https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions

Karst spring discharge modeling based on deep learning using spatially distributed input data

Andreas Wunsch[1], Tanja Liesch[1], Guillaume Cinkus[2], Nataša Ravbar[3], Zhao Chen[4], Naomi Mazzilli[5], Hervé Jourde[2], and Nico Goldscheider[1]

[1]Karlsruhe Institute of Technology (KIT), Institute of Applied Geosciences, Hydrogeology, Kaiserstr. 12, 76131 Karlsruhe, Germany
[2]HydroSciences Montpellier (HSM), Université de Montpellier, CNRS, IRD, 34090 Montpellier, France
[3]ZRC SAZU, Karst Research Institute, Titov trg 2, 6230 Postojna, Slovenia
[4]Environmental Resources Management, Siemensstr. 9, 63263 Neu-Isenburg, Germany
[5]UMR 1114 EMMAH (AU-INRAE), Université d'Avignon, 84000 Avignon, France

Correspondence: Andreas Wunsch (andreas.wunsch@kit.edu)

Abstract. Despite many existing approaches, modeling karst water resources remains challenging and often requires solid system knowledge. Artificial Neural Network approaches offer a convenient solution by establishing a simple input-output relationship on their own. However, in this context, temporal and especially spatial data availability is often an important constraint, as usually no or few climate stations within a karst spring catchment are available. Hence spatial coverage is often unsatisfying and can introduce severe uncertainties. To avoid these problems, we use 2D-Convolutional Neural Networks (CNN) to directly process gridded meteorological data followed by a 1D-CNN to perform karst spring discharge simulation. We investigate three karst spring catchments in the Alpine and Mediterranean region with different meteorologic-hydrological characteristics and hydrodynamic system properties. We compare our 2D-models both to existing modeling studies in these regions and to 1D-models, which use climate station data, as it is common practice. Our results show that our models are excellently suited to model karst spring discharge and rival the simulation results of existing approaches in the respective areas. The 2D-models learn relevant parts of the input data and by performing a spatial input sensitivity analysis we can further show their potential for karst catchment localization and delineation.

1 Introduction

Karst aquifers and karst springs are crucial for freshwater supply in many regions and 9% of the global population partly or fully rely on karst water resources (Stevanović, 2019). Karst systems in general are characterized by high heterogeneity due to the at least in large parts unknown conduit network, which controls the highly variable groundwater flow. This makes modeling difficult, hence a large variety of different approaches exists (Jeannin et al., 2021). Most of them require a certain level of background knowledge about the system in order to achieve high quality results. In contrary, deep learning approaches propose

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

a convenient possibility of modeling by being able to establish a simple input-output relationship on their own, without detailed system knowledge necessary. Even though Artificial Neural Networks (ANN) are not a standard method in karst modeling yet, different types of ANNs have been applied in modeling karst water resources for quite a long time. In fact the study of Johannet et al. (1994) was even one of the first applications of ANNs in water related research. In this study, we apply Convolutional Neural Networks (CNN), which have been shown to be fast and reliable for the closely related application of groundwater level forecasting (Wunsch et al., 2021). We directly adapt this approach of using 1D-CNNs to the modeling of karst spring discharge time series. Although 1D-CNNs are frequently used to model time series in different contexts, besides some first rudimentary experiments (Jeannin et al., 2021), they have not yet been applied to model karst spring discharge.

Especially in direct comparison to concept based modeling, where shorter time series can be sufficient to achieve satisfying modeling results, one drawback of the 1D-CNN approach, as well as most other data-driven approaches, is the dependency on high data availability and quality. In many areas climate stations are often not available within the catchment of a spring, do not match the data availability of the discharge time series (period or temporal resolution), or are more distant and thus do not truly represent the events in the catchment itself. For this, gridded climate data can provide a solution, with reasonable time periods and temporal resolutions being available from various products. Especially for karst springs it is difficult to simply extract corresponding time series from the gridded data, since this requires knowledge of the spring catchment and usually the available grid cells do not exactly match the catchment either. Besides the modeling purpose, the delineation of karst catchments is generally important to sustainably exploit but also protect karst water resources by establishing protection zones accordingly. Malard et al. (2015) explain that only few generalizable methods for karst spring catchment delineation have been proposed. Instead, delineations usually rely on classical hydrogeological methods such as assessing geology, topography, hydrology, water balance, elaborate tracer tests and geophysical investigations. There has already been an attempt by Longenecker et al. (2017) to semi-automatically derive approximate catchment boundaries by correlating karst spring discharge events with global precipitation measurement (GPM) gridded data (NASA, 2016). The authors were able to achieve reasonable results with their method, but also noticed that they could not replace conventional methods.

Anderson and Radic (2021) have already applied gridded meteorological data to streamflow modeling in western Canada and used a 2D-CNN-LSTM model that is able to directly use spatial input data. They showed that such models learn the relevant parts of the large scale gridded input data for each local or regional streamflow by themselves. We adapt and extent this approach to karst spring discharge modeling by establishing a 2D-1D-CNN model, where the 2D-CNN processes the spatially distributed input data and the 1D-CNN performs the forecasting of the spring discharge time series in accordance with the 1D-CNN approach based on climate station data. Thus, we can now directly use gridded meteorological data to potentially overcome the common data availability problems at climate stations, without the need for a prior description of the catchment area. On the contrary, we investigate the potential of this approach for catchment localization and delineation based on a modified spatial input sensitivity analysis as performed in Anderson and Radic (2021). Deriving recharge areas based on rainfall-discharge event correlation as previously done by Longenecker et al. (2017) requires (i) heterogeneous rainfall at catchment scale, (ii) precipitation data with sufficient spatial resolution that capture this heterogeneity and (iii) a karst system

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

without too much dampening of the precipitation signals. This also applies to our proposed methodology, although we expect ANNs to be superior at least for points (i) and (iii).

In total, we explore the applicability of our proposed deep learning approaches with spatially distributed input data in modeling karst spring discharge in three different study areas. We further compare the results with 1D-CNN models using conventional climate station input data. As spatially distributed inputs we use either hourly ERA5-Land reanalysis data (Muñoz Sabater, 2019) or daily E-OBS data (Cornes et al., 2018), depending on the temporal resolution of spring discharge data. Finally, we explore the potential of the 2D-approach for karst spring catchment localization and delineation by investigating the spatial input sensitivity of the trained CNN models.

2 Data and Study Areas

2.1 Overview

In this study, we investigate three different karst springs: Aubach spring in the Hochifen-Gottesacker area in Austria (Fig. 1a), springs of Unica river in Slovenia (Fig. 1b) and Lez spring in southern France (Fig. 1c). All springs show different characteristics regarding relevant system properties (e.g. catchment size, complexity of the hydrological system), environmental conditions (e.g. dominant climate, anthropogenic forcing) and data availability (see also Table A1). Furthermore, all areas are well studied, existing data was easily accessible and except Unica, several previous modeling approaches are available for comparison.

2.2 Aubach Spring, Austria

Aubach spring is a major karst spring in the Hochifen-Gottesacker karst area in the northern Alps at the border between Germany and Austria. Southern border of the area is the Schwarzwasser valley, which geologically forms the contact zone between the Helvetic Säntis nappe in the north and sedimentary rocks of the Flysch zone in the south (Goldscheider, 2005). In the northern part the dominant karst formation is the Schrattenkalk formation, a cretaceous limestone with a thickness of about 100 m. This Schrattenkalk is structured in folds, which hydrogeologically form parallel sub-catchments (Fig. 1a) that contribute to different proportions to the several springs in the valley (Goldscheider, 2005; Chen and Goldscheider, 2014). In this study we focus on one large, non-permanent spring called Aubach spring (1080 m asl, discharge up to 10 m^3/s). The Hochifen-Gottesacker area is largely influenced by seasonal snow accumulation and melting in the elevated regions (>1,600 m asl), which is also clearly reflected in the discharge of Aubach spring by increased baseflow and daily snowmelt-induced variations, especially in the months of April to June. Earlier studies by Goldscheider (2005) and Chen and Goldscheider (2014) have identified one major catchment area of Aubach spring with approximately 9 km^2 (Fig. 1a), still, to smaller parts upstream catchments can also contribute to Aubach spring discharge. This applies also to the non-karstified Flysch area directly in the South (southernmost sub-catchment in Fig. 1a), where precipitation events are only relevant during low flow conditions. Then, the surface runoff from this area sinks into an upstream estavelle and contributes via an underground connection to the

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

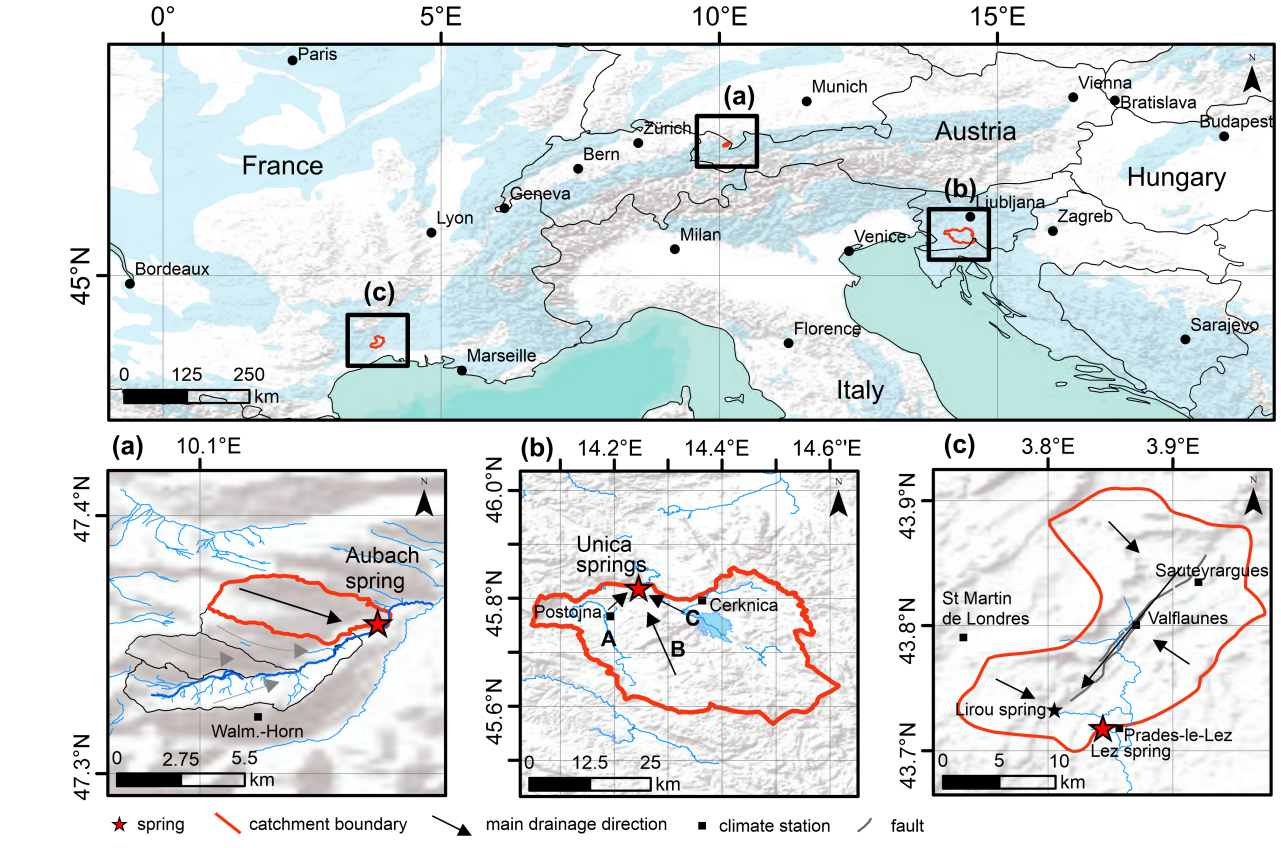


Figure 1. Overview of all three study areas, the simulated springs (red star) and their catchments (red lines). Black squares indicate locations of climate stations used for 1D-modeling (not all are shown in these maps), blue shadings in the upper map show karst areas based on WOKAM (Chen et al., 2017a) (a) Hochifen-Gottesacker karst area and Aubach spring, black lines depict minor contributing sub-catchments; (b) Unica river springs and Javorniki karst plateau (B); (c) Lez spring catchment, Lirou overflow spring (black star) and major fault Corconne-Les Matelles (grey line);

discharge of Aubach spring. During high flow conditions, the estavelle itself acts as an overflow spring and no contribution from surface runoff at Aubach spring occurs. Generally, the climate in the area can be described as cooltemperate and humid and the mean annual precipitation at the closest used climate station in this study (Walmendinger Horn) is about 2000 mm (2003-2019).

For this study Aubach spring is selected because of the good data availability. We use 8 years of hourly discharge data provided by the office of the federal state of Vorarlberg, division of water management, and we further use precipitation and temperature data from three surrounding climate stations: Oberstdorf, Walmendinger Horn (shown in Fig. 1a) and Diedamskopf. Additionally, due to the high importance of snow in the area, we run a snowmelt routine as preprocessing of the meteorological input data as described in Chen et al. (2018). This routine is a slightly modified version (after Hock, 1999) of the HBV hydro-

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

logical model snow routine (e.g. Bergström, 1975, 1995; Kollat et al., 2012; Seibert, 2000), which redistributes the precipitation time series in accordance with probable snow accumulation and snowmelt.

2.3 Unica Springs, Slovenia

The Unica springs (450 m asl) are located on the southern edge of a karst polje in SW Slovenia and are important from a biodiversity and water supply perspective. There are two permanent and several temporary springs that feed the Unica river. The joint discharge during 1989-2018 ranged from 1 to 90 m^3/s, while the mean discharge was 21 m^3/s (ARSO, 2020a). The springs are fed by three clearly distinguishable sub-catchments covering an area of about 820 km^2. The main recharge area is the highly karstified Javorniki plateau (up to 1,800 m; marked B on Fig. 1b), whose predominant lithology is Cretaceous rocks, mainly limestones, changing in places to dolomites and breccias. To a lesser extent, Jurassic and Palaeogene carbonate rocks are also present. The thickness of the unsaturated zone is estimated to be up to several hundred meters (Petrič et al., 2018, and references therein). To the east, the hydrology of the area is controlled by a strike-slip fault zone, along which a chain of karst poljes developed (between 500 and 700 m; marked C on Fig. 1b). Upper Triassic dolomites predominate, changing to Jurassic limestones and dolomites in the south and west, forming aquifers with fracture porosity, which in places have very low to moderate permeability, and in some parts a superficial river network forms. As the karst poljes follow each other in a downward series, they are connected in a common hydrological system with transitions between surface and groundwater flows and frequent flooding (Mayaud et al., 2019). In the west, the Pivka River Basin (between 500 and 700 m; marked A on Fig.1b) consists of poorly permeable Eocene flysch in the north, which conditions a surface river network. The southern part consists of Cretaceous and Jurassic carbonate rocks forming a shallow karst aquifer. Surface flow occur during high water levels, receiving additional water from intermittent springs on the western foothills of the Javorniki. The water flow of the sinking rivers in the subsurface from the A and C parts is clearly of the channel flow type. The springs were selected for this study because they drain a complex binary karst system of the so-called Classical Karst, they are well studied with a large amount of hydro-meteorological data and their hydrology is influenced by significant snow accumulation and melting. The catchment belongs to the moderate continental climate and is mostly covered with forests. For this study we use daily discharge data from the Unica-Hasberg gauging station (in the following called Unica) (ARSO, 2020a) and daily meteorological data from Postojna and Cerknica meteorological stations ranging from 1981 to 2018 (ARSO, 2020b). The meteorological stations (squares in Fig. 1b) are located on the western (Postojna) and eastern (Cerknica) part of the catchment, representing different climate regimes and are separated by the karst massif in between. For Postojna station the following parameters are available: precipitation (P), temperature (T), potential evapotranspiration (PET), relative humidity (rH), snow (S) and new snow (nS). For Cerknica station only P, S and nS are available. Average annual precipitation during 1989-2018 is about 1500 mm and on average, there are 33 days of snow cover per year in Postojna (530 m), while even longer snow cover is expected on the plateau.

2.4 Lez Spring, France

Our third study area is located 15 km north of Montpellier in France, within a large and complex karst system delimited by rivers and marly terrains. Eastern and western borders are the Vidourle and Hérault river valleys, northern and southern

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

borders are piezometric limits. At larger scale northern and southern boundaries are structural boundaries due to Cévennes and Montpellier faults, respectively. The dominant karst formations are Argovian to Kimmeridgian, and Berriasian massive limestones with 650 m to 1000 m thickness. Infiltration occurs mostly diffuse but also localized through fractures and sinkholes along the basin and through the major geologic fault of Corconne-Les Matelles in the northern part of the basin (indicated by a grey line in Fig. 1c).

The hydrogeological basin associated to the Lez spring has been estimated to be about 240 km^2 (Fig. 1c) on the basis of the hydrodynamic response to high discharge continuous pumping into the saturated zone of the aquifer (Thiéry and Bérard, 1983). However, the effective recharge catchment of the Lez spring, which corresponds to the extent of Jurassic limestone outcrops, has been estimated to be about 130 km^2 (Fleury et al., 2009; Jourde et al., 2014). The Lez karst aquifer is under anthropogenic pressure (i.e. aquifer exploitation for water supply) with pumping performed directly within the karst conduit. The discharge is measured at the spring pool and is regularly null during low water periods, when the pumping rate exceeds the natural spring discharge. Ecological water discharge towards the Lez river (160 L/s then 230 L/s after 2018) is ensured during such periods by a partial deviation of the pumped water to the river. Lirou spring (Fig. 1c) is the main of several overflow springs that activate during high flow periods (Jourde et al., 2014).

The Lez catchment is exposed to a Mediterranean climate, which is characterized by hot and dry summers, mild winters and wet autumns. Analyses by MeteoFrance show that on average 40% of the annual precipitation occurs between September and November with a high variability across years (Bicalho et al., 2012). The average annual rainfall rate for the 2008-2018 period is 904 mm.

For this study, the modeling is based on nearly 10 years of daily discharge data provided by SNO KARST (Jourde et al., 2018; SNO KARST, 2021). The temperature data is from the Prades-le-Lez climate station; we use, however, an interpolated precipitation data series that is derived from a weighted average of four rainfall stations (Fig. 1c) (similar to Fleury et al., 2009; Mazzilli et al., 2011), three of them being located on the Lez catchment (Prades-le-Lez, Valflaunès, Sauteyrargues). The fourth station (Saint-Martin-de-Londres) is located few kilometers west of the catchment. Interpolation is in principle possible in this area due to the existing topography; at the same time, interpolation based on Thiessen-polygons (compare Appendix B) also allows compensation for data gaps at single stations. We decided to apply this preprocessing, because all but Saint-Martin-de-Londres climate station show such gaps from time to time, which explains the benefit from including within-catchment precipitation.

2.5 Spatial Climate Data

Besides climate station data, we explored raster data from the E-OBS (Cornes et al., 2018), the ERA5-Land (Muñoz Sabater, 2019) and from the RADOLAN (DWD Climate Data Center (CDC)) as spatially distributed model inputs. E-OBS provides daily gridded meteorological data for Europe from 1950 to present, derived from in-situ observations, ERA5-Land provides hourly reanalysis data from 1981 to present. Both are available with a spatial resolution of 0.1° x 0.1° (approx. 8 km x 11 km for all study areas). Depending on the dataset, different sets of parameters are available. In case of E-OBS we initially provide our models with precipitation (P), mean, minimum and maximum temperature (T, Tmin, Tmax), relative humidity (rH) and

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

surface shortwave downwelling radiation (Rad). For ERA5-Land, where a significantly larger set of parameters is available, the following were used as initial inputs: total precipitation (P), 2m temperature (T), total evaporation (E), snowmelt (SMLT), snowfall (SF) and volumetric soil water of all four available layers (SWVL1: 0 - 7 cm, SWVL2: 7 - 28 cm, SWVL3: 28 - 100 cm, SWVL4: 100 - 289 cm). Relevant input parameters from both datasets are later selected through Bayesian optimization (see section 3.3). The spatial extent of the input data is chosen very generously for each spring, so that between 6 and 8 additional cells are available as input data around the respective catchments. This prevents a predefinition of the area that needs to be identified as relevant as well as reduces the influence of possible border effects due to the CNN approach using 3x3 filters (compare methodology section). The resolution of ERA5-Land and E-OBS data corresponds to the grid cell size shown in the catchment plots in Figures 1a-c, although each showing a slightly different absolute position of grid center points. We already see that the relation of grid resolution to catchment size varies strongly, with Aubach catchment being the smallest and Unica catchment being the largest. Depending on the temporal resolution of the available spring discharge measurements, we choose the spatial input data in accordance, thus E-OBS for Unica and Lez spring, ERA5-Land for Aubach spring.

Compared to the catchment size of Aubach spring (about 9 km^2), the spatial resolution (approx. 8 km x 11 km) of the gridded input data is extremely coarse. We therefore additionally explore a combination of ERA5-Land input parameters (except P) with radar based precipitation data (RADOLAN) that offers a spatial resolution of 1 km x 1 km (DWD Climate Data Center (CDC)). For each of the RADOLAN grid cells, the according parameter value of the ERA5-Land data is chosen, where the RADOLAN grid cell center lies in. This corresponds to the simplest form of downscaling. For this analysis, we reduce the spatial extent of the 2D-input data to save calculation time, but increase the number of cells compared to the ERA5 section around Aubach spring due to the higher resolution of the RADOLAN grid.

3 Methodology

3.1 Modeling Approach

In this study, we simulate karst spring discharge with deep learning models using meteorological input data. As proof of feasibility, we use meteorological data from surrounding climate stations. However, these stations often show limited data availability in terms of the number of parameters (mostly limited to precipitation and temperature, rarely more), the record length, as well as the sampling interval. Also, the spatial coverage and proximity is often unsatisfactory, which especially in mountainous regions can introduce a distinct error of parameters with high spatial variability such as precipitation.

Gridded meteorological data can offer a solution to these issues, as they usually provide good temporal coverage and resolution, a reasonable spatial resolution as well as a large-scale (e.g. continental or even global) availability. Further, especially reanalysis data include a larger parameter set. However, when the catchment of the spring is unknown, it remains unclear which cells to retrieve time series from when using 1D-time series as inputs. Based on our revised version of the approach of Anderson and Radic (2021), we demonstrate a solution by processing 2D-inputs and letting the model decide by itself, which parts of the input data are relevant to model the spring discharge.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

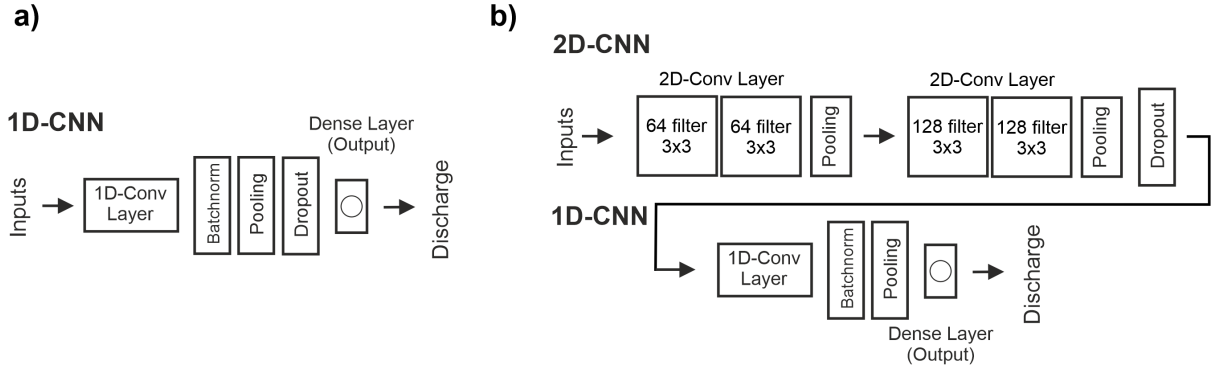


Figure 2. Model structures applied for modeling karst spring discharge based on climate station data (a) and gridded meteorological input data (b). Flatten layers are not displayed.

3.2 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (LeCun et al., 2015) are widely applied in several domains such as object recognition, image classification, and signal processing. The structure of most CNN models is based on the repetition of blocks that are made up of several layers, typically at least one convolutional layer followed by a pooling layer. The former matches the dimension of the input data (e.g. 2D for image alike data, 1D for sequences such as time series) and uses filters with a fixed size (receptive field) to produce feature maps of the input. The latter performs down-sampling of the produced feature maps and hence increases the density of information. A large variety of model structures based on such blocks, in combination with additional layers in between to prevent exploding gradients (e.g. batch normalization layers) or model overfitting (e.g. dropout layers) are possible; however CNNs usually end with one or several fully connected dense layers to produce a meaningful output.

In this study, we use 1D-CNN models as we have shown earlier that these are fast, reliable and excellently suited for modeling hydrogeological time series (Wunsch et al., 2021). Especially compared to LSTMs, which are often the method of choice, they are more stable and significantly faster, while showing similar performance for this specific application. We adapt this approach and perform karst spring discharge modeling in two different setups. One setup uses 1D-meteorological input data from surrounding climate stations, the other one uses gridded climate data as input and thus combines the 1D-CNN model with a time distributed 2D-CNN model to process the 2D-input data. The general model structure of both setups is shown in Figure 2. They basically use the same 1D-model except the position of the dropout layer. We use Bayesian hyperparameter optimization to select the 1D-filter number, batch-size and input sequence length of each model in both setups.

To reduce the dependency on the random initialization of the models, we use an ensemble with 10 members, each based on a different pseudo-random seed. Further, we implement Monte-Carlo dropout to estimate the model uncertainty from a distribution of 100 results for each calculation of each model performed in this study. Our models are implemented in Python 3.8 (van Rossum, 1995) and we use the following libraries and frameworks: Numpy (van der Walt et al., 2011), Pandas (McKin-

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

ney, 2010; Reback et al., 2020), Scikit-Learn (Pedregosa et al., 2011), Unumpy (Lebigot, 2010), Matplotlib (Hunter, 2007), BayesOpt (Nogueira, 2014), TensorFlow and its Keras API (Abadi et al., 2015; Chollet, 2015).

3.3 Model Calibration and Evaluation

We split the time series data for each site into four parts according to Table 1. Training epoch number and early stopping patience are varied manually for each model at each test site. Hyperparameters for the 1D-CNNs of both setups are optimized on the respective optimization set as stated above, maximizing the sum of Nash-Sutcliffe efficiency and R^2 (calculated as explained below). The number of optimization steps is also varied manually for each model and is always a trade-off between accuracy and computational costs. In case of many available input parameters we treat input parameter selection equally as a global optimization problem and use Bayesian optimization to simultaneously select a proper set of input parameters and hyperparameters. Thus, input optimization is used for each 2D-model, as ERA5-Land and E-OBS offer several different meteorological parameters, as well as to the 1D-model of Unica springs. For Lez spring and Aubach spring, only a smaller input parameter set is available and hence fully used. For all models we offer an additional input (Tsin), which is a sinus curve fitted to the temperature data. This parameter can provide the model with information on seasonality and on the current position in the annual cycle (Kong-A-Siou et al., 2014). Precipitation is the only parameter that is not optimized but fixed as input, because it has undoubtedly the most important influence on the discharge of a karst spring.

Table 1. Time series splitting schemes for all study areas.

	Time Interval	Training	Validation	Optimization	Testing
Aubach spring	Hourly	2012-2017	2018	2019	2020
Unica spring	Daily	1981-2012	2013+2014	2015+2016	2017+2018
Lez spring	Daily	2008-2016	2017	2018	2019

We calculate several metrics to evaluate the performance of our models: Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), squared Pearson r (R^2), root mean squared error (RMSE), Bias (Bias) as well as Kling-Gupta-Efficiency (KGE) (Gupta et al., 2009). For squared Pearson r we use the notation of the coefficient of determination (R^2), because we compare the linear fit between simulated and observed discharge, thus of a simple linear model, which makes them equal in this case.

3.4 Spatial Input Sensitivity and Catchment Delineation

Anderson and Radic (2021) show in their study that when using 2D-input data for CNNs for streamflow modeling, there is physical meaning in the way that spatial fractions of the input data are learned to be more important than others. We modify this approach and transfer it to karst spring modeling, where we demonstrate that this approach has even some potential for karst spring catchment delineation, which to date is dependent on elaborate tracer tests.

We use the Gaussian spatial perturbation approach from Anderson and Radic (2021) and modify it so that only a single channel (input parameter, e.g. precipitation) is perturbed at a time. For details of this approach we refer to the original study.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

In short it works by perturbing spatial fractions of the input data by using a 2D-Gaussian curve. The perturbed data is passed through the trained model to determine the resulting error from the specific perturbation. In this way, after many repetitions, heat maps are created that show how sensitive the trained model is to perturbations of certain areas of the input data. In the case of karst spring modeling we perturb only single channels, instead of all channels at once as in the original approach, for the following reasons. We use some very different input parameters and ideally want to separate the individual influence of each of these channels. Also these parameters show a very different spatial heterogeneity and variability, which means that the model can learn the importance for example of temperature from a large area, as long as it favors a lower simulation error, while for precipitation a smaller area will most certainly have a higher influence on the spring discharge and therefore has to be learned more accurately. Furthermore we mainly focus on the precipitation channel for catchment localization and delineation, because on the one hand we have a high spatial variability and on the other hand it is undoubtedly the major parameter that defines the catchment in karst areas.

4 Results and Discussion

4.1 Aubach Spring

Figure 3a shows the simulation results of the 1D-CNN model for the test period 2020, using only available climate station data. NSE and R^2 values both are 0.74, KGE is 0.79. Overall, this is a satisfying fit of a complex time series. We observe that peaks in winter and spring are underestimated but the snowmelt period, clearly visible by increased baseflow and daily variations from April to June, is nicely fitted, as well as the following summer peaks. A short series of discharge peaks in the end of September/beginning of October is not captured. We assume that these were caused by small-scale precipitation events that are not represented in the data of the climate stations used as inputs. Interestingly, daily variations, which are probably learned during the snowmelt period, are also visible in periods not influenced by snow (e.g. in August). From experience (Chen et al., 2017b) we know the high relevance of snow in this area and by coupling the CNN model with a snow routine data preprocessing, we are able to further improve the model performance (Fig. 3b). We now can achieve a fit with 0.77 for both NSE and R^2, KGE increases to 0.84. Our model is able to nicely fit the second largest peak of the whole dataset, which occurs in February, though, the peak is slightly overestimated, whereas other peaks still tend to be underestimated. The snowmelt period remains well simulated, but shows increasing deviations close the end of the period. The earlier noticed daily variations in summer and autumn, now no longer appear, which is most certainly an effect of the snowmelt preprocessing.

Please note that the 95% model uncertainty, estimated from 10 different models with a Monte-Carlo dropout distribution from 100 runs each (i.e. 1000 simulations in total), is very low for both modeling results (a+b) compared to the overall variability of the discharge. Major source of uncertainty is therefore probably the especially spatially limited input data, as all climate stations have a certain distance to the catchment area. Other modeling approaches (Chen and Goldscheider, 2014; Chen et al., 2017b, 2018) based on three successive and improved versions of a combined lumped parameter (SWMM) and distributed model achieve similar or higher NSE values for the simulation of Aubach spring discharge (0.92, 0.83, 0.80 respectively) but in contrary to our approach simulate three springs simultaneously. However, these results are hardly comparable with each other

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

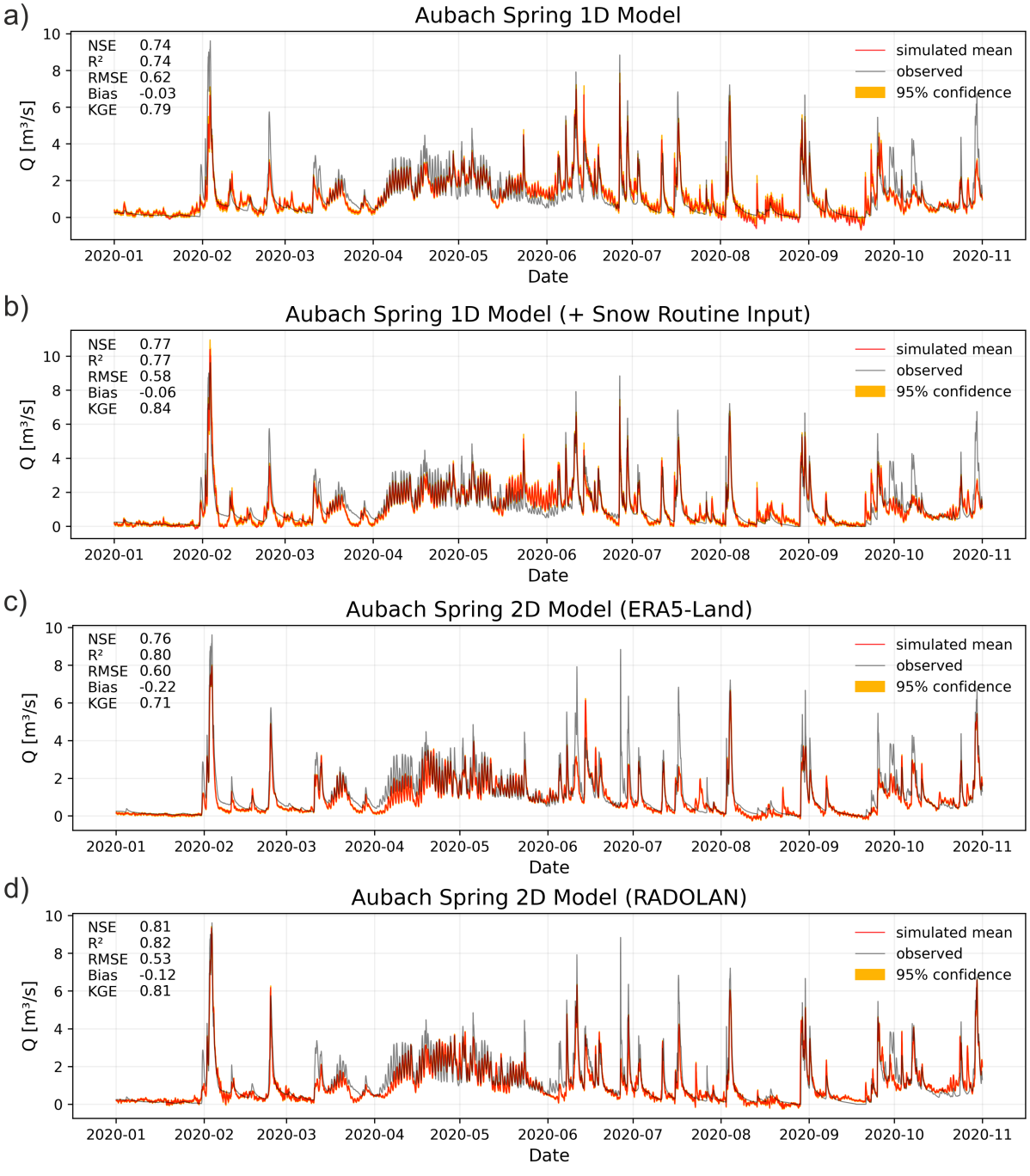


Figure 3. Simulation results for the year 2020 at Aubach spring: (a) 1D-model based on climate station inputs, (b) 1D-model with additional snow routine preprocessing, (c) 2D-model based on ERA5-Land gridded data and (d) 2D-model with combination of ERA5-Land data and RADOLAN precipitation input.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

and neither with this study. Reasons are (i) different data basis regarding the number and position of climate stations used as input data, (ii) different simulation periods as well as (iii) very different test set lengths. The shortest test set only had 40 days (in autumn), the longest (Chen et al., 2017b) used one year of data for model calibration and performed a split-sample test on the same data set. Despite our model shows a slightly lower NSE value compared to these three models, it is in the same order of magnitude and performs probably at least equally as none of the previous studies covered a complete annual cycle as contiguous test period, including high peaks in late winter and strong snowmelt influence in spring and early summer.

Figure 3c shows the results of the 2D-modeling setup using (only) ERA5-land input data. Based on Bayesian optimization, besides the fixed and not optimized input P, the following input parameters are selected: T, E, SMLT, SWVL2 and SWVL4 (for a comparison of selected parameters with other study areas see also Table A1). The performance of the 2D-model is similar to that of the 1D-models, showing a NSE (0.76) and RMSE in-between both, a larger R^2 (0.8) but a lower KGE (0.71). This performance is still very satisfying considering that the major catchment is extremely small (about 9 km^2) compared to one ERA5-Land grid cell, and that a large grid section of 14 x 14 ERA5-Land cells (1.4° x 1.4°) was used as input. We see that the major peak in February is slightly underestimated, as well as the beginning of the snowmelt period in April; however, the end of this period in May/June has improved now compared to (b). Both 1D-models are superior in estimating the peaks especially during summer, except the already mentioned peaks in September/October, which have improved using the 2D-input data. This supports the assumption that the climate stations do not represent these precipitation events, but the 2D-data does.

To account for the small area of the catchment of Aubach spring, Figure 3d shows the results of the 2D-input data, using the spatially higher resolved RADOLAN precipitation data in combination with downscaled ERA5-Land data for all other parameters. We have reduced the spatial extent of the 2D-input, but use an even higher grid cell number (22 x 22 or 22^2 km^2), with a reasonable buffer around the catchments. Additionally to P, the optimized model uses inputs from all available parameters except E and SWVL3, thus T, Tsin, SMLT, SF, SWVL1/2/4. This model shows the best performance of all four models by reaching an NSE of 0.81, R^2 of 0.82 and KGE of 0.81. Similar to the model in (c), the beginning of the snowmelt period in April remains slightly underestimated and compared to the 1D models and the peaks in summer are less well fitted. Nevertheless, we generally see a very satisfying fit of the simulation results, especially the largest peak in February is simulated almost perfectly. Compared to the 1D-approach, the main source of uncertainty for both 2D-models is probably the uncertainty of parameter values resulting from the ERA5-land grid cell sizes, which is too large compared to the catchment size. Improved downscaling of ERA5 data or other high resolved climate data for a combination with RADOLAN precipitation data might be a promising approach for simulating small catchments like this one. Model uncertainty derived from random number effects and Monte Carlo dropout is (equally to the 1D-models) satisfyingly small. In total we see that both the 1D and the 2D approach for this catchment bear significant uncertainties in terms of input data, even though the results are quite satisfactory. On the one hand the climate stations represent the true observed climate, on the other hand this is true only for a very specific point, in this case even outside the catchment and embedded into a highly variable topography. The 2D data, however, shows too large grid cell sizes for Aubach spring catchment and is itself modeled (in case of ERA5-Land). We therefore do not think that one approach is superior in terms of uncertainty for this study area.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

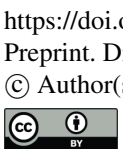

Hydrology and Earth System Sciences
Discussions

4.2 Unica Springs

Figure 4 summarizes the 1D- and 2D-model performance on the years 2017 and 2018 for Unica springs in Slovenia. Due to highly complex hydraulic behavior in this study area, which is for example related to large polje floodings and to a strongly variable water level in the system that varies also the catchment area, extracting the highly non-linear precipitation-discharge is especially challenging in this study area. We generally observe less dynamics in terms of the number of flood pulse events compared to Aubach spring. But in terms of intensity of hydrologic variability, discharge rates can vary by many orders of magnitude. This is primarily due to the large size of the catchment area, the very high degree of karstification of the carbonate rocks, and the fact that the main spring may acts as an overflow spring. Many high flow events have a quite long duration of days to even weeks resulting in a plateau-like shape. This is due to the regular flooding of the polje. After the drainage areas of the polje are completely flooded, there is a progressive back-flooding and a steady rise in the water level, which makes it impossible to accurately monitor the inflow conditions. The simulation of this quite large catchment area (820 km^2) is based on the data of only two climate stations (Postojna and Cerknica). All available input parameters from both stations except relative humidity from Postojna station and new snow from Cerknica station were used as inputs as selected by the Bayesian optimization model. The 1D-model shows solid performance overall (NSE: 0.73, R^2: 0.79, KGE: 0.63), including reaction for all major discharge events. Recessions especially in 2017 are estimated significantly too conservative and the plateau shapes of the large peaks (e.g. January 2018) are not well captured, but rather simulated as multiple peaks. Conceptually this might be even more true than the plateau-like shape, however not easy to evaluate. The peak in April 2018 is quite clearly underestimated, whereas the following low flow period (summer 2018) is slightly overestimated. One possible reason could be significant environmental changes that occurred in the catchment during 2014-18 (Kovačič et al., 2020). During this period a considerable amount of vegetation was destroyed by a series of large-scale forest disturbances. We expect the evapotranspiration changed due to changes in canopy interception, water use, and soil moisture. As a result, spring behavior has likely changed, because vegetation cover is an important element of the water balance and recharge events may have resulted in higher infiltration rates and more intense spring response, as well as more pronounced droughts. Because the period prior to 2014 was part of the training data (1981-2012), the change in spring behavior could not be predicted, resulting in an observed modeling incompatibility.

Using the 2D-input data from 18 x 21 E-OBS grid cells we were able to improve the model performance significantly (Fig. 4b), showing now a NSE of 0.83, R^2 of 0.84 and KGE of 0.80. Selected input parameters are: P (fixed), Tmax, rH and Rad. We generally observe a similar shape of the simulation as for the 1D-model but with overall reduced errors. Still, the plateau shape of some peaks is not well captured but the same conceptual understanding as for the 1D-model seems to be learned. Recessions are generally too conservative, especially the simulation of low flow periods and minor discharge events improve clearly though. These results are plausible, as minor discharge peaks are probably due to small precipitation events, which are probably not reflected in the selected climate station data, but visible in the gridded data derived from a larger set of observational points. We have, however, noticed that training the 1D-model with data starting 1961 (20 years earlier), can still significantly improve the performance, similar to the 2D results shown here. We do not provide these results as they were produced in a slightly different model setting, not fully consistent with this work. To fully exploit the gridded parameter set

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021



Hydrology and Earth System Sciences
Discussions
Open Access
EGU

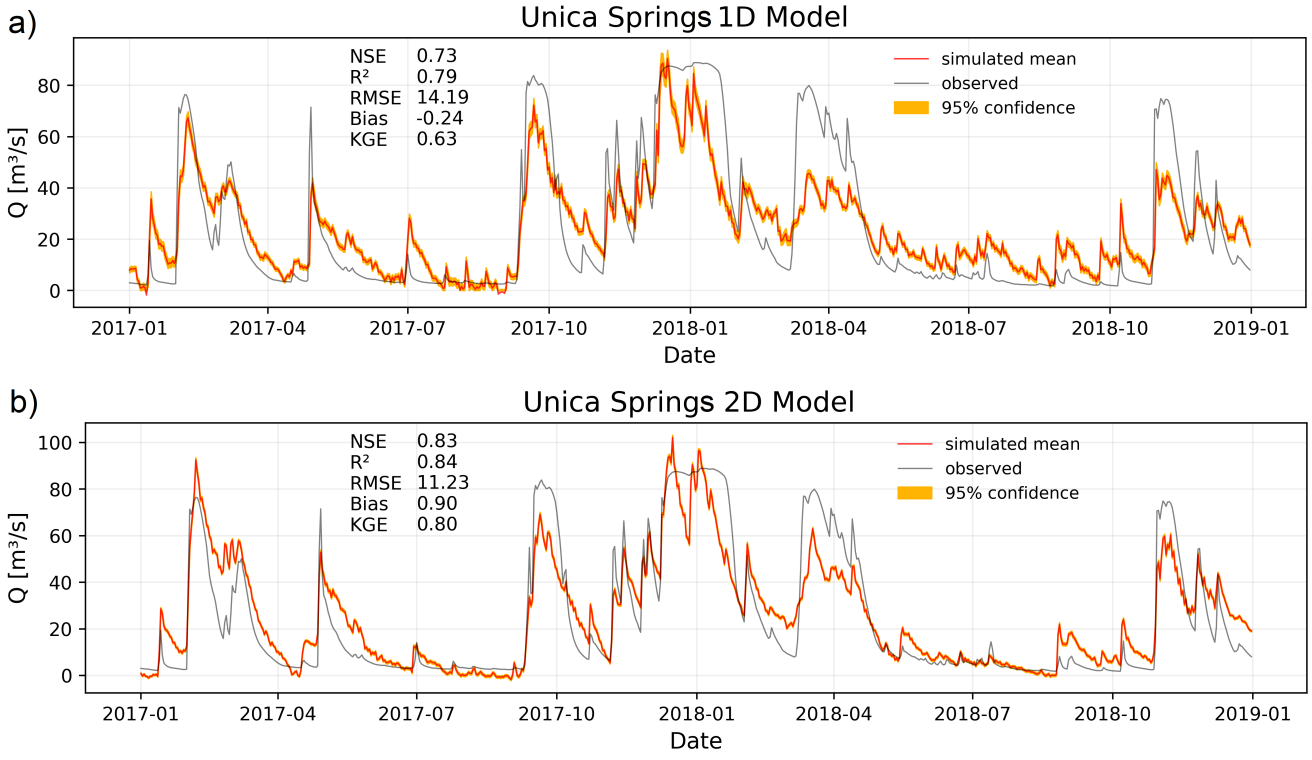


Figure 4. Simulation results for 2017-2018 at Unica springs in Slovenia using climate station input data (a) and E-OBS gridded data (b).

and maintain comparability, the results shown here are based on a start in 1981, which corresponds to the start date of rH from the E-OBS data (other parameters start earlier). As for Aubach spring, both models show a comparably low uncertainty based on random number variation and Monte-Carlo dropout, the uncertainty of the 2D-simulation is even a bit lower than for the 1D-model. The main source of uncertainty in the 1D-model is probably the spatially limited climate station data, because meteorological stations are located on the western and eastern side of the karst massive. The karst massive itself represents the orographic barrier with different temperature and precipitation regimes that are certainly not captured by the considered meteorological stations. For 2D-data the general grid uncertainty may be comparably low, due to the large size of the catchment. Several earlier modeling studies were conducted in the area of the Unica catchment (e.g. Kaufmann et al., 2016; Mayaud et al., 2019; Kaufman et al., 2020; Kovačič et al., 2020) even including ANNs (Sezen et al., 2019), but none of these directly modeled Unica springs discharge, but rather focused on other aspects like cave hydraulics or polje modeling.

4.3 Lez Spring

Lez spring represents a third class of study area, as the catchment size (around 240 km^2) is somewhere in between the two others, the climate is Mediterranean and the spring runs dry for a significant amount of time during the annual cycle due

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

to a constant exploitation through pumping. Figure 5 shows both the results for the 1D- (a) and the 2D-model (b). Despite comparably short training (daily data, starting in 2008) we observe a very high fit of the 1D-model above 0.86 for NSE, R^2 and KGE. As well the timing of the peaks, the absolute height of the peaks, as the dry periods are simulated very well, except some deviations in early 2019.

For the 2D-model we use input from 19 x 18 E-OBS grid cells and the Bayesian model selects only rH and Rad as inputs besides the fixed input P. Considering the high relevance of potential evapotranspiration in Mediterranean areas, this is probably perfectly learnable from relative humidity and radiation data. The performance of the model is basically very good, but clearly lower compared to the 1D-model, showing NSE, R^2 and KGE between 0.75 and 0.78. Generally 2018 is better simulated than 2019, which is however also slightly visible for the 1D-model. Some non-existent peaks are simulated by the model in the dry sections, after all one of them (in Oct. 2018) is also simulated by the 1D-model. These differences in performance can probably be explained by looking again at the input data. The climate stations, from which the interpolated precipitation time series is derived, are mainly located inside the catchment and additionally represent a good spatial coverage. Compared to both other study areas the 1D input uncertainty is definitely the best for Lez spring catchment. Though, it seems harder to extract the relevant data from the gridded data, which may be related to uncertainties from the grid cell size in relation to the catchment size (e.g. which may be more important than for Unica but way less than for Aubach). Nevertheless, the model uncertainty based on initializations and derived from Monte-Carlo dropout again is satisfyingly small for both model setups, especially for the dry periods.

For all study areas in this work, Lez spring is the one with the most existing modeling approaches, including several ANN studies. Please refer to Kong A Siou et al. (2011) for an overview about older modeling studies at Lez spring with different approaches. In this study, we focus on the most recent studies for comparison, all of which were conducted with ANNs (Kong A Siou et al., 2011, 2012; Kong-A-Siou et al., 2013, 2014; Darras et al., 2015; Kong-A-Siou et al., 2015; Darras et al., 2017). These studies are based on modeling Lez spring using Multi-Layer-Perceptrons (MLP) or recurrent MLPs. They use a wide variation of daily data from different time periods and also have different study goals. Darras et al. (2015, 2017) focus only on single events to evaluate flash flood forecasting. Of the other studies, especially three are suitable for a comparison with the results of our work. Kong-A-Siou et al. (2013) uses a 1-year test set in 2002-2003 and obtains the highest NSE of all studies (0.96), however the focus of the study is knowledge extraction from the ANN. Kong-A-Siou et al. (2014) simulate the period from October 2003 to August 2004 (NSE: 0.85) and Kong-A-Siou et al. (2015) investigate different years between 1997 and 2004 with strongly varying NSE values ranging from 0.1 to 0.85. Compared to our study, usually a larger database is used for training in these studies, which in itself is a plausible reason for the good performance. Also significant effort was mostly put into consideration of evapotranspiration or potential evaporation, further sometimes pumping data from the accompanying wells is also used as input. We therefore can conclude that our 1D-CNN models in can compete with the performance of these earlier ANN approaches even though we do not beat the best NSE value from these studies; however we use less input data for our 1D-model. The 2D-approach still shows very solid performance in comparison, but cannot fully compete with the 1D-model and earlier ANN approaches. Nevertheless, we also show that it offers a solid alternative if no climate stations were available.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

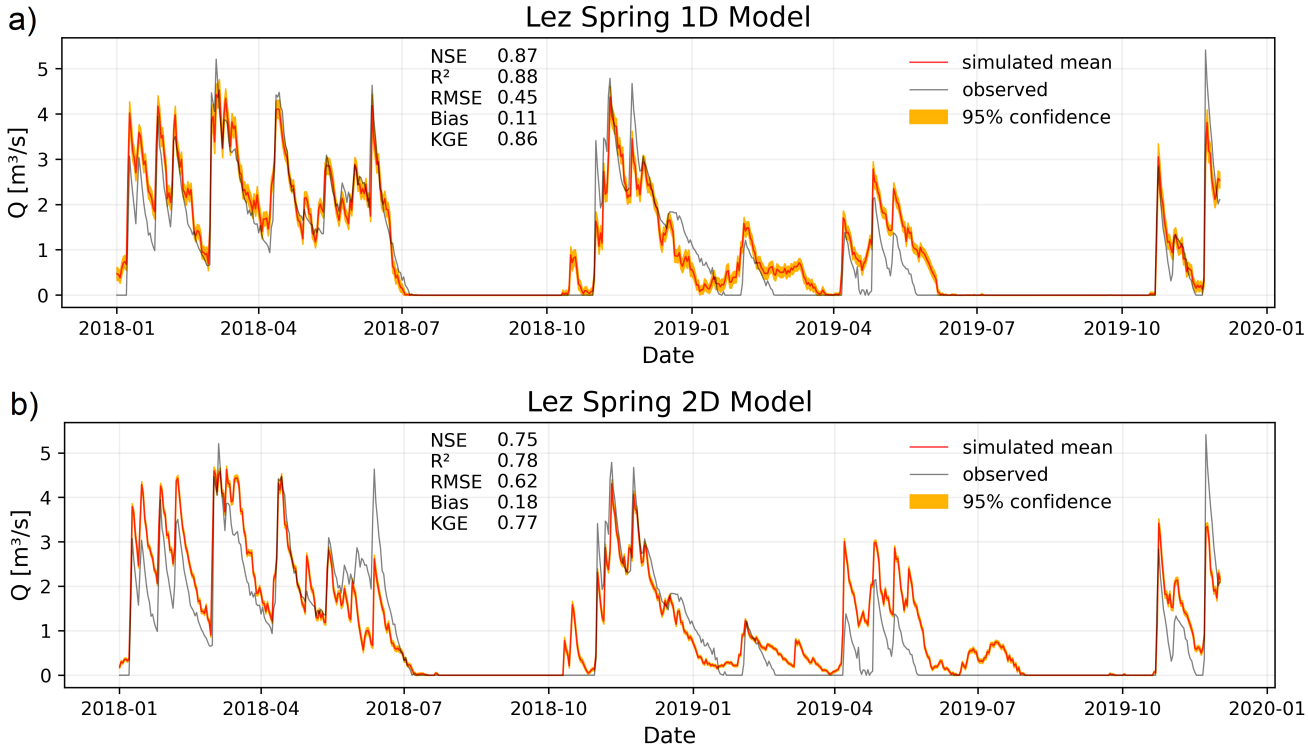


Figure 5. Simulation results for 2018-2019 at Lez spring in France using climate station input data (a) and E-OBS gridded data (b).

4.4 Spatial Input Sensitivity Results

The most important results of the spatial input sensitivity analysis from all catchments are shown in Figure 6. In case of Aubach spring modeled with ERA5-Land data (Fig. 6a) we can see that the catchment is hardly the size of one grid cell. Hence, despite the quite solid discharge modeling, we see no clear spatial meaning of the precipitation channel heatmap but the higher sensitivity in the direction the weather is approximately coming from. We also see a border effect with an almost uniform decrease in sensitivity toward the edges. This is an important reason to choose the spatial extent of the data large enough, as it is probably related to the size of the filter in the convolutional layer (3 x 3). This occurs even though clipping is used while performing convolutions to improve the informative value of the edges. Though, not all heatmaps show this pattern. In case of Lez spring (Fig. 6d) we see that the sensitive region is clearly separated from the rest, thus we cannot see a border effect even if probably present. For Aubach spring, precipitation shows only the fourth highest sensitivity (S) in terms of absolute values, while the second most sensitive parameter is snowmelt (SMLT), which shows also the best spatial agreement with the catchment area. This is plausible insofar as the discharge for a large part of the time is dominated by snowmelt and to a lesser extent directly by precipitation. We conclude that even though the modeling results are satisfying, not much meaning can be extracted from the spatial sensitivity analysis for such a small catchment. Please find heatmaps of all other parameters

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

in Appendix C. The combined approach of RADOLAN and ERA5-Land (Fig. 6b) data shows the heatmap in more detail in relation to the size of the catchment. We show only the precipitation heatmap, because it is the only parameter with a native resolution of 1 km x 1 km and we do not consider the spatial patterns of the remaining ERA5-Land-based parameters to be meaningful to interpret. We observe that the most sensitive cells are identified close to the spring and at the border between the main catchment and the southern adjacent subcatchment. The pronounced elongation from NW to SE might be a relic from spatial correlation of precipitation events. No obvious explanation can be given for the two separate sensitive areas in the SW and NE corners, however, they are less sensitive then the center cells of the map.

Heatmaps of all four selected E-OBS parameters at Unica catchment are shown in Figure 6c. In accordance with our expectation for karst areas, we see the highest sensitivity for precipitation, which also identifies the catchment area quite exactly. We think that the size of the Unica catchment in terms of catchment delineation is probably the best out of all three catchments investigated in this study, concerning the given spatial resolution of the used input data. Especially Tmax and rH show high sensitivities on larger areas, however they are usually highly spatially autocorrelated and do not show a strong spatial heterogeneity like precipitation, which makes it plausible that the model learns from larger data fractions. The model further identifies an area in the north as most sensitive for radiation. We know that radiation is a controlling factor for snowmelt. Thus, we can speculate that snowmelt in this northern area, which is already part of the Alps, might be well correlated to the catchment snowmelt. On catchment scale, however, more snow is expected in the southern part, which cannot be seen on the heatmap.

Heatmaps of the 2D-Lez spring model are shown in Figure 6d. In this area the model very strongly ignores large parts of the input data (dark blue, no visible border effects) and comparably well identifies the relevant area for the spring. This might be related to (i) the higher spatial heterogeneity of precipitation in Mediterranean climate or (ii) the lower dampening of the Lez spring karst system compared to the other study areas. Generally, we observe a slight south and east shift of the highest sensitivity compared to the catchment position. This might be related to the performance of the 2D-approach, which could not compete with the 1D-models. Maybe the model did not exactly learn the most relevant input fraction. The most sensitive parameter is precipitation, while the rH channel shows the best spatial fit. We furthermore see that the size of the catchment is about the minimum size to produce meaningful heatmaps based on this given grid resolution, which corresponds also to our interpretation of the 2D-model performance shortcomings in comparison with the 1D-approach.

In summary, we observe that the approach can produce meaningful heatmaps for at least roughly locating karst spring catchments. We notice that it works better the larger the area, especially in relation to the grid cell size, but the absolute size of the catchment itself appears to be also important. For small catchments it seems harder to extract precise catchment locations, even if spatially finer-resolved data are available. This might be related to the fact that at small scales, precipitation has a distinct spatial correlation, which can lead to higher sensitivity also in areas outside the catchment. However, one should keep in mind that these conclusions are only tendencies as we only investigated a small number of catchments. Moreover, it can be expected that more and better gridded meteorological data products will be available in the future, which might lead to better results with the proposed methodology.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

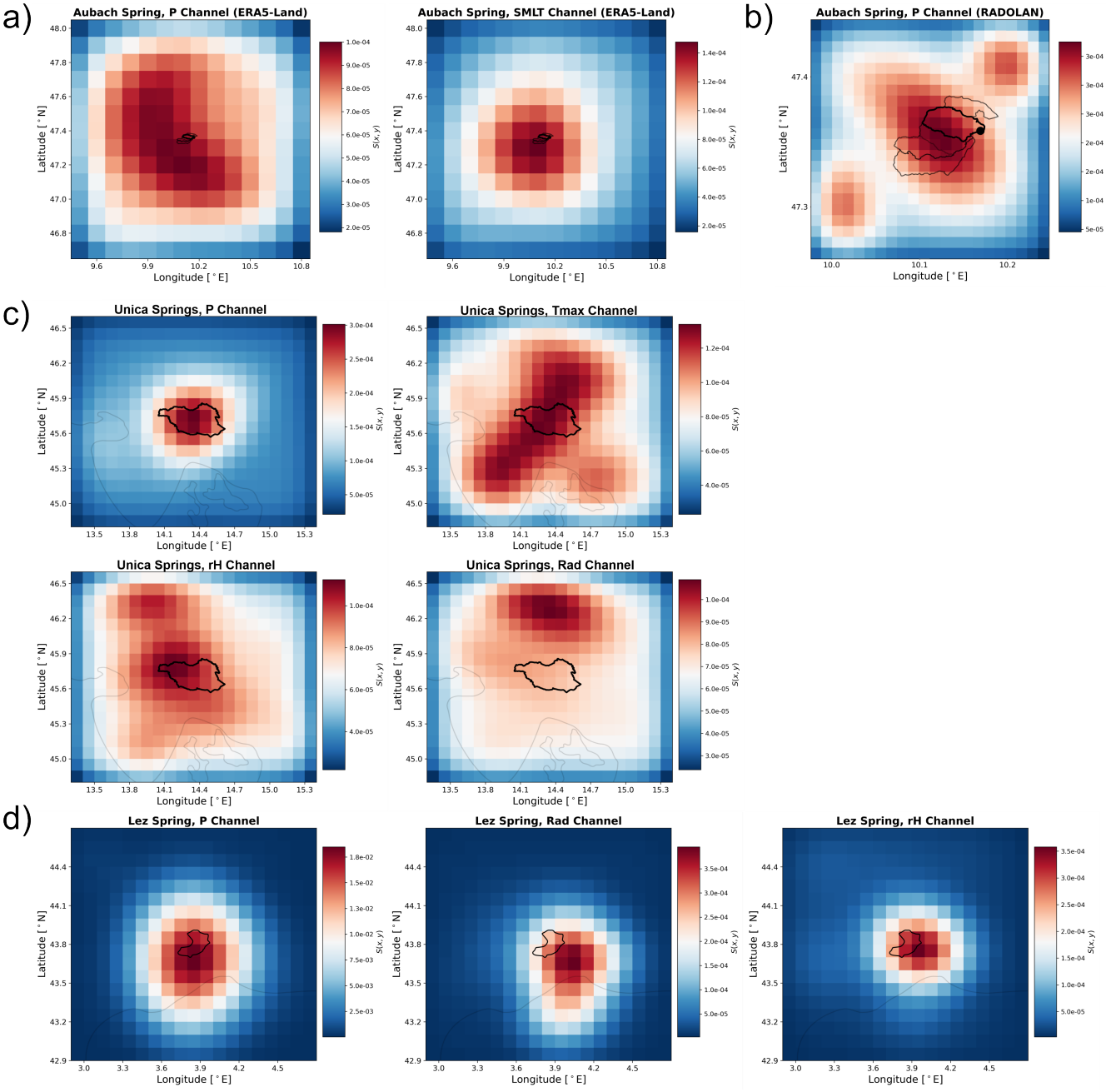


Figure 6. Heatmaps of spatial input sensitivity for Aubach spring based on ERA5-Land gridded data (a), for Aubach spring based on RADOLAN precipitation data (b), Unica springs (c) and Lez spring (d) both based on E-OBS gridded data. In case of (c) and (d), light-grey lines indicate the coastlines for orientation.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021


Hydrology and Earth System Sciences
Discussions

5 Conclusions

From the obtained insights we can conclude that karst spring discharge can be predicted accurately with the presented approaches. Their accuracy rivals that of existing models in the three study areas, with the difference that far less prior knowledge of the system under consideration is required than, e.g., for lumped parameter models. This can significantly reduce the amount of work required, provided that a mere simulation of the spring discharge is the objective; albeit, we do not gain knowledge about hydraulic processes in the aquifer as we do from lumped parameter models. We can further show that gridded climate data can provide an excellent substitute for non-existent or patchy climate station data. This does not require knowledge of the exact catchment area, which is a critical component especially for karst springs. Rather, 2D-CNNs can be used to generate a first approximation of the catchment area. However, this approach is subject to some inaccuracies and needs further development in combination with 2D-meteorological input data in a finer spatial resolution in relation to the catchment size. Additionally, a sufficient heterogeneity of precipitation in comparison to the catchment size is necessary. Then it can most certainly be used to delineate catchments quite accurately, which can then be evaluated against tracer tests and hydrogeological studies. In the present state, however, it does not replace conventional catchment delineation methods. In terms of accuracy, we do not find that one of the tested model setups is fundamentally superior. Nevertheless, we would conclude that the 2D-approach is superior to the 1D-approach with respect to the effort of data collection as well as data and parameter availability. Though, a significantly increased computational effort is necessary for the training as well as for the optimization of the models. In summary, gridded meteorological data is excellently useful to overcome missing climate station data and to get a quite good idea of the spatial extent of larger catchments in relation to the grid cell size.

Code and data availability. We provide complete model codes on Github (AndreasWunsch/CNN_KarstSpringModeling) (Wunsch, 2021). Due to redistribution restrictions from several parties we cannot provide a dataset. Nevertheless, the data is available from the respective local authorities listed in the main text and in the following. 2D-datasets (E-OBS, ERA5-Land) are fully accesible online via Copernicus (cds.climate.copernicus.eu). Aubach spring discharge and climate data from surrounding climate stations in Austria are availble on request from the office of the federal state of Vorarlberg, division of water management, Oberstdorf station data (German Meteorological Service) is available online (opendata.dwd.de). Data from Slovenia can be retrieved from ARSO (Slovenian Environment Agency)(ARSO, 2020a, b). Lez spring discharge was provided by SNO KARST (2021), climate data is available on request from MeteoFrance.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions

Appendix A: Study Area Comparison Table

Table A1. Summary and comparison of different aspects of all three study areas.

	Aubach Spring	**Unica Springs**	**Lez Springs**
Country	Austria	Slovenia	France
Climate	cooltemperate and humid	moderate continental	mediterranean
Catchment Area [km^2]	9	820	240
mean Precipitation [mm/year]	2000	1500	904
Station, Period	(Walm.-Horn, 2003-2019)	(1989-2018)	(2008-2018)
spatially distributed input datasets	ERA5-Land, RADOLAN	E-OBS	E-OBS
Offered Parameters	P, T, Tsin, E, SMLT, SF, SWVL1-4	P, T, Tmin, Tmax, Tsin, rH, Rad	P, T, Tmin, Tmax, Tsin, rH, Rad
Selected Parameters	ERA5-Model: P, T, E, SMLT, SWVL2, 4 RADOLAN-Model: P, T, Tsin, SMLT, SF, SWVL1, 2, 4	P, Tmax, rH, Rad	P, rH, Rad
Omitted Parameters	ERA5-Model: Tsin, SF, SWVL1, 3 RADOLAN-Model: E, SWVL3	T, Tmin, Tsin	T, Tmin, Tmax, Tsin

Appendix B: Lez Catchment Precipitation Interpolation

The Thiessen's polygon interpolation method consists of calculating a weighted average of the precipitation data by allocating a contribution percentage to each meteorological station, based on its influence area on the catchment. These influence areas are calculated through geometric operations. First, we draw straight-line segments between each adjacent station, then we add the perpendicular bisectors of each segment, which will define the edges of the polygons. Each meteorological station thus corresponds to a particular polygon, for which the precipitation over the surface is assumed to be the same as the measured precipitation at the station.

The weighted average of the precipitation P_{wa} at each time step is calculated as follows:

$$P_{wa} = \frac{\sum_{i=1}^{n} A_i P_i}{A} \tag{B1}$$

With n the number of meteorological stations, A_i the area (over the catchment) of the polygon corresponding to the ith station, P_i the precipitation measured at the ith station and A the area of the catchment.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions

Appendix C: Heatmaps

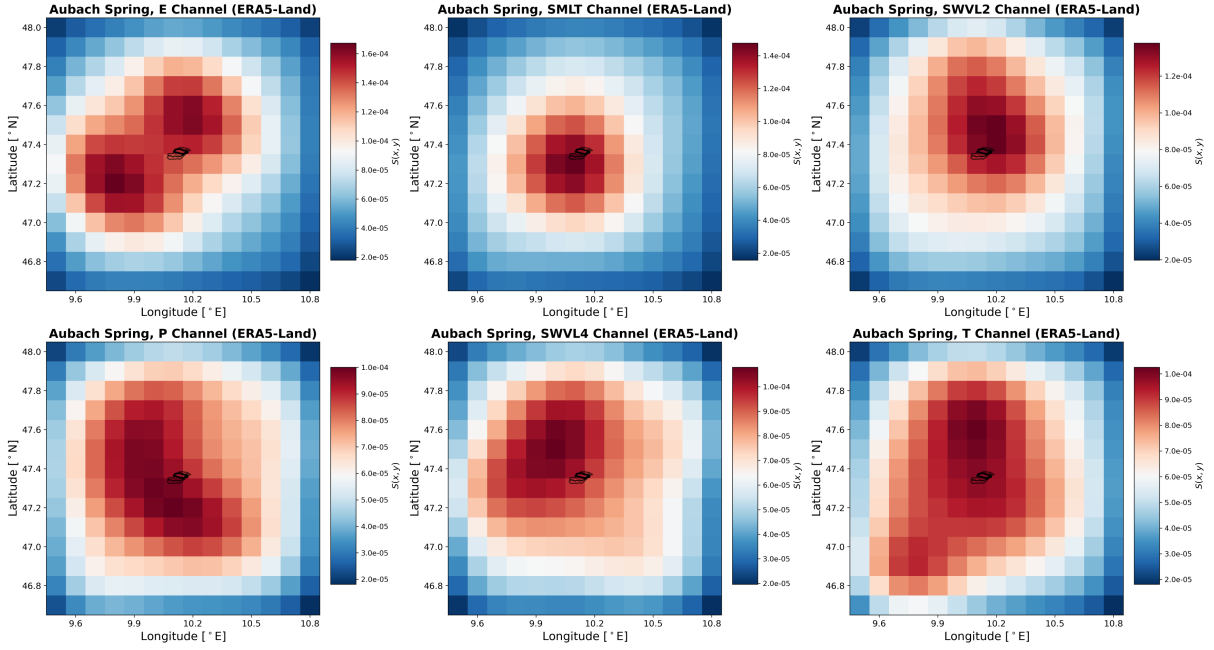


Figure C1. Spatial input sensitivity heatmaps for Aubach spring based on ERA5-Land gridded data.

Author contributions. AW, TL and NG conceptualized the study, AW and TL developed the methodology and software code, and validated the results. AW performed the experiments, and investigated and visualized the results. GC and ZC performed formal analysis, NR and GC contributed to data curation activities. AW wrote the original paper draft with contributions from GC and NR. All authors contributed to interpretation of the results, and review and editing of the paper draft. TL and NG supervised the work.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The financial support of KIT through the German Federal Ministry of Education and Research (BMBF) and the European Commission through the Partnership for Research and Innovation in the Mediterranean Area (PRIMA) program under Horizon 2020 (KARMA project, grant agreement number 01DH19022A) is gratefully acknowledged. We thank the French Ministry of Higher Education and Research for the thesis scholarship of G. Cinkus as well as the European Commission and the Agence Nationale de la Recherche (ANR) for its support of HSM and UMR through the Partnership for Research and Innovation in the Mediterranean Area (PRIMA) program under Horizon 2020 (KARMA project, ANR-18-PRIM-0005). We further acknowledge financial support by the Slovenian Research Agency within

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

485 the project Infiltration processes in forested karst aquifers under changing environment (No. J2-1743). The authors acknowledge support by the state of Baden-Württemberg through bwHPC. Muñoz Sabater, J., (2019) was downloaded from the Copernicus Climate Change Service (C3S) Climate Data Store. The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. We acknowledge the E-OBS dataset and the data providers in the ECA&D project (https://www.ecad.eu), data from MeteoFrance, DWD and the office of the
490 federal state of Vorarlberg, division of water management. Lez spring discharge data were provided by the KARST observatory network (SNO KARST) initiative from the INSU/CNRS (FRANCE), which aims to strengthen knowledge-sharing and to promote cross-disciplinary research on karst systems.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, p. 19, https://www.tensorflow.org/, 2015.

Anderson, S. and Radic, V.: Evaluation and Interpretation of Convolutional-Recurrent Networks for Regional Hydrological Modelling, Hydrology and Earth System Sciences Discussions, 2021, 1–43, https://doi.org/10.5194/hess-2021-113, 2021.

ARSO: Slovenian Environment Agency. Archive of Hydrological Data., http://vode.arso.gov.si/hidarhiv/, 2020a.

ARSO: Slovenian Environment Agency. Archive of Meteorological Data., http://www.meteo.si, 2020b.

Bergström, S.: The Development of a Snow Routine for the HBV-2 Model, Hydrology Research, 6, 73–92, https://doi.org/10/gkcqz5, 1975.

Bergström, S.: The HBV Model, in: Computer Models of Watershed Hydrology., edited by Singh, V. P., pp. 443–476, Water Resources Publications, Colorado, USA, https://www.cabdirect.org/cabdirect/abstract/19961904773, 1995.

Bicalho, C. C., Batiot-Guilhe, C., Seidel, J. L., Van Exter, S., and Jourde, H.: Hydrodynamical Changes and Their Consequences on Groundwater Hydrochemistry Induced by Three Decades of Intense Exploitation in a Mediterranean Karst System, Environ Earth Sci, 65, 2311–2319, https://doi.org/10.1007/s12665-011-1384-2, 2012.

Chen, Z. and Goldscheider, N.: Modeling Spatially and Temporally Varied Hydraulic Behavior of a Folded Karst System with Dominant Conduit Drainage at Catchment Scale, Hochifen–Gottesacker, Alps, Journal of Hydrology, 514, 41–52, https://doi.org/10.1016/j.jhydrol.2014.04.005, 2014.

Chen, Z., Auler, A. S., Bakalowicz, M., Drew, D., Griger, F., Hartmann, J., Jiang, G., Moosdorf, N., Richts, A., Stevanovic, Z., Veni, G., and Goldscheider, N.: The World Karst Aquifer Mapping Project: Concept, Mapping Procedure and Map of Europe, Hydrogeol J, 25, 771–785, https://doi.org/10/f98h6g, 2017a.

Chen, Z., Hartmann, A., and Goldscheider, N.: A New Approach to Evaluate Spatiotemporal Dynamics of Controlling Parameters in Distributed Environmental Models, Environmental Modelling & Software, 87, 1–16, https://doi.org/10.1016/j.envsoft.2016.10.005, 2017b.

Chen, Z., Hartmann, A., Wagener, T., and Goldscheider, N.: Dynamics of Water Fluxes and Storages in an Alpine Karst Catchment under Current and Potential Future Climate Conditions, Hydrol. Earth Syst. Sci., p. 17, https://doi.org/10.5194/hess-22-3807-2018, 2018.

Chollet, F.: Keras, Keras, https://github.com/keras-team/keras, 2015.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, J. Geophys. Res. Atmos., 123, 9391–9409, https://doi.org/10/gfk3hd, 2018.

Darras, T., Borrell Estupina, V., Kong-A-Siou, L., Vayssade, B., Johannet, A., and Pistre, S.: Identification of Spatial and Temporal Contributions of Rainfalls to Flash Floods Using Neural Network Modelling: Case Study on the Lez Basin (Southern France), Hydrology and Earth System Sciences, 19, 4397–4410, https://doi.org/10/f7xq6q, 2015.

Darras, T., Kong-A-Siou, L., Vayssade, B., Johannet, A., and Pistre, S.: Karst Flash Flood Forecasting Using Recurrent and Nonrecurrent Artificial Neural Network Models: The Case of the Lez Basin (Southern France), in: EuroKarst 2016, Neuchâtel, Advances in Karst Science, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-45465-8, 2017.

DWD Climate Data Center (CDC): Historical and Current Hourly RADOLAN Grids of Precipitation Depth (Binary). Version V001, https://opendata.dwd.de/climate_environment/CDC/grids_germany/hourly/radolan/.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Fleury, P., Ladouche, B., Conroux, Y., Jourde, H., and Dörfliger, N.: Modelling the Hydrologic Functions of a Karst Aquifer under Active Water Management – The Lez Spring, Journal of Hydrology, 365, 235–243, https://doi.org/10.1016/j.jhydrol.2008.11.037, 2009.

Goldscheider, N.: Fold Structure and Underground Drainage Pattern in the Alpine Karst System Hochifen-Gottesacker, Eclogae geol. Helv., 98, 1–17, https://doi.org/10.1007/s00015-005-1143-z, 2005.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the Mean Squared Error and NSE Performance Criteria: Implications for Improving Hydrological Modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hock, R.: A Distributed Temperature-Index Ice- and Snowmelt Model Including Potential Direct Solar Radiation, Journal of Glaciology, 45, 101–111, https://doi.org/10/ggnvkt, 1999.

Hunter, J. D.: Matplotlib: A 2D Graphics Environment, Computing in Science Engineering, 9, 90–95, https://doi.org/10.1109/mcse.2007.55, 2007.

Jeannin, P.-Y., Artigue, G., Butscher, C., Chang, Y., Charlier, J.-B., Duran, L., Gill, L., Hartmann, A., Johannet, A., Jourde, H., Kavousi, A., Liesch, T., Liu, Y., Lüthi, M., Malard, A., Mazzilli, N., Pardo-Igúzquiza, E., Thiéry, D., Reimann, T., Schuler, P., Wöhling, T., and Wunsch, A.: Karst Modelling Challenge 1: Results of Hydrological Modelling, Journal of Hydrology, 600, 126 508, https://doi.org/10.1016/j.jhydrol.2021.126508, 2021.

Johannet, A., Mangin, A., and D'Hulst, D.: Subterranean Water Infiltration Modelling by Neural Networks: Use of Water Source Flow, in: ICANN '94: Proceedings of the International Conference on Artificial Neural Networks Sorrento, Italy, 26–29 May 1994 Volume 1, Parts 1 and 2, pp. 1033–1036, Springer Berlin Heidelberg, Sorrento, Italy, 1994.

Jourde, H., Lafare, A., Mazzilli, N., Belaud, G., Neppel, L., Dörfliger, N., and Cernesson, F.: Flash Flood Mitigation as a Positive Consequence of Anthropogenic Forcing on the Groundwater Resource in a Karst Catchment, Environ Earth Sci, 71, 573–583, https://doi.org/10.1007/s12665-013-2678-3, 2014.

Jourde, H., Massei, N., Mazzilli, N., Binet, S., Batiot-Guilhe, C., Labat, D., Steinmann, M., Bailly-Comte, V., Seidel, J. L., Arfib, B., Charlier, J. B., Guinot, V., Jardani, A., Fournier, M., Aliouache, M., Babic, M., Bertrand, C., Brunet, P., Boyer, J. F., Bricquet, J. P., Camboulive, T., Carrière, S. D., Celle-Jeanton, H., Chalikakis, K., Chen, N., Cholet, C., Clauzon, V., Soglio, L. D., Danquigny, C., Défargue, C., Denimal, S., Emblanch, C., Hernandez, F., Gillon, M., Gutierrez, A., Sanchez, L. H., Hery, M., Houillon, N., Johannet, A., Jouves, J., Jozja, N., Ladouche, B., Leonardi, V., Lorette, G., Loup, C., Marchand, P., de Montety, V., Muller, R., Ollivier, C., Sivelle, V., Lastennet, R., Lecoq, N., Maréchal, J. C., Perotin, L., Perrin, J., Petre, M. A., Peyraube, N., Pistre, S., Plagnes, V., Probst, A., Probst, J. L., Simler, R., Stefani, V., Valdes-Lao, D., Viseur, S., and Wang, X.: SNO KARST: A French Network of Observatories for the Multidisciplinary Study of Critical Zone Processes in Karst Watersheds and Aquifers, Vadose Zone Journal, 17, 180 094, https://doi.org/10/gk9t3n, 2018.

Kaufman, G., Mayaud, C., Kogovšek, B., and Gabrovšek, F.: Understanding the Temporal Variation of Flow Direction in a Complex Karst System (Planinska Jama, Slovenia), Acta Carsologica, 49, https://doi.org/10/gkcs9h, 2020.

Kaufmann, G., Gabrovšek, F., and Turk, J.: Modelling Flow of Subterranean Pivka River in Postojnska Jama, Slovenia, Acta Carsologica, 45, https://doi.org/10.3986/ac.v45i1.3059, 2016.

Kollat, J. B., Reed, P. M., and Wagener, T.: When Are Multiobjective Calibration Trade-Offs in Hydrologic Models Meaningful?, Water Resources Research, 48, https://doi.org/10/gkcq5k, 2012.

Kong A Siou, L., Johannet, A., Borrell, V., and Pistre, S.: Complexity Selection of a Neural Network Model for Karst Flood Forecasting: The Case of the Lez Basin (Southern France), Journal of Hydrology, 403, 367–380, https://doi.org/10.1016/j.jhydrol.2011.04.015, 2011.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Hydrology and Earth System Sciences
Discussions

Kong A Siou, L., Johannet, A., Valérie, B. E., and Pistre, S.: Optimization of the Generalization Capability for Rainfall–Runoff Modeling by Neural Networks: The Case of the Lez Aquifer (Southern France), Environmental Earth Sciences, 65, 2365–2375, https://doi.org/10.1007/s12665-011-1450-9, 2012.

Kong-A-Siou, L., Cros, K., Johannet, A., Borrell-Estupina, V., and Pistre, S.: KnoX Method, or Knowledge eXtraction from Neural Network Model. Case Study on the Lez Karst Aquifer (Southern France), Journal of Hydrology, 507, 19–32, https://doi.org/10.1016/j.jhydrol.2013.10.011, 2013.

Kong-A-Siou, L., Fleury, P., Johannet, A., Borrell Estupina, V., Pistre, S., and Dörfliger, N.: Performance and Complementarity of Two Systemic Models (Reservoir and Neural Networks) Used to Simulate Spring Discharge and Piezometry for a Karst Aquifer, Journal of Hydrology, 519, 3178–3192, https://doi.org/10.1016/j.jhydrol.2014.10.041, 2014.

Kong-A-Siou, L., Johannet, A., Borrell Estupina, V., and Pistre, S.: Neural Networks for Karst Groundwater Management: Case of the Lez Spring (Southern France), Environmental Earth Sciences, 74, 7617–7632, https://doi.org/10.1007/s12665-015-4708-9, 2015.

Kovačič, G., Petrič, M., and Ravbar, N.: Evaluation and Quantification of the Effects of Climate and Vegetation Cover Change on Karst Water Sources: Case Studies of Two Springs in South-Western Slovenia, Water, 12, 3087, https://doi.org/10/gksnmz, 2020.

Lebigot, E. O.: Uncertainties: A Python Package for Calculations with Uncertainties, https://pythonhosted.org/uncertainties/numpy_guide.html, 2010.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep Learning, Nature, 521, 436–444, https://doi.org/10.1038/nature14539, 2015.

Longenecker, J., Bechtel, T., Chen, Z., Goldscheider, N., Liesch, T., and Walter, R.: Correlating Global Precipitation Measurement Satellite Data with Karst Spring Hydrographs for Rapid Catchment Delineation, Geophysical Research Letters, 44, 4926–4932, https://doi.org/10.1002/2017GL073790, 2017.

Malard, A., Jeannin, P.-Y., Vouillamoz, J., and Weber, E.: An Integrated Approach for Catchment Delineation and Conduit-Network Modeling in Karst Aquifers: Application to a Site in the Swiss Tabular Jura, Hydrogeol J, 23, 1341–1357, https://doi.org/10.1007/s10040-015-1287-5, 2015.

Mayaud, C., Gabrovšek, F., Blatnik, M., Kogovšek, B., Petrič, M., and Ravbar, N.: Understanding Flooding in Poljes: A Modelling Perspective, Journal of Hydrology, 575, 874–889, https://doi.org/10.1016/j.jhydrol.2019.04.092, 2019.

Mazzilli, N., Jourde, H., Guinot, V., Bailly-Comte, V., and Fleury, P.: Hydrological Modelling of a Karst Aquifer under Active Groundwater Management Using a Parsimonious Conceptual Model, in: H2Karst, Besançon, France, https://hal.archives-ouvertes.fr/hal-01844603, 2011.

McKinney, W.: Data Structures for Statistical Computing in Python, in: Python in Science Conference, pp. 56–61, Austin, Texas, https://doi.org/10.25080/majora-92bf1922-00a, 2010.

Muñoz Sabater, J.: ERA5-Land Hourly Data from 2001 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)., https://doi.org/10.24381/CDS.E2161BAC, 2019.

NASA: GPM - Global Precipitation Measurement, http://www.nasa.gov/mission_pages/GPM/main/index.html, 2016.

Nash, J. E. and Sutcliffe, J. V.: River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles, Journal of hydrology, 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nogueira, F.: Bayesian Optimization: Open Source Constrained Global Optimization Tool for Python, https://github.com/fmfn/BayesianOptimization, 2014.

https://doi.org/10.5194/hess-2021-403
Preprint. Discussion started: 11 August 2021

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., and Cournapeau, D.: Scikit-Learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Petrič, M., Kogovšek, J., and Ravbar, N.: Effects of the Vadose Zone on Groundwater Flow and Solute Transport Characteristics in Mountainous Karst Aquifers – the Case of the Javorniki–Snežnik Massif (SW Slovenia), AC, 47, https://doi.org/10.3986/ac.v47i1.5144, 2018.

Reback, J., McKinney, W., Jbrockmendel, Bossche, J. V. D., Augspurger, T., Cloud, P., Gfyoung, Sinhrks, Klein, A., Roeschke, M., Hawkins, S., Tratner, J., She, C., Ayd, W., Petersen, T., Garcia, M., Schendel, J., Hayden, A., MomIsBestFriend, Jancauskas, V., Battiston, P., Seabold, S., Chris-B1, H-Vetinari, Hoyer, S., Overmeire, W., Alimcmaster1, Dong, K., Whelan, C., and Mehyar, M.: Pandas-Dev/Pandas: Pandas 1.0.3, Zenodo, https://doi.org/10.5281/ZENODO.3509134, 2020.

Seibert, J.: Multi-Criteria Calibration of a Conceptual Runoff Model Using a Genetic Algorithm, Hydrology and Earth System Sciences, 4, 215–224, https://doi.org/10/crmzd7, 2000.

Sezen, C., Bezak, N., Bai, Y., and Šraj, M.: Hydrological Modelling of Karst Catchment Using Lumped Conceptual and Data Mining Models, Journal of Hydrology, 576, 98–110, https://doi.org/10.1016/j.jhydrol.2019.06.036, 2019.

SNO KARST: Time Series of Type Hydrology-Hydrogeology in Le Lez (Méditerranée) Basin - MEDYCYSS Observatory - KARST Observatory Network - OZCAR Critical Zone Network Research Infrastructure, https://doi.org/10.15148/CFD01A5B-B7FD-41AA-8884-84DBDDAC767E, 2021.

Stevanović, Z.: Karst Waters in Potable Water Supply: A Global Scale Overview, Environ Earth Sci, 78, 662, https://doi.org/10.1007/s12665-019-8670-9, 2019.

Thiéry, D. and Bérard, P.: Alimentation en eau de la ville de Montpellier - captage de la source du Lez - études des relations entre la source et son réservoir aquifère, Tech. rep., BRGM No. 83, SNG 167 LRO, http://infoterre.brgm.fr/rapports/83-SGN-167-LRO.pdf, 1983.

van der Walt, S., Colbert, S. C., and Varoquaux, G.: The NumPy Array: A Structure for Efficient Numerical Computation, Comput. Sci. Eng., 13, 22–30, https://doi.org/10.1109/mcse.2011.37, 2011.

van Rossum, G.: Python Tutorial, 1995.

Wunsch, A.: AndreasWunsch/CNN_KarstSpringModeling, GitHub repository, Zenodo, https://github.com/AndreasWunsch/CNN_KarstSpringModeling, 2021.

Wunsch, A., Liesch, T., and Broda, S.: Groundwater Level Forecasting with Artificial Neural Networks: A Comparison of Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNNs), and Non-Linear Autoregressive Networks with Exogenous Input (NARX), Hydrology and Earth System Sciences, 25, 1671–1687, https://doi.org/10.5194/hess-25-1671-2021, 2021.