPs: RCP 2.6, which aims to limit the increase in global mean temperature to 2 K by a CO_2 emission decline since 2020, RCP 4.5 which is an intermediate scenario, where the emissions start to decline after 2040, and RCP 8.5 which is a worst-case scenario in which emissions continue to rise during the entire 21st century. From each simulation, we used daily mean values of snowfall flux (PRSN, used as $\dot{p}_s$) [kg m^{-2} s^{-1}], snowmelt flux (SNM, used as $\dot{s}$) [kg m^{-2} s^{-1}], precipitation flux (PR) [kg m^{-2} s^{-1}], near-surface air temperature (tas) [K], and potential evapotranspiration (EVSPBLPOT) [kg m^{-2} s^{-1}]. We used the following preprocessing steps before bias correction. Daily snowmelt, snowfall, precipitation, and potential evapotranspiration fluxes were multiplied by 86400s to change units to mm.

For the 2015-2019 period (for which the 20CR data was not available), when the hydrochemical validation took place we used the 2 km gridded precipitation and temperature data-set (Piniewski et al., 2021) and snowfall and snow depth data from the Biebrza-Pieńczykówek meteorological station managed by the Institute of Meteorology and Water Management - National Research Institute (IMGW-PIB).

2.2.1 Bias correction

We used the quantile mapping (Gudmundsson et al., 2012) bias correction using the R software package "qmap". The following meteorological observations were used to identify parameters of bias correction: the total precipitation and air temperature from a 5km gridded data-set (Berezowski et al., 2016) (the 2km data set was not available at that time), PET from a gridded 25 km data-set (Joint Research Center, 2019), and the snowfall from the Biebrza-Pieńczykówek meteorological station. In such variable availability, we were not able to conduct bias correction of snowmelt, s, and rainfall, p_r. The snowmelt was constrained to the snowfall using the sum of uncorrected snowmelt ($\dot{s}_v$) and the sum of bias-corrected snowfall ($p_{s,v}$) in a given event v. An event was defined as a period between the start of snow accumulation and the end of snowmelt; most often there are one or two larger events in each year. Daily snowmelt [mm] in an event v was calculated as $s = \dot{s}\frac{p_{s,v}}{\dot{s}_v}$. The rainfall p_r [mm] for a given day was calculated by subtracting bias-corrected snowfall from bias-corrected precipitation. We used a maximum overlapping period for bias correction of each variable, which was 1955-2013 for precipitation and air temperature, 1957-2015 for snowfall, and 1979-2015 for PET for the 20CR data. In the case of the EURO-CORDEX data, we were additionally limited by the historical period, which was either 1951-2005 or 1970-2005 (Table S1). After conducting the bias correction we calculated the daily average value of each variable over all grid cells in the Biebrza catchment and used this data to force the hydrological simulations.

2.3 Hydrological model

We simulated the transient water fluxes in the Biebrza River catchment using HydroGeoSphere (Brunner & Simmons, 2012; Hwang et al., 2014) IHM. The 3D groundwater flow was solved using Richard's equation in prism elements and the 2D surface water flow was solved using the diffusion wave approximation of the Saint-Venant equations in triangular elements. The surface-subsurface flow coupling was realized using the first-order exchange. Evapotranspiration flux was simulated using the Kristensen and Jensen (1975) conceptual model, which takes into account interception storage, time-variable leaf area index (LAI), pounding, and soil saturation. Snowmelt and rainfall fluxes were provided as forcing data boundary conditions. The model parameters were specified spatially according to relevant geological, land-use, or vegetation units.

We simulated water mixing using the hydraulic mixing-cell (HMC) method (Partington et al., 2011). In our case the mixing was simulated only for the surface flow domain, however, simulations in groundwater are also possible (Nogueira et al., 2022). The HMC method accounts for water fluxes from various boundary conditions and groundwater discharge effectively producing a fraction of each water source in a model node. Water sources were differentiated spatially. To calculate the river water fractions we summed all fractions upstream of the floodplain area. Whereas in the floodplain area, original fractions of rainfall, snowmelt, and groundwater were used to represent the inundation components generated therein. In the first time step, the fractions are initialized using an artificial initial fraction, equal to one.

We used the parallel solver in the HGS, which split the coefficient matrix into two parts. The flow solver convergence criteria for the maximum absolute residual error was

$1 \text{x} 10^{-10}$ m^3s^{-1}, and the Newton iteration convergence criteria for the maximum absolute nodal change in the pressure head was 1 cm. In the HMC method, the maximum ratio between fractions volume was set to 2048 and above this threshold, all fractions are set to zero and the reset fraction is set to one (Partington et al., 2013).

2.3.1 Leaf area index estimation

The HydroGeoSphere model uses LAI during the estimation of evapotranspiration. Since LAI was not available in any data set covering the simulation period we used a degree-day model to simulate LAI for each meteorological data set used in this study. The model was based on observations that wheat requires about 760 degree-days for development and 500 more degree-days for maturity (Rawson & Macpherson, n.d.) and can be summarized in the following steps:

1. At the beginning of a calendar year LAI is equal to the minimum for a given vegetation
2. Growing season is defined as a day when the monthly average temperature is greater than 5°C.
3. Since the beginning of the growing season LAI increases proportionally to degree-days to reach the maximum for given vegetation at 760 degree-days.
4. LAI remains at the maximum for 500 degree-days.
5. LAI decreases linearly to reach the minimum for given vegetation on the last day of the growing season.

The maximum LAI for each vegetation was based on measurements in the study area (Dąbrowska-Zielińska et al., 2014; Suliga et al., 2015).

2.4 Error metrics

In this study, we use the same error metrics for a number of different simulated quantities, such as water levels, discharge, water source fractions, and area. We present the general form of the equations below. Whenever a given error metric is used in the text it is specified based on which quantities it was calculated for and, if applicable, to which quantity it was normalized.

The Kling-Gupta efficiency [-]:

$$\text{KGE} = \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{1}$$

where r [-] is the correlation coefficient between simulated and observed discharge, α [-] and β [-] are ratios of simulated to observed mean and standard deviation discharges respectively. The KGE ranges between $-\infty$ and 1 and the higher the value the better fit to the observation is achieved by the model.

The root mean square error [units the same as input data]:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{h_i} - h_i\right)^2}{N}} \tag{2}$$

where h_i and $\hat{h_i}$ are observed and simulated quantities respectively for a data record (e.g. time step) i out of N. The RMSE represents the magnitude of error between the observations and simulations and ranges between 0 and ∞.

The systematic error, or bias [units the same as input data]:

$$b = \sum_{i=1}^{N} \hat{h_i} - h_i \tag{3}$$

where the symbols are the same as in Eq. 2. The bias shows whether the simulated quantities overestimate (positive b) or underestimate (negative b) the observed quantities and $b = 0$ indicate no bias.

The linear correlation between two variables was quantified using Pearson's correlation coefficient (r) [-] and the fraction of variance explained between the two variables was quantified using the coefficient of determination (r^2). If two variables are time-dependent the linear correlation can be interpreted in terms of the temporal variability agreement between them.

2.5 Model grid

To prepare the model grid we processed the relevant geographical information in QGIS 3.10 software in the following steps. We simplified the geometry of the rivers by limiting the minimum node distance to 125 m along the river course for major rivers and 500 m for minor rivers. For the major rivers, the banks were limited to a 60 m buffer around the river. This forced the perpendicular river cross-section to be trapezoidal. For minor rivers, no buffer was created and the perpendicular cross-section was triangular. The catchment boundary was simplified by limiting the minimum node distance to 2000 m. The geographical data source used in these steps was the Map of the Hydrographic Division of Poland in scale 1:10 000. The feature nodes obtained from the previous steps and the nodes representing locations of the observation wells were used to generate a Delaunay triangular grid in the triangle software (Shewchuk, 1996). The triangulation constraints were the maximum triangle size of 1 km^2 and the minimum angle in a triangle of 31°. Finally, we refined the grid fourfold in the floodplain area and relaxed the nodes using an algorithm provided by Kaser et al. (2014). The triangular grid consisted of 19297 nodes and 38081 elements of which 10436 were in the floodplain. The median element area for the whole grid was 71243 m^2 and for the floodplain was 20037 m^2; the minimum element area was 1017 m^2.

The nodes elevation was obtained from a Digital Elevation Model (DEM) of Poland in the resolution of 1m and from the Shuttle Radar Topography Mission in 30 m resolution outside the Polish border (in total 0.4% of the study area). The digital elevation model was updated with the lake bathymetry. The riverbed elevation for the major rivers was obtained from 160 land-survey perpendicular cross-sections conducted by the Polish Water Authority. The distance between subsequent cross-sections was about 500 m. As a riverbed elevation, the first quartile of the elevation in the nearest cross-section was used. The minor river's riverbed was calculated by subtracting river depths from a surface elevation. The river depth was estimated based on point measurement data provided by the Biebrza National Park and from our field survey.

The grid consisted of six vertical layers in which the top four layers had gradually increasing thickens and represented the stratification of peat, sand, and glacial till formed between the Riss glaciation and Holocene. The thickness of the first layer was 0.75 m in the floodplain. The two bottom layers were thick and represented glacial till deposited during the Riss glaciation. The elevation of the lowest layer was equal to -30 m AMSL, the average lower boundary of the Quaternary sediments (Banaszuk, 2004). In total, the grid consisted of 135097 nodes and 228486 prism elements.

We defined three porous materials: glacial till, sand, and peat with different hydraulic properties. In the river valley and its proximity, we assigned the materials based on geological cross-section data (Banaszuk, 2004), whereas in the remaining parts of the upland we used data from several geological bore profiles provided by the Polish Geological Institute (Polish Geological Institute, 2014). The hydraulic properties for the surface water flow and evapotranspiration were assigned to ten land-use and vegetation classes present in the study area based on the Corine Land Cover map (Commission of the European Communities, 2013).

2.6 Model calibration

We used a screening approach to find an optimal parameter set for the hydrological model. For this purpose, we randomly sampled 800 random parameter sets using the latin hypercube algorithm. We used the latin hypercube algorithm implementation from the "tgp" R package (Gramacy & Taddy, 2010). Each set consisted of 26 base parameters, which produced 43 model parameters by applying the constraints and transformations (Table S3). The constraints were used to scale a base parameter by a factor for different material types, such as vegetation types and produce multiple model parameters. We used the logarithmic transformation for the hydraulic conductivity and gamma distribution transformation for evapotranspiration parameters (details in Table S3-S5). The calibration period was two years and ten months (2004-01-01 to 2006-10-31) followed by a one and half year warm-up period (2002-06-01 to 2003-12-31). The initial conditions for each calibration run were transferred from a steady-state simulation using parameters from our previous model version (Berezowski et al., 2019). We choose the best model base on KGE for two discharge stations and RMSE [m] for five groundwater wells heads. The locations of discharge stations were chosen at the inlet and outlet of the floodplain (Osowiec and Burzyn) and the location of the wells were chosen two in the floodplain, one in the middle and upper parts of the valley. The relation between average KGE and average RMSE for all stations forms a Pareto front with a group of the best parameter sets from which the final model was selected manually by reviewing the simulated hydrographs.

2.7 Model validation

2.7.1 Hydrological validation

We used several contemporary and archival data sources with varied temporal coverage for the validation of simulated river flow and groundwater heads (Table S2). We used the same metrics as for calibration and the RMSE was normalized by the data range for each station or well. To investigate how the hydrological model performed temporarily we calculated KGE for discharge and RMSE for river water levels per decade.

The oldest water level records (Table S2) for the study area contained only the relative water level in reference to the gauge zero level. For these records we calculated the absolute water level, i.e. in meters AMSL, using a relation between the mean absolute and relative water levels for the remaining records for a particular gauge. The disadvantage of this approach is that the temporal trend is not preserved and the RMSE is biased.

Some of the groundwater heads data were missing the absolute readings, i.e. depth instead of elevation was measured. Calculation of the absolute levels was done by using a 1x1 m digital elevation model values in the well location as the zero depth. Few groundwater wells showed a clear step in the records, which could have been due to the displacement of the reference point. We removed records with the step from the database.

2.7.2 Remote sensing validation

We validated the simulated water extent using a multi-temporal remote sensing data-set (Berezowski et al., 2020). In that data-set 161 water extent maps were developed for the 2014-2019 period using the Sentinel-1 synthetic aperture radar (SAR) for the floodplain with the average water level error of the flood extent of 0.21 m. The major drawback of this data-set was that in densely vegetated areas the flood extent was obscured and effectively these areas are labeled as not flooded even if the water level was high. Further, the data-set was not sensitive to shallow water, which limits its applicability only to an indication of deeper river water within the Biebrza flooding extent. From this data set, we selected 134 flood maps with the lowest error and used them along with the hydrological model output to calculate the following validation metrics.

Despite some drawbacks, the remote sensing data-set was a good indicator of the temporal dynamics of the flooding extent, especially for the river water zone. Therefore validation in the floodplain was calculated using the total flooding area due to simulated water depth [m^2]:

$$a_h = \sum_{m=1}^{M} \tilde{h}^m a^m \tag{4}$$

and the flooding area due to river water fraction presence [m^2]:

$$a_{\text{river}} = \sum_{m=1}^{M} \tilde{f}^m_{\text{river}} a^m \tag{5}$$

where $\tilde{h}^m = 1$ if simulated water depth in a node m is greater than 5 cm and $\tilde{h}^m = 0$ otherwise, $\tilde{f}^m_{\text{river}} = 1$ if river water fraction is greater than 0.1 in a node m and $\tilde{f}^m_{\text{river}} = 0$ otherwise, a_m is the node m contributing area, and M is the total number of nodes in the floodplain. The values of a_h and a_{river} are calculated for each time step and used to calculate the correlation coefficient with the flooded area from the remote sensing data-set. Further, we calculated the fraction of area that is indicated as flooded on the intersection of hydrological model output and remote sensing data-set:

$$i_h = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{h}^{m,t}a^m\right) \wedge \left(\tilde{a}_{rs}^{m,t}a^m\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{a}_{rs}^{m,t}a^m\right)}$$

for intersection with the simulated water depth and

$$i_{\text{river}} = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{f}_{\text{river}}^{m,t}a^m\right) \wedge \left(\tilde{a}_{rs}^{m,t}a^m\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{a}_{rs}^{m,t}a^m\right)}$$

for intersection with simulated river fraction, where $\tilde{h}^{m,t}$ and $\tilde{f}_{\text{river}}^{m,t}$ are the same as $\tilde{h}^m$ and $\tilde{f}_{\text{river}}^m$, but indexed also for time step t, $\tilde{a}_{rs}^{m,t} = 1$ if the flooded area in the remote sensing data-set in a node m is greater than 25% and $\tilde{a}_{rs}^{m,t} = 0$ otherwise, and T is a group of time steps which overlap in the hydrological simulations and remote sensing data-set. Ideally, this validation should be extended to the calculation of true-negative flooding extent. This, however, was not possible due to false negative flooding extent in the remote-sensing data-set due to vegetation cover.

2.7.3 Hydrochemical validation

To investigate whether the different water sources presence is related to the simulated water source fractions we measured the electrical conductivity (EC) [μS cm^{-1}] of 133 samples in the floodplain during winter (24-25 January 2019) and spring (27-29 March 2019). The HI991300 portable EC meter was used and the location was recorded using a handheld GNSS receiver. We chose EC because prior research by (Chormański et al., 2011) indicated that EC is effective at discriminating between river water and other sources. We used random 50% of the measurement points to establish a linear regression model explaining the EC by the river, rain, snowmelt, and groundwater fractions in the model nodes on the measurement days. The remaining 50% of the data was used for validation of the regression model using RMSE [μS cm^{-1}] and bias [μS cm^{-1}]. All measurement points were used to calculate the correlation coefficients between the water source fractions and EC.

2.8 Changes of water sources fraction in the past and future climate

Next to the simulated water sources fractions, we analyzed the mixing degree [-] (Berezowski et al., 2019), which quantifies the mixing between river and floodplain (sum of snow, rainfall, and groundwater) water fractions:

$$d = 1 - \frac{\mid f_{\text{river}} - f_{\text{floodplain}} \mid}{1 - f_{\text{initial}}} \tag{6}$$

The changes in water sources fraction and mixing degree were assessed by calculating a length [days] of a period during which they were greater than 0.75 and the water depth was greater than 1 cm, in a hydrological year for each model node m in the floodplain:

$$l_s^m = \sum_{y=1}^{Y} \begin{cases} 1 & w_s^{y,m} > 0.75 \land h^{y,m} > 0.01 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $w_{s,t}$ is a value of s water source fraction (river, snow, rainfall, or groundwater) or the mixing degree d during a day y of a all days Y in a hydrological year, and $h_{t,m}$ is water depth [m]. The total annual volume of surface water in the floodplain weighted by the water sources fractions and the mixing degree in a hydrological year was calculated by performing a weighted integration using the following equation:

$$v_s = \sum_{y=1}^{Y} \sum_{m=1}^{M} \begin{cases} h^{y,m} a^m w_s & h_t > 0.01 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The mean surface water depth ($\bar{h}$) [m] and the length of a period with water depth greater than 1 cm (l_h) [days] was calculated for each model node in each hydrological year.

For future climate simulations, we calculated the above metrics for each EURO-CORDEX simulation and calculated the ensemble mean for each RCP scenario. Next, we used the ensemble means and historical simulations forced using 20CR data to calculate trends using the slope of the regression line, where the independent variable is the hydrological year. Finally, we used the t-test to investigate whether a trend estimate is significantly different from zero.

3 Results

3.1 Bias-corrected forcing data

Each forcing data have similar statistics as meteorological observations for the period in which the quantile mapping parameters were identified (Table S6). Both for EURO-CORDEX and 20CR data the air temperature underestimated the observations mean, but had similar standard deviations. Snowfall and PET were bias-corrected near perfectly in terms of mean and standard deviation. Total precipitation was overestimated

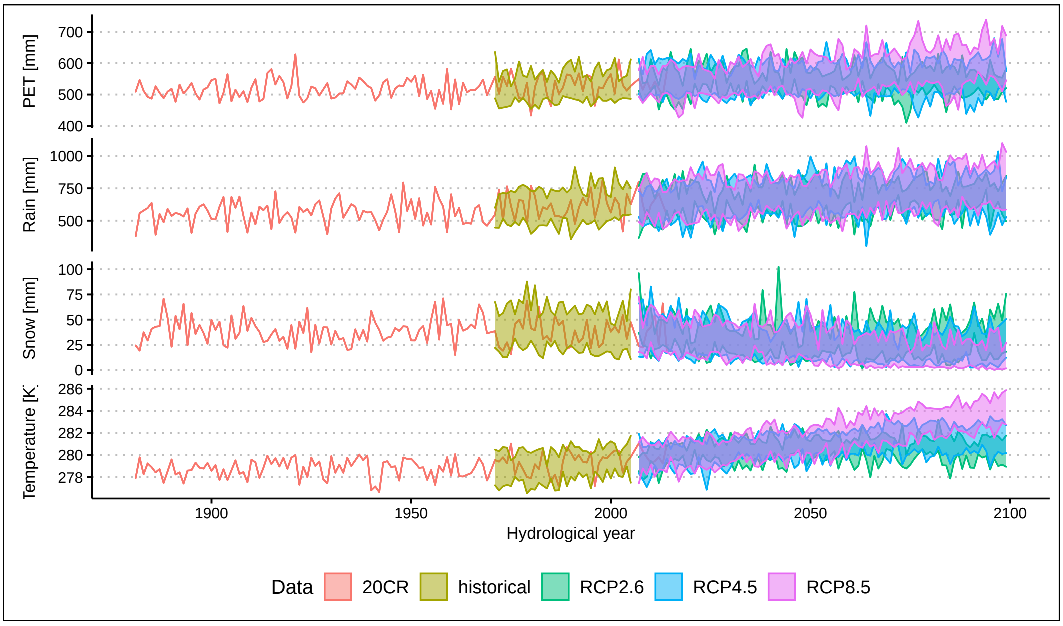

Figure 2. The 20CR and EURO-CORDEX data for the Biebrza catchment after bias correction. Temperature is the yearly mean and the remaining variables are yearly sums. The ribbons present the 2.5-97.5 percentiles range for all simulations in a given RCP or historical experiment for EURO-CORDEX data. The gap between historical and RCP ribbons is due to data presentation in hydrological years, whereas the EURO-CORDEX simulations starts and finishes as calendar years.

in reference to daily mean observations by 11.8% and 10.9% by 20CR and EURO-CORDEX mean respectively.

The 20CR data fits the EURO-CORDEX ensemble in the overlapping historical period after bias correction (Figure 2). The 20CR data show no significant trends until the end of the first half of the 20th century. In the 1950-2015 period the air temperature trend of 0.02 K year^{-1} (p=0.0008) was observed. EURO-CORDEX data presented significant trends for ensemble yearly medians for all meteorological variables except PET for the RCP 2.6. The PET trends for the remaining RCPs were 0.24 (RCP 4.5), and 0.81 mm year^{-1} (RCP 8.5). For RCP 2.6, RCP 4.5, and RCP 8.5 respectively the trends were -0.08, -0.16, and -0.31 mm year^{-1} for snowfall, 0.64, 0.81, and 1.61 mm year^{-1} for rainfall, and 0.01, 0.02, and 0.05 K year^{-1} for air temperature.

3.2 Model calibration

The hydrological model calibration results formed a clear Pareto front with a minimum RMSE of 0.19 m and maximum KGE of 0.86 (Figure S1). Out of these models we choose one with an RMSE of 0.24 m and a KGE of 0.69 as the best performing and used it for further simulations. The calibrated parameter values (Table S7) had values within the range presented in the literature for porous media materials (Wösten et al., 1999; Gnatowski et al., 2010). The parameter search space was relatively wide for all material types, yet the saturated hydraulic conductivity presented an expected pattern with greater values for sands than for glacial till and relatively low value for peat. The Manning roughness coefficient had higher values for the Biebrza River and floodplain than reported in the literature (Chow et al., 1988).

3.3 Hydrological validation

Simulated water levels and surface water discharge matched the observations well (Figure 3). Daily discharge at the Osowiec and Burzyn stations, which are located at the inlet and outlet of the floodplain were only slightly overestimated with an absolute error that was 5% of the data range (Table 1). Similar simulated discharge errors were also present for Czachy, which is a major inlet into the floodplain, and Sztabin, which is located in the upper part of the catchment. Overall fit to observations expressed by KGE for discharge showed that Burzyn and Osowiec performed better than smaller stations Czachy and Sztabin. A similar pattern was also present for correlation, which indicated that the discharge temporal variability was simulated better for Burzyn and Osowiec than for Czachy and Sztabin.

Simulated daily water levels showed a good overall fit as expressed by KGE (Table 1). The high values of the correlation coefficient and the visual comparison shows that within-year and multi-year (Figure 3) variability of water levels was simulated correctly. The water levels were overestimated by 3% for Burzyn and underestimated by 4% for Osowiec. The water levels RMSE were the same for both stations in the floodplain and were more attributed to high flows in Osowiec and low flows in Burzyn.

Table 1. Error metrics for all available observations for river gauges. RMSE and bias are in the same units as indicated in the table, remaining metrics are dimensionless. H and Q are water levels and discharge respectively, RMSE / d.r. and bias / d.r. area RMSE and bias normalized to the observations data range (d.r.), corr. is the correlation coefficient.

Station	Units	Period with observations	RMSE	RMSE / d.r.	bias	bias / d.r.	Corr.	KGE
H Burzyn	m	1930-1935, 1946-2017	0.37	9%	0.12	3%	0.83	0.68
H Osowiec	m	1881-1911, 1921-1923, 1925-1935, 1946-2017	0.37	10%	-0.15	-4%	0.79	0.67
Q Burzyn	$m^3\ s^{-1}$	1951-2017	25.88	5%	5.14	1%	0.69	0.64
Q Czachy	$m^3\ s^{-1}$	1957-2017	2.33	4%	-0.73	-1%	0.63	0.50
Q Osowiec	$m^3\ s^{-1}$	1951-2017	17.02	5%	2.79	1%	0.69	0.63
Q Sztabin	$m^3\ s^{-1}$	1951-2017	4.73	5%	0.84	1%	0.60	0.53

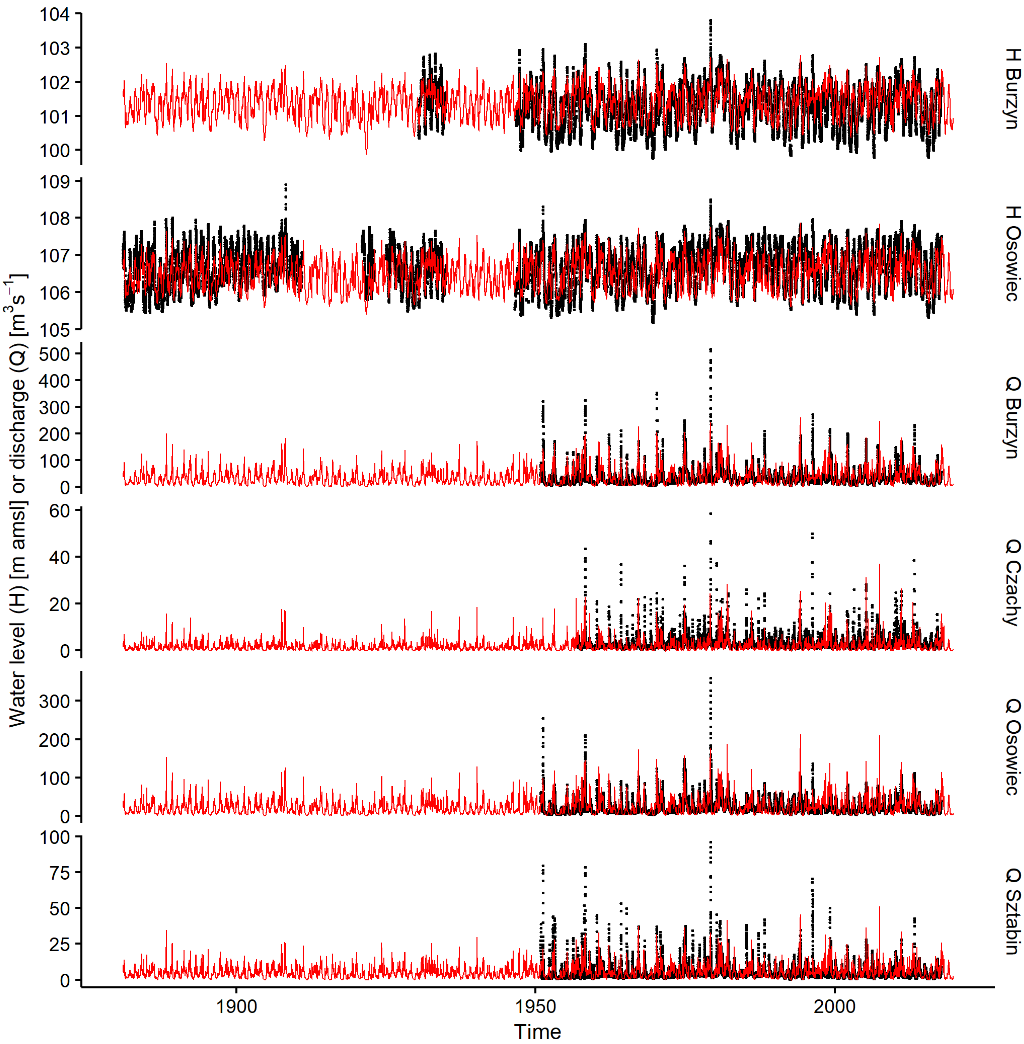

Figure 3. Water levels (H) [m AMSL] and discharges (Q) [m^3s^{-1}] for river gauges. The location of river gauges in presented in Figure 1 except for Sztabin, which is located in the upper part of Biebrza River.

At the catchment scale, the model simulated groundwater levels very well, with the r^2=0.99 (Figure 4). Clear deviation of simulated groundwater levels was observed for the household wells located in the upland. Individual well's performance varied with the location in the model grid. In the floodplain, where the grid was finer than in the remaining parts of the model, the mean RMSE for nine wells was 23% of the data range with a 9% underestimation (Table S8). Outside the floodplain, i.e. in the middle and upper parts of the Biebrza valley, the mean RMSE was 36% and 34% respectively (Table S9). In these parts of the catchment simulated groundwater levels performed worse for certain wells with RMSE up to 76% of the observed data range, although all wells preserved the temporal variability as in the observed data (Table S9 and Figures S2-S5)

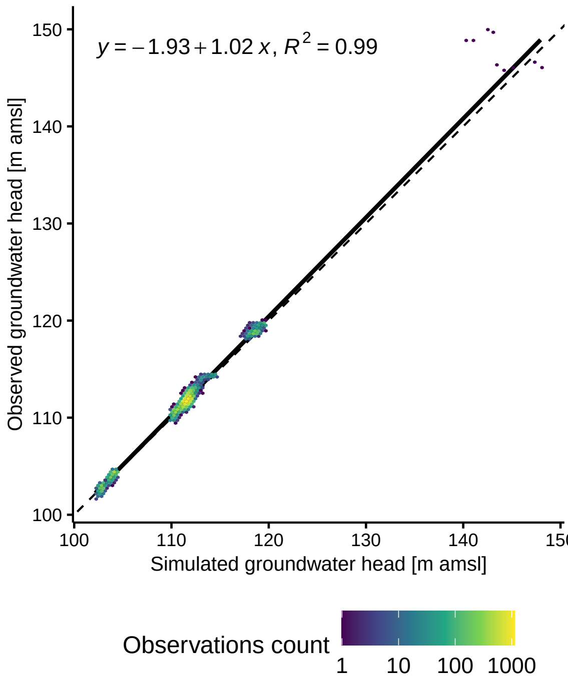

Figure 4. Validation of the simulated groundwater levels using daily observations (usually in ten days resolution) in 43 wells in the period 1994-2019 (N=18032). Solid line - regression line, dashed line - 1:1 line.

3.4 Remote sensing validation

The temporal variability of the SAR water extent correlated better to the flooding extent derived from the river water fractions (a_{river}, r=0.75) than to total extent estimated from the water depth (a_h, r=0.64) (Figure 5). Both a_{river} and a_h water extents overestimated the SAR flooding extent maps for the periods of the lowest water levels when the Biebrza River was not flooding. In these periods the remote sensing data-set was not indicating surface water extent (including between the river banks, Figure S6), while the total area of Biebrza River and oxbow lakes in the floodplain is 2.97 km^2. The Biebrza River and its tributaries were always visible in the hydrological model output. The hydrological model predicted a summer flood in 2017 that was not visible in the SAR data. Also, one summer flood in 2015 visible in SAR data was not simulated by the hydrological model. There was a good agreement in the intersection of the true positive flooding extent from the remote sensing data-set with simulated water depth i_h=0.77

and with simulated river fractions i_{river}=0.78. The lowest agreement occurred during low flow (below bankfull) periods with i_h=0.19 and i_{river}=0.16, while during higher flows (above bankfull) the agreement was higher i_h=0.82 and i_{river}=0.83.

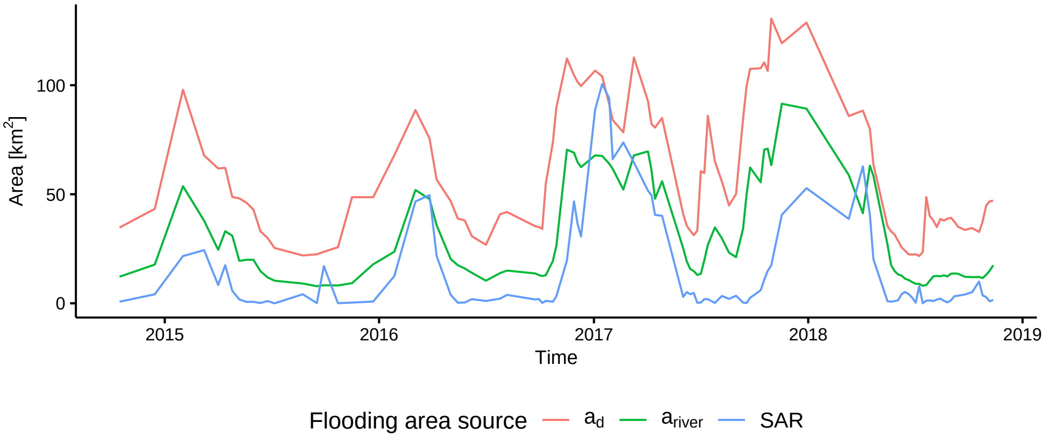

Figure 5. Total area of flooding extent from remote sensing data-set (SAR), calculated for simulated water depths > 5cm (a_d) using Eq. 4, and calculated for river water fractions > 0.1 (a_{river}) using Eq. 5. There are 134 dates in which remote sensing data-set overlapped with the simulation period are presented.

3.5 Hydrochemical validation

The correlation with EC measurements was strongly negative for snow fractions (-0.62) and moderately positive for the river fractions (0.48). A very weak correlation was observed for rainfall (-0.07) and groundwater (0.00). The linear regression model, which explained the EC measurements with the water source fraction predictors, showed that all fractions were significant (p<0.001). The validation metrics for the regression model were r^2=0.58, RMSE=91μS cm^{-1} (18% of data range), and bias b=12 μS cm^{-1} (2% of data range). The highest underestimation visible in the validation of the EC regression model was for measurements located next to an asphalt road located in a central part of the floodplain (~8.5 km from the Biebrza river) (Figure 6). The underestimated predictions are present in the direction of water flow from the road to the river, which indicates possible increased salinity due to car traffic.

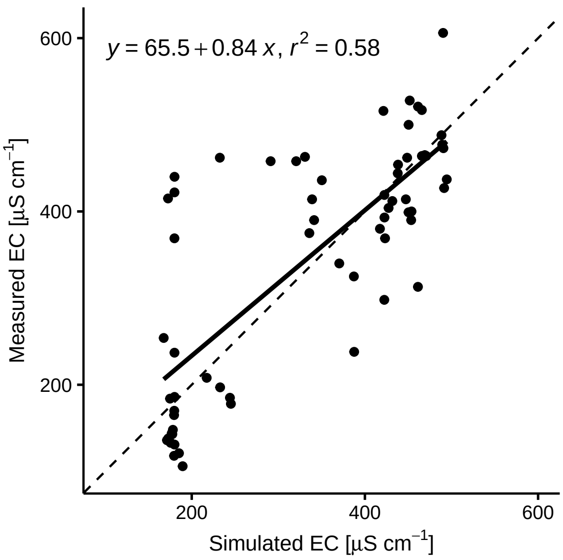

Figure 6. Validation of the EC measurements regression model in the period 2019-2021 (N=64). Solid line - regression line, dashed line - 1:1 line.

3.6 Changes in Biebrza River flow in the past and future climate

Simulated flow characteristics at the Burzyn (outlet) station showed that the 2.5-97.5% range simulations forced by the EURO-CORDEX historical experiments and the 20CR had similar characteristics (Figure 7). The mean simulated water levels overestimated the observations by 1.4% (20CR) and 2.7% (EURO-CORDEX mean) of the observed data range with the underestimated standard deviation by 26% (20CR) and 32% (EURO-CORDEX mean) (Table S10). In the case of discharge, the overestimation was 0.5% (20CR) and 0.9% (EURO-CORDEX mean) with a standard deviation overestimation of 6.4% for models forced using 20CR data and an underestimation by 6.5% (mean) for model forced using EURO-CORDEX data.

Within the 1970-2005 period no significant trends were observed in daily mean discharge or water levels for the models forced EURO-CORDEX or, 20CR data nor for the observation at the Burzyn station. However, in the 1951-2015 period, when observations overlap with the 20CR data a significant trend of 0.173 $m^3s^{-1}year^{-1}$ (p=0.031) was observed; no significant trend was observed for water levels. For this period a similar trend of 0.057 $m^3s^{-1}year^{-1}$ was observed in the model forced with the 20CR data however, it was not significant (p=0.527); in the complementary (1881-1950) period no trend (0.01 $m^3s^{-1}year^{-1}$, p=0.861) was observed. For the future climate impact simulations using the EURO-CORDEX data, significant trends (2006-2099) were observed only for RCP

Figure 7. Mean daily simulated discharge and water levels per year for the Burzyn station forced using 20CR and EURO-CORDEX data. The ribbons present the 2.5-97.5 percentiles range for all simulations in a given RCP or historical experiment for EURO-CORDEX data. The gap between historical and RCP ribbons is due to data presentation in hydrological years, whereas the EURO-CORDEX simulation starts and finishes as calendar years.

2.6 and 4.5. The trend for mean daily discharge was 0.092 $m^3s^{-1}year^{-1}$ (p=0.005) for RCP 2.6 and 0.080 $m^3s^{-1}year^{-1}$ (p<0.001) for RCP 4.5. In the case of mean daily water levels, the trend was 0.0015 m $year^{-1}$ (p=0.007) for RCP 2.6 and 0.0007 m $year^{-1}$ (p=0.032) for RCP 4.5.

3.7 Changes of water sources fraction in the past and future climate

The simulated daily mean volume of water from different sources did not show significant trends for the past climate forced with the 20CR data (Figure 8). In the simulations forced by the EURO-CORDEX data for future climate positive trends were observed for the river, rainfall, groundwater, and river-floodplain mixed water volumes in RCP 2.6 and RCP 4.5. In the RCP 8.5 significant trends were observed only for rainfall and snowmelt volume. For all RCP snowmelt volume trends were negative, however, the trend was not significant for RCP 2.6. The snowmelt water was characterized by the lowest volume in the floodplain area and was subjected to the highest relative changes in the RCPs 4.5 and 8.5.

Length of a period in which water source fractions were dominant, the river-floodplain mixing degree was high, or water depth was greater than 1 cm was stable before 1950

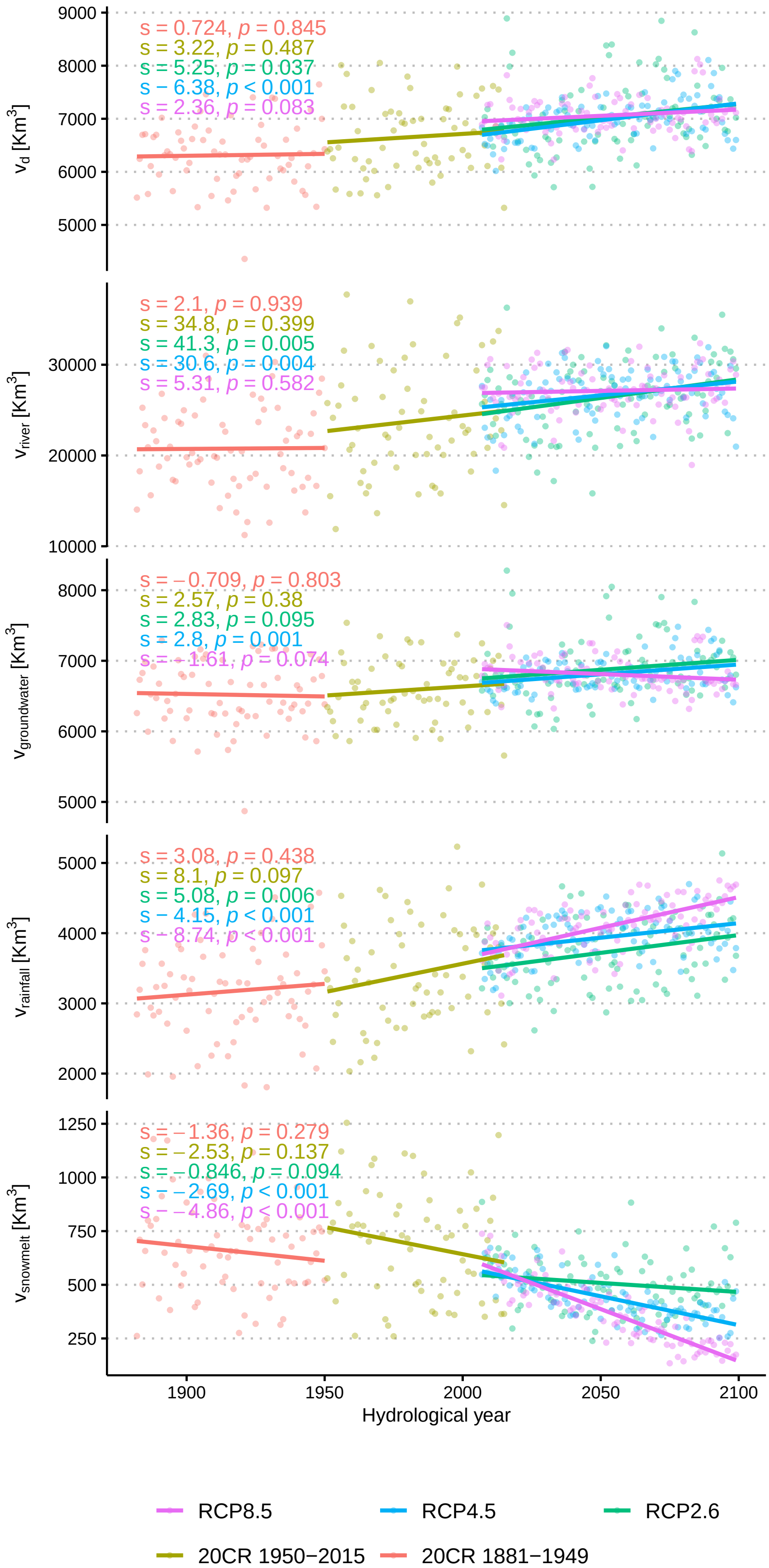

Figure 8. Simulated surface water volume daily means for river-floodplain mixing (v_d), river (v_{river}), groundwater ($v_{\mathrm{groundwater}}$), rainfall (v_{rainfall}), and snowmelt (v_{snowmelt}) in the floodplain per hydrological year. The RCPs data is presented as the ensemble means. The s symbol is a slope of a regression line (trend) [Km^3 year^{-1}] and the p symbol is a p value of a t-test for the slope estimate, 1Km^3 is 1000 m^3.

with only a few nodes showing a slight increase for l_{rainfall}(Figure 9). A similar situation was observed in the simulations for the 1950-2015 period. Therein, however, l_h and l_{river} increased in proximity to the Biebrza River. Also, l_{rainfall} showed a more distinctive patch of increased vales when compared to the latter period.

The trends for the ensemble mean in RCP 2.6 and 4.5 showed a similar pattern of increased l_h and l_{river} in the proximity of the Biebrza River and increased l_{rainfall} in the central part of the floodplain (Figure 9). The increase of l_h and l_{river} was the greatest in RCP 2.6 out of all analyzed RCP scenarios and past climate periods. The increase of $l_{\mathrm{groundwater}}$ was observed in RCP 2.6 near the valley margin, which was not visible for RCP 4.5. Unlike RCP 2.6, RCP 4.5 showed a decrease of $l_{\mathrm{groundwater}}$ and l_{snowmelt} in the central part of the floodplain.

The RCP 8.5 simulations showed that l_h and l_{river} was small and clearly smaller than in RCP 2.6 and 4.5 while the change of l_{snowmelt} was similar as in RCP 4.5 (Figure 9. The decrease of $l_{\mathrm{groundwater}}$ was the highest in RCP 8.5 and was visible in the central part of the floodplain (especially in the ditches), near the valley margin (northern part), and in the Biebrza River bed. Also, the increase of l_{rainfall} was the highest in RCP 8.5 and was present in almost the entire floodplain.

In all simulations the length of the high river-floodplain mixing period, l_d, increased with increasing l_{river}, yet, the trend in l_d was smaller than the increase of l_{river}. An exception of this was in the central part of the floodplain in all RCP scenarios, where l_{river} did not show a significant trend, but l_d showed an increase. Therein l_{rainfall} increased the most along with the $l_{\mathrm{groundwater}}$ increase in RCP 4.5 and the $l_{\mathrm{groundwater}}$ decrease in RCP 8.5. The l_d did not change nearest to the river in the RCP scenarios, whereas the l_{river} changed the most. In this area, l_d was high due to mixing at the beginning of the flood.

The trend of surface water depth above 1 cm period, l_h, resembles that of l_{river} in the area where both trends were significant, i.e. in the proximity of the river. The l_h, unlike l_{river}, increased also in the central part of the floodplain, especially in the ditches, and next to the valley margin in all RCP scenarios and a few nodes in the 1950-2015 period. The highest l_h increase in these areas was observed in the RCP 2.6, although the change in l_{snowmelt}, $l_{\mathrm{groundwater}}$, and especially in l_{rainfall} was the smallest in this scenario among all RCPs. Overall, the magnitude of l_h change was the highest in RCP 2.6 (accompanied by the highest magnitude of l_{river} change) although the area of significant changes was greater in remaining RCPs. Notably, in the areas where l_d increased, but l_{river} trend was not significant the l_h also showed an increase. Still, $l_h$increased in areas further away from the rivers where neither l_d nor l_{river} increased.

Significant trends in mean daily water depth, $\bar{h}$, were observed spatially only in the RCP scenarios for the future climate (Figure 10). The trends were the greatest in the proximity of the river, reaching some river nodes up to 6.3 mm year^{-1} in RCP 2.6, 4.0 mm year^{-1} in RCP 4.5, and 0.7 mm year^{-1} in RCP 8.5 (Figure 10). The RCP 4.5 and 8.5 scenarios predict a very small positive trend across the majority if the floodplain, whereas RCP 2.6 predicts such a trend in only remote parts of the floodplain and in ditches.

4 Discussion

4.1 Forcing data

The forcing data matched the meteorological observations in terms of mean and standard deviation, which indicates, that the biases were removed correctly. The highest deviations from observations were observed for the total precipitation and air temperature, which still, were comparable to other studies conducted in our study region. Mezghani et al. (2017) reported an RMSE of 15.5 mm month^{-1} (equivalent to about 0.51 mm day^{-1}) and the air temperature monthly mean RMSE of 1.1°C (daily minimum) and 1.6°C (daily maximum) using an ensemble of 9 EURO-CORDEX simulations. The differences between observations and bias-corrected air temperature did not have a large impact on the hydrological simulations, because the air temperature was not used to calculate PET in the hydrological model. Rather than that, the air temperature was only used to calculate the degree-days for LAI estimation. The bias-corrected PET data had very small deviations from the observations.

The predicted change of total precipitation and air temperature varies between the model applied. In general, other studies show indicated an increase in yearly precipitation, and air temperature, and a decrease in snow cover by the end of the 21st century in our study area. Warszawski et al. (2013) showed that the yearly precipitation will increase by up to 10% and air temperature by 2-6 K the in RCP 8.5 scenario. Similarly, Schneider et al. (2013) showed that winter half-year precipitation will increase by 5-15%, with no changes in the summer half-year precipitation, mean annual temperature will increase by 2-2.5 K and the snow-cover period will decrease by 20-30 days. Also, Mezghani et al. (2017) predicted an increase of precipitation by 9.7% in RCP 4.5 and by 15% in RCP 8.5 and the air temperature increase of 2 K in RCP 4.5 and 3.6 K in RCP 8.5. Except that we were not able to compare the RCP 2.6 scenario, these results are consistent with the bias-corrected data used to force hydrological simulations in our study.

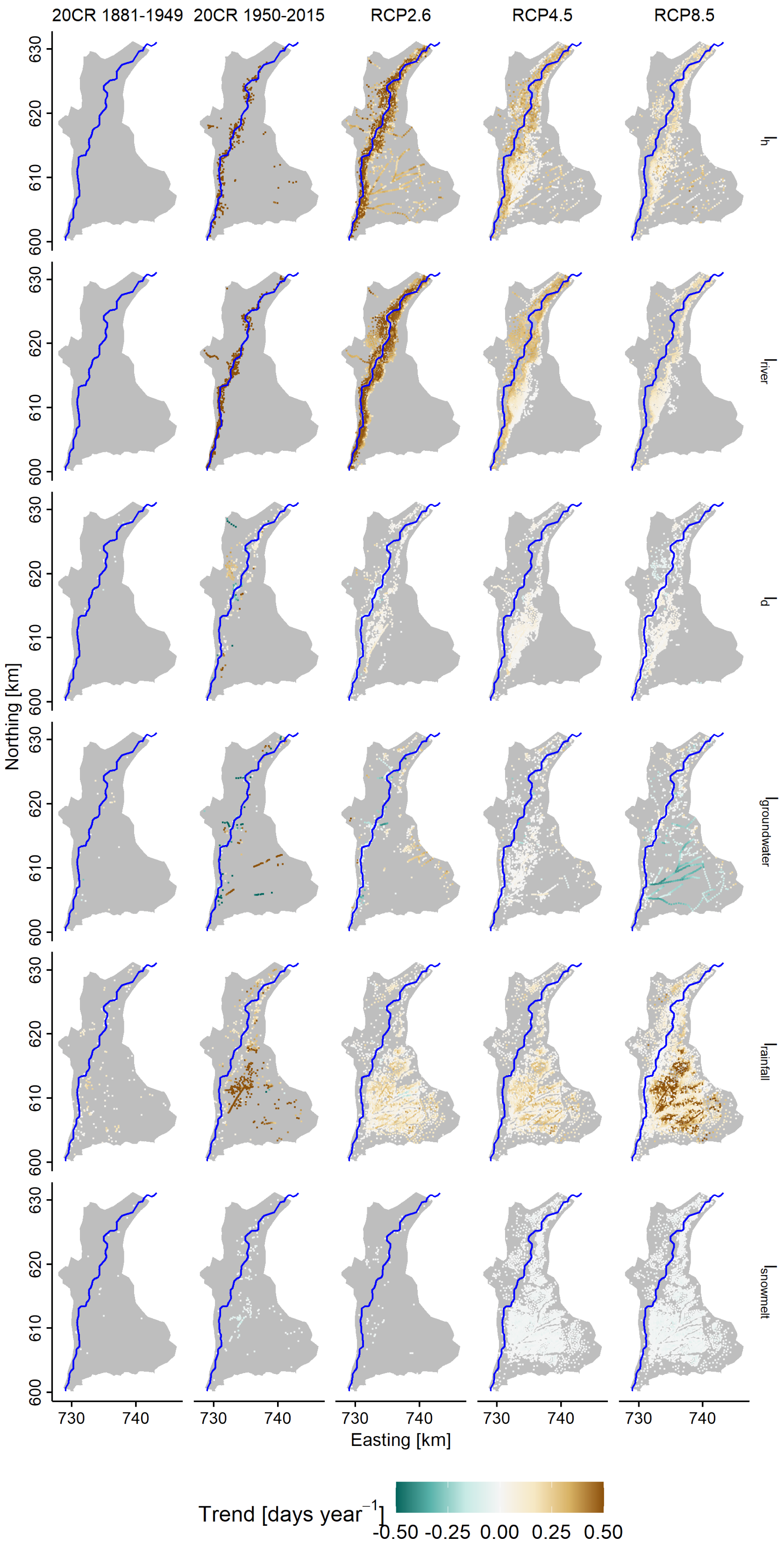

Figure 9. Changes of the period's length when water depth, h, is greater than 1 cm (l_h), river water (l_{river}) or floodplain water ($l_{\mathrm{groundwater}}$, l_{rainfall}, and l_{snowmelt}) fractions are greater than 0.75, and the river-floodplain mixing degree, d, is greater than 0.75 (l_d) annually. Only model nodes with significant trends (p<0.05) are shown. The Grey polygon is the floodplain area, the blue line is the Biebrza River; tributaries and ditches are not shown for clarity, please refer to Figure 1 to identify their course.

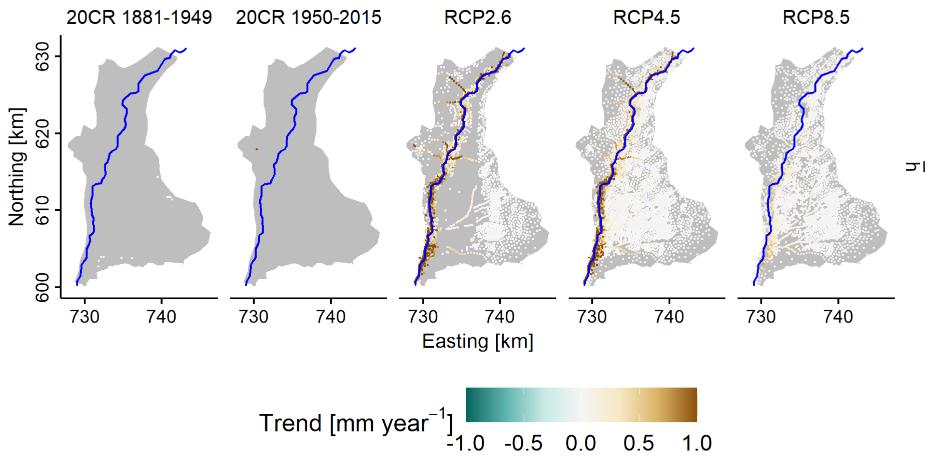

Figure 10. Changes in the annual mean daily water depth, $\bar{h}$. Only model nodes with significant trends (p<0.05) are shown. The color scale is clipped to the <-1, 1> mm year^{-1} range, and the values outside this range are colored as equal to -1, or 1 mm year^{-1}; the clipping affected 15% of the data in RCP 4.5 and 30% of the data in RCP 2.6 located in the proximity of the river. The Grey polygon is the floodplain area, the blue line is the Biebrza River; tributaries and ditches are not shown for clarity, please refer to 1 to identify their course.

4.2 Model development and calibration

An alternative model calibration strategy to the one used in our study was to calibrate the model on a coarser grid and then conduct only fine-tuning in the finer grid (von Gunten et al., 2014). We decided not to use this approach because our target grid was relatively coarse with a number of simplifications. Another approach was to calibrate the model to steady-state using average fluxes as boundary conditions, which was used in several studies involving IHMs (Partington et al., 2020). The advantage of this approach is that the steady-state simulations require shorter simulation time than transient-state simulations for one or more events or hydrological years. We, however, were focused on the dynamic process of flood development, involving interactions of water from groundwater and surface water. Therefore the steady-state calibration for average conditions could lead to unrealistic parameter estimations during flooding, especially for surface water flow parameters for the floodplain.

Still, our strategy with the screening of 800 quasi-random parameter sets was adequate for the model calibration problem. An advantage of this approach is that the approximate total computation time is known a-priori and the problem is easily parallelized on a cluster. A disadvantage is that too sparse parameter space sampling may lead to unsuccessful calibration. The calibration results showed rather high equifinality when only one optimization criterion was analyzed (either KGE or RMSE). However, selecting a model with high KGE and low RMSE considerably decreased the number of behavioral models. At the Pareto front, the relation between KGE and RMSE is non-decreasing (for RMSE < 0.5 m), meaning that the selection of a model with higher KGE results in higher RMSE, i.e., in worse groundwater simulation performance. As indicated in previous studies (e.g. McCabe et al. (2005); Rientjes et al. (2013)), this stresses the importance of using multi-objective calibration when compared to a single-objective calibration.

The porous media parameters were calibrated to realistic values when compared to literature values. This was not entirely the case for the overland flow parameters, where especially the Manning roughness coefficient was higher than expected. This was a result of the generalization of the river channels in the model grid, which resulted in wider and straighter channels than in reality. Eventually, this generalization with realistic roughness parameter values would lead to increased simulated water velocity and too-low water levels. The effect of too high roughness was too high water levels during low flow when water was in the river bed. This effect was reinforced by the high obstruction height parameter value in the river bed, which further increased the roughens for the lowest water levels. The high obstruction height was calibrated in the river bed to compensate for unnaturally-wide perpendicular cross-sections used in the model grid.

The model's purpose was to analyze hydrological conditions during flooding, focusing more on the floodplain area, rather than on the river bed. Further, our aim was to analyze multiple long-term climatic scenarios that require very long computation times. Therefore, in our opinion that the simplifications used herein and the resulting unrealistic surface water parameters in river were justified. While local-scale IHMs are often developed with very fine girds and short time steps, the regional-, country-, or continental-scale model use simplification strategies for model development. One of the strategies used in climate-change studies in regional-scale IHMs is to use aggregated water fluxes in monthly resolution (Goderniaux et al., 2009; Erler et al., 2019). Another strategy is to use a coarser grid, which preserves only key landscape features, such as bigger lakes or major river tributaries. Following this strategy, Goderniaux et al. (2009) used a model with 785 nodes per layer in a 480 km^2 catchment, Erler et al. (2019) used 33092 nodes per layer in a 6800 km^2, and Chen et al. (2019) used about 225000 nodes per layer in 10.5 million km^2 basin. This strategy also involves using relatively thick top layers, which were 1 m in Goderniaux et al. (2009) and 2.5 m in Chen et al. (2019). Our strategy with daily fluxes, 19297 nodes per layer in a 7000 km^2 catchment (refined to 10436 nodes in

220 km^2 floodplain) and about 0.75 m thick top layer makes the model comparable or higher resolution to the mentioned studies.

4.3 Model Validation

The multi-site validation presented in this study showed overall good performance of dynamic hydrological processes simulated in the model. However, a model performance degradation, such as increased RMSE for groundwater heads and decreased KGE for discharge, was observed outside the floodplain area (in the middle basin, upper basin, and upland). This was primarily a result of using a finer grid in the floodplain and a coarser grid elsewhere. Fraction of the error may be attributed to errors in the elevation of the groundwater wells or the DEM used for the model.

From the flooding perspective, the model was unable to simulate correctly the highest discharge peaks (above 250 m^3s^{-1}), which occurred five times in the 1951-2017 period. The remaining events were simulated with smaller errors both in terms of water levels and discharge. We attribute the inability to simulate the highest peak discharges primarily to the too-high roughness coefficients obtained during the calibration, which decrease the water velocity and effectively produce a smaller and wider flood. Partially, this problem may be also attributed to the coarse resolution (1°x1°) 20CR forcing data and bias-correction approach which was not able to suitably force the highest events. Therefore, our model is unsuitable for reliably predicting rare, extreme events in the past and future climate. However, it has demonstrated the capacity to predict normal hydrological behavior including flood events with shorter return periods.

The validation with remote sensing showed good agreement with the spatial and temporal dynamics of river water flooding. This was not the case for the total flooding extent (from the river and all floodplain water sources). Even though a high (5 cm water depth) threshold was used to identify the total flooding extent, the simulation provided a larger and longer-lasting extent than the remote sensing estimate. Apart from the bias in the remote sensing product caused due to spatial resolution which disabled identification of all permanent open water objects, the remote sensing validation indicates that SAR water extent estimation in a densely vegetated wetland area is problematic. Several attempts were made to the problem of vegetation, or other objects obscuring water extent by using auxiliary information such as elevation models (e.g. Mason et al. (2012)). In wetland cases, where a small surface water depth is frequently present and a flat land surface includes micro-topography features, these methods have limited applicability. A recent approach involving multiple polarimetric decomposition models for SAR data in the Biebrza wetland has shown that with a C-band (the same band as in our data set) SAR the identification of open water with vegetation emerging more than 10 cm can be difficult Gierszewska and Berezowski (2022). The solution could be flooding extent mapping in wetlands using SAR sensors with longer (P, or L) wavelengths.

The SAR flooding area correlated more strongly with the river water than with the total flooding extent. This was reflected in the number of small features visible both in SAR and river water flooding extents that diminished in the total flooding extent. This shows a potential application of SAR data in densely vegetated wetlands where they can be used to track the extent of river water flooding. This is not due to different sediment concentrations, which are used for mapping using optical sensors (Mertes, 1997; Park & Latrubesse, 2015), but due to high water depths in the river flooding zone, which can overtop vegetation. As this phenomenon was the case in Biebrza wetlands it does not necessarily have to be the case in other sites, which can still have too low water levels for detection of surface water.

The EC of water can be used as an indicator of the surface water source, as it has higher values in the river than in floodplain water (Chormański et al., 2011). Our results are in agreement with this showing that EC correlated positively with river water fractions and negatively with snowmelt water fractions. Further, the hydrochemical validation shows that all water source fractions are significant predictors of surface water EC, which indicates that the simulated fractions agree with the true water sources. Our previous study (Berezowski et al., 2019) conducted for a single flooding event on a finer grid showed that the simulated fractions agree with water sources derived from a multi-parameter hydrochemical analysis. In the current study, due to high labor intensity, we were unable to repeat the hydrochemical analysis.

We put a lot of emphasis in this study on model validation which is a key step in impact model development in climate change studies. The validation for the Burzyn station was satisfactory and the lack of the trends in observed discharge and water levels were preserved in the model simulations (although one of the trends was significant in the observations and not significant in the simulations forced by the 20CR data). This indicates that the model passes the comprehensive evaluation criteria described in Krysanova et al. (2018) for the Burzyn Station. The remaining stations, situated in the upper parts of the catchment, have in general lower correlation coefficients, however, their KGE is still comparable to the KGE of the outlet. The comprehensive evaluation can be used as an indicator of a robust impact model (Gelfan et al., 2020), therefore, our model is suitable for climate change impact study of the floodplain area.

Needless to say, our evaluation was more comprehensive than described above. This was because the simulation of water mixing, which is a product of interaction between climate, groundwater, and river flooding, requires more confidence in the modeling results than just agreement with observed water levels or discharge. Our remote sensing and EC evaluation criteria indicate that the model is suitable for the analysis of water mixing.

4.4 Changes in Biebrza River flow in the past and future climate

Our future climate impact simulations that show a positive trend (2005-2099) of the mean discharge in Biebrza River are consistent with Roudier et al. (2015), who have shown that less severe droughts and higher flooding discharges will be present in this region. Two studies conducted for nearby catchments (Guber and Narewka) close to Biebrza also indicated decreased severity of droughts in RCP 4.5 (Meresa et al., 2016) and an increase of yearly maximum flows in RCP 4.5 and 8.5 (Osuch et al., 2016). A regional study also showed that both low and high flow will increase by 2100 in RCP 4.5 and 8.5 in the Biebrza catchment, although the ensemble of simulations was inconsistent for the RCP 4.5 in the 2071–2100 period (Piniewski et al., 2017). On the other hand, our finding that no discharge trends (2006-2099) will be in the RCP 8.5 is inconsistent with Alfieri et al. (2015), who showed that the mean daily flow in Biebrza will increase by about 15% (1990-2080) in the RCP 8.5.

The RCP 2.6 ensemble means in our study are associated with the highest uncertainty, because only five EURO-CORDEX simulations were available. Therefore, our findings that the highest trend (2006-2099) in mean discharge will be observed in RCP2.6 have to be considered less robust than the results of the trends in remaining RCPs. Still, this finding is partially supported by Marx et al. (2018), who showed that the 10-20% change in the mean low flow in Biebrza River will take place under 2K air temperature increase, whereas under 1.5K and 3K scenarios the change will be between -10% and 10%.

Projections of future hydrological impact often disagree due to differences in forcing-data sources and processing, impact models used, impact indicators, and methods of comparison with the reference period (Z. W. Kundzewicz et al., 2016). All of these reasons are relevant for comparisons presented in this section. The aim of this study was not to compare the climate change impact on the Biebrza River with other studies but to investigate the impact on the water mixing using the best methods available. We used all available EURO-CORDEX simulations which provided the required forcing data. However, these simulations used often different GCMs or RCMs than in the discussed studies. Moreover, our simulations were limited by the use of data from only one RCM (but multiple GCMs), while most of the remaining studies used more than one RCM. Also, we ran continuous simulations in a daily resolution for the 1881-2099 period, which was used to calculate trends and their statistical significance for variables relevant to our study. This was not the case in the other studies discussed in this section, which calculated a relative change of low- or high-flow indicators with respect to a reference period. Finally, we used a finer spatial resolution and/or better physical representation of hydrological processes in the HydroGeoSphere model than in models used in these studies.

4.5 Changes of water sources fraction in the past and future climate

Water source fractions were stable in the 1881-2015 period in terms of the associated volume of water, which coincide with no trends in the forcing data. Since the second half of the 20th century, a shift from rainfall replacing snowmelt fractions dominance was observed in the central part of the floodplain. In parallel, river fractions and high water depths persisted longer in the proximity of the river due to the rainfall accumulation in the whole catchment. Neither of these changes was related to a significant trend in rainfall or snowmelt in the forcing data, nor resulted in a significant change in the flooding volume of these water sources.

In the RCP 2.6 volume of river and rainfall water significantly increased during the 2005-2099 period. An increase in rainfall with a significant, but eight-fold smaller decrease in snowfall, and no change in PET resulted in overall wetter conditions. This translated not only to increased river discharge, and high river fraction persistence but also to longer-lasting high groundwater fractions and surface water depth. Effectively, the period of river-floodplain water mixing was longer in the proximity of the river by forming a clear belt and resulting in a significantly increased volume of mixing water.

A similar situation was observed in RCP 4.5, but due to a two-fold higher decrease of snowfall and similar magnitude of rainfall trend as in RCP 2.6 period of high snowmelt fractions shortened during the 2005-2099 period. In addition to that, a greater decrease in PET in RCP 4.5 than in RCP 2.6 resulted in less wet conditions. As a result, a lesser trend of river discharge and water levels was observed. Since groundwater discharge is more related to overall drier or wetter conditions rather than to instantaneous fluxes of water the decrease in snowmelt water resulted in longer dominance of groundwater fractions in the river proximity in RCP 4.5. Despite drier conditions in RCP 4.5, the longer periods of groundwater, rainfall, and river fractions in this area affected in the river-floodplain mixing zone last longer in greater areas than in RCP 2.6.

A different situation was observed in RCP 8.5, where snowfall nearly ceased and the increase of rainfall was in great part balanced by the increase of PET resulting in no trends observed in discharge or water levels in the 2005-2099 period. The stability of discharges was not accompanied by the stability of the volume of water from different sources, as rainfall volume increased and snowmelt volume decreased. The most distinctive pattern of high groundwater fractions persistence decrease was observed in RCP 8.5. Such a big groundwater fraction decrease is an indicator of drought conditions locally in the central part of the floodplain. This was however balanced by longer persistence in the northern part of the floodplain and near the valley margin, resulting in no significant trend in groundwater volume. The area in which the river-floodplain water mixing period increased was similar to in RCP 4.5 but more patchy, and the magnitude of the trend was smaller. Despite this spatial pattern the trend of mixing water volume did not change significantly. Unlike in other scenarios, in RCP 8.5 longer lasting high rainfall fractions resulted in a zone of a shortened period of river-floodplain mixing in the NE part of the floodplain, near the river.

Trends in mean water depth and inundation period length did not align with the trends of water source fractions. Overall, the trends of mean water depth were rather small except in the river proximity. Also, greater variability was present in the water source fraction trends than in water depth or inundation period trends.

4.6 Implications for modeling

When taking into account distributed hydrological processes, the opposite directions of changes can be present within one floodplain. These changes either combine with each other to amplify a signal in the lumped volume or cancel each other and the signal in the lumped volume attenuates. A comparison of the lumped volume of water associated with different fractions and the accompanying spatial patterns in the floodplain area shows the advantage of using fine, distributed model output. As illustrated for the 1951-2015 period, the climatic signal was lost in the lumped output due to averaging with other effects, whereas a clear trend pattern was visible spatially. A similar situation was observed in the RCP 8.5 scenario, where the greatest spatial trends were present in groundwater fractions persistence, but the trend in groundwater volume was not significant. This is relevant if upon the lumped or distributed impact model output another process (e.g. ecological, or hydrochemical) would be modeled or a management decision would be undertaken.

This links to another advantage of using IHMs in climate change scenarios, which is revealed when data on all boundary conditions are not explicitly available temporarily and spatially. Climate influences both surface water and groundwater and thereby affects also feedback between the two domains expressed as groundwater-surface water interactions. The assumption about surface water infiltration, or groundwater discharge is difficult to make properly for simulations with long time horizons, whereas, they are required for surface water models that are not integrated with groundwater models. This is not the case for catchment-scale IHMs, where surface water and groundwater simulations are simultaneously forced by the climate data and the time-variable feedback between the two domains are preserved.

As illustrated in our study, the period of inundation with water depth above 1 cm, was also influenced by climate outside the river water flooding zone. Moreover, the trend of this period was not correlated to the distance from the river, as it increased near the river, then decreased, and increased again in the central part of the floodplain (RCP 4.5 and 8.5), or it increased in the entire floodplain (RCP 2.6). The trend in water depth change was correlated to the distance from the river, however, significant positive trends were observed both in areas dominated by the river and floodplain water. Furthermore, water depth trends in RCP 2.6 showed that large areas of the floodplain did not have any trends, while the most remote parts of the floodplain had a significant positive trend. This depicts another advantage of IHMs, which is the representation of water depths in the floodplain. Hydrodynamic models for 2D surface water routing perform very well in simulating river water flooding extent, whereas, are unable to simulate inundation from other sources, such as groundwater, without coupling with different models (Appledorn et al., 2019). Still, some studies use a surface-water-only model to analyze long-term floodplain inundation changes (Veijalainen et al., 2010; Wen et al., 2013). While in some areas a lack of groundwater coupling may not influence the results, in other areas it may be a source of bias in simulated inundation extent.

Although not discussed in this study, interactions between water sources may influence the surface water velocity field in the floodplain in reference to a situation when river water is the sole inundation source. This may further influence the sedimentation pattern in the floodplain due to the settling velocity parameter of the particles. Several climate impact studies analyze the floodplain sedimentation patterns by taking into account the major water sources, such as rivers (Park et al., 2022) or sea level (Manh et al., 2015). Whereas, as illustrated in this study, the river-floodplain mixing zone is relatively wide and it varies under climate change, which may affect sedimentation patterns.

4.7 Implications for ecological processes

Mixing of water from different sources creates biogeochemical hot spots and hot moments, such as denitrification (McClain et al., 2003). A number of studies have analyzed denitrification spatially in inundated floodplains to reveal that it is strongly affected by connectivity with river water (Forshay & Stanley, 2005; Racchetti et al., 2011; Jones et al., 2014; Scott et al., 2014). Our study shows that, that the river-floodplain water mixing volume, extent, and persistence varies with climate change, therefore denitrification patterns can also be affected. This variability was visible much better in the spatial pattern than in the lumped, water volumes. Therefore, improvement in denitrification modeling at floodplain (Hallberg et al., 2022) or catchment (Adame et al., 2019) scale or in the use of scaling relationships (O'Connor et al., 2006; Tomasek et al., 2017) could be achieved by introducing additional variables related to groundwater discharge, river water, or the river-floodplain mixing extents.

Another process, that is related to the extent of water from different sources is vegetation development (Keizer et al., 2014; Park & Latrubesse, 2015). Modeling of vegetation development under climate change may be hampered because studies using the process-based model (Politti et al., 2014), a statistical approach (Mosner et al., 2015) do not include hydrodynamic feedback between water from different sources. As shown by (Gattringer et al., 2019) predictors from an IHM improve habitat modelling in comparison to groundwater, or surface water only predictors scenarios. Our results indicated that in some scenarios the trends were not present in water levels, or discharges while they were present in the persistence of dominant water sources. Therefore we believe, that the inclusion of water sources extents predictors could improve vegetation models further. We are not aware of any study that used IHM-simulated water sources to model vegetation development or distribution.

A more recent study conducted in the Biebrza floodplain revealed that the vegetation productivity was better predicted by the zone of nutrients rich sediment deposition, located close to the river, rather than by the river water extent or the total inundation extent (Keizer et al., 2018). As mentioned in Section 4.6, sedimentation is related to water velocity, which may decrease where water sources with different momentum mix. Therefore, the mixing degree, d, which was strongly variable under climate change in this study, can potentially be a candidate for high productivity vegetation zone predictor in temperate floodplains. This, however, was not tested here and should be investigated in a future study.

4.8 Implications for management

The Biebrza floodplain, as part of a national park, has been subjected to active protection measures. An increase in water levels through the construction of dams, or vegetation removal by mowing, allowed, to some extent, to diminish the potential effect of climate change in this area (Berezowski et al., 2018). Our results together with other experiments discussed herein show that the analysis of water sources and their mixing may have a considerable ecological effect. However, at this point, more models are needed to asses this effect more precisely spatially and temporarily. Therefore, the current local management strategy could be to increase the resilience of the wetland ecosystem and implementation of adaptive management (Lawler, 2009). Except that, the local management strategies may be somewhat challenging, as tools for preserving the shape and duration of water sources' zones are limited. On the other hand, our results have shown that the spatially distributed trends in water source fractions were driven solely by climate change, as our model neglected other divers (water use, land-use change, etc.). Therefore, global actions limiting climate change impact on wetlands driven by national and international policies (Moomaw et al., 2018) seem to be an appropriate measure to limit the shift in the extent of water from different sources.

4.9 Note on hardware requirements

The simulations were run on the Tryton cluster, which has 3215 Intel Xeon Processors (E5 v3, 2.3 GHz, 12-core) with 128 GB RAM, resulting in a total of 1.792 PFLOPS. We split the simulations into 978 smaller tasks (a three-year simulation period with a two-year warm-up period), to use the resources in parallel and to fit into 72h wall time for a single simulation. The cluster resources were shared with other users therefore it took about five months to finish all computations. The total output data produced by the models accounted for about 20TB.

5 Summary and conclusions

Simulations of surface water source fractions in a natural wetland floodplain over a two-century period reveal that by 2099 the projected future climate change will significantly alter the patterns that were relatively stable in the 1881-2015 period. Our results show that analysis of the lumped output of the model was less sensitive to depict the climate change effect that was visible when the trends were analyzed spatially in the floodplain. Different future climate scenarios showed very variable impacts on water source fractions, which were often counterintuitive. In the RCP 2.6, which projected the least climate change in the study area, we observed the highest magnitude of changes related to the increase in river discharges, water levels, and river water fractions. In the RCP 8.5 scenario, which projected the greatest increase in PET and rainfall accompanied by the greatest decrease in snowfall, these trends were less significant, while only this scenario projected dry conditions exhibited by a decrease of groundwater fractions in the inundation. The trends in water source fractions had different spatial patterns and showed greater sensitivity to climate change than trends in water depth and inundation duration.

This complex hydrological impact was simulated by the IHM, which allowed us to model interactions between groundwater and surface water and limit the assumptions about hydrological fluxes in the top layer of the model to the meteorological forcing. This is the first study that simulated the climate impact on water source fractions in the inundation and the longest application of IHM in terms of the simulation period. Hydrological impact studies are always related to uncertainty, which we limited here by multi-variable verification and projection of future impact using an ensemble of 10 EURO-CORDEX simulations (only 4 in RCP 2.6).

We showed that the water source fractions are sensitive to the climate in a natural temperate zone wetland floodplain. This fact has several implications for other modeling studies, ecological processes, and management in similar wetlands. Modeling problems should be carried out using IHMs to depict proper inundation or sedimentation patterns spatially, because, even if the water sources fractions are not explicitly simulated using HMC, IHMs capture the interactions between water from different sources which produce inundation outside the river water zone and changes the velocity field. Since ecological processes, such as denitrification or vegetation development, are in part related to water sources' zonation and their mixing, these variables should be taken into account in models, especially in climate change impact studies. Finally, the managers have limited tools in shaping the surface water zonation and extent, therefore except for increasing the wetlands resilience, and adaptive management using an IHM output, global actions aimed at decreasing climate change impact should be the main priority.

Open Research Section

The IHM simulation output used for the analysis, forcing data, and historical water levels (which were not published elsewhere, see below) are available in (Berezowski, 2023). The groundwater levels data was provided by the Biebrza National Park, the data is available upon request from `https://www.biebrza.org.pl/`. Meteorological observations of snowfall, water levels in the rivers, and river discharge was provided by Institute of Meteorology and Water Management - National Research Institute (IMGW-PIB),

Poland; data is available at `https://danepubliczne.imgw.pl/`. Support for the Twentieth Century Reanalysis Project version 3 dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), by the National Oceanic and Atmospheric Administration Climate Program Office, and by the NOAA Earth System Research Laboratory Physical Sciences Laboratory; NOAA/CIRES/DOE 20th Century Reanalysis (V3) data provided by the NOAA PSL, Boulder, Colorado, USA, from their website at `https://psl.noaa.gov`. We thank to: Polish Geological Institute, National Research Institute `https://www.pgi.gov.pl/en/data-bases.html` for providing geological data, Head Office of Geodesy and Cartography (GUGiK) `https://www.geoportal.gov.pl/` for providing the Digital Elevation Model of Poland, Water Authority of Poland (Wody Polskie) for providing the Map of the Hydrographic Division of Poland in scale 1:10 000, EURO-CORDEX initiative `https://www.euro-cordex.net/060378/index.php.en`, and the Joint Research Center Agri4Cast `https://agri4cast.jrc.ec.europa.eu/dataportal/` and CORINE Land Cover `https://land.copernicus.eu/pan-european/corine-land-cover` for sharing the data required for this research.

Acknowledgments

This research was financed by grant: 2017/26/D/ST10/00665 funded by the National Science Centre, Poland. Computations were carried out using the computers of Centre of Informatics Tricity Academic Supercomputer & Network in Poland. We thank to Tomasz Bieliński and Marek Kulawiak for help in field measurement and data processing. We thank the Biebrza National Park for the permission for field measurements in the Park area.

References

Aalto, R., Maurice-Bourgoin, L., Dunne, T., Montgomery, D. R., Nittrouer, C. A., & Guyot, J.-L. (2003). Episodic sediment accumulation on Amazonian flood plains influenced by El Nino/Southern Oscillation. *Nature*, *425*(6957), 493–497.

Adame, M. F., Roberts, M. E., Hamilton, D. P., Ndehedehe, C. E., Reis, V., Lu, J., . . . Ronan, M. (2019). Tropical coastal wetlands ameliorate nitrogen export during floods. *Frontiers in Marine Science*, *6*. doi: 10.3389/fmars.2019.00671

Alfieri, L., Burek, P., Feyen, L., & Forzieri, G. (2015). Global warming increases the frequency of river floods in Europe. *Hydrology and Earth System Sciences*, *19*(5), 2247–2260. doi: 10.5194/hess-19-2247-2015

Anonymous. (1912). *Summary of water levels on the inland water ways of Russia* *from observations at water gauge posts. period: 1881-1910.* Ministry of communications, internal control, waterways, and roads.

Anonymous. (1932). *Hydrographic yearbook. Vistula basin. period: 1918-1932.* Ministerstwo komunikacji.

Anonymous. (1970). *Hydrologic yearbook of surface waters. the Vistula basin and the rivers of the coast region east of the vistula river. period 1955-1970.* Państwowy Instytut Hydrologiczno-Meteorologiczny, Wydawnictwo komunikacji i łączności.

Anonymous. (1980). *Hydrologic yearbook of surface waters. the Vistula basin and the rivers of the coast region east of the vistula river. period 1971-1980.* Instytut Meteorologii i Gospodarki Wodnej, Wydawnictwo komunikacji i łączności.

Anonymous. (2019). *Public data, period: 1951-2019.* Retrieved from `https://danepubliczne.imgw.pl/data/dane_pomiarowo_obserwacyjne/` (Accessed: 2019-04-21)

Appledorn, M. V., Baker, M. E., & Miller, A. J. (2019). River-valley morphology, basin size, and flow-event magnitude interact to produce wide variation in flooding dynamics. *Ecosphere*, *10*(1). doi: 10.1002/ecs2.2546

Arnell, N. W., & Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, *134*(3), 387–401. doi: 10.1007/s10584-014-1084-5

Banaszuk, H. (2004). *Kotlina biebrzanska i biebrzanski park narodowy.* Bialystok: Ekonomia i Srodowisko. (In Polsih)

Barthel, R., & Banzhaf, S. (2015). Groundwater and surface water interaction at the regional-scale - a review with focus on regional integrated models. *Water Resources Management*, *30*(1), 1–32. doi: 10.1007/s11269-015-1163-z

Berezowski, T. (2023). *Hydrological indicators of water zones in inundation, historical water levels, and forcing data for the 1881-2099 period in the lower biebrza valley.* doi: 10.34808/323p-nd55

Berezowski, T., Bieliński, T., & Osowicki, J. (2020). Flooding extent mapping for synthetic aperture radar time series using river gauge observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2626-2638. doi: 10.1109/JSTARS.2020.2995888

Berezowski, T., Partington, D., Chormański, J., & Batelaan, O. (2019). Spatiotemporal dynamics of the active perirheic zone in a natural wetland floodplain. *Water Resources Research*, *55*(11), 9544–9562. doi: 10.1029/2019wr024777

Berezowski, T., Szcześniak, M., Kardel, I., Michałowski, R., Okruszko, T., Mezghani, A., & Piniewski, M. (2016). Cplfd-gdpt5: High-resolution gridded daily precipitation and temperature data set for two largest polish river basins. *Earth System Science Data*, *8*(1), 127–139.

Berezowski, T., Wassen, M., Szatyłowicz, J., Chormański, J., Ignar, S., Batelaan, O., & Okruszko, T. (2018). Wetlands in flux: looking for the drivers in a central european case. *Wetlands Ecology and Management*, *26*(5), 849-863.

Boko, B. A., Konaté, M., Yalo, N., Berg, S. J., Erler, A. R., Bazié, P., ... Sudicky, E. A. (2020). High-resolution, integrated hydrological modeling of climate change impacts on a semi-arid urban watershed in niamey, niger. *Water*, *12*(2), 364. doi: 10.3390/w12020364

Brunner, P., & Simmons, C. T. (2012). Hydrogeosphere: A fully integrated, physically based hydrological model. *Ground Water*, *50*(2), 170–176.

Chen, J., Sudicky, E. A., Davison, J. H., Frey, S. K., Park, Y.-J., Hwang, H.-T., ... Peltier, W. R. (2019, oct). Towards a climate-driven simulation of coupled surface-subsurface hydrology at the continental scale: a Canadian example. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, *45*(1), 11–27. doi: 10.1080/07011784.2019.1671235

Chormański, J., Okruszko, T., Ignar, S., Batelaan, O., Rebel, K., & Wassen, M. (2011). Flood mapping with remote sensing and hydrochemistry: a new method to distinguish the origin of flood water during floods. *Ecological Engineering*, *37*(9), 1334-1349.

Chow, V. T., Maidment, D., & Mays, L. (1988). *Applied hydrology.* McGraw-Hill.

Commission of the European Communities. (2013). *Corine land-cover.* Retrieved from `http://www.eea.europa.eu/publications/COR0-landcover` (Date accessed: 2013-10-12)

Dąbrowska-Zielińska, K., Budzyńska, M., Tomaszewska, M., Bartold, M., Gatkowska, M., Malek, I., ... Napiórkowska, M. (2014). Monitoring wetlands ecosystems using ALOS PALSAR (L-band, HV) supplemented by optical data: A case study of Biebrza wetlands in northeast Poland. *Remote Sensing*, *6*(2), 1605–1633. Retrieved from `https://www.mdpi.com/2072-4292/6/2/1605` doi: 10.3390/rs6021605

Erler, A. R., Frey, S. K., Khader, O., d'Orgeville, M., Park, Y.-J., Hwang, H.-T., ... Sudicky, E. A. (2019). Evaluating climate change impacts on soil moisture and groundwater resources within a lake-affected region. *Water Resources Research*, *55*(10), 8142–8163. doi: 10.1029/2018wr023822

Eurostat. (2019). *EUROPOP2019 - Population projections at regional level (2019-2100).*

Ferguson, I. M., & Maxwell, R. M. (2010). Role of groundwater in watershed response and land surface feedbacks under climate change. *Water Resources Research*, *46*(10). doi: 10.1029/2009wr008616

Forshay, K. J., & Stanley, E. H. (2005). Rapid nitrate loss and denitrification in a temperate river floodplain. *Biogeochemistry*, *75*(1), 43-64.

Garris, H. W., Mitchell, R. J., Fraser, L. H., & Barrett, L. R. (2014). Forecasting climate change impacts on the distribution of wetland habitat in the Midwestern United states. *Global Change Biology*, *21*(2), 766–776. doi: 10.1111/gcb.12748

Gattringer, J. P., Maier, N., Breuer, L., Otte, A., Donath, T. W., Kraft, P., & Harvolk-Schöning, S. (2019). Modelling of rare flood meadow species distribution by a combined habitat surface water-groundwater model. *Ecohydrology*, *12*(6). doi: 10.1002/eco.2122

Gelfan, A., Kalugin, A., Krylenko, I., Nasonova, O., Gusev, Y., & Kovalev, E. (2020). Does a successful comprehensive evaluation increase confidence in a hydrological model intended for climate impact assessment? *Climatic Change*, *163*(3), 1165–1185. doi: 10.1007/s10584-020-02930-z

Gierszewska, M., & Berezowski, T. (2022). On the role of polarimetric decomposition and speckle filtering methods for C-band SAR wetland classification purposes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 2845-2860. doi: 10.1109/JSTARS.2022.3162641

Giuntoli, I., Vidal, J.-P., Prudhomme, C., & Hannah, D. M. (2015). Future hydrological extremes: the uncertainty from multiple global climate and global hydrological models. *Earth System Dynamics*, *6*(1), 267–285. doi: 10.5194/esd-6-267-2015

Gnatowski, T., Szatyłowicz, J., Brandyk, T., & Kechavarzi, C. (2010). Hydraulic properties of fen peat soils in Poland. *Geoderma*, *154*(3-4), 188–195.

Goderniaux, P., Brouyère, S., Fowler, H. J., Blenkinsop, S., Therrien, R., Orban, P., & Dassargues, A. (2009). Large scale surface-subsurface hydrological model to assess climate change impacts on groundwater reserves. *Journal of Hydrology*, *373*(1-2), 122–138. doi: 10.1016/j.jhydrol.2009.04.017

Gramacy, R. B., & Taddy, M. (2010). Categorical inputs, sensitivity analysis, optimization and importance tempering with tgp version 2, an R package for treed gaussian process models. *Journal of Statistical Software*, *33*(6), 1–48. doi: 10.18637/jss.v033.i06

Grygoruk, M., Kochanek, K., & Mirosław-Świątek, D. (2021). Analysis of long-term changes in inundation characteristics of near-natural temperate riparian habitats in the lower basin of the Biebrza valley, Poland. *Journal of Hydrology:*

Regional Studies, *36*, 100844. doi: 10.1016/j.ejrh.2021.100844

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Technical note: Downscaling RCM precipitation to the station scale using statistical transformation - a comparison of methods. *Hydrology and Earth System Sciences*, *16*(9), 3383–3390. doi: 10.5194/hess-16-3383-2012

Hallberg, L., Hallin, S., & Bieroza, M. (2022). Catchment controls of denitrification and nitrous oxide production rates in headwater remediated agricultural streams. *Science of The Total Environment*, *838*, 156513. doi: 10.1016/j.scitotenv.2022.156513

Hwang, H.-T., Park, Y.-J., Sudicky, E., & Forsyth, P. (2014). A parallel computational framework to solve flow and transport in integrated surface subsurface hydrologic systems. *Environmental Modelling & Software*, *61*, 39–58.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., … Yiou, P. (2014). Euro-cordex: new high-resolution climate change projections for european impact research. *Regional Environmental Change*, *14*(2), 563–578. doi: 10.1007/s10113-013-0499-2

Jacobson, R. B., Bouska, K. L., Bulliner, E. A., Lindner, G. A., & Paukert, C. P. (2022). Geomorphic controls on floodplain connectivity, ecosystem services, and sensitivity to climate change: An example from the lower Missouri River. *Water Resources Research*, *58*(6). doi: 10.1029/2021wr031204

Joint Research Center. (2019). *Agri4Cast Resources Portal.* Retrieved from `https://agri4cast.jrc.ec.europa.eu/dataportal/` (Date accessed: 2019-12-03)

Jones, C. N., Scott, D. T., Edwards, B. L., & Keim, R. F. (2014). Perirheic mixing and biogeochemical processing in flow-through and backwater floodplain wetlands. *Water Resour. Res.*, *50*(9), 7394–7405.

Karim, F., Petheram, C., Marvanek, S., Ticehurst, C., Wallace, J., & Hasan, M. (2015). Impact of climate change on floodplain inundation and hydrological connectivity between wetlands and rivers in a tropical river catchment. *Hydrological Processes*, *30*(10), 1574–1593. doi: 10.1002/hyp.10714

Kaser, D., Graf, T., Cochand, F., McLaren, R., Therrien, R., & Brunner, P. (2014). Channel representation in physically based models coupling groundwater and surface water: Pitfalls and how to avoid them. *Groundwater*, *52*(6), 827–836.

Keizer, F., der Lee, G. V., Schot, P., Kardel, I., Barendregt, A., & Wassen, M. (2018, aug). Floodplain plant productivity is better predicted by particulate nutrients than by dissolved nutrients in floodwater. *Ecological Engineering*, *119*, 54–63.

Keizer, F., Schot, P., Okruszko, T., Chormanski, J., Kardel, I., & Wassen, M. (2014). A new look at the flood pulse concept: The (ir)relevance of the moving littoral in temperate zone rivers. *Ecological Engineering*, *64*(0), 85–99.

Kollet, S., Sulis, M., Maxwell, R. M., Paniconi, C., Putti, M., Bertoldi, G., … Sudicky, E. (2017). The integrated hydrologic model intercomparison project, IH-MIP2: A second set of benchmark results to diagnose integrated hydrology and feedbacks. *Water Resources Research*, *53*(1), 867–890. doi: 10.1002/2016wr019191

Kotowski, W., Jabłońska, E., & Bartoszuk, H. (2013). Conservation management in fens: Do large tracked mowers impact functional plant diversity? *Biological Conservation*, *167*, 292-297. doi: 10.1016/j.biocon.2013.08.021

Kristensen, K. J., & Jensen, S. E. (1975). A model for estimating actual evapotranspiration from potential evapotranspiration. *Hydrology Research*, *6*(3), 170-188.

Krysanova, V., Donnelly, C., Gelfan, A., Gerten, D., Arheimer, B., Hattermann, F., & Kundzewicz, Z. W. (2018). How the performance of hydrological models relates to credibility of projections under climate change. *Hydrological Sciences Journal*, *63*(5), 696–720. doi: 10.1080/02626667.2018.1446214

Kundzewicz, Z., Krysanova, V., Benestad, R., Hov, Ø., Piniewski, M., & Otto, I. (2018, jan). Uncertainty in climate change impacts on water resources. *Environmental Science & Policy*, *79*, 1–8. doi: 10.1016/j.envsci.2017.10.008

Kundzewicz, Z. W., Krysanova, V., Dankers, R., Hirabayashi, Y., Kanae, S., Hattermann, F. F., … Schellnhuber, H.-J. (2016). Differences in flood hazard projections in Europe - their causes and consequences for decision making. *Hydrological Sciences Journal*. doi: 10.1080/02626667.2016.1241398

Laranjeiras, T. O., Naka, L. N., Leite, G. A., & Cohn-Haft, M. (2021). Effects of a major Amazonian river confluence on the distribution of floodplain forest avifauna. *Journal of Biogeography*, *48*(4), 847–860. doi: 10.1111/jbi.14042

Lawler, J. J. (2009, apr). Climate change adaptation strategies for resource management and conservation planning. *Annals of the New York Academy of Sciences*, *1162*(1), 79–98. doi: 10.1111/j.1749-6632.2009.04147.x

Manh, N. V., Dung, N. V., Hung, N. N., Kummu, M., Merz, B., & Apel, H. (2015). Future sediment dynamics in the Mekong Delta floodplains: Impacts of hydropower development, climate change and sea level rise. *Global and Planetary Change*, *127*, 22–33. doi: 10.1016/j.gloplacha.2015.01.001

Marx, A., Kumar, R., Thober, S., Rakovec, O., Wanders, N., Zink, M., … Samaniego, L. (2018, feb). Climate change alters low flows in europe under global warming of 1.5, 2, and 3c. *Hydrology and Earth System Sciences*, *22*(2), 1017–1032. doi: 10.5194/hess-22-1017-2018

Mason, D. C., Davenport, I. J., Neal, J. C., Schumann, G. J.-P., & Bates, P. D. (2012). Near real-time flood detection in urban and rural areas using high-resolution synthetic aperture radar images. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(8), 3041–3052. doi: 10.1109/tgrs.2011.2178030

McCabe, M., Franks, S., & Kalma, J. (2005). Calibration of a land surface model using multiple data sets. *Journal of Hydrology*, *302*(1-4), 209–222. doi: 10.1016/j.jhydrol.2004.07.002

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., ... Pinay, G. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, *6*(4), 301–312. doi: 10.1007/s10021-003-0161-9

Meresa, H., Osuch, M., & Romanowicz, R. (2016, may). Hydro-meteorological drought projections into the 21-st century for selected Polish catchments. *Water*, *8*(5), 206. doi: 10.3390/w8050206

Mertes, L. A. K. (1997). Documentation and significance of the perirheic zone on inundated floodplains. *Water Resour. Res.*, *33*(7), 1749–1762.

Mezghani, A., Dobler, A., Haugen, J. E., Benestad, R. E., Parding, K. M., Piniewski, M., ... Kundzewicz, Z. W. (2017). CHASE-PL climate projection dataset over Poland - bias adjustment of EURO-CORDEX simulations. *Earth System Science Data*, *9*(2), 905–925. doi: 10.5194/essd-9-905-2017

Mohanty, M. P., & Simonovic, S. P. (2021). Fidelity of reanalysis datasets in floodplain mapping: Investigating performance at inundation level over large regions. *Journal of Hydrology*, *597*, 125757. doi: 10.1016/j.jhydrol.2020.125757

Moomaw, W. R., Chmura, G. L., Davies, G. T., Finlayson, C. M., Middleton, B. A., Natali, S. M., ... Sutton-Grier, A. E. (2018). Wetlands in a changing climate: Science, policy and management. *Wetlands*, *38*(2), 183–205. doi: 10.1007/s13157-018-1023-8

Mosner, E., Weber, A., Carambia, M., Nilson, E., Schmitz, U., Zelle, B., ... Horchler, P. (2015). Climate change and floodplain vegetation - future prospects for riparian habitat availability along the Rhine river. *Ecological Engineering*, *82*, 493–511. doi: 10.1016/j.ecoleng.2015.05.013

Murray-Hudson, M., Wolski, P., & Ringrose, S. (2006, nov). Scenarios of the impact of local and upstream changes in climate and water use on hydro-ecology in the Okavango Delta, Botswana. *Journal of Hydrology*, *331*(1-2), 73–84. doi: 10.1016/j.jhydrol.2006.04.041

Natho, S., Tschikof, M., Bondar-Kunze, E., & Hein, T. (2020). Modeling the effect of enhanced lateral connectivity on nutrient retention capacity in large river floodplains: How much connected floodplain do we need? *Frontiers in Environmental Science*, *8*. doi: 10.3389/fenvs.2020.00074

Nogueira, G. E. H., Schmidt, C., Partington, D., Brunner, P., & Fleckenstein, J. H. (2022). Spatiotemporal variations in water sources and mixing spots in a riparian zone. *Hydrology and Earth System Sciences*, *26*(7), 1883–1905. doi: 10.5194/hess-26-1883-2022

O'Connor, B. L., Hondzo, M., Dobraca, D., LaPara, T. M., Finlay, J. C., & Brezonik, P. L. (2006). Quantity-activity relationship of denitrifying bacteria and environmental scaling in streams of a forested watershed. *Journal of Geophysical Research: Biogeosciences*, *111*(G4). doi: 10.1029/2006jg000254

Osuch, M., Lawrence, D., Meresa, H. K., Napiorkowski, J. J., & Romanowicz, R. J. (2016, aug). Projected changes in flood indices in selected catchments in Poland in the 21st century. *Stochastic Environmental Research and Risk Assessment*, *31*(9), 2435–2457. doi: 10.1007/s00477-016-1296-5

Paiva, R. C. D., Collischonn, W., & Buarque, D. C. (2012). Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon. *Hydrological Processes*, *27*(3), 333–346. doi: 10.1002/hyp.8425

Pałczyński, A. (1984). Natural differentiation of plant communities in relation to hydrological conditions of the biebrza valley. *Polish Ecological Studies*, *10*, 347–385.

Park, E., Ho, H. L., Binh, D. V., Kantoush, S., Poh, D., Alcantara, E., ... Lin, Y. N. (2022). Impacts of agricultural expansion on floodplain water and sediment budgets in the Mekong River. *Journal of Hydrology*, *605*, 127296. doi: 10.1016/j.jhydrol.2021.127296

Park, E., & Latrubesse, E. M. (2015). Surface water types and sediment distribution patterns at the confluence of mega rivers: The Solimões-Amazon and Negro Rivers junction. *Water Resources Research*, *51*(8), 6197–6213.

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., ... Fleckenstein, J. H. (2013). Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment. *Water Resour. Res.*, *49*(9), 5501–5519. Retrieved from `http://dx.doi.org/10.1002/wrcr.20405`

Partington, D., Brunner, P., Simmons, C., Therrien, R., Werner, A., Dandy, G., & Maier, H. (2011). A hydraulic mixing-cell method to quantify the groundwater component of streamflow within spatially distributed fully integrated surface water-groundwater flow models. *Environmental Modelling & Software*, *26*(7), 886–898.

Partington, D., Knowling, M. J., Simmons, C. T., Cook, P. G., Xie, Y., Iwanaga, T., & Bouchez, C. (2020). Worth of hydraulic and water chemistry observation data in terms of the reliability of surface water-groundwater exchange flux predictions under varied flow conditions. *Journal of Hydrology*, *590*, 125441. doi: 10.1016/j.jhydrol.2020.125441

Perra, E., Piras, M., Deidda, R., Paniconi, C., Mascaro, G., Vivoni, E. R., ... Meyer, S. (2018). Multimodel assessment of climate change-induced hydrologic impacts for a Mediterranean catchment. *Hydrology and Earth System Sciences*, *22*(7), 4125–4143. doi: 10.5194/hess-22-4125-2018

Piniewski, M., Szcześniak, M., Kardel, I., Chattopadhyay, S., & Berezowski, T. (2021). G2dc-pl+: a gridded 2 km daily climate dataset for the union of the Polish territory and the Vistula and Odra basins. *Earth System Science Data*, *13*(3), 1273–1288. doi: 10.5194/essd-13-1273-2021

Piniewski, M., Szcześniak, M., Kundzewicz, Z. W., Mezghani, A., & Hov, Ø. (2017). Changes in low and high flows in the Vistula and the Odra basins: Model projections in the European-scale context. *Hydrological Processes*, *31*(12), 2210–2225. doi: 10.1002/hyp.11176

Polish Geological Institute. (2014). *Ikar geoportal.* Retrieved from `ikar.pig.gov.pl`

Politti, E., Egger, G., Angermann, K., Rivaes, R., Blamauer, B., Klösch, M., ... Habersack, H. (2014). Evaluating climate change impacts on Alpine floodplain vegetation. *Hydrobiologia*, *737*(1), 225–243. doi: 10.1007/s10750-013-1801-5

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., ... Wisser, D. (2013). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences*, *111*(9), 3262–3267. doi: 10.1073/pnas.1222473110

Racchetti, E., Bartoli, M., Soana, E., Longhi, D., Christian, R. R., Pinardi, M., & Viaroli, P. (2011). Influence of hydrological connectivity of riverine wetlands on nitrogen removal via denitrification. *Biogeochemistry*, *103*(1), 335-354.

Ramteke, G., Singh, R., & Chatterjee, C. (2020). Assessing impacts of conservation measures on watershed hydrology using MIKE SHE model in the face of climate change. *Water Resources Management*, *34*(13), 4233–4252. doi: 10.1007/s11269-020-02669-3

Rawson, H. M., & Macpherson, H. G. (n.d.). *Irrigated wheat.* FAO. Retrieved from www.fao.org/3/X8234E/X8234E00.htm

Rientjes, T., Muthuwatta, L., Bos, M., Booij, M., & Bhatti, H. (2013). Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *Journal of Hydrology*, *505*, 276–290. doi: 10.1016/j.jhydrol.2013.10.006

Roudier, P., Andersson, J. C. M., Donnelly, C., Feyen, L., Greuell, W., & Ludwig, F. (2015, nov). Projections of future floods and hydrological droughts in Europe under a +2c global warming. *Climatic Change*, *135*(2), 341–355. doi: 10.1007/s10584-015-1570-4

Scaroni, A. E., Nyman, J. A., & Lindau, C. W. (2011). Comparison of denitrification characteristics among three habitat types of a large river floodplain: Atchafalaya River Basin, Louisiana. *Hydrobiologia*, *658*(1), 17-25.

Schneider, C., Laizé, C. L. R., Acreman, M. C., & Flörke, M. (2013). How will climate change modify river flow regimes in Europe? *Hydrology and Earth System Sciences*, *17*(1), 325–339. doi: 10.5194/hess-17-325-2013

Scott, D. T., Keim, R. F., Edwards, B. L., Jones, C. N., & Kroes, D. E. (2014). Floodplain biogeochemical processing of floodwaters in the Atchafalaya River Basin during the Mississippi River flood of 2011. *Journal of Geophysical Research: Biogeosciences*, *119*(4), 537–546. (2013JG002477)

Sebben, M. L., Werner, A. D., Liggett, J. E., Partington, D., & Simmons, C. T. (2013). On the testing of fully integrated surface-subsurface hydrological models. *Hydrological Processes*, *27*(8), 1276-1285.

Shewchuk, J. (1996). Triangle: Engineering a 2d quality mesh generator and delaunay triangulator. In M. Lin & D. Manocha (Eds.), *Lecture notes in computer science* (Vol. 1148, p. 203-222). Springer Berlin Heidelberg.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., ... Wyszyński, P. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876-2908. doi: 10.1002/qj.3598

Statistics Poland. (2021). *Population by sex, feminization rate, population density. Teritorial units: Podlaskie and Warminsko-Mazurskie. as of day 31 XII 2021. accessed: 18 XII 2022.* Retrieved from swaid.stat.gov.pl

Suliga, J., Chormański, J., Szporak-Wasilewska, S., Kleniewska, M., Berezowski, T., van Griensven, A., & Verbeiren, B. (2015). Derivation from the Landsat 7 NDVI and ground truth validation of LAI and interception storage capacity for wetland ecosystems in Biebrza Valley, Poland. In C. M. U. Neale & A. Maltese

(Eds.), *Remote sensing for agriculture, ecosystems, and hydrology XVII* (Vol. 9637, p. 96371Z). SPIE. doi: 10.1117/12.2194975

Sulis, M., Paniconi, C., Marrocu, M., Huard, D., & Chaumont, D. (2012). Hydrologic response to multimodel climate output using a physically based model of groundwater/surface water interactions. *Water Resources Research*, *48*(12). doi: 10.1029/2012wr012304

Sulis, M., Paniconi, C., Rivard, C., Harvey, R., & Chaumont, D. (2011). Assessment of climate change impacts at the catchment scale with a detailed hydrological model of surface-subsurface interactions and comparison with a land surface model. *Water Resources Research*, *47*(1). doi: 10.1029/2010wr009167

Thompson, J. R., Crawley, A., & Kingston, D. G. (2016). GCM-related uncertainty for river flows and inundation under climate change: the inner niger delta. *Hydrological Sciences Journal*, *61*(13), 2325–2347. doi: 10.1080/02626667.2015.1117173

Thompson, J. R., Gavin, H., Refsgaard, A., Sørenson, H. R., & Gowing, D. J. (2008). Modelling the hydrological impacts of climate change on UK lowland wet grassland. *Wetlands Ecology and Management*, *17*(5), 503–523. doi: 10.1007/s11273-008-9127-1

Tomasek, A., Kozarek, J. L., Hondzo, M., Lurndahl, N., Sadowsky, M. J., Wang, P., & Staley, C. (2017). Environmental drivers of denitrification rates and denitrifying gene abundances in channels and riparian areas. *Water Resources Research*, *53*(8), 6523–6538. doi: 10.1002/2016wr019566

Veijalainen, N., Lotsari, E., Alho, P., Vehviläinen, B., & Käyhkö, J. (2010). National scale assessment of climate change impacts on flooding in Finland. *Journal of Hydrology*, *391*(3-4), 333–350. doi: 10.1016/j.jhydrol.2010.07.035

von Gunten, D., Wohling, T., Haslauer, C., Merchan, D., Causape, J., & Cirpka, O. (2014). Efficient calibration of a distributed pde-based hydrological model using grid coarsening. *Journal of Hydrology*, *519*, 3290–3304.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2013). The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework. *Proceedings of the National Academy of Sciences*, *111*(9), 3228–3232. doi: 10.1073/pnas.1312330110

Wassen, M. J., Okruszko, T., Kardel, I., Chormanski, J., Swiatek, D., Mioduszewski, W., . . . Meire, P. (2006). Eco-hydrological functioning of the biebrza wetlands: Lessons for the conservation and restoration of deteriorated wetlands rid c-7306-2008. *Wetlands: Functioning, Biodiversity Conservation, and Restoration*, *191*, 285–310.

Wen, L., Macdonald, R., Morrison, T., Hameed, T., Saintilan, N., & Ling, J. (2013). From hydrodynamic to hydrological modelling: Investigating long-term hydrological regimes of key wetlands in the Macquarie Marshes, a semi-arid lowland floodplain in Australia. *Journal of Hydrology*, *500*, 45–61. doi: 10.1016/j.jhydrol.2013.07.015

Wösten, J., Lilly, A., Nemes, A., & Bas, C. L. (1999). Development and use of a database of hydraulic properties of European soils. *Geoderma*, *90*(3-4), 169–185.

Yuan, X., Lu, Y., Jiang, L., Liang, S., Jiang, Y., & Xiao, F. (2021). Runoff responses to climate change in China's Buyuan River basin. *River Research and Applications*, *37*(8), 1134–1144. doi: 10.1002/rra.3785

Zhang, Y., Wang, Y., Chen, Y., Liang, F., & Liu, H. (2019). Assessment of future flash flood inundations in coastal regions under climate change scenarios—a case study of Hadahe River basin in northeastern China. *Science of The Total Environment*, *693*, 133550. doi: 10.1016/j.scitotenv.2019.07.356

Zulkafli, Z., Buytaert, W., Manz, B., Rosas, C. V., Willems, P., Lavado-Casimiro, W., ... Santini, W. (2016). Projected increases in the annual flood pulse of the Western Amazon. *Environmental Research Letters*, *11*(1), 014013. doi: 10.1088/1748-9326/11/1/014013

1412 ## Supplement

Table S1. Forcing data sources used in this study. The EURO-CORDEX data were available for different RCP simulations. Each RCP simulation period was 2006-2100 followed by historical simulations of different lengths. Observations indicated here were used only for hydrological model forcing, not for bias correction.

Data (Institute-GCM)	Period	RCP
CNRM-CERFACS-CNRM-CM5	1970-2100	4.5, 8.5
ICHEC-EC-EARTH	1970-2100	2.6, 4.5, 8.5
MOHC-HadGEM2-ES	1970-2100	2.6, 4.5, 8.5
MPI-M-MPI-ESM-LR	1970-2100	2.6, 4.5, 8.5
NCC-NorESM1-M	1970-2100	2.6, 4.5, 8.5
CCCma-CanESM2	1951-2100	4.5, 8.5
CSIRO-QCCCE-CSIRO-Mk3-6-0	1951-2100	4.5, 8.5
IPSL-IPSL-CM5A-MR	1951-2100	4.5, 8.5
MIROC-MIROC5	1951-2100	2.6, 4.5, 8.5
NOAA-GFDL-GFDL-ESM2M	1951-2100	4.5, 8.5
Observations	2005-2019	-
20CR	1880-2005	-

Table S2. Data sources for hydrological validation. Hydrological variables are H - water levels in rivers, Q - discharge, and G - groundwater head. Periods of missing data are not indicated in this table. Unreferenced data sources are available upon request from the authority.

Data source	Variables	Stations	Period	Frequency
Russian hydrological yearbook (Anonymous, 1912)	H	Osowiec	1881-1910	daily
Polish hydrological yearbook Anonymous (1932, 1970, 1980)	H, Q	Q and H: Burzyn, Osowiec, Q: Czachy, Rudzki, Sztabin	1918-1980	daily
IMGW unpublished data archive	H	Osowiec	1924	daily
IMGW Public data repository Anonymous (2019)	H, Q	Q and H: Burzyn, Osowiec, Q: Czachy, Rudzki, Sztabin	1951-2019	daily
Biebrza National Park database	G	41 groundwater wells in the national park	1994-2019	10-days (median)
Household wells measurements	G	2 wells in the Biebrza catchment	1999-2002	once per year

Table S3. Calibration parameters ranges with their constrains and transformations

Porous media parameters	units	class	min.	max.	Constrains and transformation
Van Genuchten model inverse of the air-entry pressure head, α	m^{-1}	glacial till	0.008	0.03	-
		peat	1.2	2.6	
		sand	0.008	0.03	
Van Genuchten model pore-size distribution index, β	-	glacial till	1.3	3	-
		peat	1.3	1.65	
		sand	1.3	3	
Saturated hydraulic conductivity	ms^{-1}	glacial till	1.00E-07	5.00E-03	Logarithmic, with base=10, transformation.
		peat	1.00E-07	5.00E-04	
		sand	1.00E-07	5.00E-03	
Porosity	-	glacial till	0.32	0.45	-
		peat	0.8	0.92	
		sand	0.32	0.45	

Table S4. Calibration parameters ranges with their constrains and transformations

Evapotranspiration parameters	units	class	min.	max.	Constrains and transformation
transpiration fitting parameter, c1	-	Ten vegetation classes	0.001	1.3	One parameter value was selected randomly [0-1] for all vegetation types and scaled using an inversion of maximum leaf area index for a given vegetation type.
Lower limit of soil saturation for transpiration, e1	-	upland, wetland	0.133	1	The evaporation limiting saturations: e1 and e2 parameters were derived simultaneously for each vegetation type from the gamma distribution using: $1 - (\mathrm{g}(p, s) / \mathrm{g}(1, s))$, where g is a function returning quantiles of gamma distribution, p is the probability of 0.05 for e1 and 0.6 for e1, and s [0-1] is shape parameter of gamma distribution provided during the calibration. The rate parameter of the gamma distribution is 1.
Upper limit of soil saturation for transpiration, e2	-	upland, wetland	0.08	0.951	
field capacity, fc	-	upland	0.3	0.951	The transpiration limiting saturations parameters: wp, fc, ox, and aox parameters were derived simultaneously for each vegetation type from the gamma distribution using: $1 - (\mathrm{g}(p, s) / \mathrm{g}(1, s))$, where g is a function returning quantiles of gamma distribution, p is the probability of 0.001 for wp, 0.05 for fc, 0.6 for ox and 0.99 for aox, and s [0-1] is shape parameter of gamma distribution provided during the calibration. The rate parameter of the gamma distribution is 1.
		wetland	0.3	0.87	
wilting point, wp	-	upland	0.09	0.41	-
		wetland	0.09	0.33	
oxic limit, ox	-	upland	0.46	1	-
		wetland	0.46	0.99	
anoxic limit, aox	-	upland, wetland	0.56	1	-

Table S5. Calibration parameters ranges with their constrains and transformations

Surface water flow parameter	units	class	min.	max.	Constrains and transformation
Manning roughness coefficient	$ms^{-\frac{1}{3}}$	Lower Biebrza	0.06	0.25	-
		Major rivers	0.015	0.05	
		Other rivers	0.02	0.05	
		Upland	0.015	0.05	
		Upper Biebrza	0.02	0.2	
		Flood-plain	0.02	0.2	
obstruction height	m	Major rivers	0.05	0.4	-
		Other	0.05	0.4	
		Flood-plain	0.01	0.4	

Table S6. Daily mean values of observations and bias-corrected 20CR and EURO-CORDEX data. A summary is presented for the period 1970-2005 except the PET, which was summarized for 1979-2005. The 20CR diff. row presents the observations subtracted from 20CR values. The EURO-CORDEX mean diff. row presents the mean difference of observations subtracted from each EURO-CORDEX simulations values.

Data source	Total precipitation [mm]		Snowfall [mm]		PET [mm]		Air temperature [K]	
	mean	sd	mean	sd	mean	sd	mean	sd
Observations	1.84	3.40	0.24	1.09	1.70	1.42	280.20	8.48
20CR	2.06	3.78	0.23	1.04	1.73	1.45	279.24	8.56
CNRM-CERFACS-CNRM-CM5	2.08	3.65	0.24	1.05	1.69	1.41	279.13	8.46
ICHEC-EC-EARTH	2.08	3.62	0.24	1.03	1.69	1.41	279.14	8.45
MOHC-HadGEM2-ES	2.06	3.52	0.24	1.02	1.70	1.42	279.18	8.45
MPI-M-MPI-ESM-LR	2.08	3.61	0.24	1.04	1.70	1.42	279.22	8.40
NCC-NorESM1-M	2.07	3.47	0.24	1.04	1.69	1.41	279.12	8.50
CCCma-CanESM2	2.00	3.69	0.23	1.01	1.71	1.42	279.25	8.76
CSIRO-QCCCE-CSIRO-Mk3-6-0	2.01	3.81	0.24	1.04	1.70	1.42	278.90	8.86
IPSL-IPSL-CM5A-MR	2.03	3.68	0.24	1.10	1.70	1.41	279.12	8.71
MIROC-MIROC5	1.99	3.67	0.24	1.09	1.71	1.42	278.92	8.78
NOAA-GFDL-GFDL-ESM2M	2.01	3.79	0.24	1.10	1.70	1.42	278.95	8.92
20CR diff.	0.22	0.39	0.00	-0.04	0.02	0.03	-0.96	0.08
EURO-CORDEX mean diff.	0.20	0.26	0.00	-0.04	0.00	0.00	-1.11	0.15

Table S7. Calibrated parameters for the best model.

Parameter	Units	Material	Value
Porous media parameters			
Van Genuchten model inverse of the air-entry pressure head, α	m^{-1}	glacial till	0.0136
		peat	2.054
		sand	0.025
Van Genuchten model pore-size distribution index, β	-	glacial till	1.735
		peat	1.535
		sand	2.632
Saturated hydraulic conductivity	ms^{-1}	glacial till	6.56E-07
		peat	4.52E-07
		sand	2.24E-03
Porosity	-	glacial till	0.36
		peat	0.86
		sand	0.39
Evapotranspiration parameters			
transpiration fitting parameter, c1	-	Ten vegetation classes	0.06 to 0.21
Lower limit of soil saturation for transpiration, e1	-	upland	0.996
		wetland	0.858
Upper limit of soil saturation for transpiration, e2	-	upland	0.889
		wetland	0.636
field capacity, fc	-	upland	0.922
		wetland	0.623
wilting point, wp	-	upland	0.376
		wetland	0.206
oxic limit, ox	-	upland	0.999
		wetland	0.846
anoxic limit, aox	-	upland	1
		wetland	0.936
Surface water flow parameter			
Manning roughness coefficient	$ms^{-\frac{1}{3}}$	Lower Biebrza	0.191
		Major rivers	0.042
		Other rivers	0.020
		Upland	0.019
		Upper Biebrza	0.152
		Floodplain	0.128
obstruction height	m	Major rivers	0.312
		Other	0.370
		Floodplain	0.064

Table S8. Error metrics for groundwater wells observations in the floodplain. RMSE / d.r. and bias / d.r. area RMSE and bias normalized to the observations data range (d.r.). Errors for individual wells in the middle and upper Biebrza basins are presented in Table S9.

Well	Period with observations	RMSE [m]	RMSE / d.r.	bias [m]	bias / d.r.
BPN116	1998-2010	0.28	21%	-0.15	-11%
BPN121	1998-2010, 2016-2019	0.25	19%	-0.06	-4%
BPN122	1998-2010, 2016-2019	0.33	22%	0.16	10%
BPN123	1998-2010	0.23	22%	-0.08	-8%
BPN124	2010-2019	0.25	28%	-0.11	-12%
BPN125	2010-2012, 2014-2019	0.27	22%	-0.14	-11%
BPN126	2010-2019	0.25	18%	-0.07	-5%
BPN167	2010-2019	0.31	26%	-0.19	-16%
BPN168	2010-2019	0.38	32%	-0.28	-24%
mean		0.28	23%	-0.10	-9%

Table S9. Error metrics for groundwater wells observations in the middle and upper basins. RMSE and bias are in the same units as indicated in the table, remaining metrics are dimensionless. RMSE / d.r. and bias / d.r. area RMSE and bias normalized to the observations data range (d.r.).

Well	Period with observations	RMSE	RMSE / d.r.	bias	bias / d.r.
BPN132	1998-2019	0.34	24%	0.05	4%
BPN133	1998-2019	0.34	23%	0.06	4%
BPN134	1998-2019	0.33	25%	0.10	8%
BPN135	1994-2015	0.28	25%	-0.14	-13%
BPN136	1994-2019	0.24	21%	0.04	4%
BPN137	1994-2019	0.29	23%	-0.09	-7%
BPN139	1994-2019	0.33	26%	-0.16	-12%
BPN140	1994-2019	0.43	33%	-0.32	-24%
BPN141	1994-2019	0.33	26%	-0.14	-11%
BPN142	1994-2019	0.35	26%	-0.13	-10%
BPN143	1994-2019	0.36	27%	-0.18	-13%
BPN144	1994-2019	0.40	26%	-0.23	-15%
BPN145	1994-2015	0.54	44%	-0.47	-38%
BPN147	1994-2019	0.60	43%	-0.55	-39%
BPN150	1996-2019	0.39	29%	-0.29	-21%
BPN152	1996-2019	0.79	50%	-0.73	-46%
BPN179	2010-2019	0.69	76%	-0.50	-55%
BPN182	1996-2019	0.83	58%	-0.78	-55%
BPN184	1996-2019	0.80	60%	-0.75	-56%
BPN186	1998-2019	0.71	48%	-0.63	-43%
BPN189	1996-2019	0.38	26%	-0.27	-19%
BPN190	1996-2019	0.31	20%	-0.15	-9%
BPN191	1994-2019	0.33	22%	-0.20	-13%
BPN207	2012-2017	0.38	33%	-0.32	-28%
BPN208	2012-2015	0.46	55%	-0.42	-50%
BPN209	2012-2017	0.34	34%	-0.26	-25%
BPN210	2012-2015	0.52	60%	-0.48	-55%
BPN211	2012-2017	0.51	47%	0.48	44%
BPN213	2012-2017	0.27	22%	-0.16	-13%
Middle basin mean		0.44	36%	-0.26	-21%
BPN155	1998-2019	0.59	37%	-0.45	-28%
BPN156	1998-2019	0.45	36%	-0.36	-29%
BPN158	1998-2019	0.32	29%	-0.17	-16%
Upper basin mean		0.46	34%	-0.33	-24%

Table S10. Statistics for the daily water levels and discharge at the Burzyn station in the 1970-2005 period, when both 20CR, and EURO-CORDEX forcing data to overlap with observations. The 20CR diff. row presents the observations subtracted from values simulated using model forced with 20CR data. The EURO-CORDEX mean diff. row presents the mean difference of observations subtracted from values simulated using models forced with EURO-CORDEX data.

Source	Water level [m amsl]		Discharge [m^3s^{-1}]	
	mean	sd	mean	sd
Observations	101.36	0.62	38.07	31.91
20CR	101.41	0.46	40.35	33.95
CCCma-CanESM2	101.42	0.46	40.88	31.30
CNRM-CERFACS-CNRM-CM5	101.50	0.44	45.51	30.93
CSIRO-QCCCE-CSIRO-Mk3-6-0	101.40	0.40	36.78	26.19
ICHEC-EC-EARTH	101.47	0.42	42.63	30.71
IPSL-IPSL-CM5A-MR	101.48	0.40	42.25	28.46
MIROC-MIROC5	101.40	0.45	38.73	29.85
MOHC-HadGEM2-ES	101.50	0.45	45.71	32.43
MPI-M-MPI-ESM-LR	101.53	0.40	46.85	31.58
NCC-NorESM1-M	101.51	0.40	44.72	28.74
NOAA-GFDL-GFDL-ESM2M	101.46	0.40	40.59	28.13
20CR diff.	0.05	-0.16	2.28	2.04
EURO-CORDEX mean diff.	0.11	-0.20	4.40	-2.08

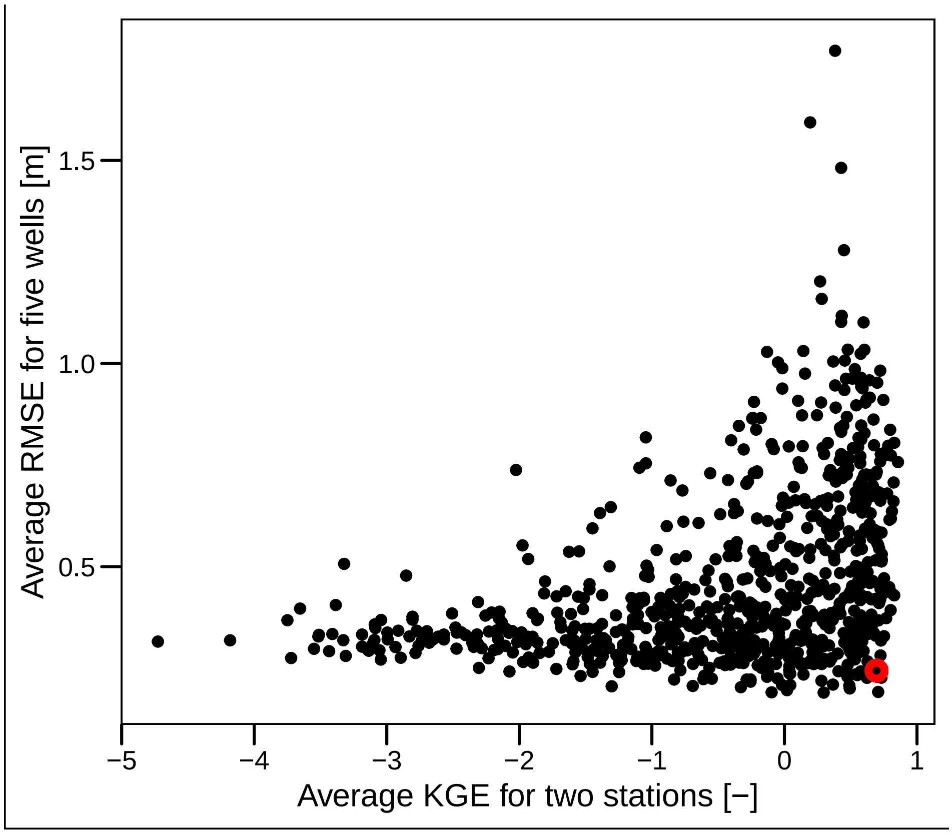

Figure S1. Average RMSE for five groundwater wells and average KGE for two stations calculated for 800 calibration runs. The red point indicates the selected model.

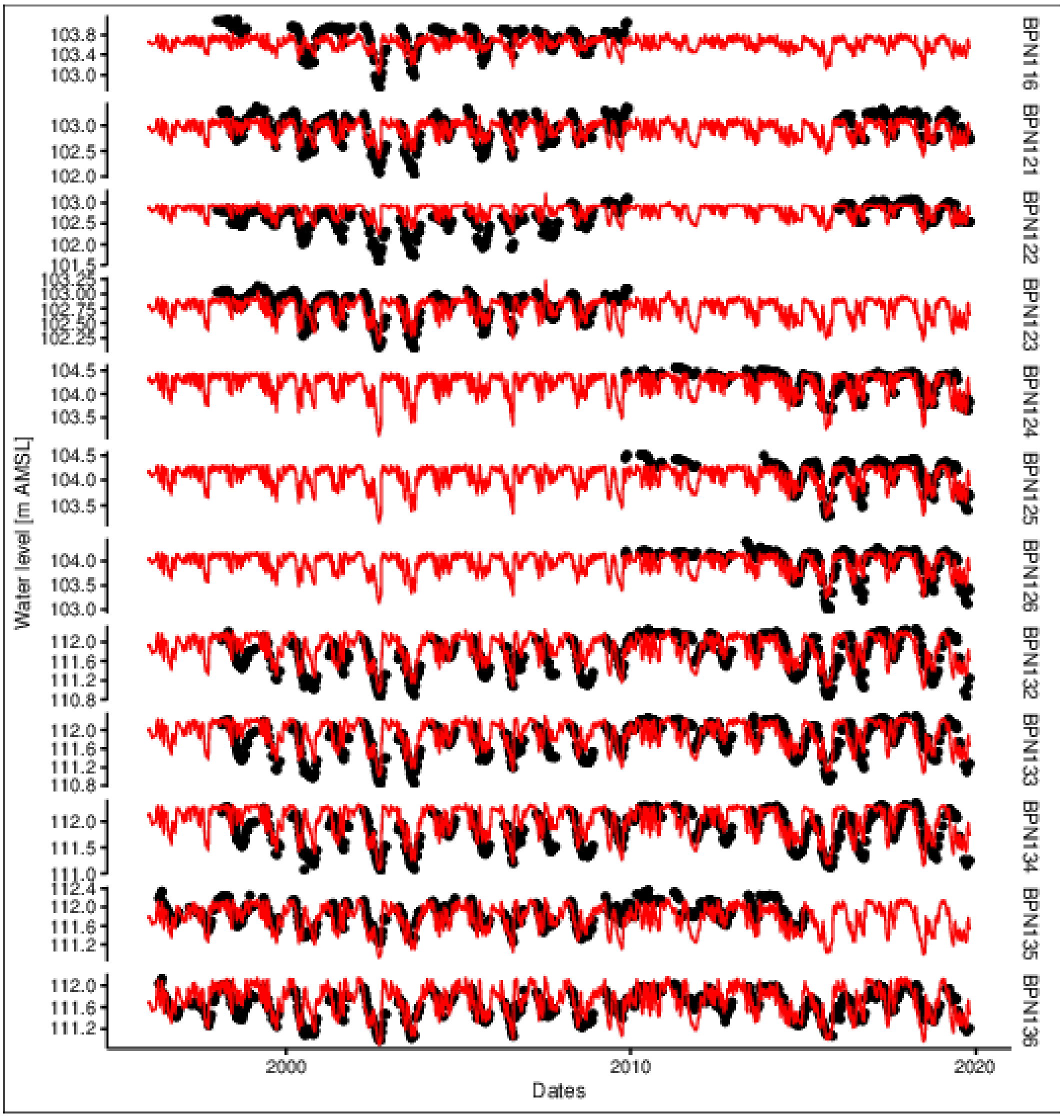

Figure S2. A Simulated (black) and observed (red) water levels for groundwater wells.

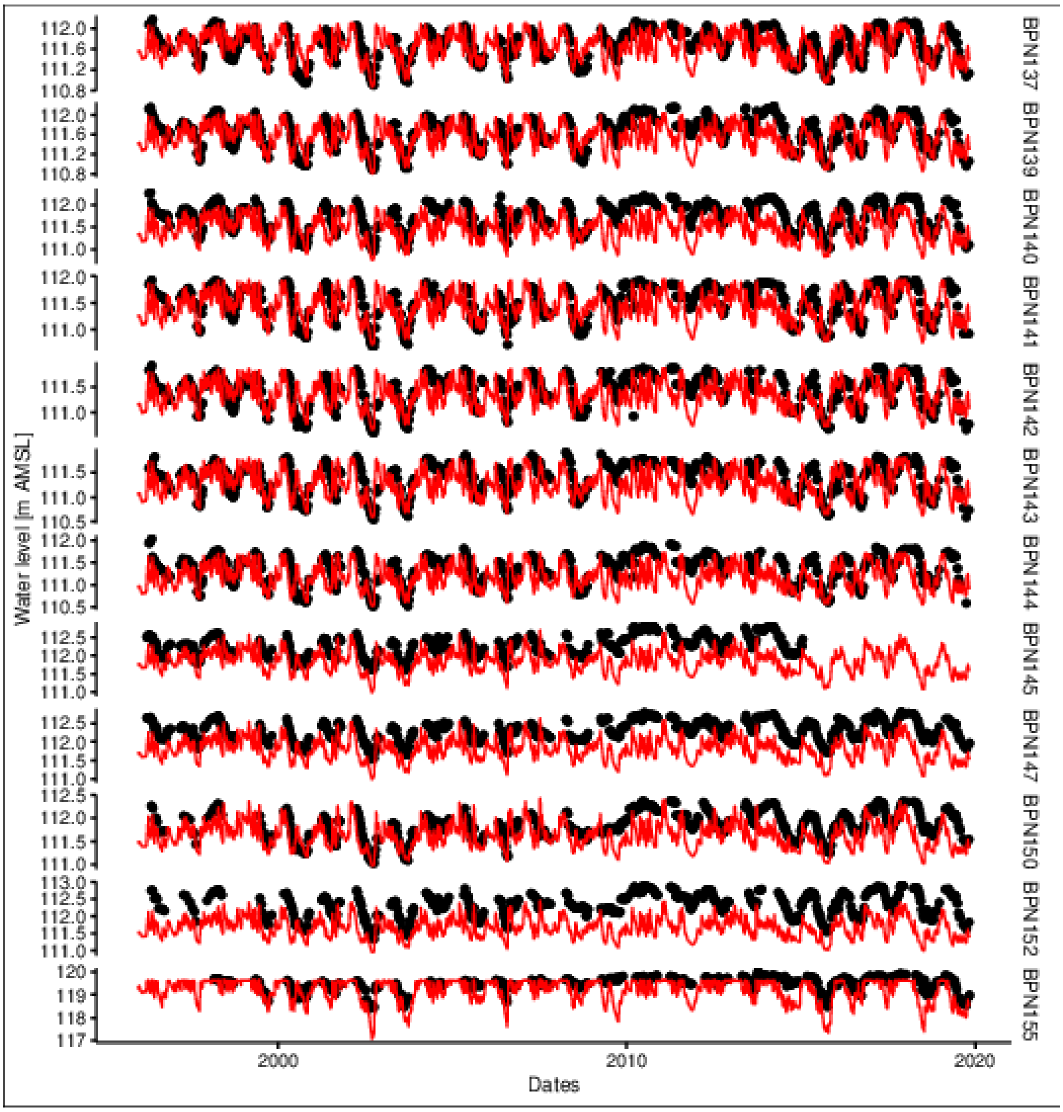

Figure S3. A Simulated (black) and observed (red) water levels for groundwater wells.

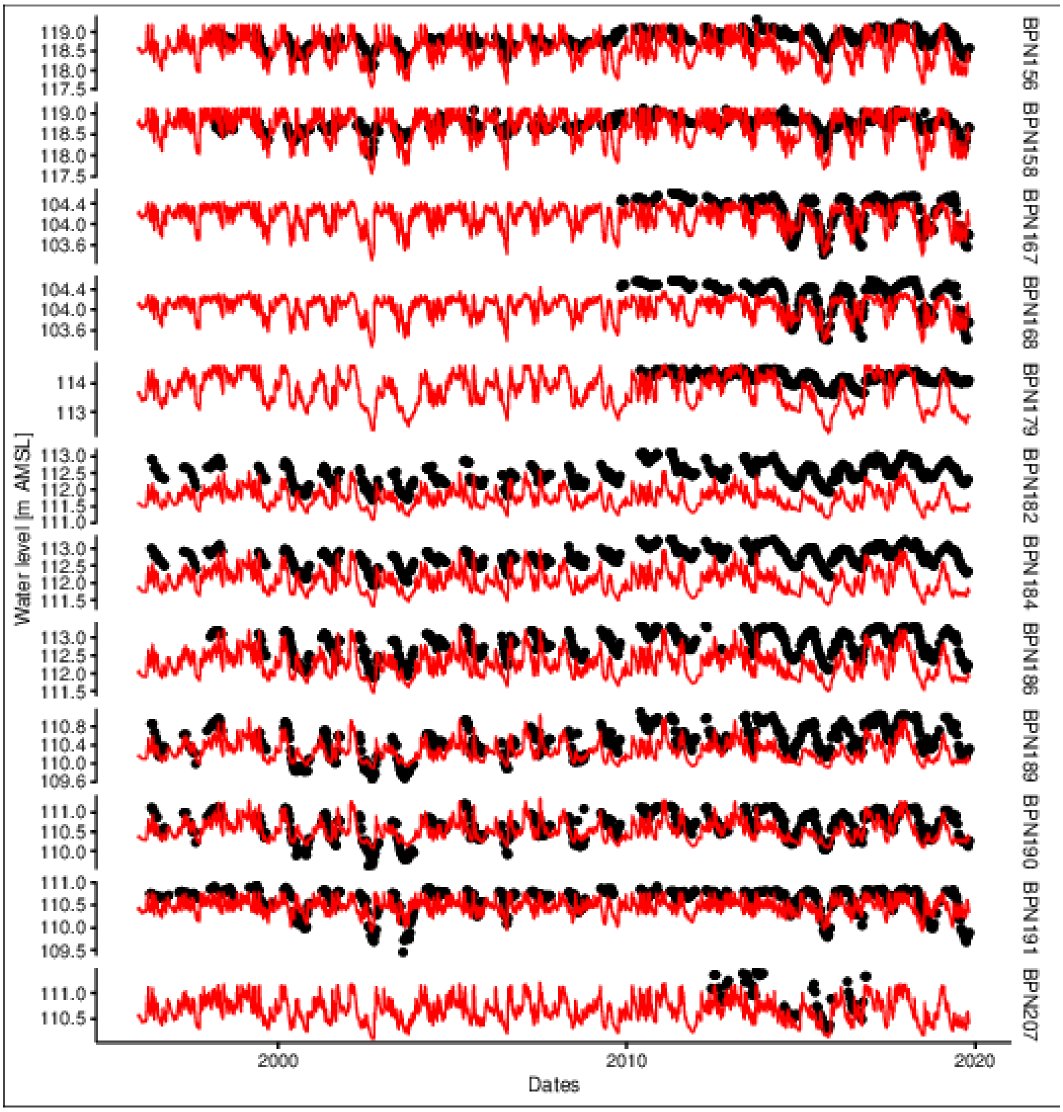

Figure S4. A Simulated (black) and observed (red) water levels for groundwater wells.

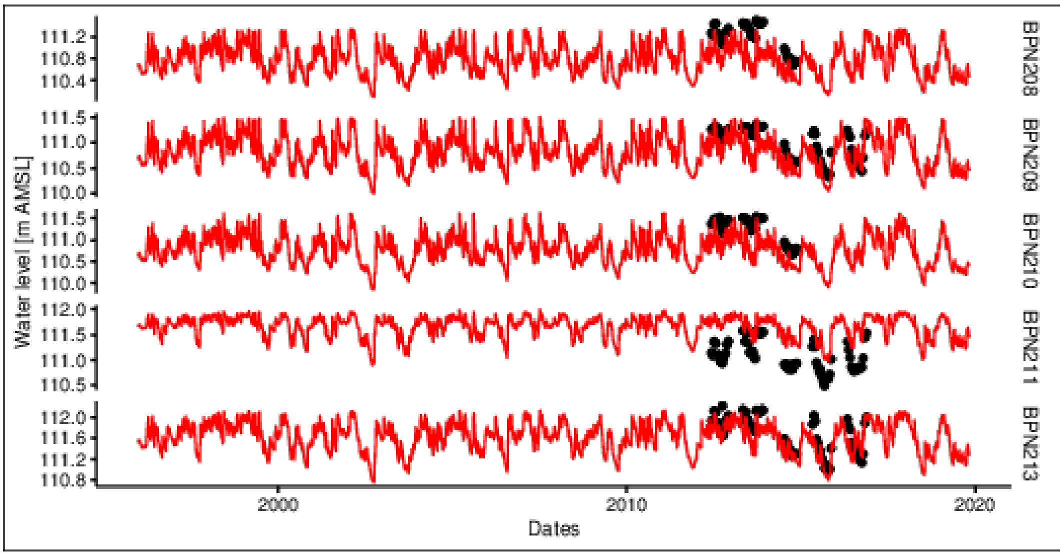

Figure S5. A Simulated (black) and observed (red) water levels for groundwater wells.

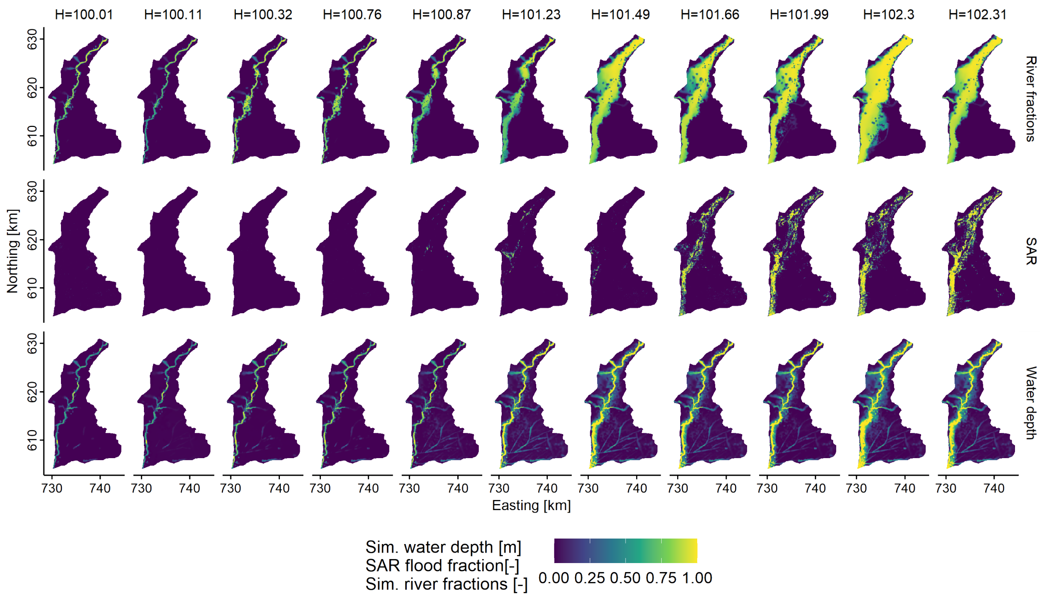

Figure S6. Flooding extent from remote sensing data-set (SAR), simulated HGS surface water depth, and HMC river water fractions for 11 increasing outlet water levels. Water depths > 1 m were plotted as equal to 1 m in this plot to have consistency in the color scale.

Impact of climate change on water sources and river-floodplain mixing in the natural wetland floodplain of Biebrza River

Tomasz Berezowski [1]*, and Daniel Partington [2]

[1]Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Gabriela Narutowicza 11/12, 80-233 Gdansk, Poland
[2]National Centre for Groundwater Research & Training (NCGRT), College of Science & Engineering, Flinders University, Adelaide, SA, Australia

Key Points:

- The extent of water sources in inundation will strongly vary in the future climate.
- The volume of water sources will vary considerably less than the extent in the floodplain.
- The shifted extents of water may have implications for floodplain management and ecology.

*This research was financed by grant: 2017/26/D/ST10/00665 funded by the National Science Centre, Poland

Corresponding author: Tomasz Berezowski, `tomberez@eti.pg.edu.pl`

Abstract

The origins of river and floodplain waters (groundwater, rainfall, and snowmelt) and their extent during overbank flow events strongly impact ecological processes such as denitrification and vegetation development. However, the long-term sensitivity of floodplain water signatures to climate change remains elusive. We examined how the integrated hydrological model HydroGeoSphere and the Hydraulic Mixing-Cell method could help us understand the long-term impact of climate change on water signatures and their spatial distribution in the protected Biebrza River Catchment in northeastern Poland. Our model relied on 20th century Reanalysis Data from 1881 to 2015 and an ensemble of EURO-CORDEX simulations for RCP 2.6, 4.5, and 8.5 from 2006 to 2099. The historical component of the simulations was subjected to extensive multiple-variable validation from 1881 to 2019. The results show that the extents of water sources were rather stable in the floodplain in the 1881-2015 period. The projected future impacts were variable with each analyzed RCP, but in all cases, different significant trends were present for the spatial distribution of water sources and for the river-floodplain mixing. However, the total volume of water from different sources was less sensitive to climate change than the dominant sources and spatial distribution of water. The simulation results highlight the impact of climate change on the extent of water sources in temperate zone wetlands with significant implications for ecological processes and management. These results also underscore the urgent need to leverage such modeling studies to inform protective and preservation strategies of floodplain wetlands.

Plain Language Summary

In this study, we used a hydrological model that was capable to simulate volumes of water from rain, snowmelt, groundwater discharge, and river flooding to investigate how these volumes will vary with the climatic conditions. For the study site, we selected the Biebrza River wetland floodplain, where former research highlighted the presence of these water sources in inundation during flooding. It was also known that the water sources have different chemical (e.g. nutrients) and physical (e.g. sediments) compositions and they correlate with the vegetation in the wetland. Hence, any change in the extent of these water sources (driven e.g. by climate change) may affect vegetation. Our research indicated that indeed the spatial extent of water sources will strongly vary with the future climate projection while the less detailed floodplain-wise volume of the water sources will not vary that much. We also showed that the direction of change in the water sources' extent will be different given the analyzed climate scenario. These results should be taken into account especially by the natural conservation managers to prepare for the changes.

1 Introduction

Mixing of river and floodplain water during floods, also known as perirheic mixing (Mertes, 1997), has great significance for ecological and hydrochemical processes. This significance in floodplain ecology is reflected by the floodplain vegetation zonation, which is related to the differences in the chemical or sediment composition of water from river and groundwater, rain and snowmelt inundation in the floodplain (Chormański et al., 2011; Keizer et al., 2014). Similar relations are present in the Amazon floodplain, where the mixing of sediment-rich and sediment-poor water near the confluences is related to vegetation (Park & Latrubesse, 2015), and avifauna (Laranjeiras et al., 2021). Also, in the Amazon floodplain, the river-floodplain water frontier is controlling the crevasse splays occurrence (Aalto et al., 2003). The hydrochemical significance of water mixing is mainly due to nitrate removal by denitrification. This process occurs in the flow-through wetlands, where nitrate- and oxygen-rich water from a river mixes with the oxygen-poor floodplain water. Although this effect was reported in several floodplains, including Atchafalaya (Jones et al., 2014; Scott et al., 2014), Po (Racchetti et al., 2011), and Wisconsin (Forshay & Stanley, 2005), to achieve considerable nitrate removal a significant floodplain area has be connected to the river (Natho et al., 2020). As we have shown previously for a natural temperate zone wetland floodplain - Biebrza River, the river-floodplain water mixing, or the active perirheic zone, is very dynamic in space and time (Berezowski et al., 2019). In that study, we used state-of-the-art modeling tools for a single flood event study, hence we were not able to assess the active perirheic zone's long-term variability and the role of the changing climate.

Hydrological impact models of climate change predict a shift of the highest and lowest discharges at the end of the twenty-first century for several regions of the world (Prudhomme et al., 2013; Giuntoli et al., 2015; Arnell & Gosling, 2016). These regions include the major floodplain and wetlands, where the shift in flooding pattern may influence ecological processes such as vegetation development (Murray-Hudson et al., 2006; Garris et al., 2014; Zulkafli et al., 2016; Thompson et al., 2016). The hydrological shifts in the future will also lead to changes in floodplain connectivity in unregulated floodplains. This may result in increased nitrate removal by denitrification, as simulated for the Lower Missouri River (Jacobson et al., 2022). Nitrate removal varies in floodplain habitats with different contact with river water (Scaroni et al., 2011). Since, the zonation of water sources within the flooding extent is relevant for vegetation development and denitrification, more precise quantification of these ecological processes in the scope of climate change could be achieved by analyzing water sources' zonation. This remains a gap in the literature.

Modeling of climate change impact on floodplain inundation is usually done using either 1D or 2D hydrodynamic models. Such models require a precise definition of boundary conditions for which coupling with catchment-based hydrological models is often used (Thompson et al., 2008; Karim et al., 2015; Zhang et al., 2019). Another approach is to drive a hydrodynamic model using boundary conditions, such as surface runoff, from hydrological components of general circulation models (GCM), or climate reanalysis (Mohanty & Simonovic, 2021). In either case, the surface water in the floodplain lacks or has limited, feedback with parts of the catchment that are not represented by the hydrodynamic model, which includes groundwater, tributary inflow, or surface runoff. These feedbacks are important in the proper modeling of floodplain inundation, as those minor water sources produce the inundation in remote parts of the floodplain and determine the river-floodplain water frontier (Berezowski et al., 2019) and groundwater mixing zone (Nogueira et al., 2022). Therefore, to achieve full feedback between all water sources integrated hydrological models (IHMs) are required (Sebben et al., 2013). The computational complexity of these models often requires some simplifications or limiting the simulation area (Barthel & Banzhaf, 2015) to achieve feasible run times. Also, the application of IHMs to climate change impact research is limited in scenarios and analysis periods lengths (e.g. Ferguson and Maxwell (2010); Sulis et al. (2011); Erler et al. (2019)), while using a GCM ensemble reduces uncertainty related to future climate projections impact on hydrology Z. Kundzewicz et al. (2018). Currently, this research area remains relatively unexplored, as only a few studies run such models with long-term forcing data from GCMs ensembles, such as the Intergovernmental Panel on Climate Change (IPCC) emission scenarios (Goderniaux et al., 2009; Sulis et al., 2012; Perra et al., 2018; Boko et al., 2020; Ramteke et al., 2020; Yuan et al., 2021) and no such models have analyzed the extent of water from different sources.

Except for the GCM ensemble, the credibility of the modeling results in climate change studies is achieved by comprehensive model validation and using multiple impact models. The latter is especially important in large-scale (regional and continental) cases, where some parts of the study area are ungauged (Krysanova et al., 2018). Further, it seems that using IHMs ensembles may not be crucial, since contrary to many conceptual models used in climate change impact studies, they are almost entirely physically based and perform similarly (Kollet et al., 2017). On the other hand, a comparison of conceptual, physically based, and fully integrated hydrological models in a climate change impact study revealed that the models showed the same direction of change for most of the indicators, however, a fully coupled IHM indicated an the opposite trend in mean annual evapotranspiration when compared to remaining models (Perra et al., 2018). Either way, using an IHM ensemble would be beneficial for the results by increasing credibility, although this comes with an associated computational burden.

Comprehensive validation of an IHM, should, therefore, be of greatest concern in climate change impact studies. Most often the observations of river discharge are used for the validation of impact models, while IHMs, due to the simulation of surface water hydrodynamics can be further validated against water levels. The spatial aspect of validation can be achieved by using multiple gauges, however, flooding water extent can serve this purpose as well. The latter is often achieved using multi-temporal remote sensing data providing spatiotemporal insight into model performance (e.g. Paiva et al. (2012)), however, in vegetated areas such validation can be problematic, due to obscuring by vegetation canopy. IHMs are usually also validated against groundwater levels, which gives further insight into processes relevant to catchment functioning that are not depicted by surface water. Also, if transport or water mixing is simulated, IHMs can be validated against hydrochemical parameters. This list of validation variables for IHMs does not include all the possibilities. Instead, it indicates that, contrary to conceptual models, the physically based IHMs can be validated comprehensively to minimize uncertainty related to the simulated complex interactions, such as mixing of water from different sources.

To examine the impact of climate change on spatiotemporal water signatures during flooding in a natural temperate zone wetlands, this research aims to employ a robust IHM for the Bierbza catchment to investigate the long-term variability of the extent and mixing of water from different sources during flooding. The model for the Biebrza will be run for a historical period using 20th Century Reanalysis data and a GCM ensemble for representative concentration pathways (RCP) 2.6, 4.5, and 8.5 scenarios for the future. With this model, the aims of the research are:

- To determine if the past climate and future climates under RCPs 2.6, 4.5 and 8.5 will drive any significant changes in the spatial distribution and dominance of water sources in the Biebrza floodplain.
- To determine if the volume of water in the floodplain will significantly change under past climate and possible future climates with RCPs 2.6, 4.5 and 8.5.
- To highlight the implications for ecological processes, modeling, and management strategies under climate change.

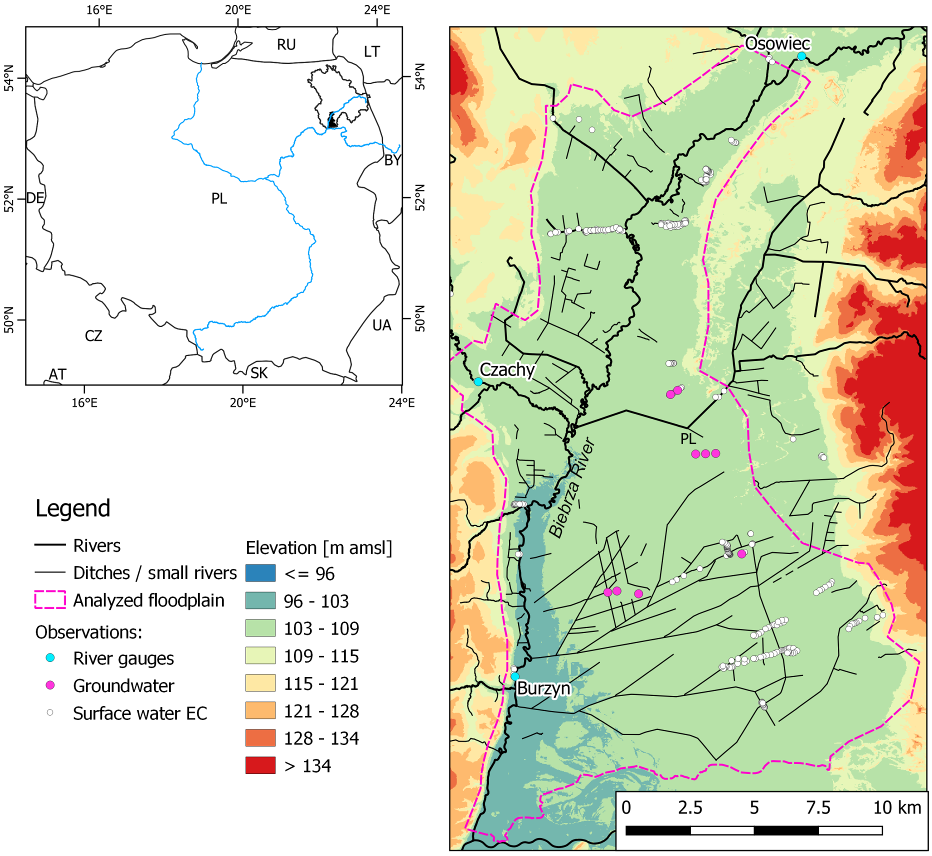

Figure 1. The floodplain area and the measurements points (right panel). Location of the study area in Poland (left panel) with the major rivers (blue lines), Biebrza river catchment (black outline), and the floodplain (black patch). The legend concern only the right panel.

2 Methods

2.1 Study area

The Biebrza catchment (22.7° E, 53.7°N) is of medium size, 7091 km ^{2} and the lower Biebrza valley (hereinafter referred to as floodplain), where we focus our analysis comprises 297 km^2 (Figure 1).

We chose the Biebrza valley as the study area because of its natural character and ecological significance. The major river engineering work was conducted in the area in the first half of the 19th century to establish a waterway between Biebrza and Neman Rivers. Next, in the middle of the 19th century parts of the Wetlands located in the lower and middle parts of the valley were meliorated. In the 20th century, only minor melioration work was conducted except in the middle part of the Biebrza valley (Banaszuk, 2004). Currently, the anthropogenic pressure is low, as the population density in the region where the Biebrza River catchment is located is the lowest in Poland (58 people per km^2) (Statistics Poland, 2021). The future population projections for this region predict a 32% decline between 2020 and 2100 (Eurostat, 2019). The Biebrza valley was grazed and mowed in the past and aquatic vegetation in the river was occasionally removed (Berezowski et al., 2018). Since the establishment of the Biebrza National Park in 1993 mowing and grazing is continued as an active protection measure (Kotowski et al., 2013). Currently, the Biebrza National Park is one of the largest active protection areas in Europe (59223 ha), with the Biebrza Wetlands listed as Ramsar and Natura 2000 sites.

Long-term average discharge in Biebrza River has been 38.1 m^3s^{-1} (1970-2005), with a minimum of 4.33 m^3s^{-1} and maximum of 517 m^3s^{-1}. The river flooding area reaches up to 52.5 km^2 and inundation can last on average between 121 to 193 days depending on location (Grygoruk et al., 2021). The average annual precipitation over the period 1970-2005 in the catchment has been 672 mm, of which 88 mm was snow, whereas the yearly potential evapotranspiration (PET) was 621 mm.

Wetland vegetation in the floodplain exhibits zonation related to flooding (Pałczyński, 1984). The Phragmition belt is located around the river up to about 500-900 m, further away up to 2500 m from the river, Magnocaricion vegetation is present, and further again, Fen vegetation, such as Calamagrostion neglectae, Caricion diandrae, or Caricion demissae is located up to the valley margin.

The Quaternary deposits are 130-212 m deep and the majority consist of glacial till with minor sand layers deposited during the Riss glaciation. Middle and lower parts of the Biebrza valley have a sand layer deposited during the Weichselian glaciation on top of which the Holocene sand and peat layers are present (Banaszuk, 2004).

Given undisturbed vegetation, unregulated river, natural hydrology, and low contamination in relation to European standards the Biebrza wetlands may be considered as a reference site for similar fen wetlands (Wassen et al., 2006).

2.2 Forcing data

Hydrological simulations for over two hundred years period required several sources of forcing data (Table S1). The criteria for selecting a data source were daily (or higher) temporal resolution and availability of the required forcing variables (precipitation, snow cover dynamics, air temperature, and PET).

For the historical 1880-2015 period we used the 20th century climate reanalysis (20CR) data (Slivinski et al., 2019). Out of this data-set, we used ensemble mean of water equivalent of accumulated snow depth (WEASD) [kg m^{-2}], daily mean of 3-hour accumulated precipitation amount (APCP) [kg m^{-2}], air temperature at 2m (air2m) [K], and potential evaporation rate (PEVPR) [W m^{-2}]. We used the following preprocessing steps before bias correction. The difference of WEASD between subsequent days was calculated. Then, the negative values were multiplied by -1 and used as uncorrected daily snowmelt, $\dot{s}$, [mm] and the positive values were used as uncorrected daily snowfall ($\dot{p}_s$) [mm]. For PET, the PEPVR values were multiplied by 0.01152 to change units to mm.

For the future period, we used the EURO-CORDEX data (Jacob et al., 2014) from ten simulations using different GCMs (Table S1). Each simulation used the SMHI-RCA4 regional climate model (RCM). We selected all available simulations from the EURO-CORDEX archive that had the required forcing data for the hydrological model. Only four out of ten simulations had the required forcing data for RCP 2.6. To investigate the effect of greenhouse gases emission scenarios on water sources mixing in the floodplain we used the following RCPs: RCP 2.6, which aims to limit the increase in global mean temperature to 2 K by a CO_2 emission decline since 2020, RCP 4.5 which is an intermediate scenario, where the emissions start to decline after 2040, and RCP 8.5 which is a worst-case scenario in which emissions continue to rise during the entire 21st century. From each simulation, we used daily mean values of snowfall flux (PRSN, used as $\dot{p}_s$) [kg m^{-2} s^{-1}], snowmelt flux (SNM, used as $\dot{s}$) [kg m^{-2} s^{-1}], precipitation flux (PR) [kg m^{-2} s^{-1}], near-surface air temperature (tas) [K], and potential evapotranspiration (EVSPBLPOT) [kg m^{-2} s^{-1}]. We used the following preprocessing steps before bias correction. Daily snowmelt, snowfall, precipitation, and potential evapotranspiration fluxes were multiplied by 86400s to change units to mm.

For the 2015-2019 period (for which the 20CR data was not available), when the hydrochemical validation took place we used the 2 km gridded precipitation and temperature data-set (Piniewski et al., 2021) and snowfall and snow depth data from the Biebrza-Pieńczykówek meteorological station managed by the Institute of Meteorology and Water Management - National Research Institute (IMGW-PIB).

2.2.1 Bias correction

We used the quantile mapping (Gudmundsson et al., 2012) bias correction using the R software package "qmap". The following meteorological observations were used to identify parameters of bias correction: the total precipitation and air temperature from a 5km gridded data-set (Berezowski et al., 2016) (the 2km data set was not available at that time), PET from a gridded 25 km data-set (Joint Research Center, 2019), and the snowfall from the Biebrza-Pieńczykówek meteorological station. In such variable availability, we were not able to conduct bias correction of snowmelt, s, and rainfall, p_r. The snowmelt was constrained to the snowfall using the sum of uncorrected snowmelt ($\dot{s}_v$) and the sum of bias-corrected snowfall ($p_{s,v}$) in a given event v. An event was defined as a period between the start of snow accumulation and the end of snowmelt; most often there are one or two larger events in each year. Daily snowmelt [mm] in an event v was calculated as $s = \dot{s}\frac{p_{s,v}}{\dot{s}_v}$. The rainfall p_r [mm] for a given day was calculated by subtracting bias-corrected snowfall from bias-corrected precipitation. We used a maximum overlapping period for bias correction of each variable, which was 1955-2013 for precipitation and air temperature, 1957-2015 for snowfall, and 1979-2015 for PET for the 20CR data. In the case of the EURO-CORDEX data, we were additionally limited by the historical period, which was either 1951-2005 or 1970-2005 (Table S1). After conducting the bias correction we calculated the daily average value of each variable over all grid cells in the Biebrza catchment and used this data to force the hydrological simulations.

2.3 Hydrological model

We simulated the transient water fluxes in the Biebrza River catchment using HydroGeoSphere (Brunner & Simmons, 2012; Hwang et al., 2014) IHM. The 3D groundwater flow was solved using Richard's equation in prism elements and the 2D surface water flow was solved using the diffusion wave approximation of the Saint-Venant equations in triangular elements. The surface-subsurface flow coupling was realized using the first-order exchange. Evapotranspiration flux was simulated using the Kristensen and Jensen (1975) conceptual model, which takes into account interception storage, time-variable leaf area index (LAI), pounding, and soil saturation. Snowmelt and rainfall fluxes were provided as forcing data boundary conditions. The model parameters were specified spatially according to relevant geological, land-use, or vegetation units.

We simulated water mixing using the hydraulic mixing-cell (HMC) method (Partington et al., 2011). In our case the mixing was simulated only for the surface flow domain, however, simulations in groundwater are also possible (Nogueira et al., 2022). The HMC method accounts for water fluxes from various boundary conditions and groundwater discharge effectively producing a fraction of each water source in a model node. Water sources were differentiated spatially. To calculate the river water fractions we summed all fractions upstream of the floodplain area. Whereas in the floodplain area, original fractions of rainfall, snowmelt, and groundwater were used to represent the inundation components generated therein. In the first time step, the fractions are initialized using an artificial initial fraction, equal to one.

We used the parallel solver in the HGS, which split the coefficient matrix into two parts. The flow solver convergence criteria for the maximum absolute residual error was

$1\text{x}10^{-10}$ m^3s^{-1}, and the Newton iteration convergence criteria for the maximum absolute nodal change in the pressure head was 1 cm. In the HMC method, the maximum ratio between fractions volume was set to 2048 and above this threshold, all fractions are set to zero and the reset fraction is set to one (Partington et al., 2013).

2.3.1 Leaf area index estimation

The HydroGeoSphere model uses LAI during the estimation of evapotranspiration. Since LAI was not available in any data set covering the simulation period we used a degree-day model to simulate LAI for each meteorological data set used in this study. The model was based on observations that wheat requires about 760 degree-days for development and 500 more degree-days for maturity (Rawson & Macpherson, n.d.) and can be summarized in the following steps:

1. At the beginning of a calendar year LAI is equal to the minimum for a given vegetation
2. Growing season is defined as a day when the monthly average temperature is greater than 5°C.
3. Since the beginning of the growing season LAI increases proportionally to degree-days to reach the maximum for given vegetation at 760 degree-days.
4. LAI remains at the maximum for 500 degree-days.
5. LAI decreases linearly to reach the minimum for given vegetation on the last day of the growing season.

The maximum LAI for each vegetation was based on measurements in the study area (Dąbrowska-Zielińska et al., 2014; Suliga et al., 2015).

2.4 Error metrics

In this study, we use the same error metrics for a number of different simulated quantities, such as water levels, discharge, water source fractions, and area. We present the general form of the equations below. Whenever a given error metric is used in the text it is specified based on which quantities it was calculated for and, if applicable, to which quantity it was normalized.

The Kling-Gupta efficiency [-]:

$$\mathrm{KGE} = \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{1}$$

where r [-] is the correlation coefficient between simulated and observed discharge, α [-] and β [-] are ratios of simulated to observed mean and standard deviation discharges respectively. The KGE ranges between $-\infty$ and 1 and the higher the value the better fit to the observation is achieved by the model.

The root mean square error [units the same as input data]:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{h_i} - h_i\right)^2}{N}} \tag{2}$$

where h_i and $\hat{h_i}$ are observed and simulated quantities respectively for a data record (e.g. time step) i out of N. The RMSE represents the magnitude of error between the observations and simulations and ranges between 0 and ∞.

The systematic error, or bias [units the same as input data]:

$$b = \sum_{i=1}^{N} \hat{h_i} - h_i \tag{3}$$

where the symbols are the same as in Eq. 2. The bias shows whether the simulated quantities overestimate (positive b) or underestimate (negative b) the observed quantities and $b = 0$ indicate no bias.

The linear correlation between two variables was quantified using Pearson's correlation coefficient (r) [-] and the fraction of variance explained between the two variables was quantified using the coefficient of determination (r^2). If two variables are time-dependent the linear correlation can be interpreted in terms of the temporal variability agreement between them.

2.5 Model grid

To prepare the model grid we processed the relevant geographical information in QGIS 3.10 software in the following steps. We simplified the geometry of the rivers by limiting the minimum node distance to 125 m along the river course for major rivers and 500 m for minor rivers. For the major rivers, the banks were limited to a 60 m buffer around the river. This forced the perpendicular river cross-section to be trapezoidal. For minor rivers, no buffer was created and the perpendicular cross-section was triangular. The catchment boundary was simplified by limiting the minimum node distance to 2000 m. The geographical data source used in these steps was the Map of the Hydrographic Division of Poland in scale 1:10 000. The feature nodes obtained from the previous steps and the nodes representing locations of the observation wells were used to generate a Delaunay triangular grid in the triangle software (Shewchuk, 1996). The triangulation constraints were the maximum triangle size of 1 km^2 and the minimum angle in a triangle of 31°. Finally, we refined the grid fourfold in the floodplain area and relaxed the nodes using an algorithm provided by Kaser et al. (2014). The triangular grid consisted of 19297 nodes and 38081 elements of which 10436 were in the floodplain. The median element area for the whole grid was 71243 m^2 and for the floodplain was 20037 m^2; the minimum element area was 1017 m^2.

The nodes elevation was obtained from a Digital Elevation Model (DEM) of Poland in the resolution of 1m and from the Shuttle Radar Topography Mission in 30 m resolution outside the Polish border (in total 0.4% of the study area). The digital elevation model was updated with the lake bathymetry. The riverbed elevation for the major rivers was obtained from 160 land-survey perpendicular cross-sections conducted by the Polish Water Authority. The distance between subsequent cross-sections was about 500 m. As a riverbed elevation, the first quartile of the elevation in the nearest cross-section was used. The minor river's riverbed was calculated by subtracting river depths from a surface elevation. The river depth was estimated based on point measurement data provided by the Biebrza National Park and from our field survey.

The grid consisted of six vertical layers in which the top four layers had gradually increasing thickens and represented the stratification of peat, sand, and glacial till formed between the Riss glaciation and Holocene. The thickness of the first layer was 0.75 m in the floodplain. The two bottom layers were thick and represented glacial till deposited during the Riss glaciation. The elevation of the lowest layer was equal to -30 m AMSL, the average lower boundary of the Quaternary sediments (Banaszuk, 2004). In total, the grid consisted of 135097 nodes and 228486 prism elements.

We defined three porous materials: glacial till, sand, and peat with different hydraulic properties. In the river valley and its proximity, we assigned the materials based on geological cross-section data (Banaszuk, 2004), whereas in the remaining parts of the upland we used data from several geological bore profiles provided by the Polish Geological Institute (Polish Geological Institute, 2014). The hydraulic properties for the surface water flow and evapotranspiration were assigned to ten land-use and vegetation classes present in the study area based on the Corine Land Cover map (Commission of the European Communities, 2013).

2.6 Model calibration

We used a screening approach to find an optimal parameter set for the hydrological model. For this purpose, we randomly sampled 800 random parameter sets using the latin hypercube algorithm. We used the latin hypercube algorithm implementation from the "tgp" R package (Gramacy & Taddy, 2010). Each set consisted of 26 base parameters, which produced 43 model parameters by applying the constraints and transformations (Table S3). The constraints were used to scale a base parameter by a factor for different material types, such as vegetation types and produce multiple model parameters. We used the logarithmic transformation for the hydraulic conductivity and gamma distribution transformation for evapotranspiration parameters (details in Table S3-S5). The calibration period was two years and ten months (2004-01-01 to 2006-10-31) followed by a one and half year warm-up period (2002-06-01 to 2003-12-31). The initial conditions for each calibration run were transferred from a steady-state simulation using parameters from our previous model version (Berezowski et al., 2019). We choose the best model base on KGE for two discharge stations and RMSE [m] for five groundwater wells heads. The locations of discharge stations were chosen at the inlet and outlet of the floodplain (Osowiec and Burzyn) and the location of the wells were chosen two in the floodplain, one in the middle and upper parts of the valley. The relation between average KGE and average RMSE for all stations forms a Pareto front with a group of the best parameter sets from which the final model was selected manually by reviewing the simulated hydrographs.

2.7 Model validation

2.7.1 Hydrological validation

We used several contemporary and archival data sources with varied temporal coverage for the validation of simulated river flow and groundwater heads (Table S2). We used the same metrics as for calibration and the RMSE was normalized by the data range for each station or well. To investigate how the hydrological model performed temporarily we calculated KGE for discharge and RMSE for river water levels per decade.

The oldest water level records (Table S2) for the study area contained only the relative water level in reference to the gauge zero level. For these records we calculated the absolute water level, i.e. in meters AMSL, using a relation between the mean absolute and relative water levels for the remaining records for a particular gauge. The disadvantage of this approach is that the temporal trend is not preserved and the RMSE is biased.

Some of the groundwater heads data were missing the absolute readings, i.e. depth instead of elevation was measured. Calculation of the absolute levels was done by using a 1x1 m digital elevation model values in the well location as the zero depth. Few groundwater wells showed a clear step in the records, which could have been due to the displacement of the reference point. We removed records with the step from the database.

2.7.2 Remote sensing validation

We validated the simulated water extent using a multi-temporal remote sensing data-set (Berezowski et al., 2020). In that data-set 161 water extent maps were developed for the 2014-2019 period using the Sentinel-1 synthetic aperture radar (SAR) for the floodplain with the average water level error of the flood extent of 0.21 m. The major drawback of this data-set was that in densely vegetated areas the flood extent was obscured and effectively these areas are labeled as not flooded even if the water level was high. Further, the data-set was not sensitive to shallow water, which limits its applicability only to an indication of deeper river water within the Biebrza flooding extent. From this data set, we selected 134 flood maps with the lowest error and used them along with the hydrological model output to calculate the following validation metrics.

Despite some drawbacks, the remote sensing data-set was a good indicator of the temporal dynamics of the flooding extent, especially for the river water zone. Therefore validation in the floodplain was calculated using the total flooding area due to simulated water depth [m^2]:

$$a_h = \sum_{m=1}^{M} \tilde{h}^m a^m \tag{4}$$

and the flooding area due to river water fraction presence [m^2]:

$$a_{\mathrm{river}} = \sum_{m=1}^{M} \tilde{f}^m_{\mathrm{river}} a^m \tag{5}$$

where $\tilde{h}^m = 1$ if simulated water depth in a node m is greater than 5 cm and $\tilde{h}^m = 0$ otherwise, $\tilde{f}^m_{\mathrm{river}} = 1$ if river water fraction is greater than 0.1 in a node m and $\tilde{f}^m_{\mathrm{river}} = 0$ otherwise, a_m is the node m contributing area, and M is the total number of nodes in the floodplain. The values of a_h and a_{river} are calculated for each time step and used to calculate the correlation coefficient with the flooded area from the remote sensing data-set. Further, we calculated the fraction of area that is indicated as flooded on the intersection of hydrological model output and remote sensing data-set:

$$i_h = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{h}^{m,t}a^m\right) \wedge \left(\tilde{a}_{rs}^{m,t}a^m\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{a}_{rs}^{m,t}a^m\right)}$$

for intersection with the simulated water depth and

$$i_{\text{river}} = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{f}_{\text{river}}^{m,t}a^m\right) \wedge \left(\tilde{a}_{rs}^{m,t}a^m\right)}{\sum_{m=1}^{M}\sum_{t=1}^{T}\left(\tilde{a}_{rs}^{m,t}a^m\right)}$$

for intersection with simulated river fraction, where $\tilde{h}^{m,t}$ and $\tilde{f}_{\text{river}}^{m,t}$ are the same as $\tilde{h}^m$ and $\tilde{f}_{\text{river}}^m$, but indexed also for time step t, $\tilde{a}_{rs}^{m,t} = 1$ if the flooded area in the remote sensing data-set in a node m is greater than 25% and $\tilde{a}_{rs}^{m,t} = 0$ otherwise, and T is a group of time steps which overlap in the hydrological simulations and remote sensing data-set. Ideally, this validation should be extended to the calculation of true-negative flooding extent. This, however, was not possible due to false negative flooding extent in the remote-sensing data-set due to vegetation cover.

2.7.3 Hydrochemical validation

To investigate whether the different water sources presence is related to the simulated water source fractions we measured the electrical conductivity (EC) [μS cm^{-1}] of 133 samples in the floodplain during winter (24-25 January 2019) and spring (27-29 March 2019). The HI991300 portable EC meter was used and the location was recorded using a handheld GNSS receiver. We chose EC because prior research by (Chormański et al., 2011) indicated that EC is effective at discriminating between river water and other sources. We used random 50% of the measurement points to establish a linear regression model explaining the EC by the river, rain, snowmelt, and groundwater fractions in the model nodes on the measurement days. The remaining 50% of the data was used for validation of the regression model using RMSE [μS cm^{-1}] and bias [μS cm^{-1}]. All measurement points were used to calculate the correlation coefficients between the water source fractions and EC.

2.8 Changes of water sources fraction in the past and future climate

Next to the simulated water sources fractions, we analyzed the mixing degree [-] (Berezowski et al., 2019), which quantifies the mixing between river and floodplain (sum of snow, rainfall, and groundwater) water fractions:

$$d = 1 - \frac{\mid f_{\text{river}} - f_{\text{floodplain}} \mid}{1 - f_{\text{initial}}} \tag{6}$$

The changes in water sources fraction and mixing degree were assessed by calculating a length [days] of a period during which they were greater than 0.75 and the water depth was greater than 1 cm, in a hydrological year for each model node m in the floodplain:

$$l_s^m = \sum_{y=1}^{Y} \begin{cases} 1 & w_s^{y,m} > 0.75 \land h^{y,m} > 0.01 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $w_{s,t}$ is a value of s water source fraction (river, snow, rainfall, or groundwater) or the mixing degree d during a day y of a all days Y in a hydrological year, and $h_{t,m}$ is water depth [m]. The total annual volume of surface water in the floodplain weighted by the water sources fractions and the mixing degree in a hydrological year was calculated by performing a weighted integration using the following equation:

$$v_s = \sum_{y=1}^{Y} \sum_{m=1}^{M} \begin{cases} h^{y,m} a^m w_s & h_t > 0.01 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The mean surface water depth ($\bar{h}$) [m] and the length of a period with water depth greater than 1 cm (l_h) [days] was calculated for each model node in each hydrological year.

For future climate simulations, we calculated the above metrics for each EURO-CORDEX simulation and calculated the ensemble mean for each RCP scenario. Next, we used the ensemble means and historical simulations forced using 20CR data to calculate trends using the slope of the regression line, where the independent variable is the hydrological year. Finally, we used the t-test to investigate whether a trend estimate is significantly different from zero.

3 Results

3.1 Bias-corrected forcing data

Each forcing data have similar statistics as meteorological observations for the period in which the quantile mapping parameters were identified (Table S6). Both for EURO-CORDEX and 20CR data the air temperature underestimated the observations mean, but had similar standard deviations. Snowfall and PET were bias-corrected near perfectly in terms of mean and standard deviation. Total precipitation was overestimated

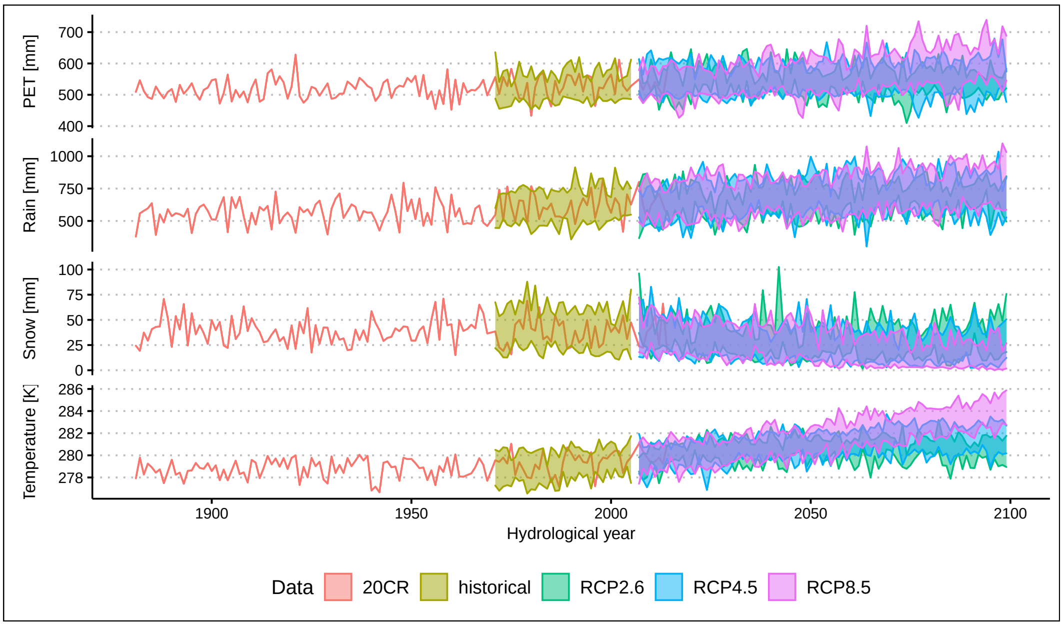

Figure 2. The 20CR and EURO-CORDEX data for the Biebrza catchment after bias correction. Temperature is the yearly mean and the remaining variables are yearly sums. The ribbons present the 2.5-97.5 percentiles range for all simulations in a given RCP or historical experiment for EURO-CORDEX data. The gap between historical and RCP ribbons is due to data presentation in hydrological years, whereas the EURO-CORDEX simulations starts and finishes as calendar years.

in reference to daily mean observations by 11.8% and 10.9% by 20CR and EURO-CORDEX mean respectively.

The 20CR data fits the EURO-CORDEX ensemble in the overlapping historical period after bias correction (Figure 2). The 20CR data show no significant trends until the end of the first half of the 20th century. In the 1950-2015 period the air temperature trend of 0.02 K year^{-1} (p=0.0008) was observed. EURO-CORDEX data presented significant trends for ensemble yearly medians for all meteorological variables except PET for the RCP 2.6. The PET trends for the remaining RCPs were 0.24 (RCP 4.5), and 0.81 mm year^{-1} (RCP 8.5). For RCP 2.6, RCP 4.5, and RCP 8.5 respectively the trends were -0.08, -0.16, and -0.31 mm year^{-1} for snowfall, 0.64, 0.81, and 1.61 mm year^{-1} for rainfall, and 0.01, 0.02, and 0.05 K year^{-1} for air temperature.

3.2 Model calibration

The hydrological model calibration results formed a clear Pareto front with a minimum RMSE of 0.19 m and maximum KGE of 0.86 (Figure S1). Out of these models we choose one with an RMSE of 0.24 m and a KGE of 0.69 as the best performing and used it for further simulations. The calibrated parameter values (Table S7) had values within the range presented in the literature for porous media materials (Wösten et al., 1999; Gnatowski et al., 2010). The parameter search space was relatively wide for all material types, yet the saturated hydraulic conductivity presented an expected pattern with greater values for sands than for glacial till and relatively low value for peat. The Manning roughness coefficient had higher values for the Biebrza River and floodplain than reported in the literature (Chow et al., 1988).

3.3 Hydrological validation

Simulated water levels and surface water discharge matched the observations well (Figure 3). Daily discharge at the Osowiec and Burzyn stations, which are located at the inlet and outlet of the floodplain were only slightly overestimated with an absolute error that was 5% of the data range (Table 1). Similar simulated discharge errors were also present for Czachy, which is a major inlet into the floodplain, and Sztabin, which is located in the upper part of the catchment. Overall fit to observations expressed by KGE for discharge showed that Burzyn and Osowiec performed better than smaller stations Czachy and Sztabin. A similar pattern was also present for correlation, which indicated that the discharge temporal variability was simulated better for Burzyn and Osowiec than for Czachy and Sztabin.

Simulated daily water levels showed a good overall fit as expressed by KGE (Table 1). The high values of the correlation coefficient and the visual comparison shows that within-year and multi-year (Figure 3) variability of water levels was simulated correctly. The water levels were overestimated by 3% for Burzyn and underestimated by 4% for Osowiec. The water levels RMSE were the same for both stations in the floodplain and were more attributed to high flows in Osowiec and low flows in Burzyn.

Table 1. Error metrics for all available observations for river gauges. RMSE and bias are in the same units as indicated in the table, remaining metrics are dimensionless. H and Q are water levels and discharge respectively, RMSE / d.r. and bias / d.r. area RMSE and bias normalized to the observations data range (d.r.), corr. is the correlation coefficient.

Station	Units	Period with observations	RMSE	RMSE / d.r.	bias	bias / d.r.	Corr.	KGE
H Burzyn	m	1930-1935, 1946-2017	0.37	9%	0.12	3%	0.83	0.68
H Osowiec	m	1881-1911, 1921-1923, 1925-1935, 1946-2017	0.37	10%	-0.15	-4%	0.79	0.67
Q Burzyn	$m^3\ s^{-1}$	1951-2017	25.88	5%	5.14	1%	0.69	0.64
Q Czachy	$m^3\ s^{-1}$	1957-2017	2.33	4%	-0.73	-1%	0.63	0.50
Q Osowiec	$m^3\ s^{-1}$	1951-2017	17.02	5%	2.79	1%	0.69	0.63
Q Sztabin	$m^3\ s^{-1}$	1951-2017	4.73	5%	0.84	1%	0.60	0.53

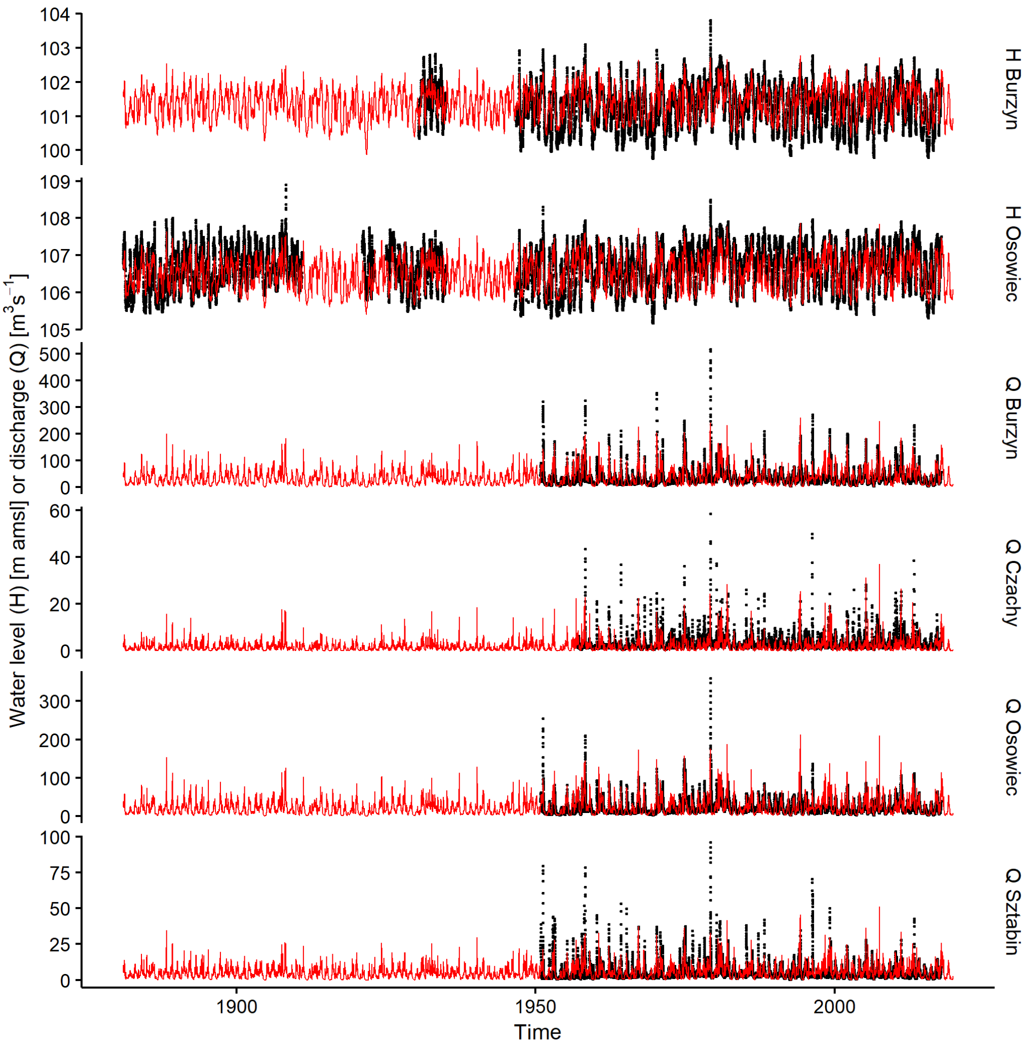

Figure 3. Water levels (H) [m AMSL] and discharges (Q) [m^3s^{-1}] for river gauges. The location of river gauges in presented in Figure 1 except for Sztabin, which is located in the upper part of Biebrza River.

At the catchment scale, the model simulated groundwater levels very well, with the r^2=0.99 (Figure 4). Clear deviation of simulated groundwater levels was observed for the household wells located in the upland. Individual well's performance varied with the location in the model grid. In the floodplain, where the grid was finer than in the remaining parts of the model, the mean RMSE for nine wells was 23% of the data range with a 9% underestimation (Table S8). Outside the floodplain, i.e. in the middle and upper parts of the Biebrza valley, the mean RMSE was 36% and 34% respectively (Table S9). In these parts of the catchment simulated groundwater levels performed worse for certain wells with RMSE up to 76% of the observed data range, although all wells preserved the temporal variability as in the observed data (Table S9 and Figures S2-S5)

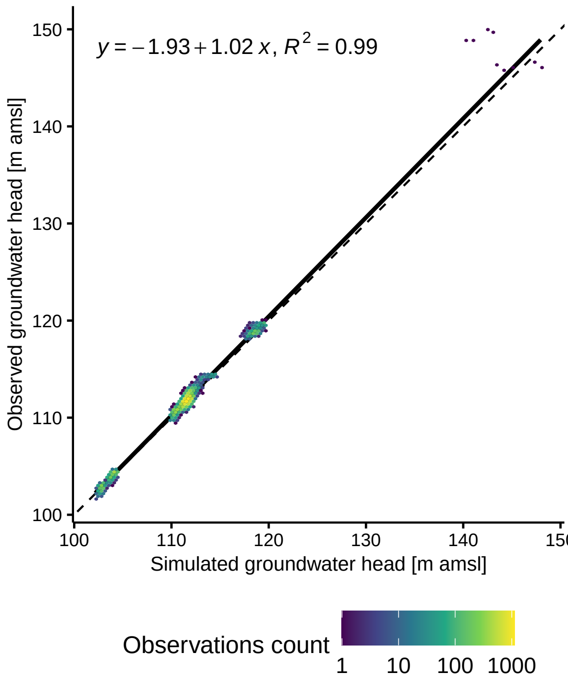

Figure 4. Validation of the simulated groundwater levels using daily observations (usually in ten days resolution) in 43 wells in the period 1994-2019 (N=18032). Solid line - regression line, dashed line - 1:1 line.

3.4 Remote sensing validation

The temporal variability of the SAR water extent correlated better to the flooding extent derived from the river water fractions (a_{river}, r=0.75) than to total extent estimated from the water depth (a_h, r=0.64) (Figure 5). Both a_{river} and a_h water extents overestimated the SAR flooding extent maps for the periods of the lowest water levels when the Biebrza River was not flooding. In these periods the remote sensing data-set was not indicating surface water extent (including between the river banks, Figure S6), while the total area of Biebrza River and oxbow lakes in the floodplain is 2.97 km^2. The Biebrza River and its tributaries were always visible in the hydrological model output. The hydrological model predicted a summer flood in 2017 that was not visible in the SAR data. Also, one summer flood in 2015 visible in SAR data was not simulated by the hydrological model. There was a good agreement in the intersection of the true positive flooding extent from the remote sensing data-set with simulated water depth i_h=0.77

and with simulated river fractions i_{river}=0.78. The lowest agreement occurred during low flow (below bankfull) periods with i_h=0.19 and i_{river}=0.16, while during higher flows (above bankfull) the agreement was higher i_h=0.82 and i_{river}=0.83.

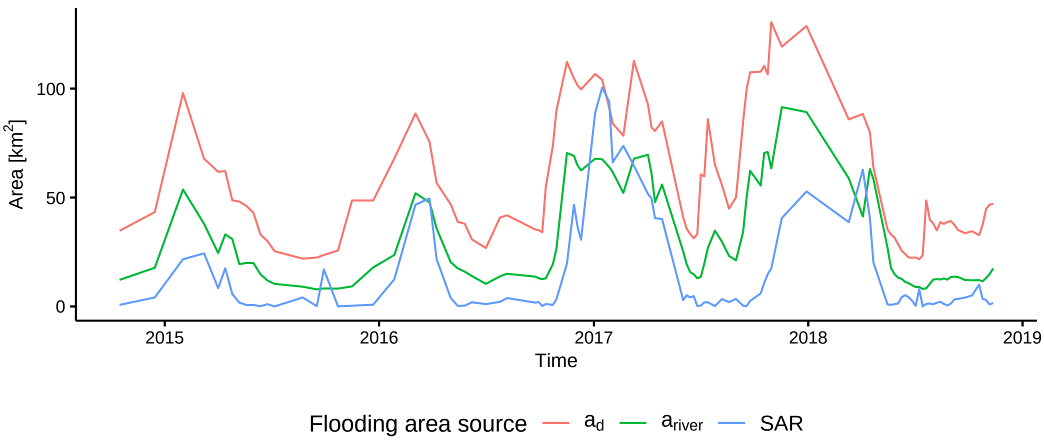

Figure 5. Total area of flooding extent from remote sensing data-set (SAR), calculated for simulated water depths > 5cm (a_d) using Eq. 4, and calculated for river water fractions > 0.1 (a_{river}) using Eq. 5. There are 134 dates in which remote sensing data-set overlapped with the simulation period are presented.

3.5 Hydrochemical validation

The correlation with EC measurements was strongly negative for snow fractions (-0.62) and moderately positive for the river fractions (0.48). A very weak correlation was observed for rainfall (-0.07) and groundwater (0.00). The linear regression model, which explained the EC measurements with the water source fraction predictors, showed that all fractions were significant (p<0.001). The validation metrics for the regression model were r^2=0.58, RMSE=91μS cm^{-1} (18% of data range), and bias b=12 μS cm^{-1} (2% of data range). The highest underestimation visible in the validation of the EC regression model was for measurements located next to an asphalt road located in a central part of the floodplain (~8.5 km from the Biebrza river) (Figure 6). The underestimated predictions are present in the direction of water flow from the road to the river, which indicates possible increased salinity due to car traffic.

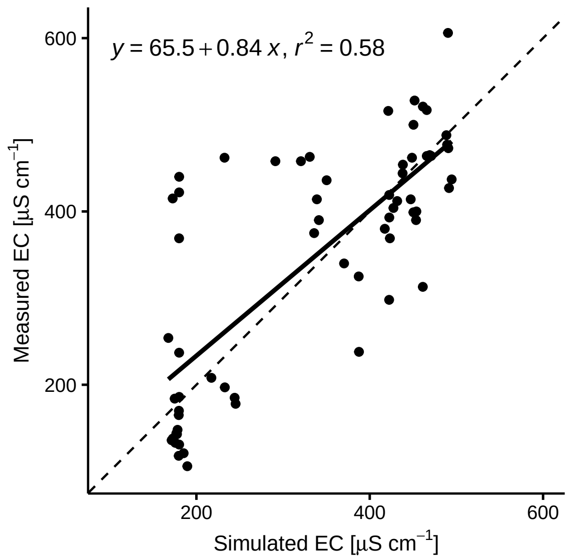

Figure 6. Validation of the EC measurements regression model in the period 2019-2021 (N=64). Solid line - regression line, dashed line - 1:1 line.

3.6 Changes in Biebrza River flow in the past and future climate

Simulated flow characteristics at the Burzyn (outlet) station showed that the 2.5-97.5% range simulations forced by the EURO-CORDEX historical experiments and the 20CR had similar characteristics (Figure 7). The mean simulated water levels overestimated the observations by 1.4% (20CR) and 2.7% (EURO-CORDEX mean) of the observed data range with the underestimated standard deviation by 26% (20CR) and 32% (EURO-CORDEX mean) (Table S10). In the case of discharge, the overestimation was 0.5% (20CR) and 0.9% (EURO-CORDEX mean) with a standard deviation overestimation of 6.4% for models forced using 20CR data and an underestimation by 6.5% (mean) for model forced using EURO-CORDEX data.

Within the 1970-2005 period no significant trends were observed in daily mean discharge or water levels for the models forced EURO-CORDEX or, 20CR data nor for the observation at the Burzyn station. However, in the 1951-2015 period, when observations overlap with the 20CR data a significant trend of 0.173 $m^3s^{-1}year^{-1}$ (p=0.031) was observed; no significant trend was observed for water levels. For this period a similar trend of 0.057 $m^3s^{-1}year^{-1}$ was observed in the model forced with the 20CR data however, it was not significant (p=0.527); in the complementary (1881-1950) period no trend (0.01 $m^3s^{-1}year^{-1}$, p=0.861) was observed. For the future climate impact simulations using the EURO-CORDEX data, significant trends (2006-2099) were observed only for RCP

Figure 7. Mean daily simulated discharge and water levels per year for the Burzyn station forced using 20CR and EURO-CORDEX data. The ribbons present the 2.5-97.5 percentiles range for all simulations in a given RCP or historical experiment for EURO-CORDEX data. The gap between historical and RCP ribbons is due to data presentation in hydrological years, whereas the EURO-CORDEX simulation starts and finishes as calendar years.

2.6 and 4.5. The trend for mean daily discharge was 0.092 m^{3}s^{-1}year^{-1} (p=0.005) for RCP 2.6 and 0.080 m^{3}s^{-1}year^{-1} (p<0.001) for RCP 4.5. In the case of mean daily water levels, the trend was 0.0015 m year^{-1} (p=0.007) for RCP 2.6 and 0.0007 m year^{-1} (p=0.032) for RCP 4.5.

3.7 Changes of water sources fraction in the past and future climate

The simulated daily mean volume of water from different sources did not show significant trends for the past climate forced with the 20CR data (Figure 8). In the simulations forced by the EURO-CORDEX data for future climate positive trends were observed for the river, rainfall, groundwater, and river-floodplain mixed water volumes in RCP 2.6 and RCP 4.5. In the RCP 8.5 significant trends were observed only for rainfall and snowmelt volume. For all RCP snowmelt volume trends were negative, however, the trend was not significant for RCP 2.6. The snowmelt water was characterized by the lowest volume in the floodplain area and was subjected to the highest relative changes in the RCPs 4.5 and 8.5.

Length of a period in which water source fractions were dominant, the river-floodplain mixing degree was high, or water depth was greater than 1 cm was stable before 1950

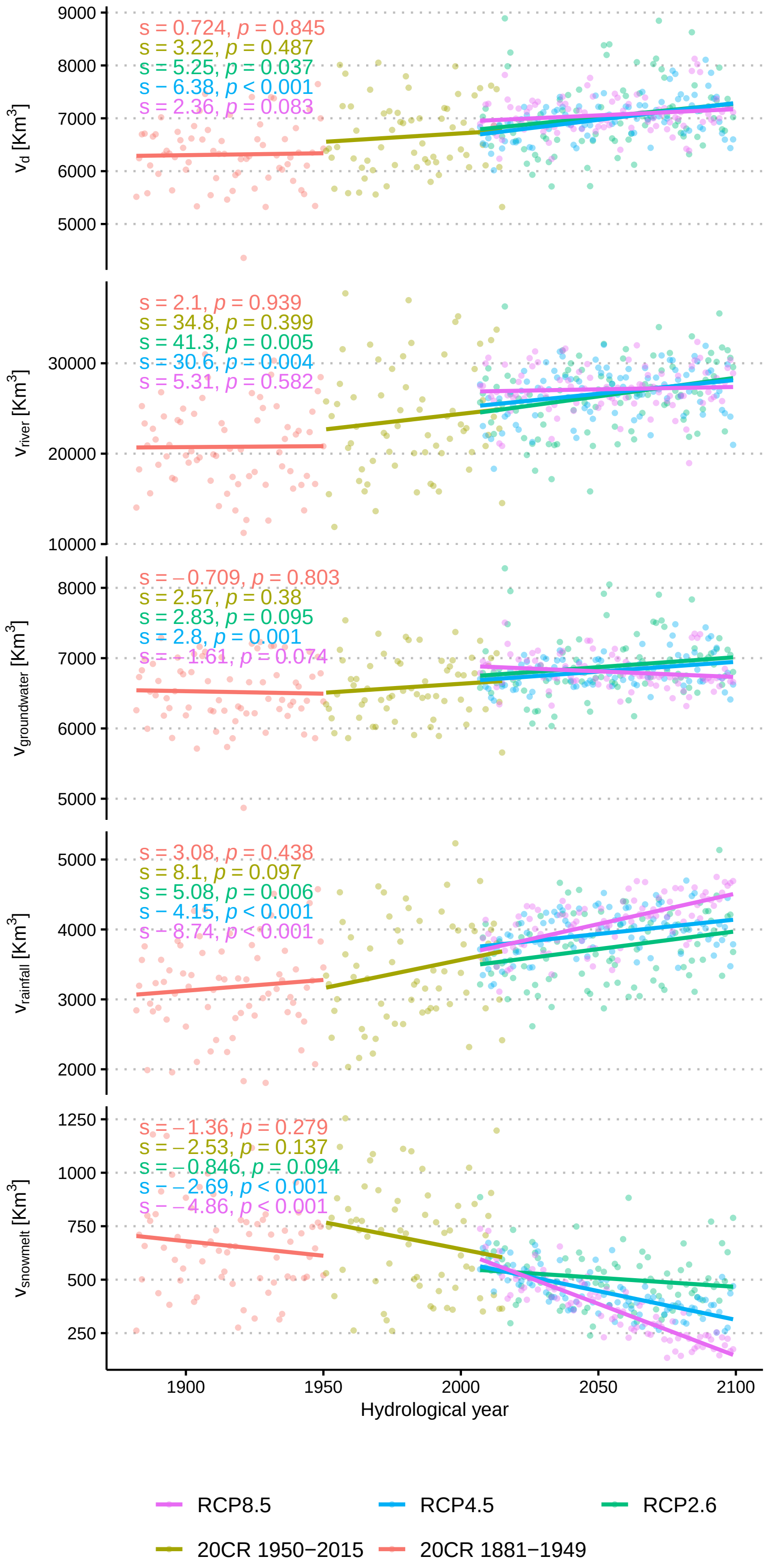

Figure 8. Simulated surface water volume daily means for river-floodplain mixing (v_d), river (v_{river}), groundwater ($v_{\mathrm{groundwater}}$), rainfall (v_{rainfall}), and snowmelt (v_{snowmelt}) in the floodplain per hydrological year. The RCPs data is presented as the ensemble means. The s symbol is a slope of a regression line (trend) [Km^3 year^{-1}] and the p symbol is a p value of a t-test for the slope estimate, 1Km^3 is 1000 m^3.

with only a few nodes showing a slight increase for l_{rainfall}(Figure 9). A similar situation was observed in the simulations for the 1950-2015 period. Therein, however, l_h and l_{river} increased in proximity to the Biebrza River. Also, l_{rainfall} showed a more distinctive patch of increased vales when compared to the latter period.

The trends for the ensemble mean in RCP 2.6 and 4.5 showed a similar pattern of increased l_h and l_{river} in the proximity of the Biebrza River and increased l_{rainfall} in the central part of the floodplain (Figure 9). The increase of l_h and l_{river} was the greatest in RCP 2.6 out of all analyzed RCP scenarios and past climate periods. The increase of $l_{\mathrm{groundwater}}$ was observed in RCP 2.6 near the valley margin, which was not visible for RCP 4.5. Unlike RCP 2.6, RCP 4.5 showed a decrease of $l_{\mathrm{groundwater}}$ and l_{snowmelt} in the central part of the floodplain.

The RCP 8.5 simulations showed that l_h and l_{river} was small and clearly smaller than in RCP 2.6 and 4.5 while the change of l_{snowmelt} was similar as in RCP 4.5 (Figure 9. The decrease of $l_{\mathrm{groundwater}}$ was the highest in RCP 8.5 and was visible in the central part of the floodplain (especially in the ditches), near the valley margin (northern part), and in the Biebrza River bed. Also, the increase of l_{rainfall} was the highest in RCP 8.5 and was present in almost the entire floodplain.

In all simulations the length of the high river-floodplain mixing period, l_d, increased with increasing l_{river}, yet, the trend in l_d was smaller than the increase of l_{river}. An exception of this was in the central part of the floodplain in all RCP scenarios, where l_{river} did not show a significant trend, but l_d showed an increase. Therein l_{rainfall} increased the most along with the $l_{\mathrm{groundwater}}$ increase in RCP 4.5 and the $l_{\mathrm{groundwater}}$ decrease in RCP 8.5. The l_d did not change nearest to the river in the RCP scenarios, whereas the l_{river} changed the most. In this area, l_d was high due to mixing at the beginning of the flood.

The trend of surface water depth above 1 cm period, l_h, resembles that of l_{river} in the area where both trends were significant, i.e. in the proximity of the river. The l_h, unlike l_{river}, increased also in the central part of the floodplain, especially in the ditches, and next to the valley margin in all RCP scenarios and a few nodes in the 1950-2015 period. The highest l_h increase in these areas was observed in the RCP 2.6, although the change in l_{snowmelt}, $l_{\mathrm{groundwater}}$, and especially in l_{rainfall} was the smallest in this scenario among all RCPs. Overall, the magnitude of l_h change was the highest in RCP 2.6 (accompanied by the highest magnitude of l_{river} change) although the area of significant changes was greater in remaining RCPs. Notably, in the areas where l_d increased, but l_{river} trend was not significant the l_h also showed an increase. Still, $l_h$increased in areas further away from the rivers where neither l_d nor l_{river} increased.

Significant trends in mean daily water depth, $\bar{h}$, were observed spatially only in the RCP scenarios for the future climate (Figure 10). The trends were the greatest in the proximity of the river, reaching some river nodes up to 6.3 mm year^{-1} in RCP 2.6, 4.0 mm year^{-1} in RCP 4.5, and 0.7 mm year^{-1} in RCP 8.5 (Figure 10). The RCP 4.5 and 8.5 scenarios predict a very small positive trend across the majority if the floodplain, whereas RCP 2.6 predicts such a trend in only remote parts of the floodplain and in ditches.

4 Discussion

4.1 Forcing data

The forcing data matched the meteorological observations in terms of mean and standard deviation, which indicates, that the biases were removed correctly. The highest deviations from observations were observed for the total precipitation and air temperature, which still, were comparable to other studies conducted in our study region. Mezghani et al. (2017) reported an RMSE of 15.5 mm month^{-1} (equivalent to about 0.51 mm day^{-1}) and the air temperature monthly mean RMSE of 1.1°C (daily minimum) and 1.6°C (daily maximum) using an ensemble of 9 EURO-CORDEX simulations. The differences between observations and bias-corrected air temperature did not have a large impact on the hydrological simulations, because the air temperature was not used to calculate PET in the hydrological model. Rather than that, the air temperature was only used to calculate the degree-days for LAI estimation. The bias-corrected PET data had very small deviations from the observations.

The predicted change of total precipitation and air temperature varies between the model applied. In general, other studies show indicated an increase in yearly precipitation, and air temperature, and a decrease in snow cover by the end of the 21st century in our study area. Warszawski et al. (2013) showed that the yearly precipitation will increase by up to 10% and air temperature by 2-6 K the in RCP 8.5 scenario. Similarly, Schneider et al. (2013) showed that winter half-year precipitation will increase by 5-15%, with no changes in the summer half-year precipitation, mean annual temperature will increase by 2-2.5 K and the snow-cover period will decrease by 20-30 days. Also, Mezghani et al. (2017) predicted an increase of precipitation by 9.7% in RCP 4.5 and by 15% in RCP 8.5 and the air temperature increase of 2 K in RCP 4.5 and 3.6 K in RCP 8.5. Except that we were not able to compare the RCP 2.6 scenario, these results are consistent with the bias-corrected data used to force hydrological simulations in our study.

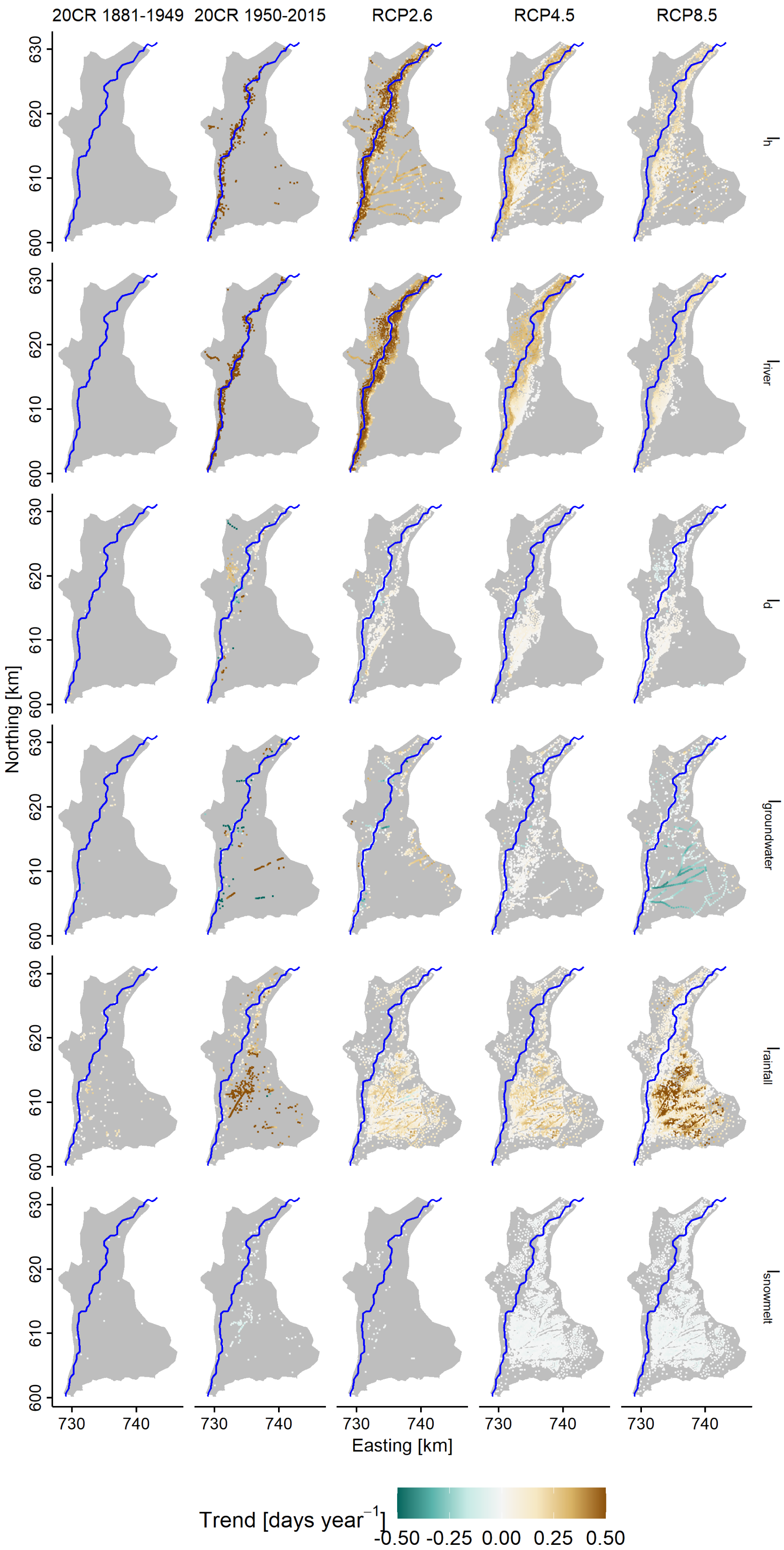

Figure 9. Changes of the period's length when water depth, h, is greater than 1 cm (l_h), river water (l_{river}) or floodplain water ($l_{\mathrm{groundwater}}$, l_{rainfall}, and l_{snowmelt}) fractions are greater than 0.75, and the river-floodplain mixing degree, d, is greater than 0.75 (l_d) annually. Only model nodes with significant trends ($p<0.05$) are shown. The Grey polygon is the floodplain area, the blue line is the Biebrza River; tributaries and ditches are not shown for clarity, please refer to Figure 1 to identify their course.

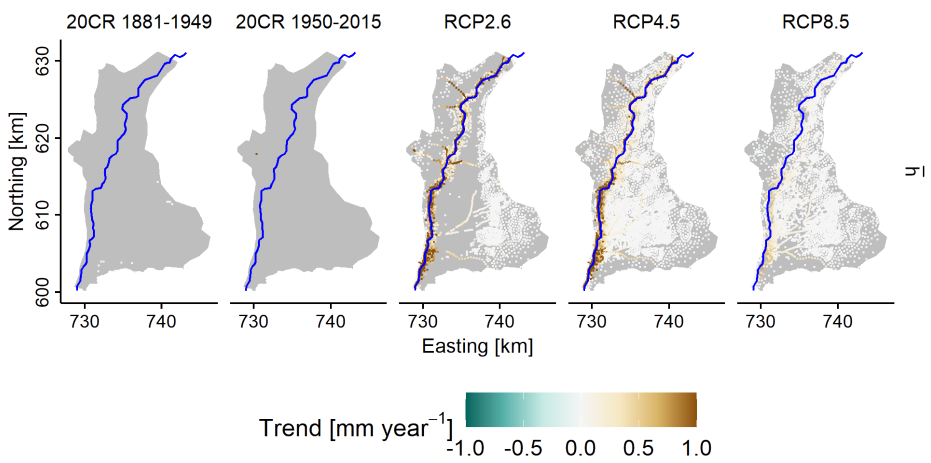

Figure 10. Changes in the annual mean daily water depth, $\bar{h}$. Only model nodes with significant trends (p<0.05) are shown. The color scale is clipped to the <-1, 1> mm year^{-1} range, and the values outside this range are colored as equal to -1, or 1 mm year^{-1}; the clipping affected 15% of the data in RCP 4.5 and 30% of the data in RCP 2.6 located in the proximity of the river. The Grey polygon is the floodplain area, the blue line is the Biebrza River; tributaries and ditches are not shown for clarity, please refer to 1 to identify their course.

4.2 Model development and calibration

An alternative model calibration strategy to the one used in our study was to calibrate the model on a coarser grid and then conduct only fine-tuning in the finer grid (von Gunten et al., 2014). We decided not to use this approach because our target grid was relatively coarse with a number of simplifications. Another approach was to calibrate the model to steady-state using average fluxes as boundary conditions, which was used in several studies involving IHMs (Partington et al., 2020). The advantage of this approach is that the steady-state simulations require shorter simulation time than transient-state simulations for one or more events or hydrological years. We, however, were focused on the dynamic process of flood development, involving interactions of water from groundwater and surface water. Therefore the steady-state calibration for average conditions could lead to unrealistic parameter estimations during flooding, especially for surface water flow parameters for the floodplain.

Still, our strategy with the screening of 800 quasi-random parameter sets was adequate for the model calibration problem. An advantage of this approach is that the approximate total computation time is known a-priori and the problem is easily parallelized on a cluster. A disadvantage is that too sparse parameter space sampling may lead to unsuccessful calibration. The calibration results showed rather high equifinality when only one optimization criterion was analyzed (either KGE or RMSE). However, selecting a model with high KGE and low RMSE considerably decreased the number of behavioral models. At the Pareto front, the relation between KGE and RMSE is non-decreasing (for RMSE < 0.5 m), meaning that the selection of a model with higher KGE results in higher RMSE, i.e., in worse groundwater simulation performance. As indicated in previous studies (e.g. McCabe et al. (2005); Rientjes et al. (2013)), this stresses the importance of using multi-objective calibration when compared to a single-objective calibration.

The porous media parameters were calibrated to realistic values when compared to literature values. This was not entirely the case for the overland flow parameters, where especially the Manning roughness coefficient was higher than expected. This was a result of the generalization of the river channels in the model grid, which resulted in wider and straighter channels than in reality. Eventually, this generalization with realistic roughness parameter values would lead to increased simulated water velocity and too-low water levels. The effect of too high roughness was too high water levels during low flow when water was in the river bed. This effect was reinforced by the high obstruction height parameter value in the river bed, which further increased the roughens for the lowest water levels. The high obstruction height was calibrated in the river bed to compensate for unnaturally-wide perpendicular cross-sections used in the model grid.

The model's purpose was to analyze hydrological conditions during flooding, focusing more on the floodplain area, rather than on the river bed. Further, our aim was to analyze multiple long-term climatic scenarios that require very long computation times. Therefore, in our opinion that the simplifications used herein and the resulting unrealistic surface water parameters in river were justified. While local-scale IHMs are often developed with very fine girds and short time steps, the regional-, country-, or continental-scale model use simplification strategies for model development. One of the strategies used in climate-change studies in regional-scale IHMs is to use aggregated water fluxes in monthly resolution (Goderniaux et al., 2009; Erler et al., 2019). Another strategy is to use a coarser grid, which preserves only key landscape features, such as bigger lakes or major river tributaries. Following this strategy, Goderniaux et al. (2009) used a model with 785 nodes per layer in a 480 km^2 catchment, Erler et al. (2019) used 33092 nodes per layer in a 6800 km^2, and Chen et al. (2019) used about 225000 nodes per layer in 10.5 million km^2 basin. This strategy also involves using relatively thick top layers, which were 1 m in Goderniaux et al. (2009) and 2.5 m in Chen et al. (2019). Our strategy with daily fluxes, 19297 nodes per layer in a 7000 km^2 catchment (refined to 10436 nodes in

220 km^2 floodplain) and about 0.75 m thick top layer makes the model comparable or higher resolution to the mentioned studies.

4.3 Model Validation

The multi-site validation presented in this study showed overall good performance of dynamic hydrological processes simulated in the model. However, a model performance degradation, such as increased RMSE for groundwater heads and decreased KGE for discharge, was observed outside the floodplain area (in the middle basin, upper basin, and upland). This was primarily a result of using a finer grid in the floodplain and a coarser grid elsewhere. Fraction of the error may be attributed to errors in the elevation of the groundwater wells or the DEM used for the model.

From the flooding perspective, the model was unable to simulate correctly the highest discharge peaks (above 250 m^3s^{-1}), which occurred five times in the 1951-2017 period. The remaining events were simulated with smaller errors both in terms of water levels and discharge. We attribute the inability to simulate the highest peak discharges primarily to the too-high roughness coefficients obtained during the calibration, which decrease the water velocity and effectively produce a smaller and wider flood. Partially, this problem may be also attributed to the coarse resolution (1°x1°) 20CR forcing data and bias-correction approach which was not able to suitably force the highest events. Therefore, our model is unsuitable for reliably predicting rare, extreme events in the past and future climate. However, it has demonstrated the capacity to predict normal hydrological behavior including flood events with shorter return periods.

The validation with remote sensing showed good agreement with the spatial and temporal dynamics of river water flooding. This was not the case for the total flooding extent (from the river and all floodplain water sources). Even though a high (5 cm water depth) threshold was used to identify the total flooding extent, the simulation provided a larger and longer-lasting extent than the remote sensing estimate. Apart from the bias in the remote sensing product caused due to spatial resolution which disabled identification of all permanent open water objects, the remote sensing validation indicates that SAR water extent estimation in a densely vegetated wetland area is problematic. Several attempts were made to the problem of vegetation, or other objects obscuring water extent by using auxiliary information such as elevation models (e.g. Mason et al. (2012)). In wetland cases, where a small surface water depth is frequently present and a flat land surface includes micro-topography features, these methods have limited applicability. A recent approach involving multiple polarimetric decomposition models for SAR data in the Biebrza wetland has shown that with a C-band (the same band as in our data set) SAR the identification of open water with vegetation emerging more than 10 cm can be difficult Gierszewska and Berezowski (2022). The solution could be flooding extent mapping in wetlands using SAR sensors with longer (P, or L) wavelengths.

The SAR flooding area correlated more strongly with the river water than with the total flooding extent. This was reflected in the number of small features visible both in SAR and river water flooding extents that diminished in the total flooding extent. This shows a potential application of SAR data in densely vegetated wetlands where they can be used to track the extent of river water flooding. This is not due to different sediment concentrations, which are used for mapping using optical sensors (Mertes, 1997; Park & Latrubesse, 2015), but due to high water depths in the river flooding zone, which can overtop vegetation. As this phenomenon was the case in Biebrza wetlands it does not necessarily have to be the case in other sites, which can still have too low water levels for detection of surface water.

The EC of water can be used as an indicator of the surface water source, as it has higher values in the river than in floodplain water (Chormański et al., 2011). Our results are in agreement with this showing that EC correlated positively with river water fractions and negatively with snowmelt water fractions. Further, the hydrochemical validation shows that all water source fractions are significant predictors of surface water EC, which indicates that the simulated fractions agree with the true water sources. Our previous study (Berezowski et al., 2019) conducted for a single flooding event on a finer grid showed that the simulated fractions agree with water sources derived from a multi-parameter hydrochemical analysis. In the current study, due to high labor intensity, we were unable to repeat the hydrochemical analysis.

We put a lot of emphasis in this study on model validation which is a key step in impact model development in climate change studies. The validation for the Burzyn station was satisfactory and the lack of the trends in observed discharge and water levels were preserved in the model simulations (although one of the trends was significant in the observations and not significant in the simulations forced by the 20CR data). This indicates that the model passes the comprehensive evaluation criteria described in Krysanova et al. (2018) for the Burzyn Station. The remaining stations, situated in the upper parts of the catchment, have in general lower correlation coefficients, however, their KGE is still comparable to the KGE of the outlet. The comprehensive evaluation can be used as an indicator of a robust impact model (Gelfan et al., 2020), therefore, our model is suitable for climate change impact study of the floodplain area.

Needless to say, our evaluation was more comprehensive than described above. This was because the simulation of water mixing, which is a product of interaction between climate, groundwater, and river flooding, requires more confidence in the modeling results than just agreement with observed water levels or discharge. Our remote sensing and EC evaluation criteria indicate that the model is suitable for the analysis of water mixing.

4.4 Changes in Biebrza River flow in the past and future climate

Our future climate impact simulations that show a positive trend (2005-2099) of the mean discharge in Biebrza River are consistent with Roudier et al. (2015), who have shown that less severe droughts and higher flooding discharges will be present in this region. Two studies conducted for nearby catchments (Guber and Narewka) close to Biebrza also indicated decreased severity of droughts in RCP 4.5 (Meresa et al., 2016) and an increase of yearly maximum flows in RCP 4.5 and 8.5 (Osuch et al., 2016). A regional study also showed that both low and high flow will increase by 2100 in RCP 4.5 and 8.5 in the Biebrza catchment, although the ensemble of simulations was inconsistent for the RCP 4.5 in the 2071–2100 period (Piniewski et al., 2017). On the other hand, our finding that no discharge trends (2006-2099) will be in the RCP 8.5 is inconsistent with Alfieri et al. (2015), who showed that the mean daily flow in Biebrza will increase by about 15% (1990-2080) in the RCP 8.5.

The RCP 2.6 ensemble means in our study are associated with the highest uncertainty, because only five EURO-CORDEX simulations were available. Therefore, our findings that the highest trend (2006-2099) in mean discharge will be observed in RCP2.6 have to be considered less robust than the results of the trends in remaining RCPs. Still, this finding is partially supported by Marx et al. (2018), who showed that the 10-20% change in the mean low flow in Biebrza River will take place under 2K air temperature increase, whereas under 1.5K and 3K scenarios the change will be between -10% and 10%.

Projections of future hydrological impact often disagree due to differences in forcing-data sources and processing, impact models used, impact indicators, and methods of comparison with the reference period (Z. W. Kundzewicz et al., 2016). All of these reasons are relevant for comparisons presented in this section. The aim of this study was not to compare the climate change impact on the Biebrza River with other studies but to investigate the impact on the water mixing using the best methods available. We used all available EURO-CORDEX simulations which provided the required forcing data. However, these simulations used often different GCMs or RCMs than in the discussed studies. Moreover, our simulations were limited by the use of data from only one RCM (but multiple GCMs), while most of the remaining studies used more than one RCM. Also, we ran continuous simulations in a daily resolution for the 1881-2099 period, which was used to calculate trends and their statistical significance for variables relevant to our study. This was not the case in the other studies discussed in this section, which calculated a relative change of low- or high-flow indicators with respect to a reference period. Finally, we used a finer spatial resolution and/or better physical representation of hydrological processes in the HydroGeoSphere model than in models used in these studies.

4.5 Changes of water sources fraction in the past and future climate

Water source fractions were stable in the 1881-2015 period in terms of the associated volume of water, which coincide with no trends in the forcing data. Since the second half of the 20th century, a shift from rainfall replacing snowmelt fractions dominance was observed in the central part of the floodplain. In parallel, river fractions and high water depths persisted longer in the proximity of the river due to the rainfall accumulation in the whole catchment. Neither of these changes was related to a significant trend in rainfall or snowmelt in the forcing data, nor resulted in a significant change in the flooding volume of these water sources.

In the RCP 2.6 volume of river and rainfall water significantly increased during the 2005-2099 period. An increase in rainfall with a significant, but eight-fold smaller decrease in snowfall, and no change in PET resulted in overall wetter conditions. This translated not only to increased river discharge, and high river fraction persistence but also to longer-lasting high groundwater fractions and surface water depth. Effectively, the period of river-floodplain water mixing was longer in the proximity of the river by forming a clear belt and resulting in a significantly increased volume of mixing water.

A similar situation was observed in RCP 4.5, but due to a two-fold higher decrease of snowfall and similar magnitude of rainfall trend as in RCP 2.6 period of high snowmelt fractions shortened during the 2005-2099 period. In addition to that, a greater decrease in PET in RCP 4.5 than in RCP 2.6 resulted in less wet conditions. As a result, a lesser trend of river discharge and water levels was observed. Since groundwater discharge is more related to overall drier or wetter conditions rather than to instantaneous fluxes of water the decrease in snowmelt water resulted in longer dominance of groundwater fractions in the river proximity in RCP 4.5. Despite drier conditions in RCP 4.5, the longer periods of groundwater, rainfall, and river fractions in this area affected in the river-floodplain mixing zone last longer in greater areas than in RCP 2.6.

A different situation was observed in RCP 8.5, where snowfall nearly ceased and the increase of rainfall was in great part balanced by the increase of PET resulting in no trends observed in discharge or water levels in the 2005-2099 period. The stability of discharges was not accompanied by the stability of the volume of water from different sources, as rainfall volume increased and snowmelt volume decreased. The most distinctive pattern of high groundwater fractions persistence decrease was observed in RCP 8.5. Such a big groundwater fraction decrease is an indicator of drought conditions locally in the central part of the floodplain. This was however balanced by longer persistence in the northern part of the floodplain and near the valley margin, resulting in no significant trend in groundwater volume. The area in which the river-floodplain water mixing period increased was similar to in RCP 4.5 but more patchy, and the magnitude of the trend was smaller. Despite this spatial pattern the trend of mixing water volume did not change significantly. Unlike in other scenarios, in RCP 8.5 longer lasting high rainfall fractions resulted in a zone of a shortened period of river-floodplain mixing in the NE part of the floodplain, near the river.

Trends in mean water depth and inundation period length did not align with the trends of water source fractions. Overall, the trends of mean water depth were rather small except in the river proximity. Also, greater variability was present in the water source fraction trends than in water depth or inundation period trends.

4.6 Implications for modeling

When taking into account distributed hydrological processes, the opposite directions of changes can be present within one floodplain. These changes either combine with each other to amplify a signal in the lumped volume or cancel each other and the signal in the lumped volume attenuates. A comparison of the lumped volume of water associated with different fractions and the accompanying spatial patterns in the floodplain area shows the advantage of using fine, distributed model output. As illustrated for the 1951-2015 period, the climatic signal was lost in the lumped output due to averaging with other effects, whereas a clear trend pattern was visible spatially. A similar situation was observed in the RCP 8.5 scenario, where the greatest spatial trends were present in groundwater fractions persistence, but the trend in groundwater volume was not significant. This is relevant if upon the lumped or distributed impact model output another process (e.g. ecological, or hydrochemical) would be modeled or a management decision would be undertaken.

This links to another advantage of using IHMs in climate change scenarios, which is revealed when data on all boundary conditions are not explicitly available temporarily and spatially. Climate influences both surface water and groundwater and thereby affects also feedback between the two domains expressed as groundwater-surface water interactions. The assumption about surface water infiltration, or groundwater discharge is difficult to make properly for simulations with long time horizons, whereas, they are required for surface water models that are not integrated with groundwater models. This is not the case for catchment-scale IHMs, where surface water and groundwater simulations are simultaneously forced by the climate data and the time-variable feedback between the two domains are preserved.

As illustrated in our study, the period of inundation with water depth above 1 cm, was also influenced by climate outside the river water flooding zone. Moreover, the trend of this period was not correlated to the distance from the river, as it increased near the river, then decreased, and increased again in the central part of the floodplain (RCP 4.5 and 8.5), or it increased in the entire floodplain (RCP 2.6). The trend in water depth change was correlated to the distance from the river, however, significant positive trends were observed both in areas dominated by the river and floodplain water. Furthermore, water depth trends in RCP 2.6 showed that large areas of the floodplain did not have any trends, while the most remote parts of the floodplain had a significant positive trend. This depicts another advantage of IHMs, which is the representation of water depths in the floodplain. Hydrodynamic models for 2D surface water routing perform very well in simulating river water flooding extent, whereas, are unable to simulate inundation from other sources, such as groundwater, without coupling with different models (Appledorn et al., 2019). Still, some studies use a surface-water-only model to analyze long-term floodplain inundation changes (Veijalainen et al., 2010; Wen et al., 2013). While in some areas a lack of groundwater coupling may not influence the results, in other areas it may be a source of bias in simulated inundation extent.

Although not discussed in this study, interactions between water sources may influence the surface water velocity field in the floodplain in reference to a situation when river water is the sole inundation source. This may further influence the sedimentation pattern in the floodplain due to the settling velocity parameter of the particles. Several climate impact studies analyze the floodplain sedimentation patterns by taking into account the major water sources, such as rivers (Park et al., 2022) or sea level (Manh et al., 2015). Whereas, as illustrated in this study, the river-floodplain mixing zone is relatively wide and it varies under climate change, which may affect sedimentation patterns.

4.7 Implications for ecological processes

Mixing of water from different sources creates biogeochemical hot spots and hot moments, such as denitrification (McClain et al., 2003). A number of studies have analyzed denitrification spatially in inundated floodplains to reveal that it is strongly affected by connectivity with river water (Forshay & Stanley, 2005; Racchetti et al., 2011; Jones et al., 2014; Scott et al., 2014). Our study shows that, that the river-floodplain water mixing volume, extent, and persistence varies with climate change, therefore denitrification patterns can also be affected. This variability was visible much better in the spatial pattern than in the lumped, water volumes. Therefore, improvement in denitrification modeling at floodplain (Hallberg et al., 2022) or catchment (Adame et al., 2019) scale or in the use of scaling relationships (O'Connor et al., 2006; Tomasek et al., 2017) could be achieved by introducing additional variables related to groundwater discharge, river water, or the river-floodplain mixing extents.

Another process, that is related to the extent of water from different sources is vegetation development (Keizer et al., 2014; Park & Latrubesse, 2015). Modeling of vegetation development under climate change may be hampered because studies using the process-based model (Politti et al., 2014), a statistical approach (Mosner et al., 2015) do not include hydrodynamic feedback between water from different sources. As shown by (Gattringer et al., 2019) predictors from an IHM improve habitat modelling in comparison to groundwater, or surface water only predictors scenarios. Our results indicated that in some scenarios the trends were not present in water levels, or discharges while they were present in the persistence of dominant water sources. Therefore we believe, that the inclusion of water sources extents predictors could improve vegetation models further. We are not aware of any study that used IHM-simulated water sources to model vegetation development or distribution.

A more recent study conducted in the Biebrza floodplain revealed that the vegetation productivity was better predicted by the zone of nutrients rich sediment deposition, located close to the river, rather than by the river water extent or the total inundation extent (Keizer et al., 2018). As mentioned in Section 4.6, sedimentation is related to water velocity, which may decrease where water sources with different momentum mix. Therefore, the mixing degree, d, which was strongly variable under climate change in this study, can potentially be a candidate for high productivity vegetation zone predictor in temperate floodplains. This, however, was not tested here and should be investigated in a future study.

4.8 Implications for management

The Biebrza floodplain, as part of a national park, has been subjected to active protection measures. An increase in water levels through the construction of dams, or vegetation removal by mowing, allowed, to some extent, to diminish the potential effect of climate change in this area (Berezowski et al., 2018). Our results together with other experiments discussed herein show that the analysis of water sources and their mixing may have a considerable ecological effect. However, at this point, more models are needed to asses this effect more precisely spatially and temporarily. Therefore, the current local management strategy could be to increase the resilience of the wetland ecosystem and implementation of adaptive management (Lawler, 2009). Except that, the local management strategies may be somewhat challenging, as tools for preserving the shape and duration of water sources' zones are limited. On the other hand, our results have shown that the spatially distributed trends in water source fractions were driven solely by climate change, as our model neglected other divers (water use, land-use change, etc.). Therefore, global actions limiting climate change impact on wetlands driven by national and international policies (Moomaw et al., 2018) seem to be an appropriate measure to limit the shift in the extent of water from different sources.

4.9 Note on hardware requirements

The simulations were run on the Tryton cluster, which has 3215 Intel Xeon Processors (E5 v3, 2.3 GHz, 12-core) with 128 GB RAM, resulting in a total of 1.792 PFLOPS. We split the simulations into 978 smaller tasks (a three-year simulation period with a two-year warm-up period), to use the resources in parallel and to fit into 72h wall time for a single simulation. The cluster resources were shared with other users therefore it took about five months to finish all computations. The total output data produced by the models accounted for about 20TB.

5 Summary and conclusions

Simulations of surface water source fractions in a natural wetland floodplain over a two-century period reveal that by 2099 the projected future climate change will significantly alter the patterns that were relatively stable in the 1881-2015 period. Our results show that analysis of the lumped output of the model was less sensitive to depict the climate change effect that was visible when the trends were analyzed spatially in the floodplain. Different future climate scenarios showed very variable impacts on water source fractions, which were often counterintuitive. In the RCP 2.6, which projected the least climate change in the study area, we observed the highest magnitude of changes related to the increase in river discharges, water levels, and river water fractions. In the RCP 8.5 scenario, which projected the greatest increase in PET and rainfall accompanied by the greatest decrease in snowfall, these trends were less significant, while only this scenario projected dry conditions exhibited by a decrease of groundwater fractions in the inundation. The trends in water source fractions had different spatial patterns and showed greater sensitivity to climate change than trends in water depth and inundation duration.

This complex hydrological impact was simulated by the IHM, which allowed us to model interactions between groundwater and surface water and limit the assumptions about hydrological fluxes in the top layer of the model to the meteorological forcing. This is the first study that simulated the climate impact on water source fractions in the inundation and the longest application of IHM in terms of the simulation period. Hydrological impact studies are always related to uncertainty, which we limited here by multivariable verification and projection of future impact using an ensemble of 10 EURO-CORDEX simulations (only 4 in RCP 2.6).

We showed that the water source fractions are sensitive to the climate in a natural temperate zone wetland floodplain. This fact has several implications for other modeling studies, ecological processes, and management in similar wetlands. Modeling problems should be carried out using IHMs to depict proper inundation or sedimentation patterns spatially, because, even if the water sources fractions are not explicitly simulated using HMC, IHMs capture the interactions between water from different sources which produce inundation outside the river water zone and changes the velocity field. Since ecological processes, such as denitrification or vegetation development, are in part related to water sources' zonation and their mixing, these variables should be taken into account in models, especially in climate change impact studies. Finally, the managers have limited tools in shaping the surface water zonation and extent, therefore except for increasing the wetlands resilience, and adaptive management using an IHM output, global actions aimed at decreasing climate change impact should be the main priority.

Open Research Section

The IHM simulation output used for the analysis, forcing data, and historical water levels (which were not published elsewhere, see below) are available in (Berezowski, 2023). The groundwater levels data was provided by the Biebrza National Park, the data is available upon request from `https://www.biebrza.org.pl/`. Meteorological observations of snowfall, water levels in the rivers, and river discharge was provided by Institute of Meteorology and Water Management - National Research Institute (IMGW-PIB), Poland; data is available at `https://danepubliczne.imgw.pl/`. Support for the Twentieth Century Reanalysis Project version 3 dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), by the National Oceanic and Atmospheric Administration Climate Program Office, and by the NOAA Earth System Research Laboratory Physical Sciences Laboratory; NOAA/CIRES/DOE 20th Century Reanalysis (V3) data provided by the NOAA PSL, Boulder, Colorado, USA, from their website at `https://psl.noaa.gov`. We thank to: Polish Geological Institute, National Research Institute `https://www.pgi.gov.pl/en/data-bases.html` for providing geological data, Head Office of Geodesy and Cartography (GUGiK) `https://www.geoportal.gov.pl/` for providing the Digital Elevation Model of Poland, Water Authority of Poland (Wody Polskie) for providing the Map of the Hydrographic Division of Poland in scale 1:10 000, EURO-CORDEX initiative `https://www.euro-cordex.net/060378/index.php.en`, and the Joint Research Center Agri4Cast `https://agri4cast.jrc.ec.europa.eu/dataportal/` and CORINE Land Cover `https://land.copernicus.eu/pan-european/corine-land-cover` for sharing the data required for this research.

Acknowledgments

This research was financed by grant: 2017/26/D/ST10/00665 funded by the National Science Centre, Poland. Computations were carried out using the computers of Centre of Informatics Tricity Academic Supercomputer & Network in Poland. We thank to Tomasz Bieliński and Marek Kulawiak for help in field measurement and data processing. We thank the Biebrza National Park for the permission for field measurements in the Park area.

References

Aalto, R., Maurice-Bourgoin, L., Dunne, T., Montgomery, D. R., Nittrouer, C. A., & Guyot, J.-L. (2003). Episodic sediment accumulation on Amazonian flood plains influenced by El Nino/Southern Oscillation. *Nature*, *425*(6957), 493–497.

Adame, M. F., Roberts, M. E., Hamilton, D. P., Ndehedehe, C. E., Reis, V., Lu, J., … Ronan, M. (2019). Tropical coastal wetlands ameliorate nitrogen export during floods. *Frontiers in Marine Science*, *6*. doi: 10.3389/fmars.2019.00671

Alfieri, L., Burek, P., Feyen, L., & Forzieri, G. (2015). Global warming increases the frequency of river floods in Europe. *Hydrology and Earth System Sciences*, *19*(5), 2247–2260. doi: 10.5194/hess-19-2247-2015

Anonymous. (1912). *Summary of water levels on the inland water ways of Russia* *from observations at water gauge posts. period: 1881-1910.* Ministry of communications, internal control, waterways, and roads.

Anonymous. (1932). *Hydrographic yearbook. Vistula basin. period: 1918-1932.* Ministerstwo komunikacji.

Anonymous. (1970). *Hydrologic yearbook of surface waters. the Vistula basin and the rivers of the coast region east of the vistula river. period 1955-1970.* Państwowy Instytut Hydrologiczno-Meteorologiczny, Wydawnictwo komunikacji i łączności.

Anonymous. (1980). *Hydrologic yearbook of surface waters. the Vistula basin and the rivers of the coast region east of the vistula river. period 1971-1980.* Instytut Meteorologii i Gospodarki Wodnej, Wydawnictwo komunikacji i łączności.

Anonymous. (2019). *Public data, period: 1951-2019.* Retrieved from `https://danepubliczne.imgw.pl/data/dane_pomiarowo_obserwacyjne/` (Accessed: 2019-04-21)

Appledorn, M. V., Baker, M. E., & Miller, A. J. (2019). River-valley morphology, basin size, and flow-event magnitude interact to produce wide variation in flooding dynamics. *Ecosphere*, *10*(1). doi: 10.1002/ecs2.2546

Arnell, N. W., & Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, *134*(3), 387–401. doi: 10.1007/s10584-014-1084-5

Banaszuk, H. (2004). *Kotlina biebrzanska i biebrzanski park narodowy.* Bialystok: Ekonomia i Srodowisko. (In Polsih)

Barthel, R., & Banzhaf, S. (2015). Groundwater and surface water interaction at the regional-scale - a review with focus on regional integrated models. *Water Resources Management*, *30*(1), 1–32. doi: 10.1007/s11269-015-1163-z

Berezowski, T. (2023). *Hydrological indicators of water zones in inundation, historical water levels, and forcing data for the 1881-2099 period in the lower biebrza valley.* doi: 10.34808/323p-nd55

Berezowski, T., Bieliński, T., & Osowicki, J. (2020). Flooding extent mapping for synthetic aperture radar time series using river gauge observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2626-2638. doi: 10.1109/JSTARS.2020.2995888

Berezowski, T., Partington, D., Chormański, J., & Batelaan, O. (2019). Spatiotemporal dynamics of the active perirheic zone in a natural wetland floodplain. *Water Resources Research*, *55*(11), 9544–9562. doi: 10.1029/2019wr024777

Berezowski, T., Szcześniak, M., Kardel, I., Michałowski, R., Okruszko, T., Mezghani, A., & Piniewski, M. (2016). Cplfd-gdpt5: High-resolution gridded daily precipitation and temperature data set for two largest polish river basins. *Earth System Science Data*, *8*(1), 127–139.

Berezowski, T., Wassen, M., Szatyłowicz, J., Chormański, J., Ignar, S., Batelaan, O., & Okruszko, T. (2018). Wetlands in flux: looking for the drivers in a central european case. *Wetlands Ecology and Management*, *26*(5), 849-863.

Boko, B. A., Konaté, M., Yalo, N., Berg, S. J., Erler, A. R., Bazié, P., ... Sudicky, E. A. (2020). High-resolution, integrated hydrological modeling of climate change impacts on a semi-arid urban watershed in niamey, niger. *Water*, *12*(2), 364. doi: 10.3390/w12020364

Brunner, P., & Simmons, C. T. (2012). Hydrogeosphere: A fully integrated, physically based hydrological model. *Ground Water*, *50*(2), 170–176.

Chen, J., Sudicky, E. A., Davison, J. H., Frey, S. K., Park, Y.-J., Hwang, H.-T., ... Peltier, W. R. (2019, oct). Towards a climate-driven simulation of coupled surface-subsurface hydrology at the continental scale: a Canadian example. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, *45*(1), 11–27. doi: 10.1080/07011784.2019.1671235

Chormański, J., Okruszko, T., Ignar, S., Batelaan, O., Rebel, K., & Wassen, M. (2011). Flood mapping with remote sensing and hydrochemistry: a new method to distinguish the origin of flood water during floods. *Ecological Engineering*, *37*(9), 1334-1349.

Chow, V. T., Maidment, D., & Mays, L. (1988). *Applied hydrology.* McGraw-Hill.

Commission of the European Communities. (2013). *Corine land-cover.* Retrieved from http://www.eea.europa.eu/publications/COR0-landcover (Date accessed: 2013-10-12)

Dąbrowska-Zielińska, K., Budzyńska, M., Tomaszewska, M., Bartold, M., Gatkowska, M., Malek, I., ... Napiórkowska, M. (2014). Monitoring wetlands ecosystems using ALOS PALSAR (L-band, HV) supplemented by optical data: A case study of Biebrza wetlands in northeast Poland. *Remote Sensing*, *6*(2), 1605–1633. Retrieved from https://www.mdpi.com/2072-4292/6/2/1605 doi: 10.3390/rs6021605

Erler, A. R., Frey, S. K., Khader, O., d'Orgeville, M., Park, Y.-J., Hwang, H.-T., ... Sudicky, E. A. (2019). Evaluating climate change impacts on soil moisture and groundwater resources within a lake-affected region. *Water Resources Research*, *55*(10), 8142–8163. doi: 10.1029/2018wr023822

Eurostat. (2019). *EUROPOP2019 - Population projections at regional level (2019-2100).*

Ferguson, I. M., & Maxwell, R. M. (2010). Role of groundwater in watershed response and land surface feedbacks under climate change. *Water Resources Research*, *46*(10). doi: 10.1029/2009wr008616

Forshay, K. J., & Stanley, E. H. (2005). Rapid nitrate loss and denitrification in a temperate river floodplain. *Biogeochemistry*, *75*(1), 43-64.

Garris, H. W., Mitchell, R. J., Fraser, L. H., & Barrett, L. R. (2014). Forecasting climate change impacts on the distribution of wetland habitat in the Midwestern United states. *Global Change Biology*, *21*(2), 766–776. doi: 10.1111/gcb.12748

Gattringer, J. P., Maier, N., Breuer, L., Otte, A., Donath, T. W., Kraft, P., & Harvolk-Schöning, S. (2019). Modelling of rare flood meadow species distribution by a combined habitat surface water-groundwater model. *Ecohydrology*, *12*(6). doi: 10.1002/eco.2122

Gelfan, A., Kalugin, A., Krylenko, I., Nasonova, O., Gusev, Y., & Kovalev, E. (2020). Does a successful comprehensive evaluation increase confidence in a hydrological model intended for climate impact assessment? *Climatic Change*, *163*(3), 1165–1185. doi: 10.1007/s10584-020-02930-z

Gierszewska, M., & Berezowski, T. (2022). On the role of polarimetric decomposition and speckle filtering methods for C-band SAR wetland classification purposes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 2845-2860. doi: 10.1109/JSTARS.2022.3162641

Giuntoli, I., Vidal, J.-P., Prudhomme, C., & Hannah, D. M. (2015). Future hydrological extremes: the uncertainty from multiple global climate and global hydrological models. *Earth System Dynamics*, *6*(1), 267–285. doi: 10.5194/esd-6-267-2015

Gnatowski, T., Szatyłowicz, J., Brandyk, T., & Kechavarzi, C. (2010). Hydraulic properties of fen peat soils in Poland. *Geoderma*, *154*(3-4), 188–195.

Goderniaux, P., Brouyère, S., Fowler, H. J., Blenkinsop, S., Therrien, R., Orban, P., & Dassargues, A. (2009). Large scale surface-subsurface hydrological model to assess climate change impacts on groundwater reserves. *Journal of Hydrology*, *373*(1-2), 122–138. doi: 10.1016/j.jhydrol.2009.04.017

Gramacy, R. B., & Taddy, M. (2010). Categorical inputs, sensitivity analysis, optimization and importance tempering with tgp version 2, an R package for treed gaussian process models. *Journal of Statistical Software*, *33*(6), 1–48. doi: 10.18637/jss.v033.i06

Grygoruk, M., Kochanek, K., & Mirosław-Świątek, D. (2021). Analysis of long-term changes in inundation characteristics of near-natural temperate riparian habitats in the lower basin of the Biebrza valley, Poland. *Journal of Hydrology:*

Regional Studies, *36*, 100844. doi: 10.1016/j.ejrh.2021.100844

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Technical note: Downscaling RCM precipitation to the station scale using statistical transformation - a comparison of methods. *Hydrology and Earth System Sciences*, *16*(9), 3383–3390. doi: 10.5194/hess-16-3383-2012

Hallberg, L., Hallin, S., & Bieroza, M. (2022). Catchment controls of denitrification and nitrous oxide production rates in headwater remediated agricultural streams. *Science of The Total Environment*, *838*, 156513. doi: 10.1016/j.scitotenv.2022.156513

Hwang, H.-T., Park, Y.-J., Sudicky, E., & Forsyth, P. (2014). A parallel computational framework to solve flow and transport in integrated surface subsurface hydrologic systems. *Environmental Modelling & Software*, *61*, 39–58.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., ... Yiou, P. (2014). Euro-cordex: new high-resolution climate change projections for european impact research. *Regional Environmental Change*, *14*(2), 563–578. doi: 10.1007/s10113-013-0499-2

Jacobson, R. B., Bouska, K. L., Bulliner, E. A., Lindner, G. A., & Paukert, C. P. (2022). Geomorphic controls on floodplain connectivity, ecosystem services, and sensitivity to climate change: An example from the lower Missouri River. *Water Resources Research*, *58*(6). doi: 10.1029/2021wr031204

Joint Research Center. (2019). *Agri4Cast Resources Portal.* Retrieved from `https://agri4cast.jrc.ec.europa.eu/dataportal/` (Date accessed: 2019-12-03)

Jones, C. N., Scott, D. T., Edwards, B. L., & Keim, R. F. (2014). Perirheic mixing and biogeochemical processing in flow-through and backwater floodplain wetlands. *Water Resour. Res.*, *50*(9), 7394–7405.

Karim, F., Petheram, C., Marvanek, S., Ticehurst, C., Wallace, J., & Hasan, M. (2015). Impact of climate change on floodplain inundation and hydrological connectivity between wetlands and rivers in a tropical river catchment. *Hydrological Processes*, *30*(10), 1574–1593. doi: 10.1002/hyp.10714

Kaser, D., Graf, T., Cochand, F., McLaren, R., Therrien, R., & Brunner, P. (2014). Channel representation in physically based models coupling groundwater and surface water: Pitfalls and how to avoid them. *Groundwater*, *52*(6), 827–836.

Keizer, F., der Lee, G. V., Schot, P., Kardel, I., Barendregt, A., & Wassen, M. (2018, aug). Floodplain plant productivity is better predicted by particulate nutrients than by dissolved nutrients in floodwater. *Ecological Engineering*, *119*, 54–63.

Keizer, F., Schot, P., Okruszko, T., Chormanski, J., Kardel, I., & Wassen, M. (2014). A new look at the flood pulse concept: The (ir)relevance of the moving littoral in temperate zone rivers. *Ecological Engineering*, *64* (0), 85–99.

Kollet, S., Sulis, M., Maxwell, R. M., Paniconi, C., Putti, M., Bertoldi, G., ... Sudicky, E. (2017). The integrated hydrologic model intercomparison project, IH-MIP2: A second set of benchmark results to diagnose integrated hydrology and feedbacks. *Water Resources Research*, *53* (1), 867–890. doi: 10.1002/2016wr019191

Kotowski, W., Jabłońska, E., & Bartoszuk, H. (2013). Conservation management in fens: Do large tracked mowers impact functional plant diversity? *Biological Conservation*, *167*, 292-297. doi: 10.1016/j.biocon.2013.08.021

Kristensen, K. J., & Jensen, S. E. (1975). A model for estimating actual evapotranspiration from potential evapotranspiration. *Hydrology Research*, *6* (3), 170-188.

Krysanova, V., Donnelly, C., Gelfan, A., Gerten, D., Arheimer, B., Hattermann, F., & Kundzewicz, Z. W. (2018). How the performance of hydrological models relates to credibility of projections under climate change. *Hydrological Sciences Journal*, *63* (5), 696–720. doi: 10.1080/02626667.2018.1446214

Kundzewicz, Z., Krysanova, V., Benestad, R., Hov, Ø., Piniewski, M., & Otto, I. (2018, jan). Uncertainty in climate change impacts on water resources. *Environmental Science & Policy*, *79*, 1–8. doi: 10.1016/j.envsci.2017.10.008

Kundzewicz, Z. W., Krysanova, V., Dankers, R., Hirabayashi, Y., Kanae, S., Hattermann, F. F., ... Schellnhuber, H.-J. (2016). Differences in flood hazard projections in Europe - their causes and consequences for decision making. *Hydrological Sciences Journal*. doi: 10.1080/02626667.2016.1241398

Laranjeiras, T. O., Naka, L. N., Leite, G. A., & Cohn-Haft, M. (2021). Effects of a major Amazonian river confluence on the distribution of floodplain forest avifauna. *Journal of Biogeography*, *48* (4), 847–860. doi: 10.1111/jbi.14042

Lawler, J. J. (2009, apr). Climate change adaptation strategies for resource management and conservation planning. *Annals of the New York Academy of Sciences*, *1162* (1), 79–98. doi: 10.1111/j.1749-6632.2009.04147.x

Manh, N. V., Dung, N. V., Hung, N. N., Kummu, M., Merz, B., & Apel, H. (2015). Future sediment dynamics in the Mekong Delta floodplains: Impacts of hydropower development, climate change and sea level rise. *Global and Planetary Change*, *127*, 22–33. doi: 10.1016/j.gloplacha.2015.01.001

Marx, A., Kumar, R., Thober, S., Rakovec, O., Wanders, N., Zink, M., ... Samaniego, L. (2018, feb). Climate change alters low flows in europe under global warming of 1.5, 2, and 3c. *Hydrology and Earth System Sciences*, *22*(2), 1017–1032. doi: 10.5194/hess-22-1017-2018

Mason, D. C., Davenport, I. J., Neal, J. C., Schumann, G. J.-P., & Bates, P. D. (2012). Near real-time flood detection in urban and rural areas using high-resolution synthetic aperture radar images. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(8), 3041–3052. doi: 10.1109/tgrs.2011.2178030

McCabe, M., Franks, S., & Kalma, J. (2005). Calibration of a land surface model using multiple data sets. *Journal of Hydrology*, *302*(1-4), 209–222. doi: 10.1016/j.jhydrol.2004.07.002

McClain, M. E., Boyer, E. W., Dent, C. L., Gergel, S. E., Grimm, N. B., Groffman, P. M., ... Pinay, G. (2003). Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems*, *6*(4), 301–312. doi: 10.1007/s10021-003-0161-9

Meresa, H., Osuch, M., & Romanowicz, R. (2016, may). Hydro-meteorological drought projections into the 21-st century for selected Polish catchments. *Water*, *8*(5), 206. doi: 10.3390/w8050206

Mertes, L. A. K. (1997). Documentation and significance of the perirheic zone on inundated floodplains. *Water Resour. Res.*, *33*(7), 1749–1762.

Mezghani, A., Dobler, A., Haugen, J. E., Benestad, R. E., Parding, K. M., Piniewski, M., ... Kundzewicz, Z. W. (2017). CHASE-PL climate projection dataset over Poland - bias adjustment of EURO-CORDEX simulations. *Earth System Science Data*, *9*(2), 905–925. doi: 10.5194/essd-9-905-2017

Mohanty, M. P., & Simonovic, S. P. (2021). Fidelity of reanalysis datasets in floodplain mapping: Investigating performance at inundation level over large regions. *Journal of Hydrology*, *597*, 125757. doi: 10.1016/j.jhydrol.2020.125757

Moomaw, W. R., Chmura, G. L., Davies, G. T., Finlayson, C. M., Middleton, B. A., Natali, S. M., ... Sutton-Grier, A. E. (2018). Wetlands in a changing climate: Science, policy and management. *Wetlands*, *38*(2), 183–205. doi: 10.1007/s13157-018-1023-8

Mosner, E., Weber, A., Carambia, M., Nilson, E., Schmitz, U., Zelle, B., ... Horchler, P. (2015). Climate change and floodplain vegetation - future prospects for riparian habitat availability along the Rhine river. *Ecological Engineering*, *82*, 493–511. doi: 10.1016/j.ecoleng.2015.05.013

Murray-Hudson, M., Wolski, P., & Ringrose, S. (2006, nov). Scenarios of the impact of local and upstream changes in climate and water use on hydro-ecology in the Okavango Delta, Botswana. *Journal of Hydrology*, *331*(1-2), 73–84. doi: 10.1016/j.jhydrol.2006.04.041

Natho, S., Tschikof, M., Bondar-Kunze, E., & Hein, T. (2020). Modeling the effect of enhanced lateral connectivity on nutrient retention capacity in large river floodplains: How much connected floodplain do we need? *Frontiers in Environmental Science*, *8*. doi: 10.3389/fenvs.2020.00074

Nogueira, G. E. H., Schmidt, C., Partington, D., Brunner, P., & Fleckenstein, J. H. (2022). Spatiotemporal variations in water sources and mixing spots in a riparian zone. *Hydrology and Earth System Sciences*, *26*(7), 1883–1905. doi: 10.5194/hess-26-1883-2022

O'Connor, B. L., Hondzo, M., Dobraca, D., LaPara, T. M., Finlay, J. C., & Brezonik, P. L. (2006). Quantity-activity relationship of denitrifying bacteria and environmental scaling in streams of a forested watershed. *Journal of Geophysical Research: Biogeosciences*, *111*(G4). doi: 10.1029/2006jg000254

Osuch, M., Lawrence, D., Meresa, H. K., Napiorkowski, J. J., & Romanowicz, R. J. (2016, aug). Projected changes in flood indices in selected catchments in Poland in the 21st century. *Stochastic Environmental Research and Risk Assessment*, *31*(9), 2435–2457. doi: 10.1007/s00477-016-1296-5

Paiva, R. C. D., Collischonn, W., & Buarque, D. C. (2012). Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon. *Hydrological Processes*, *27*(3), 333–346. doi: 10.1002/hyp.8425

Pałczyński, A. (1984). Natural differentiation of plant communities in relation to hydrological conditions of the biebrza valley. *Polish Ecological Studies*, *10*, 347–385.

Park, E., Ho, H. L., Binh, D. V., Kantoush, S., Poh, D., Alcantara, E., . . . Lin, Y. N. (2022). Impacts of agricultural expansion on floodplain water and sediment budgets in the Mekong River. *Journal of Hydrology*, *605*, 127296. doi: 10.1016/j.jhydrol.2021.127296

Park, E., & Latrubesse, E. M. (2015). Surface water types and sediment distribution patterns at the confluence of mega rivers: The Solimões-Amazon and Negro Rivers junction. *Water Resources Research*, *51*(8), 6197–6213.

Partington, D., Brunner, P., Frei, S., Simmons, C. T., Werner, A. D., Therrien, R., . . . Fleckenstein, J. H. (2013). Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment. *Water Resour. Res.*, *49*(9), 5501–5519. Retrieved from `http://dx.doi.org/10.1002/wrcr.20405`

Partington, D., Brunner, P., Simmons, C., Therrien, R., Werner, A., Dandy, G., & Maier, H. (2011). A hydraulic mixing-cell method to quantify the groundwater component of streamflow within spatially distributed fully integrated surface water-groundwater flow models. *Environmental Modelling & Software*, *26*(7), 886–898.

Partington, D., Knowling, M. J., Simmons, C. T., Cook, P. G., Xie, Y., Iwanaga, T., & Bouchez, C. (2020). Worth of hydraulic and water chemistry observation data in terms of the reliability of surface water-groundwater exchange flux predictions under varied flow conditions. *Journal of Hydrology*, *590*, 125441. doi: 10.1016/j.jhydrol.2020.125441

Perra, E., Piras, M., Deidda, R., Paniconi, C., Mascaro, G., Vivoni, E. R., ... Meyer, S. (2018). Multimodel assessment of climate change-induced hydrologic impacts for a Mediterranean catchment. *Hydrology and Earth System Sciences*, *22*(7), 4125–4143. doi: 10.5194/hess-22-4125-2018

Piniewski, M., Szcześniak, M., Kardel, I., Chattopadhyay, S., & Berezowski, T. (2021). G2dc-pl+: a gridded 2 km daily climate dataset for the union of the Polish territory and the Vistula and Odra basins. *Earth System Science Data*, *13*(3), 1273–1288. doi: 10.5194/essd-13-1273-2021

Piniewski, M., Szcześniak, M., Kundzewicz, Z. W., Mezghani, A., & Hov, Ø. (2017). Changes in low and high flows in the Vistula and the Odra basins: Model projections in the European-scale context. *Hydrological Processes*, *31*(12), 2210–2225. doi: 10.1002/hyp.11176

Polish Geological Institute. (2014). *Ikar geoportal.* Retrieved from `ikar.pig.gov.pl`

Politti, E., Egger, G., Angermann, K., Rivaes, R., Blamauer, B., Klösch, M., ... Habersack, H. (2014). Evaluating climate change impacts on Alpine floodplain vegetation. *Hydrobiologia*, *737*(1), 225–243. doi: 10.1007/s10750-013-1801-5

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., ... Wisser, D. (2013). Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences*, *111*(9), 3262–3267. doi: 10.1073/pnas.1222473110

Racchetti, E., Bartoli, M., Soana, E., Longhi, D., Christian, R. R., Pinardi, M., & Viaroli, P. (2011). Influence of hydrological connectivity of riverine wetlands on nitrogen removal via denitrification. *Biogeochemistry*, *103*(1), 335-354.

Ramteke, G., Singh, R., & Chatterjee, C. (2020). Assessing impacts of conservation measures on watershed hydrology using MIKE SHE model in the face of climate change. *Water Resources Management*, *34*(13), 4233–4252. doi: 10.1007/s11269-020-02669-3

Rawson, H. M., & Macpherson, H. G. (n.d.). *Irrigated wheat.* FAO. Retrieved from www.fao.org/3/X8234E/X8234E00.htm

Rientjes, T., Muthuwatta, L., Bos, M., Booij, M., & Bhatti, H. (2013). Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *Journal of Hydrology*, *505*, 276–290. doi: 10.1016/j.jhydrol.2013.10.006

Roudier, P., Andersson, J. C. M., Donnelly, C., Feyen, L., Greuell, W., & Ludwig, F. (2015, nov). Projections of future floods and hydrological droughts in Europe under a +2c global warming. *Climatic Change*, *135*(2), 341–355. doi: 10.1007/s10584-015-1570-4

Scaroni, A. E., Nyman, J. A., & Lindau, C. W. (2011). Comparison of denitrification characteristics among three habitat types of a large river floodplain: Atchafalaya River Basin, Louisiana. *Hydrobiologia*, *658*(1), 17-25.

Schneider, C., Laizé, C. L. R., Acreman, M. C., & Flörke, M. (2013). How will climate change modify river flow regimes in Europe? *Hydrology and Earth System Sciences*, *17*(1), 325–339. doi: 10.5194/hess-17-325-2013

Scott, D. T., Keim, R. F., Edwards, B. L., Jones, C. N., & Kroes, D. E. (2014). Floodplain biogeochemical processing of floodwaters in the Atchafalaya River Basin during the Mississippi River flood of 2011. *Journal of Geophysical Research: Biogeosciences*, *119*(4), 537–546. (2013JG002477)

Sebben, M. L., Werner, A. D., Liggett, J. E., Partington, D., & Simmons, C. T. (2013). On the testing of fully integrated surface-subsurface hydrological models. *Hydrological Processes*, *27*(8), 1276-1285.

Shewchuk, J. (1996). Triangle: Engineering a 2d quality mesh generator and delaunay triangulator. In M. Lin & D. Manocha (Eds.), *Lecture notes in computer science* (Vol. 1148, p. 203-222). Springer Berlin Heidelberg.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., … Wyszyński, P. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876-2908. doi: 10.1002/qj.3598

Statistics Poland. (2021). *Population by sex, feminization rate, population density. Teritorial units: Podlaskie and Warminsko-Mazurskie. as of day 31 XII 2021. accessed: 18 XII 2022.* Retrieved from swaid.stat.gov.pl

Suliga, J., Chormański, J., Szporak-Wasilewska, S., Kleniewska, M., Berezowski, T., van Griensven, A., & Verbeiren, B. (2015). Derivation from the Landsat 7 NDVI and ground truth validation of LAI and interception storage capacity for wetland ecosystems in Biebrza Valley, Poland. In C. M. U. Neale & A. Maltese

(Eds.), *Remote sensing for agriculture, ecosystems, and hydrology XVII* (Vol. 9637, p. 96371Z). SPIE. doi: 10.1117/12.2194975

Sulis, M., Paniconi, C., Marrocu, M., Huard, D., & Chaumont, D. (2012). Hydrologic response to multimodel climate output using a physically based model of groundwater/surface water interactions. *Water Resources Research*, *48*(12). doi: 10.1029/2012wr012304

Sulis, M., Paniconi, C., Rivard, C., Harvey, R., & Chaumont, D. (2011). Assessment of climate change impacts at the catchment scale with a detailed hydrological model of surface-subsurface interactions and comparison with a land surface model. *Water Resources Research*, *47*(1). doi: 10.1029/2010wr009167

Thompson, J. R., Crawley, A., & Kingston, D. G. (2016). GCM-related uncertainty for river flows and inundation under climate change: the inner niger delta. *Hydrological Sciences Journal*, *61*(13), 2325–2347. doi: 10.1080/02626667.2015.1117173

Thompson, J. R., Gavin, H., Refsgaard, A., Sørenson, H. R., & Gowing, D. J. (2008). Modelling the hydrological impacts of climate change on UK lowland wet grassland. *Wetlands Ecology and Management*, *17*(5), 503–523. doi: 10.1007/s11273-008-9127-1

Tomasek, A., Kozarek, J. L., Hondzo, M., Lurndahl, N., Sadowsky, M. J., Wang, P., & Staley, C. (2017). Environmental drivers of denitrification rates and denitrifying gene abundances in channels and riparian areas. *Water Resources Research*, *53*(8), 6523–6538. doi: 10.1002/2016wr019566

Veijalainen, N., Lotsari, E., Alho, P., Vehviläinen, B., & Käyhkö, J. (2010). National scale assessment of climate change impacts on flooding in Finland. *Journal of Hydrology*, *391*(3-4), 333–350. doi: 10.1016/j.jhydrol.2010.07.035

von Gunten, D., Wohling, T., Haslauer, C., Merchan, D., Causape, J., & Cirpka, O. (2014). Efficient calibration of a distributed pde-based hydrological model using grid coarsening. *Journal of Hydrology*, *519*, 3290–3304.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., & Schewe, J. (2013). The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework. *Proceedings of the National Academy of Sciences*, *111*(9), 3228–3232. doi: 10.1073/pnas.1312330110

Wassen, M. J., Okruszko, T., Kardel, I., Chormanski, J., Swiatek, D., Mioduszewski, W., ... Meire, P. (2006). Eco-hydrological functioning of the biebrza wetlands: Lessons for the conservation and restoration of deteriorated wetlands rid c-7306-2008. *Wetlands: Functioning, Biodiversity Conservation, and Restoration*, *191*, 285–310.

Wen, L., Macdonald, R., Morrison, T., Hameed, T., Saintilan, N., & Ling, J. (2013). From hydrodynamic to hydrological modelling: Investigating long-term hydrological regimes of key wetlands in the Macquarie Marshes, a semi-arid lowland floodplain in Australia. *Journal of Hydrology*, *500*, 45–61. doi: 10.1016/j.jhydrol.2013.07.015

Wösten, J., Lilly, A., Nemes, A., & Bas, C. L. (1999). Development and use of a database of hydraulic properties of European soils. *Geoderma*, *90*(3-4), 169–185.

Yuan, X., Lu, Y., Jiang, L., Liang, S., Jiang, Y., & Xiao, F. (2021). Runoff responses to climate change in China's Buyuan River basin. *River Research and Applications*, *37*(8), 1134–1144. doi: 10.1002/rra.3785

Zhang, Y., Wang, Y., Chen, Y., Liang, F., & Liu, H. (2019). Assessment of future flash flood inundations in coastal regions under climate change scenarios—a case study of Hadahe River basin in northeastern China. *Science of The Total Environment*, *693*, 133550. doi: 10.1016/j.scitotenv.2019.07.356

Zulkafli, Z., Buytaert, W., Manz, B., Rosas, C. V., Willems, P., Lavado-Casimiro, W., . . . Santini, W. (2016). Projected increases in the annual flood pulse of the Western Amazon. *Environmental Research Letters*, *11*(1), 014013. doi: 10.1088/1748-9326/11/1/014013

1412 # Supplement

Table S1. Forcing data sources used in this study. The EURO-CORDEX data were available for different RCP simulations. Each RCP simulation period was 2006-2100 followed by historical simulations of different lengths. Observations indicated here were used only for hydrological model forcing, not for bias correction.

Data (Institute-GCM)	Period	RCP
CNRM-CERFACS-CNRM-CM5	1970-2100	4.5, 8.5
ICHEC-EC-EARTH	1970-2100	2.6, 4.5, 8.5
MOHC-HadGEM2-ES	1970-2100	2.6, 4.5, 8.5
MPI-M-MPI-ESM-LR	1970-2100	2.6, 4.5, 8.5
NCC-NorESM1-M	1970-2100	2.6, 4.5, 8.5
CCCma-CanESM2	1951-2100	4.5, 8.5
CSIRO-QCCCE-CSIRO-Mk3-6-0	1951-2100	4.5, 8.5
IPSL-IPSL-CM5A-MR	1951-2100	4.5, 8.5
MIROC-MIROC5	1951-2100	2.6, 4.5, 8.5
NOAA-GFDL-GFDL-ESM2M	1951-2100	4.5, 8.5
Observations	2005-2019	-
20CR	1880-2005	-

Table S2. Data sources for hydrological validation. Hydrological variables are H - water levels in rivers, Q - discharge, and G - groundwater head. Periods of missing data are not indicated in this table. Unreferenced data sources are available upon request from the authority.

Data source	Variables	Stations	Period	Frequency
Russian hydrological yearbook (Anonymous, 1912)	H	Osowiec	1881-1910	daily
Polish hydrological yearbook Anonymous (1932, 1970, 1980)	H, Q	Q and H: Burzyn, Osowiec, Q: Czachy, Rudzki, Sztabin	1918-1980	daily
IMGW unpublished data archive	H	Osowiec	1924	daily
IMGW Public data repository Anonymous (2019)	H, Q	Q and H: Burzyn, Osowiec, Q: Czachy, Rudzki, Sztabin	1951-2019	daily
Biebrza National Park database	G	41 groundwater wells in the national park	1994-2019	10-days (median)
Household wells measurements	G	2 wells in the Biebrza catchment	1999-2002	once per year

Table S3. Calibration parameters ranges with their constrains and transformations

Porous media parameters	units	class	min.	max.	Constrains and transformation
Van Genuchten model inverse of the air-entry pressure head, α	m^{-1}	glacial till	0.008	0.03	-
		peat	1.2	2.6	
		sand	0.008	0.03	
Van Genuchten model pore-size distribution index, β	-	glacial till	1.3	3	-
		peat	1.3	1.65	
		sand	1.3	3	
Saturated hydraulic conductivity	ms^{-1}	glacial till	1.00E-07	5.00E-03	Logarithmic, with base=10, transformation.
		peat	1.00E-07	5.00E-04	
		sand	1.00E-07	5.00E-03	
Porosity	-	glacial till	0.32	0.45	-
		peat	0.8	0.92	
		sand	0.32	0.45	

Table S4. Calibration parameters ranges with their constrains and transformations

Evapotranspiration parameters	units	class	min.	max.	Constrains and transformation
transpiration fitting parameter, c1	-	Ten vegetation classes	0.001	1.3	One parameter value was selected randomly [0-1] for all vegetation types and scaled using an inversion of maximum leaf area index for a given vegetation type.
Lower limit of soil saturation for transpiration, e1	-	upland, wetland	0.133	1	The evaporation limiting saturations: e1 and e2 parameters were derived simultaneously for each vegetation type from the gamma distribution using: $1 - (\mathrm{g}(p,s)/\mathrm{g}(1,s))$, where g is a function returning quantiles of gamma distribution, p is the probability of 0.05 for e1 and 0.6 for e1, and s [0-1] is shape parameter of gamma distribution provided during the calibration. The rate parameter of the gamma distribution is 1.
Upper limit of soil saturation for transpiration, e2	-	upland, wetland	0.08	0.951	
field capacity, fc	-	upland	0.3	0.951	The transpiration limiting saturations parameters: wp, fc, ox, and aox parameters were derived simultaneously for each vegetation type from the gamma distribution using: $1 - (\mathrm{g}(p,s)/\mathrm{g}(1,s))$, where g is a function returning quantiles of gamma distribution, p is the probability of 0.001 for wp, 0.05 for fc, 0.6 for ox and 0.99 for aox, and s [0-1] is shape parameter of gamma distribution provided during the calibration. The rate parameter of the gamma distribution is 1.
		wetland	0.3	0.87	
wilting point, wp	-	upland	0.09	0.41	-
		wetland	0.09	0.33	
oxic limit, ox	-	upland	0.46	1	-
		wetland	0.46	0.99	
anoxic limit, aox	-	upland, wetland	0.56	1	-

Table S5. Calibration parameters ranges with their constrains and transformations

Surface water flow parameter	units	class	min.	max.	Constrains and transformation
Manning roughness coefficient	$ms^{-\frac{1}{3}}$	Lower Biebrza	0.06	0.25	-
		Major rivers	0.015	0.05	
		Other rivers	0.02	0.05	
		Upland	0.015	0.05	
		Upper Biebrza	0.02	0.2	
		Flood-plain	0.02	0.2	
obstruction height	m	Major rivers	0.05	0.4	-
		Other	0.05	0.4	
		Flood-plain	0.01	0.4	

Table S6. Daily mean values of observations and bias-corrected 20CR and EURO-CORDEX data. A summary is presented for the period 1970-2005 except the PET, which was summarized for 1979-2005. The 20CR diff. row presents the observations subtracted from 20CR values. The EURO-CORDEX mean diff. row presents the mean difference of observations subtracted from each EURO-CORDEX simulations values.

Data source	Total precipitation [mm]		Snowfall [mm]		PET [mm]		Air temperature [K]	
	mean	sd	mean	sd	mean	sd	mean	sd
Observations	1.84	3.40	0.24	1.09	1.70	1.42	280.20	8.48
20CR	2.06	3.78	0.23	1.04	1.73	1.45	279.24	8.56
CNRM-CERFACS-CNRM-CM5	2.08	3.65	0.24	1.05	1.69	1.41	279.13	8.46
ICHEC-EC-EARTH	2.08	3.62	0.24	1.03	1.69	1.41	279.14	8.45
MOHC-HadGEM2-ES	2.06	3.52	0.24	1.02	1.70	1.42	279.18	8.45
MPI-M-MPI-ESM-LR	2.08	3.61	0.24	1.04	1.70	1.42	279.22	8.40
NCC-NorESM1-M	2.07	3.47	0.24	1.04	1.69	1.41	279.12	8.50
CCCma-CanESM2	2.00	3.69	0.23	1.01	1.71	1.42	279.25	8.76
CSIRO-QCCCE-CSIRO-Mk3-6-0	2.01	3.81	0.24	1.04	1.70	1.42	278.90	8.86
IPSL-IPSL-CM5A-MR	2.03	3.68	0.24	1.10	1.70	1.41	279.12	8.71
MIROC-MIROC5	1.99	3.67	0.24	1.09	1.71	1.42	278.92	8.78
NOAA-GFDL-GFDL-ESM2M	2.01	3.79	0.24	1.10	1.70	1.42	278.95	8.92
20CR diff.	0.22	0.39	0.00	-0.04	0.02	0.03	-0.96	0.08
EURO-CORDEX mean diff.	0.20	0.26	0.00	-0.04	0.00	0.00	-1.11	0.15

Table S7. Calibrated parameters for the best model.

Parameter	Units	Material	Value
Porous media parameters			
Van Genuchten model inverse of the air-entry pressure head, α	m^{-1}	glacial till	0.0136
		peat	2.054
		sand	0.025
Van Genuchten model pore-size distribution index, β	-	glacial till	1.735
		peat	1.535
		sand	2.632
Saturated hydraulic conductivity	ms^{-1}	glacial till	6.56E-07
		peat	4.52E-07
		sand	2.24E-03
Porosity	-	glacial till	0.36
		peat	0.86
		sand	0.39
Evapotranspiration parameters			
transpiration fitting parameter, c1	-	Ten vegetation classes	0.06 to 0.21
Lower limit of soil saturation for transpiration, e1	-	upland	0.996
		wetland	0.858
Upper limit of soil saturation for transpiration, e2	-	upland	0.889
		wetland	0.636
field capacity, fc	-	upland	0.922
		wetland	0.623
wilting point, wp	-	upland	0.376
		wetland	0.206
oxic limit, ox	-	upland	0.999
		wetland	0.846
anoxic limit, aox	-	upland	1
		wetland	0.936
Surface water flow parameter			
Manning roughness coefficient	$ms^{-\frac{1}{3}}$	Lower Biebrza	0.191
		Major rivers	0.042
		Other rivers	0.020
		Upland	0.019
		Upper Biebrza	0.152
		Floodplain	0.128
obstruction height	m	Major rivers	0.312
		Other	0.370
		Floodplain	0.064

Table S8. Error metrics for groundwater wells observations in the floodplain. RMSE / d.r. and bias / d.r. area RMSE and bias normalized to the observations data range (d.r.). Errors for individual wells in the middle and upper Biebrza basins are presented in Table S9.

Well	Period with observations	RMSE [m]	RMSE / d.r.	bias [m]	bias / d.r.
BPN116	1998-2010	0.28	21%	-0.15	-11%
BPN121	1998-2010, 2016-2019	0.25	19%	-0.06	-4%
BPN122	1998-2010, 2016-2019	0.33	22%	0.16	10%
BPN123	1998-2010	0.23	22%	-0.08	-8%
BPN124	2010-2019	0.25	28%	-0.11	-12%
BPN125	2010-2012, 2014-2019	0.27	22%	-0.14	-11%
BPN126	2010-2019	0.25	18%	-0.07	-5%
BPN167	2010-2019	0.31	26%	-0.19	-16%
BPN168	2010-2019	0.38	32%	-0.28	-24%
mean		0.28	23%	-0.10	-9%

Table S9. Error metrics for groundwater wells observations in the middle and upper basins. RMSE and bias are in the same units as indicated in the table, remaining metrics are dimensionless. RMSE / d.r. and bias / d.r. area RMSE and bias normalized to the observations data range (d.r.).

Well	Period with observations	RMSE	RMSE / d.r.	bias	bias / d.r.
BPN132	1998-2019	0.34	24%	0.05	4%
BPN133	1998-2019	0.34	23%	0.06	4%
BPN134	1998-2019	0.33	25%	0.10	8%
BPN135	1994-2015	0.28	25%	-0.14	-13%
BPN136	1994-2019	0.24	21%	0.04	4%
BPN137	1994-2019	0.29	23%	-0.09	-7%
BPN139	1994-2019	0.33	26%	-0.16	-12%
BPN140	1994-2019	0.43	33%	-0.32	-24%
BPN141	1994-2019	0.33	26%	-0.14	-11%
BPN142	1994-2019	0.35	26%	-0.13	-10%
BPN143	1994-2019	0.36	27%	-0.18	-13%
BPN144	1994-2019	0.40	26%	-0.23	-15%
BPN145	1994-2015	0.54	44%	-0.47	-38%
BPN147	1994-2019	0.60	43%	-0.55	-39%
BPN150	1996-2019	0.39	29%	-0.29	-21%
BPN152	1996-2019	0.79	50%	-0.73	-46%
BPN179	2010-2019	0.69	76%	-0.50	-55%
BPN182	1996-2019	0.83	58%	-0.78	-55%
BPN184	1996-2019	0.80	60%	-0.75	-56%
BPN186	1998-2019	0.71	48%	-0.63	-43%
BPN189	1996-2019	0.38	26%	-0.27	-19%
BPN190	1996-2019	0.31	20%	-0.15	-9%
BPN191	1994-2019	0.33	22%	-0.20	-13%
BPN207	2012-2017	0.38	33%	-0.32	-28%
BPN208	2012-2015	0.46	55%	-0.42	-50%
BPN209	2012-2017	0.34	34%	-0.26	-25%
BPN210	2012-2015	0.52	60%	-0.48	-55%
BPN211	2012-2017	0.51	47%	0.48	44%
BPN213	2012-2017	0.27	22%	-0.16	-13%
Middle basin mean		0.44	36%	-0.26	-21%
BPN155	1998-2019	0.59	37%	-0.45	-28%
BPN156	1998-2019	0.45	36%	-0.36	-29%
BPN158	1998-2019	0.32	29%	-0.17	-16%
Upper basin mean		0.46	34%	-0.33	-24%

Table S10. Statistics for the daily water levels and discharge at the Burzyn station in the 1970-2005 period, when both 20CR, and EURO-CORDEX forcing data to overlap with observations. The 20CR diff. row presents the observations subtracted from values simulated using model forced with 20CR data. The EURO-CORDEX mean diff. row presents the mean difference of observations subtracted from values simulated using models forced with EURO-CORDEX data.

Source	Water level [m amsl]		Discharge [m^3s^{-1}]	
	mean	sd	mean	sd
Observations	101.36	0.62	38.07	31.91
20CR	101.41	0.46	40.35	33.95
CCCma-CanESM2	101.42	0.46	40.88	31.30
CNRM-CERFACS-CNRM-CM5	101.50	0.44	45.51	30.93
CSIRO-QCCCE-CSIRO-Mk3-6-0	101.40	0.40	36.78	26.19
ICHEC-EC-EARTH	101.47	0.42	42.63	30.71
IPSL-IPSL-CM5A-MR	101.48	0.40	42.25	28.46
MIROC-MIROC5	101.40	0.45	38.73	29.85
MOHC-HadGEM2-ES	101.50	0.45	45.71	32.43
MPI-M-MPI-ESM-LR	101.53	0.40	46.85	31.58
NCC-NorESM1-M	101.51	0.40	44.72	28.74
NOAA-GFDL-GFDL-ESM2M	101.46	0.40	40.59	28.13
20CR diff.	0.05	-0.16	2.28	2.04
EURO-CORDEX mean diff.	0.11	-0.20	4.40	-2.08

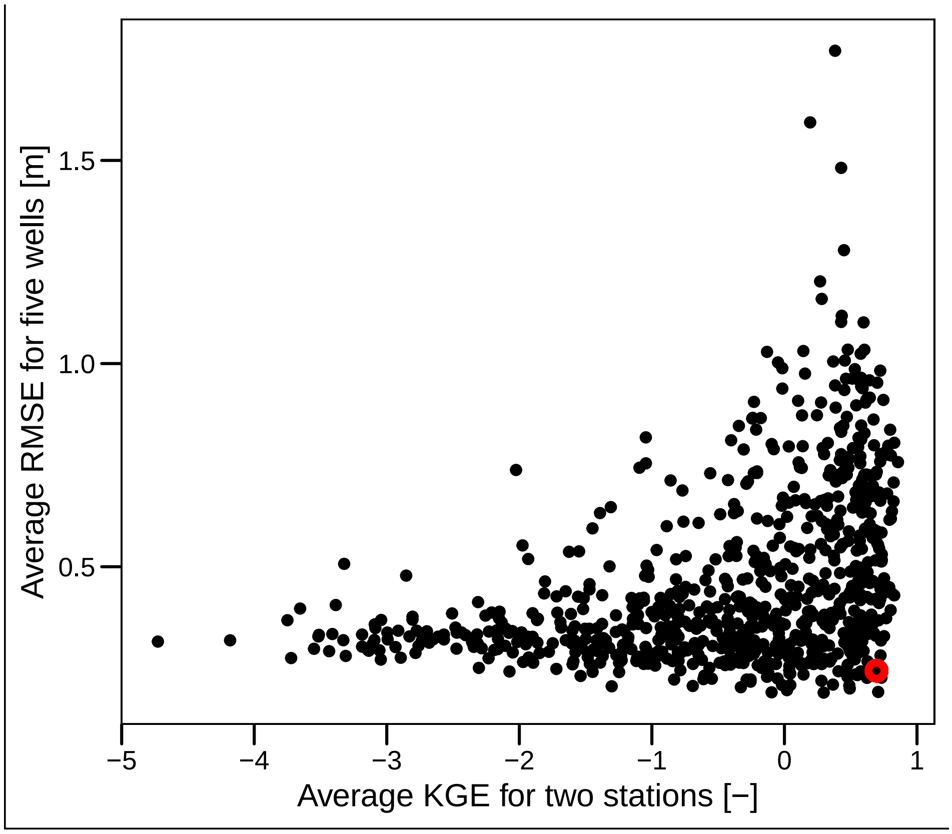

Figure S1. Average RMSE for five groundwater wells and average KGE for two stations calculated for 800 calibration runs. The red point indicates the selected model.

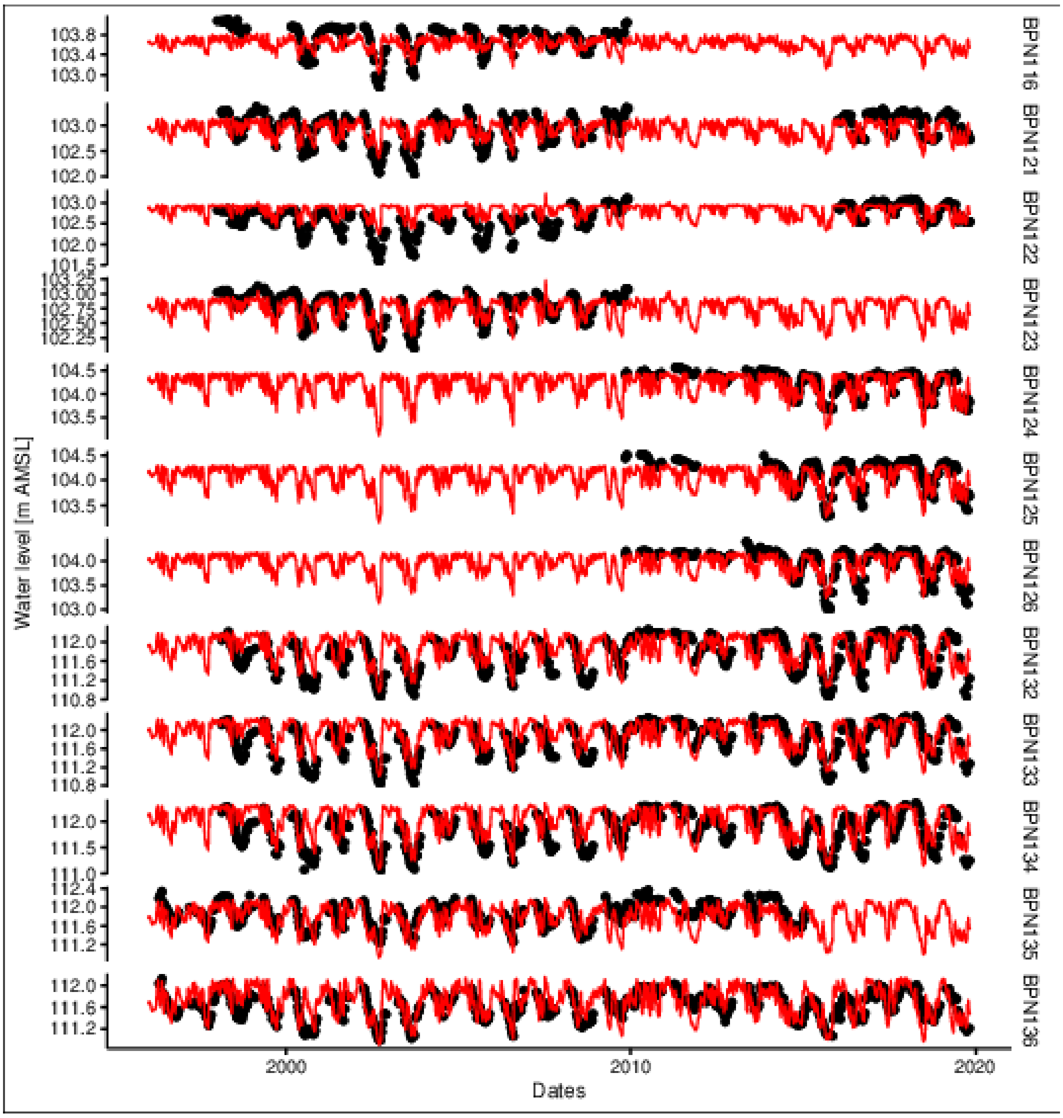

Figure S2. A Simulated (black) and observed (red) water levels for groundwater wells.

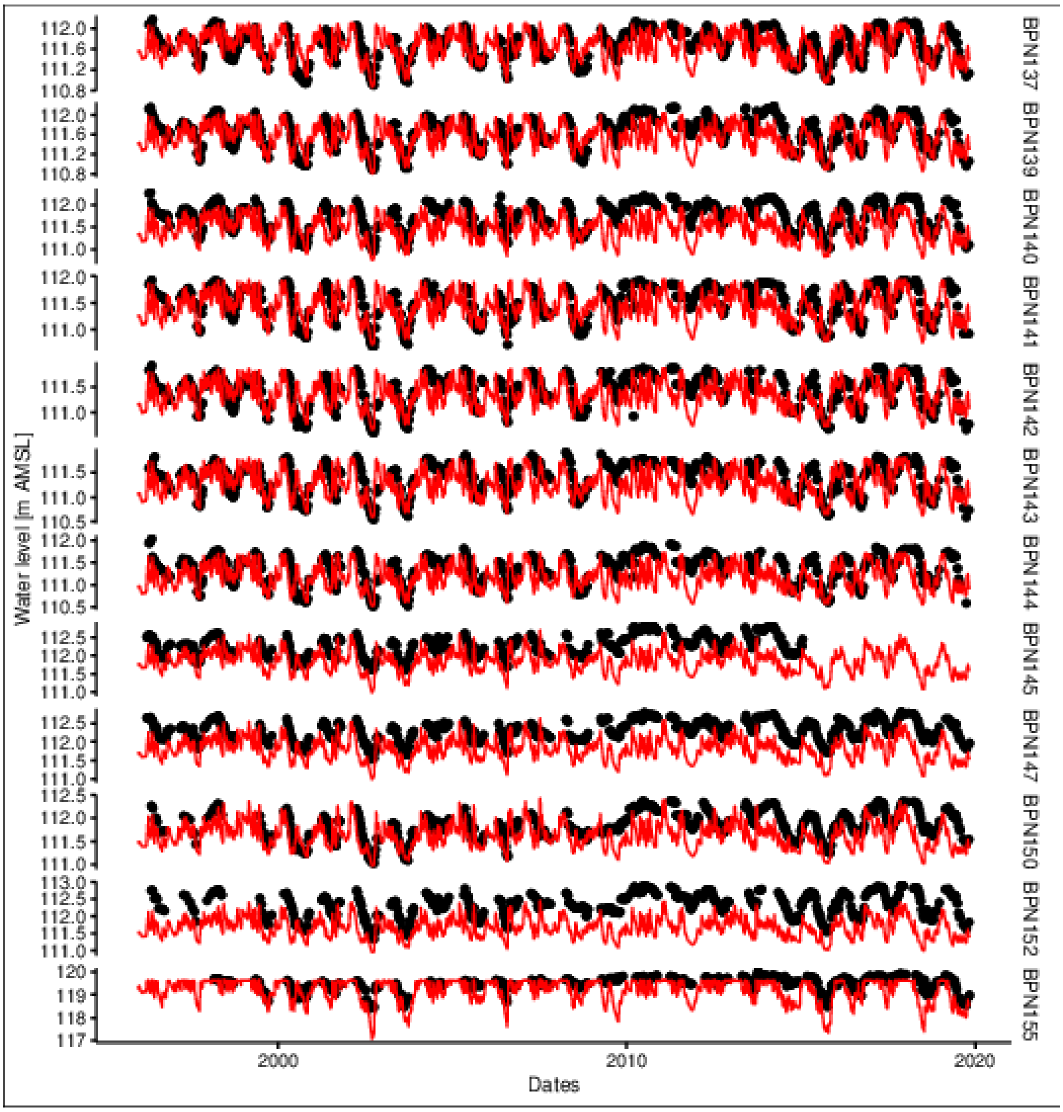

Figure S3. A Simulated (black) and observed (red) water levels for groundwater wells.

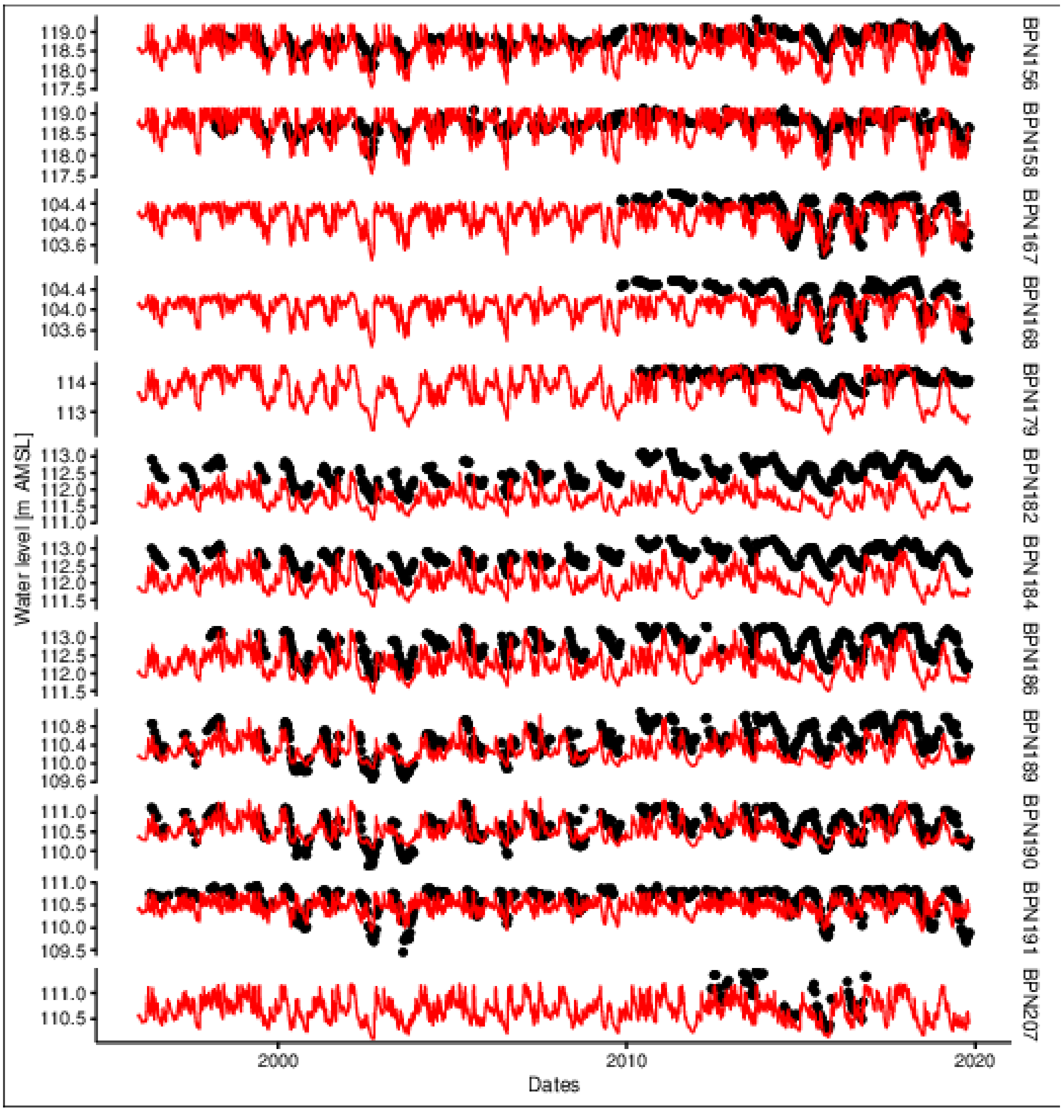

Figure S4. A Simulated (black) and observed (red) water levels for groundwater wells.

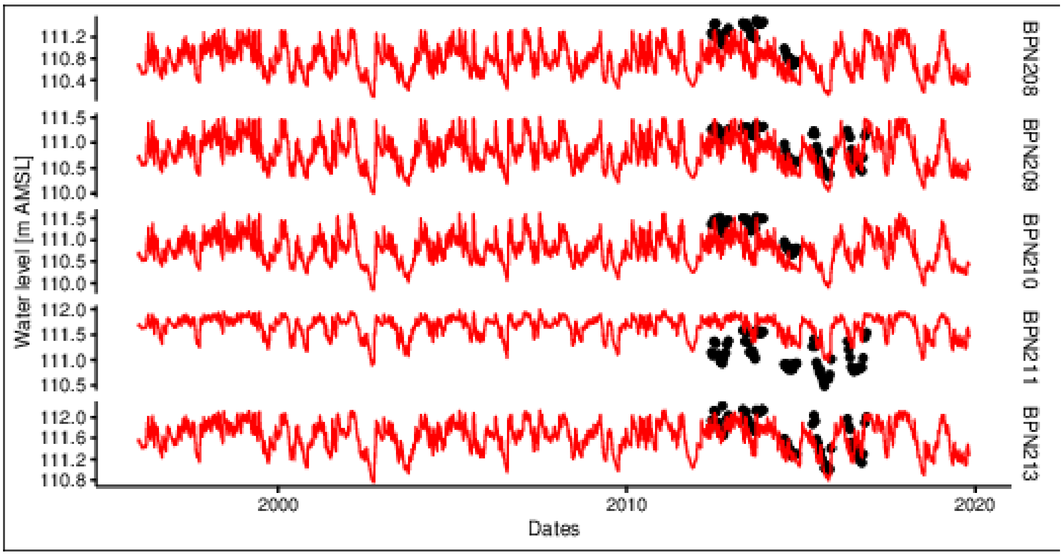

Figure S5. A Simulated (black) and observed (red) water levels for groundwater wells.

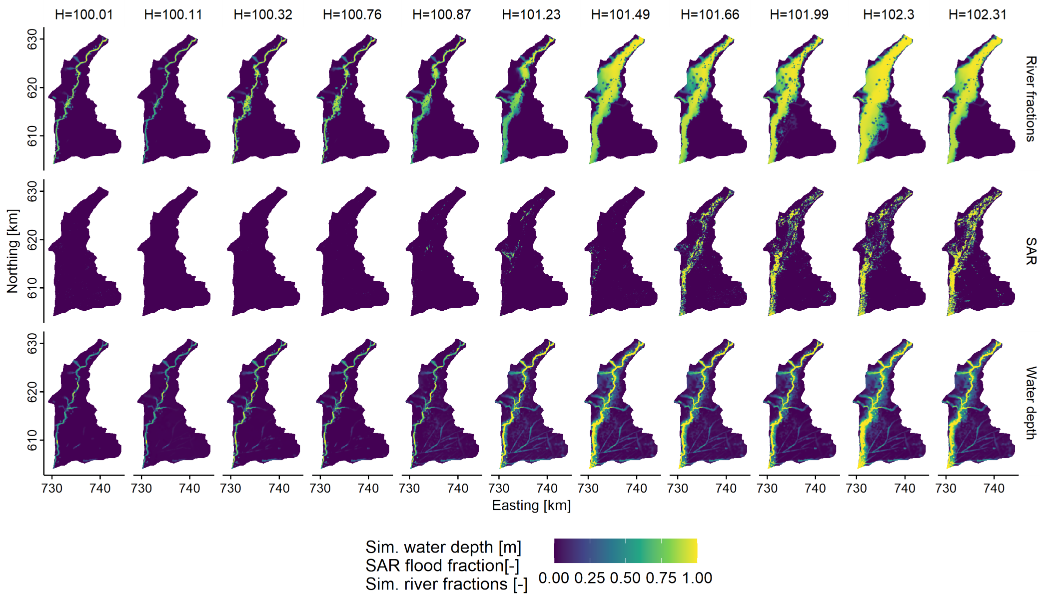

Figure S6. Flooding extent from remote sensing data-set (SAR), simulated HGS surface water depth, and HMC river water fractions for 11 increasing outlet water levels. Water depths > 1 m were plotted as equal to 1 m in this plot to have consistency in the color scale.