Can discharge be used to inversely correct precipitation?

Ashish Manoj J[1,a], Ralf Loritz[1], Hoshin Gupta[2], Erwin Zehe[1]

1 Chair of Hydrology, Institute of Water and Environment (IWU), Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

2 Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson 85721, United States

[a] *Correspondence to: Ashish Manoj J (ashish.jaseetha@kit.edu)*

Abstract. This study explores the feasibility of using the information contained in observed streamflow discharge measurements to inversely correct catchment-average precipitation time series provided by reanalysis products. We explore this possibility by training LSTM models to predict precipitation. The first model uses discharge as an input feature along with other meteorological factors, while the second model uses only the meteorological factors. Although the model provided with discharge information showed better mean performance, a detailed analysis of various time series measures across the continental scale revealed underestimation biases when compared with the original reanalysis product used for training. However, an out-of-sample test showed that the inversely estimated precipitation is better able to reproduce small-scale, high-impact events that are poorly represented in the original reanalysis product. Further, using the inversely generated precipitation time series for classical hydrological "forward" modeling resulted in improved estimates for streamflow and soil moisture. Given the notable disconnect between reanalysis products and extreme events, particularly in data-scarce regions worldwide, our findings have implications for achieving better estimates of precipitation associated with high-impact events.

1 Introduction

The performance of hydrological models has traditionally been constrained by the availability and quality of observations covering various aspects of the water cycle. Among those, precipitation and streamflow observations are pivotal, as they represent cause-and-effect in the context of system dynamics. Long-term experimental data from well-studied research catchments, and data from operational monitoring networks, have thus long been the cornerstone of the hydrological sciences (Tetzlaff et al., 2017). The relevance of observed data and research observatories cannot be overemphasised, particularly due to the invalidity of stationarity assumptions (Milly et al., 2008) in the face of anthropogenic climate change and its impacts on water-related hazards and availability.

As the availability and quality of observations crucially constrain the "realism" of a hydrological model and thus the accuracy of predictions, data scarcity impedes accurate modelling and inference of hydrological processes. Global reanalysis products (Muñoz-Sabater et al., 2021; ONOGI et al., 2007; Rienecker et al., 2011) can potentially, if of sufficient quality, complement

the few existing ground-based observations by offering a valuable alternative when exhaustive local observations are not available. Further, they play a pivotal role in hydro-climatic research (Alexopoulos et al., 2023; Gu et al., 2023), by providing a consistent, long-term view of the state of the global climate system via the assimilation of measurements and monitoring data into numerical weather models.

35 While previous studies (Essou et al., 2016; Tarek et al., 2020) have already shown the value of using reanalysis data as estimates for meteorological forcing data in regions with little or sparse ground-based weather station data, serious concerns about their quality remain when used in the context of hydrological modelling. The main issues include (Tarek et al., 2020) (i) regional variations in data quality and (ii) limited representation of local hydro-meteorological processes, with both of these impacting/biasing model structures and simulated states and fluxes. Systematic biases are also critical obstacles to the broader

40 applicability of such products (Clerc-schwarzenbach et al., 2024). In the case of ERA5-Land, a component of the Copernicus Climate Change Service (C3S) provided by the European Centre for Medium Weather Forecasting (Muñoz-Sabater et al., 2021), there is a tendency to significantly overestimate potential evapotranspiration (Clerc-schwarzenbach et al., 2024; Kratzert et al., 2023; Xu et al., 2024). Deficiencies have also been documented in the representation of convective storms (Essou et al., 2016; Taszarek et al., 2021) with subsequent underestimation of precipitation magnitudes and intensities (Manoj

45 J et al., 2024).

Because precipitation forcings data plays a crucial role in rainfall-runoff modelling, several methods (Yumnam et al., 2022) have been suggested for correcting precipitation data. These range from the use of storm multipliers (Sun and Bertrand-Krajewski, 2013) to station-wise correction of data using a gauge-based precipitation network (Cornes et al., 2018). However, gauge-based methods require a sufficient number of weather stations (Agarwal et al., 2020), which is often not the case for

50 most regions around the world. As seen from previous experience, it is usually the rainstorm events occurring in poorly observed areas that lead to high impacts (Borga et al., 2008). This is particularly true for flash floods in response to convective storm activity (Manoj J et al., 2024; Meyer et al., 2022; Villinger et al., 2022) and well related to the classical "Predictions in Ungauged Basins - PUB problem" (Sivapalan et al., 2003). To overcome this problem, and in line with Kirchner's (2008) work on "doing hydrology backwards", this paper explores options for inverse estimation of precipitation using the information

55 contained in observed streamflow.

While the classical "forward rainfall-runoff generation problem" has received considerable attention over various decades (Montanari et al., 2013; Sivapalan et al., 2003), a smaller subset of studies (Brocca et al., 2013; Kirchner, 2009; Kretzschmar et al., 2014; Krier et al., 2012; Teuling et al., 2010) has investigated the feasibility of tackling the inverse problem more efficiently. Kirchner (Kirchner, 2009) reported an early and successful attempt to infer catchment average rainfall and

60 evaporation time series from streamflow fluctuations and inspired several investigations examining the advantages and limitations of doing 'hydrology backwards' in diverse catchments (Krier et al., 2012; Teuling et al., 2010). Although these studies have established a robust mathematical foundation for addressing the inverse hydrological problem, they were limited

to smaller, well-monitored research catchments. This raises questions about the applicability of this approach to larger catchments as well as to smaller, non-experimental ones.

65 Note that inversions of the catchment water balance are inherently ill-posed, making it near impossible to find a unique solution (Bishop, 2006). Adopting the concept of micro- and macro-states from statistical mechanics (Zehe and Blöschl, 2004), we argue that the exact micro-state, i.e. the "true" space-time pattern of precipitation in the catchment, is neither uniquely identifiable nor observable. However, we conjecture that the catchment-average precipitation can be inversely identified based on the corresponding streamflow response. As streamflow remains a non-linear convolution of the catchment-average effective

70 precipitation, we propose that machine learning is well suited to this problem. Deep learning has recently revolutionised almost all fields of the natural sciences and engineering, showing great promise in solving a wide range of inverse problems, especially those related to imaging (Ongie et al., 2020). It has also been argued that such models can provide meaningful and general benchmarks for hypothesis testing (Klotz et al., 2022; Nearing and Gupta, 2015) and afford powerful avenues for generalisation using large datasets (Ralf Loritz et al., 2024).

75 The overall objective of this study is to 'do hydrology backwards' using a regional-scale long short-term memory (LSTM) network model trained on large-scale hydrological datasets using the ERA5 Land precipitation product (Muñoz-Sabater et al., 2021) as a target. The underlying research question is, "How much information about the catchment-average effective precipitation is effectively encoded in the variability of the streamflow time series observed at the outlet?" To answer this question, we first investigate whether the approach can accurately replicate the spatial gradients of the original forcing reanalysis dataset across European catchments for an unseen testing period. We then examine how the inverse model performs

80 when moving to much smaller (50-200 km^2) out-of-sample catchments. Here, we compare LSTM-based inverse estimates during flood events to the original reanalysis product (ERA5 Land) and rain gauge-based observational estimates over the same region (E-OBS: Cornes et al., 2018). Finally, we use the HBV conceptual hydrological model (BERGSTRÖM And FORSMAN, 1973) and the spatially-distributed, process-based CATFLOW model (Zehe et al., 2001) to assess the quality of

85 the precipitation estimates for forward modelling of streamflow and soil moisture dynamics, respectively.

2 Data and Methods

2.1 Model Configuration

LSTMs (Hochreiter, 1998) are a special type of recurrent neural network that makes use of cell states and so-called 'gates' to control the information flow through the network. The LSTM model used in this study extends upon the work of

90 (Kratzert et al., 2018) and (Espinoza et al., 2024). The LSTM architecture, which is commonly used for streamflow simulation in hydrology (Kratzert et al., 2018), uses a sequence of meteorological variables, such as precipitation and temperature as dynamic inputs, along with catchment attributes as static features, to predict the corresponding streamflow. In our setting, to establish an inverse model, we use the same general model architecture as in previous studies (Espinoza et al., 2024; Ralf

Loritz et al., 2024). The key difference is that future streamflow is now used along with other dynamic and static data as inputs
95   in order to estimate the precipitation forcings of the catchments. To account for the time lag between precipitation and discharge response observed at the catchment outlet, the model was provided with a 7-day lead time series for discharge. We explored ranges of hyperparameter settings on a smaller subset of the training dataset to establish relatively stable hyperparameter configurations (Fig. S1 in Supplementary Information), finally setting them according to (Espinoza et al., 2024) with a reduced number (5) of training epochs. Table S1 in the Supplementary information indicates the values used for
100   the LSTM network hyperparameters. Mean squared error was used as the training loss function. The codes for model building and training can be found online (Manoj J, 2024). The LSTM was trained as a regional model (single network trained on all available catchments) based on the openly available datasets detailed in the next section (Section 2.2).

For forward hydrological modelling using the inversely-generated precipitation timeseries estimates, we use two hydrological models - the lumped conceptual HBV model (Hydrologiska Byråns Vattenbalansavdelning: BERGSTRÖM and FORSMAN,
105   1973) and the spatially distributed process-based CATFLOW model (Zehe et al., 2001). The HBV model (Seibert, 2005) used in this paper requires precipitation (ERA5 Land/LSTM simulated), potential evapotranspiration, and air temperature (ERA5 Land) as inputs. We follow the recommendations of Clerc-Schwarzenbach et al. (2024), similar to that of Loritz et al (2024), for the calculation of potential evapotranspiration, and use the temperature-based Hargreaves formula detailed by Adam et al. (2006). For the CATFLOW model, we utilise the catchments, hillslope structures and soil/plant parameters from Manoj J et al
110   (2024).

2.2 Data sets

This study utilized the Caravan dataset (Kratzert et al., 2023) to investigate our hypothesis regarding the inverse identifiably of precipitation from information about discharge dynamics. We trained our model on European catchments from the GRDC-Caravan (Färber et al., 2023) community extension and the original Caravan dataset, which includes catchments
115   from CAMELS-GB (Coxon et al., 2020). The Caravan dataset uses the ERA5 Land (Muñoz-Sabater et al., 2021) as meteorological forcing, while the catchment attributes include data from HydroATLAS (Linke et al., 2019). The discharge data is tapped from relevant state and national authorities and is accessible as open datasets. Figure S2 in the Supplementary information depicts the study catchments (1804 in total) in the training dataset. The LSTM model was trained to predict daily ERA5 Land precipitation time series from discharge data along with meteorological forcings (temperature, surface net solar
120   radiation and surface net thermal radiation – all from ERA5 Land) and five catchment static attributes (area, p_mean, ele_mt_sav, frac_snow, pet_mm_syr: Kratzert et al., 2023).

We chose a training period of around 25 years between 01 October 1980 to 30 September 2005. Following the best practices in data-based modelling, the model was tested on an unseen testing period between 2006 and 2020 (2015 for CAMELS-GB catchments due to data unavailability). To investigate its generalizability across scales, we also tested the model on four
125   catchments (Fig. S3) that were not included in the original training set (Section 2.3.2). For the out-of-sample test, we made

use of data from the Caravan Spain (Casado Rodríguez, 2023) and Caravan Switzerland (Höge et al., 2023) extensions, in addition to data from local data providers in Germany (Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg—LUBW) and Luxembourg (Nijzink et al., 2024). The observational E-OBS precipitation product (v27.0 - Cornes et al., 2018), which uses the station network of the European Climate Assessment & Dataset (ECA&D) project, was

130   used as another benchmark for the evaluation of model performance.

2.3 Experimental Design
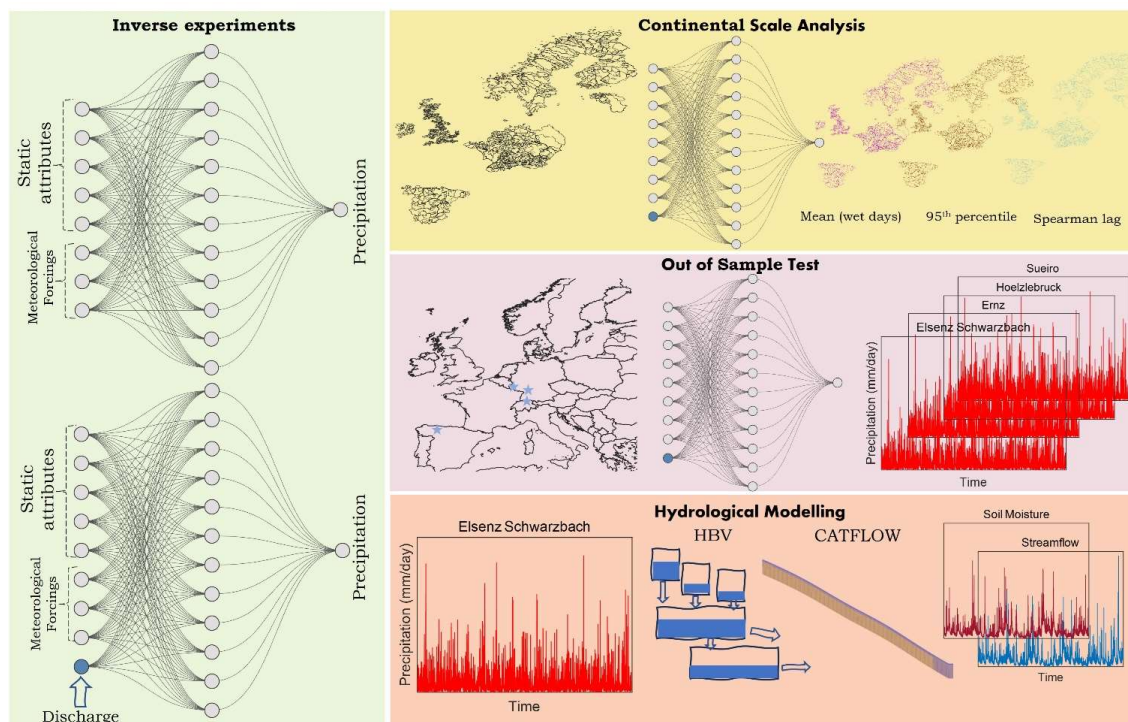
Figure 1 Schematic representation of our methodological approach.

2.3.1 Exploring information about precipitation in streamflow

135   To shed light on the value of discharge for inversely predicting precipitation, we conducted a virtual experiment in which two LSTM models (Fig. 1) were trained using the same catchments and training period. The first model (without_discharge) used only meteorological time series (air temperature, solar and thermal radiation) and static attributes (area, p_mean, ele_mt_sav, frac_snow, pet_mm_syr: Kratzert et al., 2023), while the second model (with_discharge) included lagged discharge as an additional input variable. Both models were trained to predict daily catchment average precipitation

140 sums (ERA5 Land). Therefore, we only deal with spatially averaged timeseries for precipitation, assuming that these values represent the effective precipitation over the entire catchment.

We then used the trained regional-scale model (with_discharge) to predict the precipitation time series inversely for all the test catchments over the unseen testing period and evaluated those using the wet day mean, 95th percentile and Spearman rank autocorrelation (one day lag) for each catchment, and then compared them to the values from ERA5 Land (used for

145 training the model) and E-OBS (observational product) at the continental scale.

2.3.2 Out of sample precipitation inversions and their quality

We further tested the feasibility of knowledge transfer to out-of-sample catchments and used the same regional-scale model to inversely predict the intensity of driving rainstorms for selected flood events in four hydro-climatically diverse and much smaller catchments (not included in the original training dataset). These catchments (Table 1 & Fig. S3) were chosen

150 based on the severity of the flooding and on the apparent inability of ERA5 Land forcings to accurately represent the storms that triggered the flood events.

Table 1 Attributes for the four catchments used for out-of-sample testing.

Catchment	Country	Area (km^2)	Mean precipitation (mm/day)	Mean potential evapotranspiration (mm/year)	Mean elevation (m)
Elsenz-Schwarzbach	Germany	196.5	2.51	812.85	246.7
Ernz	Luxembourg	69.3	2.31	724.04	345.5
Sueiro	Spain	132.5	3.31	873	381
Hoelzlebruck	Germany	47.1	4.14	658	980

2.3.3 The potential of inverted precipitation for forward modelling

155 To evaluate the value of generated precipitation data for forward modeling of streamflow, we calibrated the HBV conceptual hydrological model (Hydrologiska Byråns Vattenbalansavdelning: BERGSTRÖM and FORSMAN, 1973) over the Elsenz Schwarzbach catchment using both the original ERA5 Land and the LSTM-generated precipitation timeseries and compared the evaluation period performance of both model versions. For both runs, the model was calibrated using a training period from 01 January 2000 to 31 December 2010, while the models were evaluated between 01 January 2011 to 31 December

160 2016.

Complementary to streamflow modelling, the performance of a hydrological model can also be judged by how well it replicates the catchment dynamics of a region. Soil moisture is a key variable controlling the partitioning of net radiation into sensible and latent heat (Seneviratne et al., 2010) or overland flow during a rainstorm (Zehe and Blöschl, 2004). We thus used each

165   precipitation estimate (LSTM and ERA5 Land) to run the process-based hillslope scale model CATFLOW, using a setup from Manoj J et al. (2024) used for predictions of local floods in 2016. Here, we focused on one of the headwater sub-catchments (Catchment W32 in Fig. S5) within the Elsenz Schwarzbach (Fig. S3). The model simulated the period from 01 January 2008 to 31 December 2015 using each of the ERA5 Land and LSTM precipitation estimates, and the corresponding spatially averaged soil moisture states were compared against several soil moisture reanalysis products (due to the unavailability of

170   observed data). These include a) ERA5 Land: Muñoz-Sabater et al., 2021) b) GLDAS (NASA Global Land Data Assimilation System, GLDAS-2.2 GRACE DA: Li et al., 2019) and c) MERRA (Modern-Era Retrospective analysis for Research and Applications version 2 – tavg1_2d_lnd_Nx: Global Modeling And Assimilation Office, 2015).

3 Results

3.1 The information contained in streamflow about precipitation
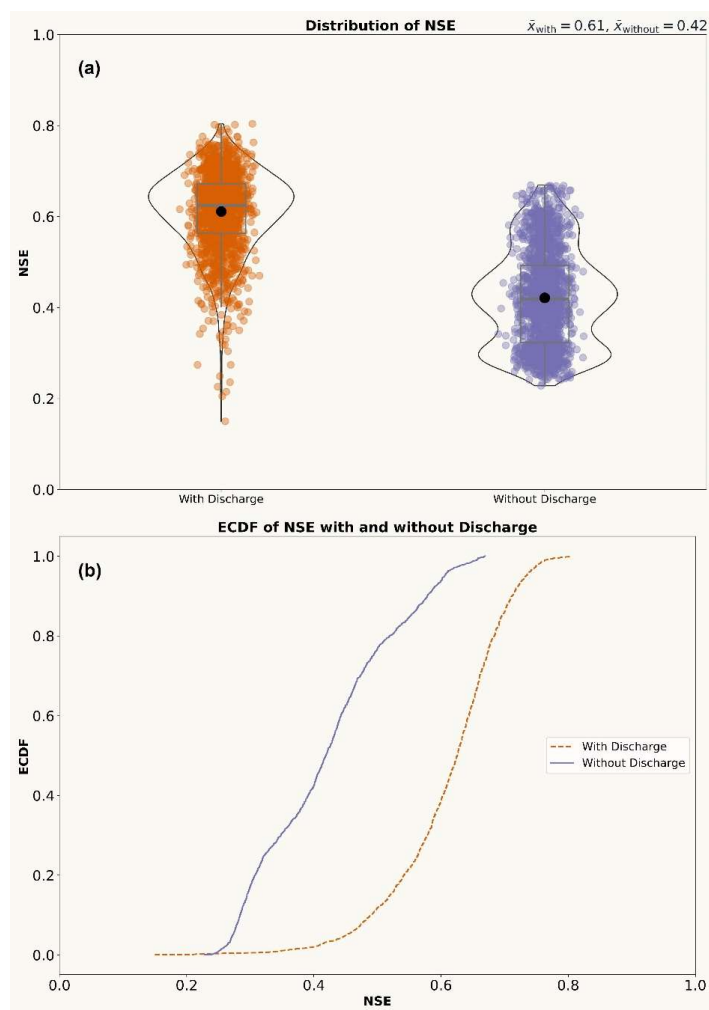
Figure 2 Comparison of the mean performance of the two regional scale LSTM models (with_discharge and without_discharge). (a) Figure 2A depicts violin plots with included boxplots showing the distribution of performance (quantified by comparing the LSTM model simulated precipitation series to the original ERA5 Land timeseries over the testing period: NSE) (b) Figure 2B displays cumulative distribution plots for the performance of the two models.

Figure 2A shows a comparison of mean performance of the LSTM estimates of precipitation to the original ERA5 product. These results reveal that the streamflow response contains useful information about the underlying causes of precipitation. The model "*with_discharge*" outperforms the model "*without_discharge*" not only on average but also with respect to the best-performing catchments. The median NSE metric value (Nash and Sutcliffe, 1970) for the regional LSTM model across the 1804 study catchments is about 20% higher when discharge is used as an additional predictor than when it is not. Further, the distribution of performance for the model *with_discharge* is unimodal with moderate skewness, whereas the distribution for the model *without_discharge* displays a bimodal distribution with heavy tails. Figure 2B shows the empirical cumulative probability plots, which reveal a steeper curve for the *without_discharge* model, indicating a lower variance in performances, in contrast with the flatter curve for the *with_discharge* model, which indicates a higher variance. Further, the *with_discharge* curve is clearly shifted towards a consistently higher performance across all quantiles. However, it is also observed that in a few outlier cases, discharge information has worsened the performance – likely due to the poor quality of streamflow data in these catchments.

3.2 Unraveling the Continental Scale Characteristics
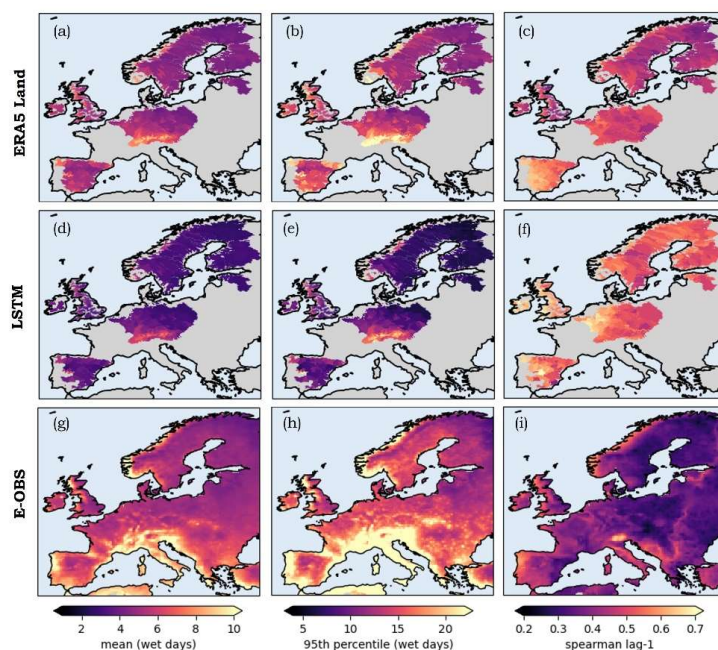
Figure 3 The spatial patterns of the different time series metrics (wet day mean, 95th percentile and Spearman autocorrelation values) over the study catchments for the three different sources ERA5 Land (top row): a) to c), LSTM model (middle row): (d) to (f) and E-OBS (bottom row): (g) to (i) from 2006 to 2020 (2015 for CAMELS-GB catchments).

200 To examine the characteristics of the simulated time series over the testing period in detail, we computed three timeseries measures (mean on wet days, 95th percentile limits, and Spearman autocorrelation values) across all the catchments, and show the results in Fig. 3.

The continental-scale analysis reveals distinct patterns for the major European climatic regions. Similar to the ERA5 Land and E-OBS products, the LSTM model (middle row) preserves spatial gradients for the wet day (daily precipitation >1 mm) mean

205 precipitation. Higher daily average values are observed towards the Alps, the Carpathian Mountain ranges, and the coast of Norway, consistent with the climatology of these regions. However, the model systematically underestimates absolute values, as evident from the scatterplot shown in Fig. 4.

A comparison with the total daily means (including both rainy and non-rainy days; Fig. S4) shows that this underestimation is particularly severe while considering only rainy days. For the 95th percentile of wet days, we again see a robust representation

210 of the spatial differences, along with an underestimation of the magnitudes (Fig. 3b-h). The Spearman autocorrelation coefficient values (with one day lag: Fig 3i-h) indicate that while the model underestimates the mean and 95th percentile limits, it overestimates the lag coefficient (which indicates the persistence in the precipitation time series) compared to the ERA5 Land time series. In addition, we also see that the ERA5 Land largely matches with the precipitation field's characteristics (wet day mean and 95th percentile limit) as in the observational E-OBS product.

215 The higher autocorrelation values for both LSTM and ERA5 Land may arise from model products incorporating catchment persistence, unlike the gridded observational E-OBS data. In the case of the LSTM model, the even higher values are likely due to the inclusion of highly auto correlated streamflow data, which adds redundancy or a longer memory.
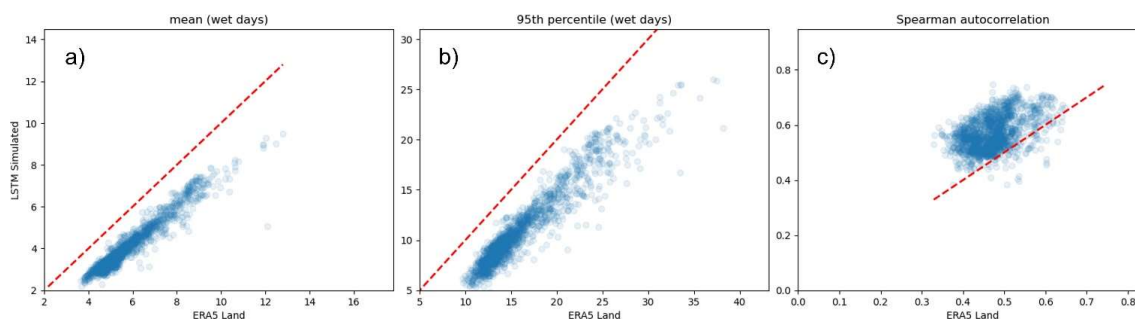
Figure 4 Scatterplots for the three timeseries measures (wet day mean, 95th percentile and spearman lag) between ERA5 Land and
220 **LSTM Simulated. Each point represents a single catchment within the dataset. A 1:1 line (shown as a red dotted line in Fig) indicates overestimation/underestimation bias.**

225   **3.3 Out of sample predictions**
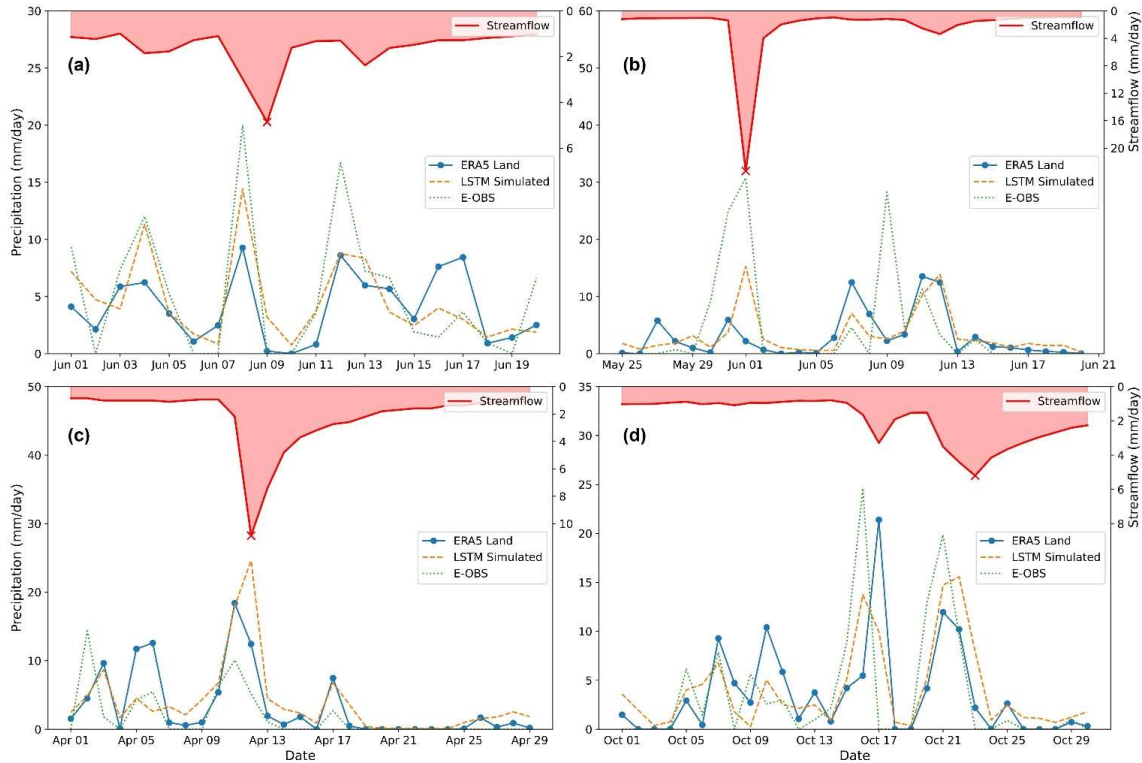
Figure 5 Precipitation estimates for flood events at the four out of sample catchments: (a) – Elsenz Schwarzbach, (b) – Ernz, (c) – Sueiro, (d) – Hoelzlebruck. The red line indicates the observed daily streamflow (with the day of the flood indicated by a cross). The orange curve denotes the precipitation amount predicted by the *with_discharge* LSTM model, while the blue line depicts the original
230   **ERA5 Land time series, and the green line indicates the estimate from the gauge based E-OBS product.**

Figure 5 shows predicted event precipitation values over time for the four out-of-sample catchments. Again, we compare the inversely modelled values to the original ERA5 Land (used for training) and the gauge-based E-OBS product. Table 2 lists the peak storm precipitation values reported by the different products along with the recorded flood values (both normalised to the catchment area in mm/day). Also shown are the storm runoff coefficients for the respective events based on the different

235   precipitation estimates and discharge data.

Figure 5A represents the summer flood in June 2016 in the Elsenz Schwarzbach catchment in Germany. This annual flood event was triggered by a series of convective rainfall events caused by persistent atmospheric conditions in Germany during the summer of 2016. Localised rainfall totals exceeded 100 mm in some catchments (Bronstert et al., 2018), triggering widespread flash floods. Our previous work (Manoj J et al., 2024) indicated that ERA5 Land could not accurately replicate the

240   characteristics of the convective storm that caused this annual flood event over the Elsenz Schwarzbach catchment. A
        comparison of LSTM-simulated precipitation values revealed closer estimates to those reported in the observational E-OBS
        product. When considering the E-OBS as the ground truth, the relative underestimation error in precipitation reduced from
        around 100% (ERA5 Land) to 40% (LSTM). The runoff coefficient for the event also decreased from 35% (ERA5 Land) to
        around 23% (LSTM), which is consistent with estimates from Manoj et al. (2024).

245   Next, the LSTM model was used to estimate precipitation for another convective episode over the Ernz Catchment in
        Luxembourg (Fig. 5B) in the summer of 2018. There was a noticeable improvement in the precipitation time series for both
        timing and peak storm values compared to ERA5 Land. While ERA5 Land completely missed this storm, the LSTM model
        was able to represent the sharp rise and descent of the curve. However, the runoff coefficients and peak storm values (Table
        2) indicate that the LSTM model underestimates the true precipitation amount. In the third catchment (Sueiro: camelses_1414
250   from Caravan Spain extension), the LSTM estimate for storm forcing was higher than ERA5 Land and E-OBS (Fig. 5C). The
        corresponding runoff coefficients underline the reliability of the storm prediction from LSTM (0.37) compared to E-OBS
        (1.05).

        In the Hoelzlebruck catchment (camelsch_4003 from Caravan Switzerland extension), two consecutive events occurred in
        October 2014. ERA5 Land was better than the LSTM model in capturing the initial event magnitude, while the LSTM model
255   had better timing accuracy for the events. For the second event, which was the annual flood event, the LSTM model, which
        incorporated streamflow information, was again able to reduce the relative errors in precipitation magnitudes (Fig. 5D)

Table 2 Event characteristics for the four out of sample catchments

Event Characteristics		Elsenz-Schwarbach	Ernz	Sueiro	Hoelzlebruck
Precipitation (mm/day)	ERA5 Land	10.62	9.15	39.8	28.55
	LSTM	16.45	23.49	64.83	44.68
	E-OBS	20.03	49.43	22.54	42.33
Discharge (mm/day)		3.75	27.12	23.68	20.85
Runoff Coefficient (-)	ERA5 Land	0.35	2.96	.60	0.73
	LSTM	0.23	1.15	.37	0.47
	E-OBS	0.19	0.55	1.05	0.49

260

3.4 Forward Hydrological Modelling

The precipitation estimates generated by the LSTM model were then used to run two classical hydrological models (HBV and CATFLOW) in a forward manner. Figure 6 illustrates that the HBV model, which utilized the inverted precipitation estimates, performed slightly better (NSE = 0.64) during the evaluation period compared to the model driven by the ERA5 Land (NSE = 0.57). To gain a better understanding of the differences between the models, we visually examined the results for three individual flood events, as shown in Fig. 6A-C.

265
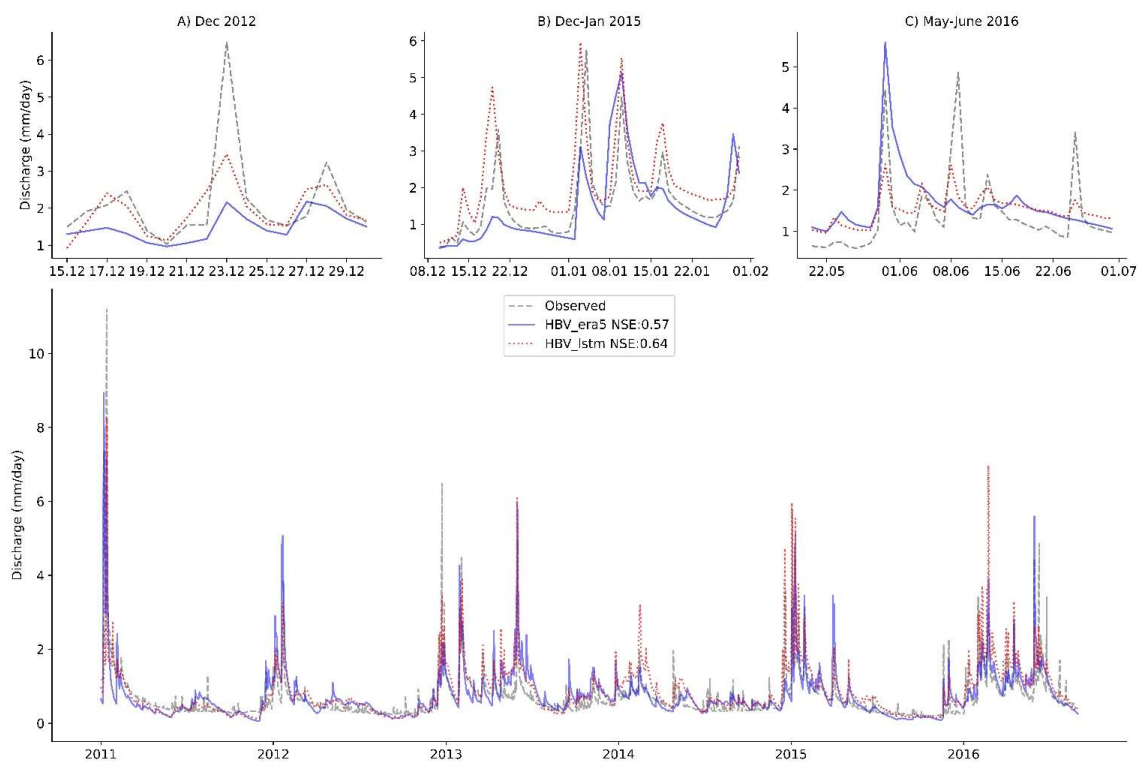
Figure 6 Observed and simulated runoff (using the HBV model) at the Elsenz Schwarzbach catchment. The blue line denotes the streamflow simulated using the ERA5 Land precipitation product, while the red curve depicts the simulations using the inversely-estimate precipitation obtained using the regional LSTM model. Moreover, three rainfall-runoff events are highlighted and displayed separately.

270

275

During the winter flood of December 2012 (23 Dec 2012, Fig. 6A), the model driven by ERA5 Land significantly underestimated both the peak and the volume of the flood event. When using LSTM-simulated precipitation, the relative peak

280   error decreased by nearly 25%. Similarly, the model runs using LSTM precipitation more accurately captured the pre-event conditions (18 Dec 2012) and the post-event conditions (28 Dec 2012). This aligns with findings from other studies (Berghuijs et al., 2019; Manoj J et al., 2023) that emphasize the importance of initial conditions for floods across Europe.

In the winter of 2015 (Fig. 6B), the model using LSTM precipitation once again demonstrated better performance (albeit with overestimation errors). During the convective summer storm event in 2016 (Fig. 6C), neither model run successfully captured

285   the flashy runoff response. Although the model that utilized ERA5 Land input predicted an earlier flood event in May 2016 with an overestimation bias, it did not accurately depict the dynamics of the annual flood event occurring a few days later. In contrast, the model with LSTM-generated precipitation generally performed better in capturing both the magnitude and timing of the smaller storm peaks as well as the annual flood event on June 8, 2016.

To understand the evolution of soil moisture dynamics while using the LSTM-based precipitation estimates, we conducted a

290   hillslope-scale CATFLOW model simulation (Loritz et al., 2017; Manoj J et al., 2024) in one of the headwater catchments in Elsenz Schwarzbach (ERA5 Land vs LSTM). The pairwise correlation values, as shown in Fig. 7, indicate that the use of the LSTM-based precipitation estimates does not lead to a loss of information regarding soil moisture dynamics in the region. In fact, we observe a slight increase in correlation when comparing the inversely derived precipitation estimates (referred to as CATFLOW_lstm) to MERRA and GLDAS, in contrast with the correlation obtained for the run with ERA5 Land (referred to

295   as CATFLOW_era5). As expected, the correlation value for the ERA5 Land run is slightly higher when assessed against soil moisture from the same ERA5 Land dataset, which may be attributed to model biases arising from using the same dataset for both precipitation and soil moisture.
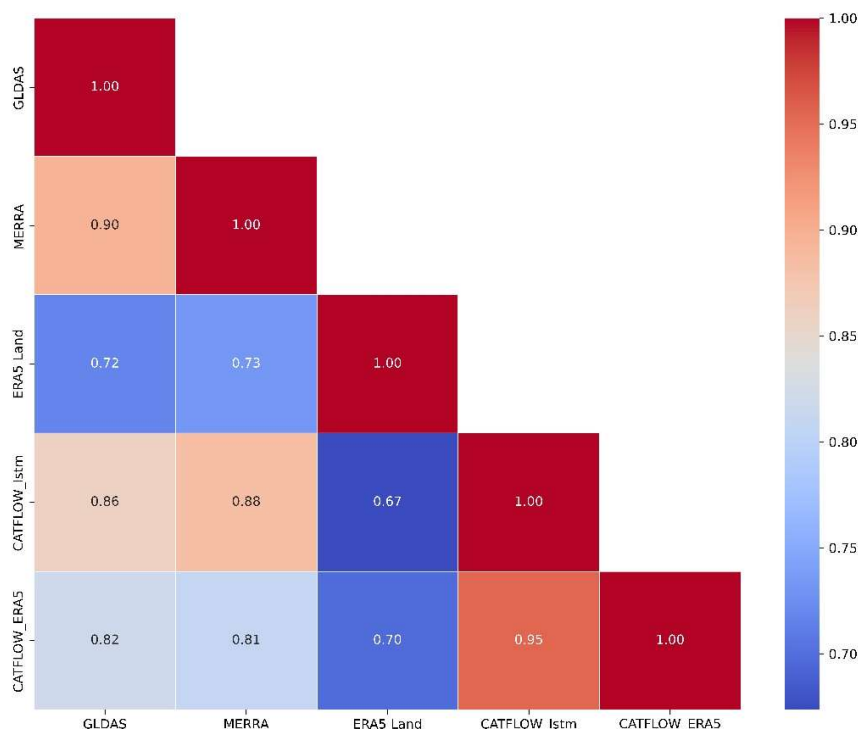
Figure 7 Correlation matrix plot illustrating the pairwise correlations between the different soil moisture estimates - a) GLDAS
300 **(NASA Global Land Data Assimilation System, GLDAS-2.2 GRACE DA: Li et al., 2019), b) MERRA (Modern-Era Retrospective analysis for Research and Applications version 2 – tavg1_2d_lnd_Nx: Global Modeling And Assimilation Office, 2015), c) ERA5 Land: (Muñoz-Sabater et al., 2021), d) CATFLOW_lstm: model run using inversely estimated precipitation estimate from the LSTM model and e) CATFLOW_ERA5: model run using precipitation estimate from ERA5 Land product.**

4 Discussion

305 ### 4.1 Better precipitation estimation using discharge

Overall, our study reiterates that streamflow data can be exploited to obtain useful information about the nature of catchment-scale effective precipitation: we can thus invert the cause using the effect as input to an LSTM. This is in line with, and steps beyond, previous studies (Brocca et al., 2013; Kirchner, 2009; Kretzschmar et al., 2014; Krier et al., 2012; Teuling et al., 2010) that explored the possibility of doing hydrology backwards using experimental catchments. Here, we successfully expanded
310 this idea to large samples, cutting across the wide range of hydro-climatic conditions that characterise Europe. We found a largely 'normal' distribution of performance, with a few outliers, the latter indicating possible poor quality of discharge data.

The spatially averaged NSE performance metric of 0.61 indicates that 61% of the variance of catchment-averaged effective rainfall is explained by the model over the unseen evaluation period. Note that an NSE value of 0 signifies a model that performs no better than using the long-term average while increasing values indicate progressively better performance.

315 Interestingly, the LSTM model systematically underestimates rainfall values, which is consistent with findings from other studies (Wang et al., 2024). In part, this could be attributed to the use of mean squared error as the loss function during training (Gupta et al., 2009). Also, in most cases, only a part of the event runoff comes from the direct runoff, while the rest is baseflow and is hence not directly related to the event precipitation. Hence, this consistent underestimation also seems physically plausible. However, more research has to be carried out to disentangle the impact of such catchment losses and the interplay

320 between various factors involved in total runoff production on the inversion problem.

Our main goal was to see if we could leverage the potential of streamflow data for a better representation of flooding and other high-impact short-time-scale events in the reanalysis product. Hence, we used the same regional-scale model to inversely predict the forcing storm precipitation values for flood events in four hydroclimatically diverse catchments. Moreover, inverted precipitation could significantly enhance paleo-hydrological studies by using reconstructed streamflow data from

325 dendrochronology to infer paleo-precipitation (Razavi et al., 2015).

Initially introduced to model the land surface components of global climate models (MANABE, 1969), land surface models are, today, commonly utilised in studies examining energy and water balance processes and assessing land-atmosphere feedback effects. The development of these models (Blyth et al., 2021) has been influenced by collaboration between climatic, hydrological, agricultural, and ecosystem communities, intercomparison projects and relevant technical integrations.

330 Furthermore, such modelled and reanalysis products have delivered commendable outcomes when capturing climate and weather extremes on regional scales. However, they remain largely unusable for estimating the impacts of smaller scale hydrological events (Seneviratne et al., 2021). This is primarily due to the mischaracterization of precipitation forcings for such spatial scales in these models. Novel data-driven approaches, as shown in this study, have the potential to better meteorological products by bridging the gap between the precipitation and discharge modelling worlds.

335 **4.2 Catchment as a functional unit**

In the introduction, we argued that the catchment scale is crucial for improving our understanding of the factors that drive the water cycle and representing them more accurately in reanalysis products. Our findings across the four catchments highlight the benefit of using streamflow variations to rectify precipitation estimates. By leveraging the generalisation capabilities of the data-driven LSTM model, we successfully transferred knowledge across different scales, indicating important implications for

340 addressing the ever-evolving challenge of predictions in ungauged basins (PUB: Hrachowitz et al., 2013)

Although this approach can only be applied after the event has taken place, it has implications for generating coherent long-term statistical records for catchment forcings, which could be used for the design of small- to medium-purpose water resource projects. Using the streamflow fluctuations, it would be possible to identify localised rainfall cells or snowfall events that are

poorly captured by traditional rain gauges (Kretzschmar et al., 2014). The approach also has potential for evaluating long-term

345   rainfall estimates from Global Circulation Models for specific catchments using information about hydrological conditions (Fujihara et al., 2008).

While the LSTM-based precipitation estimates improved the representation of most events, there were still instances where the original ERA5 Land provided better accuracy for peak flood magnitudes; this highlights the need for a blended approach that incorporates additional information rather than completely replacing one product with another. In regions around the

350   world, the wealth of streamflow information remains underutilised in this aspect. For Germany alone (R Loritz et al., 2024), there are more than 1500 streamflow gauges, which represent a significantly higher representative area compared to precipitation stations.

The forward exercise using the HBV model showed that the precipitation estimates after inversion enhanced mean performance for streamflow simulation and helped improve the modelling of extreme individual floods. The ability to match the hydrograph

355   differed between the different seasons. Compared to the storage-controlled winter floods (Dunne and Kirkby, 1978) , summer floods in these regions are usually driven by Hortonian flow (Horton, 1932) in response to high-intensity rainfall during convective storms. Previous studies (Kirchner, 2009; Krier et al., 2012) have discussed such storage-controlled dynamics and their impact on the inversion problem.

Previous experiences at the event scale (Beauchamp et al., 2013; Zehe and Blöschl, 2004) have also shown that inferring the

360   antecedent soil moisture conditions remains a key challenge for accurate and reliable flood simulations. By utilising the process-based CATFLOW model for soil moisture simulations in a small headwater catchment, we achieved high correlation values using the inverse precipitation estimate. This suggests that the approach can help represent the catchment's overall water dynamics and has the potential for reliable flood design estimations at the event scale, particularly in data-scarce regions.

4.3 Limitations

365   It is important to stress that, as for any data-driven study, the results of our work are contingent on the quality of the training dataset. While we are aware of better regional products for individual countries, ERA5 Land provides consistent global coverage, and a permissive data sharing policy makes it one of the obvious choices for a continental scale modelling exercise. To evaluate the applicability of the commonly used LSTM network architecture, we decided to use the same architecture previously employed in hydrological studies instead of creating an experimental design with modified individual layers and

370   training functions for inverse modelling. It is evident that exploring the impacts of different loss functions and deep learning model architectures like transformers would help advance the methodology discussed in this paper. This approach could also shed light on best-suited algorithms for the problem but is beyond the scope of the present work. The choice of Mean Squared Error (MSE) as the training function and Nash Sutcliffe Efficiency (NSE) as a performance metric is motivated by its success and applications in the forward problem (streamflow prediction), but this adds its own biases to the modelling exercise.

375   Additionally, we tried to complement this by calculating various other time series metrics commonly used in

hydrometeorological studies. The four events for out-of-sample tests across various catchments were chosen based on the severity of the floods and ERA5 Land's inability to capture the characteristics of the driving storms. The choice of the hydrological models and calibration period also adds uncertainty to the forward simulations.

5 Conclusions

380   Our main hypothesis was supported by the findings, which demonstrated that discharge has unused potential and can be inversely assimilated to adjust precipitation estimates derived from reanalysis products, while machine learning models are key to expanding this effort to large data sets spanning the scale of entire continents. The continental-scale analysis revealed that while the spatial gradients are well represented across Europe for the various time series attributes, there remain significant underestimation biases compared to the original reanalysis product. Insights from the out-of-sample catchments provided

385   valuable information about the applicability of our method for estimating flood forcings and the generalizability of the model. Additionally, we have shown that LSTM-based inverted precipitation estimates can improve forward modelling of both streamflow and soil moisture dynamics, illustrating how the information gained can be integrated into existing modelling strategies.

390   **Code Availability**

The codes used to conduct the LSTM analysis in this study are based on the publicly available HY^{2}DL python library (https://github.com/KIT-HYD/Hy2DL) and can be accessed at https://doi.org/10.5281/zenodo.14161112 (Manoj J, 2024).

Data Availability

The Caravan dataset and related community extensions are publicly available at https://doi.org/10.5281/zenodo.10968468

395   (Kratzert et al., 2023) and https://github.com/kratzert/Caravan/discussions/10. The datasets generated as part of this publication can be found at https://doi.org/10.5281/zenodo.14161112 (Manoj J, 2024).

Author Contribution

AMJ designed the study and carried out all analysis and model simulations. Funding was acquired by EZ. The initial draft was prepared by AMJ, with all authors contributing to review and editing. RL, HG and EZ jointly supervised the work. All authors

400   have read and agreed to the current version of the paper.

Competing interests

At least one of the (co-)authors is a member of the editorial board of Hydrology and Earth System Sciences.

Acknowledgements

The authors acknowledge support by the state of Baden-Württemberg through bwHPC (High Performance Computing
405   Cluster). AMJ would like to thank Mr Eduardo Acuña Espinoza for helpful discussions regarding the HY^{2}DL python library
for deep learning methods.

Financial Support

AMJ would like to thank the German Research Foundation (DFG) for financial support (Implementation of an InfraStructure
for dAta-BasEd Learning in environmental sciences: ISABEL - 496155047).

410

References

Agarwal, A., Marwan, N., Maheswaran, R., Ozturk, U., Kurths, J., Merz, B., 2020. Optimal design of hydrometric station
networks based on complex network analysis. Hydrol. Earth Syst. Sci. 24, 2235–2251. https://doi.org/10.5194/hess-24-
2235-2020

415   Alexopoulos, M.J., Müller-Thomy, H., Nistahl, P., Šraj, M., Bezak, N., 2023. Validation of precipitation reanalysis products
for rainfall-runoff modelling in Slovenia. Hydrol. Earth Syst. Sci. 27, 2559–2578. https://doi.org/10.5194/hess-27-2559-
2023

Beauchamp, J., Leconte, R., Trudel, M., Brissette, F., 2013. Estimation of the summer-fall PMP and PMF of a northern
watershed under a changed climate. Water Resour. Res. 49, 3852–3862. https://doi.org/10.1002/wrcr.20336

420   Berghuijs, W.R., Harrigan, S., Molnar, P., Slater, L.J., Kirchner, J.W., 2019. The Relative Importance of Different Flood-
Generating Mechanisms Across Europe. Water Resour. Res. 55, 4582–4593. https://doi.org/10.1029/2019WR024841

BERGSTROM, S., FORSMAN, A., 1973. Development of a Conceptual Deterministic Rainfall-Runoff Model. Nord. Hydrol.
4, 147–170. https://doi.org/10.2166/nh.1973.0012

Bishop, C.M., 2006. Pattern recognition and machine learning. New York : Springer, [2006] ©2006.

425   Blyth, E.M., Arora, V.K., Clark, D.B., Dadson, S.J., De Kauwe, M.G., Lawrence, D.M., Melton, J.R., Pongratz, J., Turton,
R.H., Yoshimura, K., Yuan, H., 2021. Advances in Land Surface Modelling. Curr. Clim. Chang. Reports 7, 45–71.
https://doi.org/10.1007/s40641-021-00171-5

Borga, M., Gaume, E., Creutin, J.D., Marchi, L., 2008. Surveying flash floods: gauging the ungauged extremes. Hydrol.

Process. 22, 3883–3885. https://doi.org/10.1002/hyp.7111

430   Brocca, L., Moramarco, T., Melone, F., Wagner, W., 2013. A new method for rainfall estimation through soil moisture
          observations. Geophys. Res. Lett. 40, 853–858. https://doi.org/10.1002/grl.50173

Bronstert, A., Agarwal, A., Boessenkool, B., Crisologo, I., Fischer, M., Heistermann, M., Köhn-Reich, L., López-Tarazón,
          J.A., Moran, T., Ozturk, U., Reinhardt-Imjela, C., Wendi, D., 2018. Forensic hydro-meteorological analysis of an
          extreme flash flood: The 2016-05-29 event in Braunsbach, SW Germany. Sci. Total Environ. 630, 977–991.
435   https://doi.org/10.1016/j.scitotenv.2018.02.241

Casado Rodríguez, J., 2023. CAMELS-ES: Catchment Attributes and Meteorology for Large-Sample Studies – Spain.
          https://doi.org/10.5281/zenodo.8428374

Clerc-schwarzenbach, F.M., Selleri, G., Neri, M., Toth, E., Meerveld, I. Van, Seibert, J., 2024. HESS Opinions : A few camels
          or a whole caravan ? 1–29.

440   Cornes, R.C., van der Schrier, G., van den Besselaar, E.J.M., Jones, P.D., 2018. An Ensemble Version of the E-OBS
          Temperature   and   Precipitation   Data   Sets.   J.   Geophys.   Res.   Atmos.   123,   9391–9409.
          https://doi.org/10.1029/2017JD028200

Coxon, G., Addor, N., Bloomfield, J.P., Freer, J., Fry, M., Hannaford, J., Howden, N.J.K., Lane, R., Lewis, M., Robinson,
          E.L., Wagener, T., Woods, R., 2020. CAMELS-GB: hydrometeorological time series and landscape attributes for 671
445   catchments in Great Britain. Earth Syst. Sci. Data 12, 2459–2483. https://doi.org/10.5194/essd-12-2459-2020

Dunne, T., Kirkby, M.J., 1978. Field studies of hillslope flow processes 227–293.

Espinoza, E.A., Loritz, R., Chaves, M.Á., Bäuerle, N., Ehret, U., 2024. To bucket or not to bucket? Analyzing the performance
          and interpretability of hybrid hydrological models with dynamic parameterization. Hydrol. Earth Syst. Sci. 28, 2705–
          2719. https://doi.org/10.5194/hess-28-2705-2024

450   Essou, G.R.C., Sabarly, F., Lucas-Picher, P., Brissette, F., Poulin, A., 2016. Can precipitation and temperature from
          meteorological   reanalyses   be   used   for   hydrological   modeling?   J.   Hydrometeorol.   17,   1929–1950.
          https://doi.org/10.1175/JHM-D-15-0138.1

Färber, C., Plessow, H., Kratzert, F., Addor, N., Shalev, G., Looser, U., 2023. GRDC-Caravan: extending the original dataset
          with data from the Global Runoff Data Centre. https://doi.org/10.5281/zenodo.10074416

455   Fujihara, Y., Simonovic, S.P., Topaloğlu, F., Tanaka, K., Watanabe, T., 2008. An inverse-modelling approach to assess the
          impacts of climate change in the Seyhan River basin, Turkey. Hydrol. Sci. J. 53, 1121–1136.
          https://doi.org/10.1623/hysj.53.6.1121

Global Modeling And Assimilation Office, 2015. MERRA-2 tavg1_2d_lnd_Nx: 2d, 1-Hourly, Time-Averaged, Single-Level,
          Assimilation, Land Surface Diagnostics V5.12.4. https://doi.org/10.5067/RKPHT8KC1Y1T

460   Gu, L., Yin, J., Wang, S., Chen, J., Qin, H., Yan, X., He, S., Zhao, T., 2023. How well do the multi-satellite and atmospheric
          reanalysis   products   perform   in   hydrological   modelling.   J.   Hydrol.   617,   128920.
          https://doi.org/10.1016/j.jhydrol.2022.128920

Gupta, H. V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. J. Hydrol. 377, 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hochreiter, S., 1998. The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions. Int. J. Uncertainty, Fuzziness Knowledge-Based Syst. 06, 107–116. https://doi.org/10.1142/S0218488598000094

Höge, M., Kauzlaric, M., Siber, R., Schönenberger, U., Horton, P., Schwanbeck, J., Floriancic, M.G., Viviroli, D., Wilhelm, S., Sikorska-Senoner, A.E., Addor, N., Brunner, M., Pool, S., Zappa, M., Fenicia, F., 2023. Catchment attributes and hydro-meteorological time series for large-sample studies across hydrologic Switzerland (CAMELS-CH) 1–20.

Horton, R.E., 1932. The role of infiltration in the hydrology cycle. Trans. Am. Geophys. Union 446–460.

Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J.E., Gelfan, A., Gupta, H. V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., Cudennec, C., 2013. A decade of Predictions in Ungauged Basins (PUB)-a review. Hydrol. Sci. J. 58, 1198–1255. https://doi.org/10.1080/02626667.2013.803183

Kirchner, J.W., 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. Water Resour. Res. 45, 1–34. https://doi.org/10.1029/2008WR006912

Klotz, D., Kratzert, F., Gauch, M., Keefe Sampson, A., Brandstetter, J., Klambauer, G., Hochreiter, S., Nearing, G., 2022. Uncertainty estimation with deep learning for rainfall-runoff modeling. Hydrol. Earth Syst. Sci. 26, 1673–1693. https://doi.org/10.5194/hess-26-1673-2022

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018. Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. Hydrol. Earth Syst. Sci 22, 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

Kratzert, F., Nearing, G., Addor, N., Erickson, T., Gauch, M., Gilon, O., Gudmundsson, L., Hassidim, A., Klotz, D., Nevo, S., Shalev, G., Matias, Y., 2023. Caravan - A global community dataset for large-sample hydrology. Sci. Data 10, 61. https://doi.org/10.1038/s41597-023-01975-w

Kretzschmar, A., Tych, W., Chappell, N.A., 2014. Reversing hydrology: Estimation of sub-hourly rainfall time-series from streamflow. Environ. Model. Softw. 60, 290–301. https://doi.org/10.1016/j.envsoft.2014.06.017

Krier, R., Matgen, P., Goergen, K., Pfister, L., Hoffmann, L., Kirchner, J.W., Uhlenbrook, S., Savenije, H.H.G., 2012. Inferring catchment precipitation by doing hydrology backward: A test in 24 small and mesoscale catchments in Luxembourg. Water Resour. Res. 48, 1–15. https://doi.org/10.1029/2011WR010657

Li, B., Rodell, M., Kumar, S., Beaudoing, H.K., Getirana, A., Zaitchik, B.F., de Goncalves, L.G., Cossetin, C., Bhanja, S., Mukherjee, A., Tian, S., Tangdamrongsub, N., Long, D., Nanteza, J., Lee, J., Policelli, F., Goni, I.B., Daira, D., Bila, M., de Lannoy, G., Mocko, D., Steele-Dunne, S.C., Save, H., Bettadpur, S., 2019. Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges. Water Resour. Res. 55, 7564–7586. https://doi.org/10.1029/2018WR024618

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., Thieme, M., 2019. Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. Sci. Data 6, 283. https://doi.org/10.1038/s41597-019-0300-6

500 Loritz, R, Dolich, A., Acuña Espinoza, E., Ebeling, P., Guse, B., Götte, J., Hassler, S.K., Hauffe, C., Heidbüchel, I., Kiesel, J., Mälicke, M., Müller-Thomy, H., Stölzle, M., Tarasova, L., 2024. CAMELS-DE: hydro-meteorological time series and attributes for 1555 catchments in Germany. Earth Syst. Sci. Data Discuss. 2024, 1–30. https://doi.org/10.5194/essd-2024-318

Loritz, R., Hassler, S.K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J., Zehe, E., 2017. Picturing and modeling
505   catchments by representative hillslopes. Hydrol. Earth Syst. Sci. 21, 1225–1249. https://doi.org/10.5194/hess-21-1225-2017

Loritz, Ralf, Wu, C.H., Klotz, D., Gauch, M., Kratzert, F., Bassiouni, M., 2024. Generalizing Tree–Level Sap Flow Across the European Continent. Geophys. Res. Lett. 51. https://doi.org/10.1029/2023GL107350

MANABE, S., 1969. Climate and the Ocean Circulation 1 . Mon. Weather Rev. 97, 739–774. https://doi.org/10.1175/1520-
510   0493(1969)097<0739:catoc>2.3.co;2

Manoj J, A., 2024. Ash-Manoj/lstm_backward: Datasets and Code for Manoj J et al (2024). https://doi.org/10.5281/zenodo.14161112

Manoj J, A., Loritz, R., Villinger, F., Mälicke, M., Koopaeidar, M., Göppert, H., Zehe, E., 2024. Toward Flash Flood Modeling Using Gradient Resolving Representative Hillslopes. Water Resour. Res. 60. https://doi.org/10.1029/2023WR036420

515 Manoj J, A., Pérez Ciria, T., Chiogna, G., Salzmann, N., Agarwal, A., 2023. Characterising the coincidence of soil moisture – precipitation extremes as a possible precursor to European floods. J. Hydrol. 620, 129445. https://doi.org/10.1016/j.jhydrol.2023.129445

Meyer, J., Neuper, M., Mathias, L., Zehe, E., Pfister, L., 2022. Atmospheric conditions favouring extreme precipitation and flash floods in temperate regions of Europe. Hydrol. Earth Syst. Sci. 26, 6163–6183. https://doi.org/10.5194/hess-26-
520   6163-2022

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Climate change: Stationarity is dead: Whither water management? Science (80-.). 319, 573–574. https://doi.org/10.1126/science.1151915

Montanari, A., Young, G., Savenije, H.H.G., Hughes, D., Wagener, T., Ren, L.L., Koutsoyiannis, D., Cudennec, C., Toth, E.,
525   Grimaldi, S., Blöschl, G., Sivapalan, M., Beven, K., Gupta, H., Hipsey, M., Schaefli, B., Arheimer, B., Boegh, E., Schymanski, S.J., Di Baldassarre, G., Yu, B., Hubert, P., Huang, Y., Schumann, A., Post, D.A., Srinivasan, V., Harman, C., Thompson, S., Rogger, M., Viglione, A., McMillan, H., Characklis, G., Pang, Z., Belyaev, V., 2013. "Panta Rhei-Everything Flows": Change in hydrology and society-The IAHS Scientific Decade 2013-2022. Hydrol. Sci. J. 58, 1256–1275. https://doi.org/10.1080/02626667.2013.809088

530 Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M.,

Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13, 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Nash, J.E., Sutcliffe, J. V, 1970. River flow forecasting through conceptual models. Part I - a discussion of principles. J. Hydrol. 27, 282–290.

Nearing, G.S., Gupta, H. V., 2015. The quantity and quality of information in hydrologic models. Water Resour. Res. 51, 524–538. https://doi.org/10.1002/2014WR015895

Nijzink, J., Loritz, R., Gourdol, L., Zoccatelli, D., Iffly, J.F., Pfister, L., 2024. CAMELS-LUX: Highly Resolved Hydro-Meteorological and Atmospheric Data for Physiographically Characterized Catchments around Luxembourg. https://doi.org/10.5281/zenodo.13846620

Ongie, G., Jalal, A., Metzler, C.A., Baraniuk, R.G., Dimakis, A.G., Willett, R., 2020. Deep Learning Techniques for Inverse Problems in Imaging. IEEE J. Sel. Areas Inf. Theory 1, 39–56. https://doi.org/10.1109/jsait.2020.2991563

ONOGI, K., TSUTSUI, J., KOIDE, H., SAKAMOTO, M., KOBAYASHI, S., HATSUSHIKA, H., MATSUMOTO, T., YAMAZAKI, N., KAMAHORI, H., TAKAHASHI, K., KADOKURA, S., WADA, K., KATO, K., OYAMA, R., OSE, T., MANNOJI, N., TAIRA, R., 2007. The JRA-25 Reanalysis. J. Meteorol. Soc. Japan. Ser. II 85, 369–432. https://doi.org/10.2151/jmsj.85.369

Razavi, S., Elshorbagy, A., Wheater, H., Sauchyn, D., 2015. Toward understanding nonstationarity in climate and hydrology through tree ring proxy records. Water Resour. Res. 51, 1813–1830. https://doi.org/10.1002/2014WR015696

Rienecker, M.M., Suarez, M.J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M.G., Schubert, S.D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R.D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C.R., Reichle, R., Robertson, F.R., Ruddick, A.G., Sienkiewicz, M., Woollen, J., 2011. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. J. Clim. 24, 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Seibert, J., 2005. HBV light. HBV Light version 2 User's Man.

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I., Orlowsky, B., Teuling, A.J., 2010. Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Science Rev. 99, 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Seneviratne, S.I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. Di, Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S.M., Wehner, M., Zhou, B., 2021. Weather and Climate Extreme Events in a Changing Climate. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Clim. Chang. 2021 Phys. Sci. Basis. Contrib. Work. Gr. I to Sixth Assess. Rep. Intergov. Panel Clim. Chang. 366.

Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiondo, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S., Zehe, E., 2003. IAHS Decade on

565   Predictions in Ungauged Basins (PUB), 2003-2012: Shaping an exciting future for the hydrological sciences. Hydrol. Sci. J. 48, 857–880. https://doi.org/10.1623/hysj.48.6.857.51421

Sun, S., Bertrand-Krajewski, J.L., 2013. Separately accounting for uncertainties in rainfall and runoff: Calibration of event-based conceptual hydrological models in small urban catchments using Bayesian method. Water Resour. Res. 49, 5381–5394. https://doi.org/10.1002/wrcr.20444

570   Tarek, M., Brissette, F.P., Arsenault, R., 2020. Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrol. Earth Syst. Sci. 24, 2527–2544. https://doi.org/10.5194/hess-24-2527-2020

Taszarek, M., Allen, J.T., Marchio, M., Brooks, H.E., 2021. Global climatology and trends in convective environments from ERA5 and rawinsonde data. npj Clim. Atmos. Sci. 4, 1–11. https://doi.org/10.1038/s41612-021-00190-x

575   Tetzlaff, D., Carey, S.K., McNamara, J.P., Laudon, H., Soulsby, C., 2017. The essential value of long-term experimental data for hydrology and water management. Water Resour. Res. 53, 2598–2604. https://doi.org/10.1002/2017WR020838

Teuling, A.J., Lehner, I., Kirchner, J.W., Seneviratne, S.I., 2010. Catchments as simple dynamical systems: Experience from a Swiss prealpine catchment. Water Resour. Res. 46, 1–15. https://doi.org/10.1029/2009WR008777

Villinger, F., Loritz, R., Zehe, E., 2022. Torrents in small rural Catchments and the Potential of physics-based Models for their
580   Simulation. Hydrol. und Wasserbewirtschaftung 66, 284–285.

Wang, Y.V., Kim, S.H., Lyu, G., Lee, C.L., Ryu, S., Lee, G., Min, K.H., Kafatos, M.C., 2024. Nowcasting Heavy Rainfall with Convolutional Long Short-Term Memory Networks: A Pixelwise Modeling Approach. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 17, 8424–8433. https://doi.org/10.1109/JSTARS.2024.3383397

Xu, C., Wang, W., Hu, Y., Liu, Y., 2024. Evaluation of ERA5, ERA5-Land, GLDAS-2.1, and GLEAM potential
585   evapotranspiration data over mainland China. J. Hydrol. Reg. Stud. 51, 101651. https://doi.org/10.1016/j.ejrh.2023.101651

Yumnam, K., Kumar Guntu, R., Rathinasamy, M., Agarwal, A., 2022. Quantile-based Bayesian Model Averaging approach towards merging of precipitation products. J. Hydrol. 604, 127206. https://doi.org/10.1016/j.jhydrol.2021.127206

Zehe, E., Blöschl, G., 2004. Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions.
590   Water Resour. Res. 40. https://doi.org/10.1029/2003WR002869

Zehe, E., Maurer, T., Ihringer, J., Plate, E., 2001. Modelling water flow and mass transport in a Loess catchment. Phys. Chem. Earth, Part B 26, 487–507. https://doi.org/10.1016/S0378-3774(99)00083-9