Simulation of extreme rainfall and streamflow events in small Mediterranean watersheds with a one-way coupled atmospheric-hydrologic modelling system

Corrado Camera[1], Adriana Bruggeman[2], George Zittis[3], Ioannis Sofokleous[2], and Joël Arnault[4]

[1] Dipartimento di Scienze della Terra 'A. Desio', Università degli Studi di Milano, Milano, 20133, Italy
[2] Energy Environment and Water Research Center, The Cyprus Institute, Nicosia, 2121, Cyprus
[3] Climate and Atmosphere Research Center, The Cyprus Institute, Nicosia, 2121, Cyprus
[4] Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Garmisch-Partenkirchen, 82467, Germany

Correspondence to: Corrado Camera (corrado.camera@unimi.it)

Abstract.

Few studies evaluate the hydrologic performance of coupled atmospheric-hydrologic models when forced with observed rainfall and even fewer when forced with modelled precipitation. This information is crucial for the study of floods and in general for the use of the models for water management purposes. This study's objectives were: (i) to calibrate the one-way coupled WRF-hydro model for simulating extreme events in Cyprus with observed precipitation; and (ii) to evaluate the model performance when forced with WRF-downscaled (1×1 km2) re-analysis precipitation data (ERA-Interim). Streamflow was modelled during extreme rainfall events that occurred in January 1989 and November 1994 over 22 mountain watersheds. In six watersheds, Nash-Sutcliffe Efficiencies (NSE) larger than 0.5 were obtained for both events. The WRF-modelled rainfall showed an average NSE of 0.83 for January 1989 and 0.49 for November 1994. Nevertheless, hydrologic simulations of the two events with the WRF-modelled rainfall and the calibrated WRF-Hydro returned negative streamflow NSE for 13 watersheds in January 1989 and for 18 watersheds in November 1994. These results indicate that small differences in amounts or shifts in time or space of modelled rainfall, in comparison with observed precipitation, can strongly modify the hydrologic response of small watersheds to extreme events. Thus, the calibration of WRF-Hydro for small watersheds depends on the availability of observed rainfall with high temporal and spatial resolution. However, the use of modelled precipitation input data will remain important for studying the effect of future extremes on flooding and water resources.

1 Introduction

Atmospheric and hydrologic processes are strictly related, since they share the land surface as a common interface for moisture and heat fluxes. Precipitation is the primary cause of all surface hydrologic processes, such as overland, subsurface and river flow. Conversely, soil moisture and surface water distributions affect near surface atmospheric conditions and processes, such as the temperature distribution, the structure of the atmospheric boundary layer, the formation of shallow clouds and precipitation amounts (Lin and Cheng, 2016; Zittis et al., 2014 and references therein). In recent years, the scientific community has made ever-increasing efforts to improve the simulation skills of both atmospheric and hydrologic models, leading also to

the development of coupled modelling systems. Since the beginning of the 21st century, the main research interest in developing such models has been the evaluation of the feedbacks between the hydrologic cycle and the atmospheric processes,

35   to get a deeper understanding of regional climate change and its impacts (Ning et al., 2019).

The Weather Research and Forecasting hydrologic modeling system WRF-Hydro (Gochis et al., 2015) is an example of such a modelling system. It consists of a set of routines extending the hydrologic physics options in the Noah Land Surface Model (Noah LSM, Ek et al., 2003) and Noah with Multi-Parameterization Land Surface Model (Noah-MP LSM, Niu et al., 2011), which are the most commonly used land surface schemes of WRF (Constantinidou et al., 2019; Skamarock and Klemp, 2008).

40   In relation to WRF, WRF-Hydro can be run in an uncoupled (one-way coupled) mode or in a fully-coupled (two-way coupled) mode. In the first case, WRF-Hydro is run with user's specified atmospheric forcing, which can be observations, reanalyses, previously calculated model outputs or a mixture of the three (e.g., observed precipitation and WRF-derived temperature, wind speed, humidity, radiation etc). As a result, hydrologic outputs are influenced by the atmospheric variables but not vice versa. In the second case, WRF-Hydro enhanced hydrologic routines update the land surface states and fluxes in the LSM grid, which

45   are then used by the atmospheric component of the model.

As summarized by (Rummler et al., 2019), WRF-Hydro is mainly used in its uncoupled mode for model calibration and flood forecasting (e.g., Lahmers et al., 2019; Maidment, 2017; Silver et al., 2017; Verri et al., 2017; Givati et al., 2016; Yucel et al., 2015). Conversely, the fully-coupled mode is usually adopted to investigate land-atmosphere feedbacks (Arnault et al., 2016, 2019; Rummler et al., 2019; Senatore et al., 2015; Wehbe et al., 2019; Zhang et al., 2019).

50   Focusing on the use of the model for the simulation of flood events, (Yucel et al., 2015) calibrated WRF-Hydro over one watershed and two heavy rainfall events in northern Turkey, using 4-km WRF rainfall as input. The calibrated model parameters were then applied to three other watersheds and 10 heavy rainfall events. Their main aim was to quantify the performance improvement of the calibrated WRF-Hydro model against its use with default parameterization and test parameter transferability. In addition, they tested the model with WRF, WRF with data assimilation, and EUMETSAT precipitation

55   derived input. They obtained the best results with the calibrated model, forced by WRF with data assimilation precipitation. They suggest that this model configuration allows parameter transferability to ungauged catchments.

(Givati et al., 2016) calibrated uncoupled WRF-Hydro based on gridded observations of two high intensity rainfall events that occurred in 2013 over the Ayalon basin in Israel. The calibrated model was subsequently run with WRF-derived precipitation resulting from both uncoupled and fully-coupled simulations. The study demonstrated that both precipitation and streamflow

60   as derived from the fully-coupled model were superior to one-way coupled results, suggesting a possible application of fully coupled systems for early flood warning applications. Still, the authors suggested further research with a similar study set-up but over areas characterized by different precipitation and hydrologic regimes.

Silver et al. (2017) focused on five extreme events occurring over seven watersheds located in Israel and Jordan. They proposed a procedure for parameterizing the model scaling coefficients related to infiltration partitioning and soil hydraulic conductivity,

65   as well as for defining topographic categories. The procedure was based on soil physical properties and terrain characteristics

2

only. They demonstrated that their method leads to better streamflow predictions than trial and error calibration and is as good as expert knowledge parameterization.

Verri et al. (2017) calibrated an uncoupled WRF/WRF-Hydro modelling system over the Ofanto river basin, in southern Italy. Focus was on two three-month periods, each characterized by a heavy rainfall event and covering different seasons. WRF was run with 16-km horizontal resolution and 6-h fields forced by ECMWF-IFS (European Centre for Medium-Range Weather Forecasts – Integrated Forecasting System) as initial and boundary conditions. In addition, they presented a WRF rainfall correction approach based on rainfall observations, an objective analysis and a least square melding scheme and demonstrated that it improved river discharge simulation. The study also showed that optimal, calibrated values of infiltration partitioning and baseflow coefficients differ in the two events, suggesting a seasonal dependence.

Nowadays, uncoupled WRF-Hydro is the core of the National Water Model (NWM, https://ral.ucar.edu/projects/supporting-the-noaa-national-water-model), running over the Conterminous United States and furnishing streamflow forecasts for 2.7 million river reaches. The NWM flood forecasting skills has been strengthened within the framework of the National Flood Interoperability Experiment (Maidment, 2017). The NWM and WRF-Hydro remain under constant development. An example is the study of (Lahmers et al., 2019), who added channel infiltration processes to the modelling system to improve streamflow simulations in the arid southwestern United States.

From this review, it appears that few studies focus on the evaluation of the hydrologic output of WRF-Hydro when forced with observed rainfall and just a few more when forced with modelled rainfall. Model performance loss due to differences between observed and modelled rainfall is rarely discussed. Also, little attention has been given to small watersheds (area below 100 km^2). This study aims to address this gap. The focus is on two extreme events that occurred over 22 small watersheds, located in the Troodos Mountains of Cyprus, between 1-16 January 1989 and 11-26 November 1994. The main objectives are: (i) to calibrate the uncoupled WRF-Hydro model for simulating extreme events in Cyprus with observed precipitation; and (ii) to evaluate the model performance when forced with WRF-downscaled (1×1 km^2) re-analysis precipitation data (ERA-Interim).
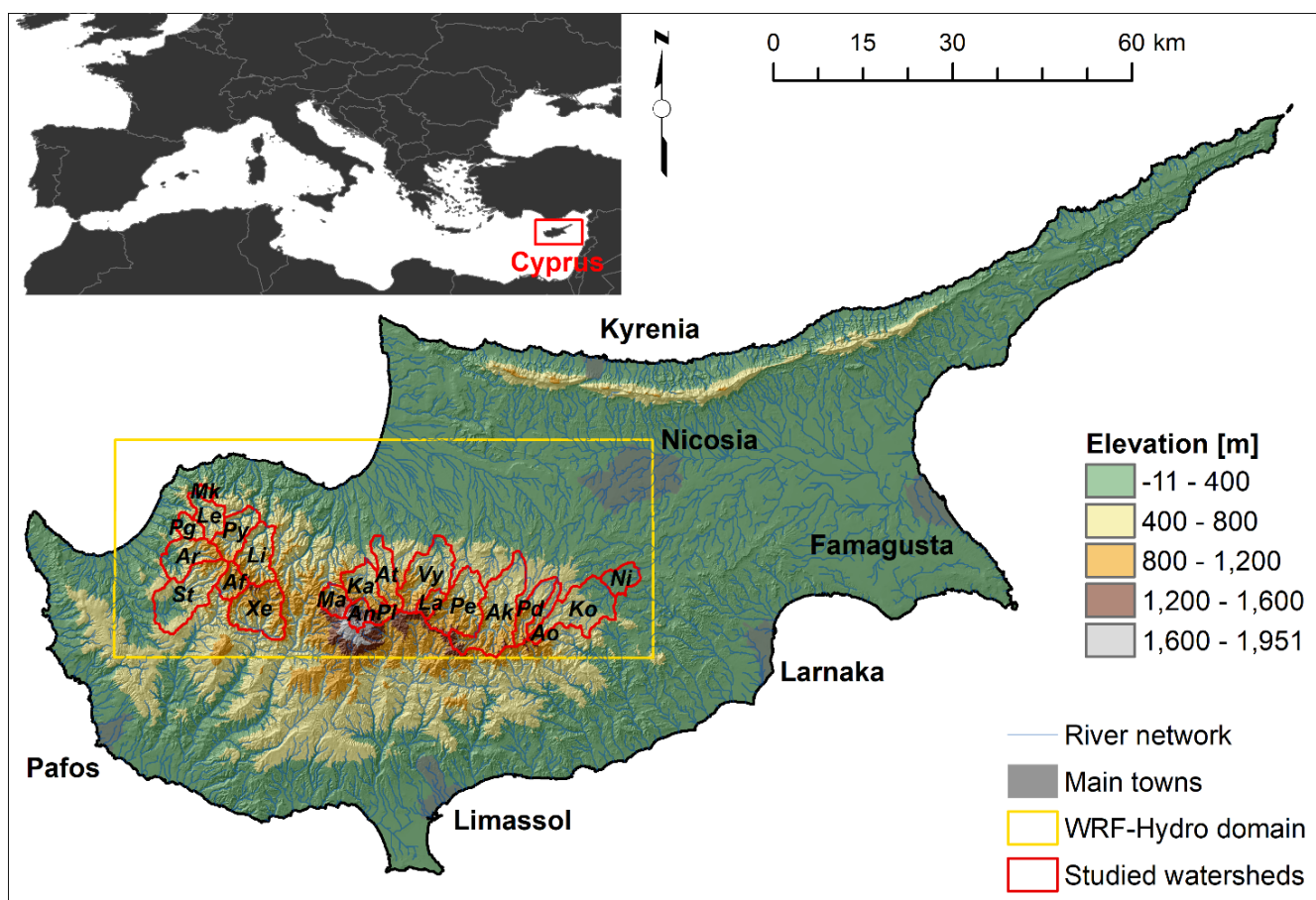

2 Study area

This study focuses on 22 watersheds located on the northern slope of the Troodos Mountains, Cyprus (Figure 1). The bedrock geology of the region is characterized by an ophiolitic complex. The highest peak of Troodos is Mt. Olympus (1952 m a.sl.). At high elevations (above 1400 m a.s.l.), ultramafic rocks are the dominant lithology (harzburgite, serpentinite, pyroxenite, wehrlite and dunite). Moving downhill, dominant rock types show a transition from gabbro to diabase, pillow lavas and sedimentary formations, therefore stratigraphically from the lower to the higher lithotype. Between gabbro and pillow lavas, diabase is present in the form of sheeted dykes and it constitutes the largest area of Troodos outcrop. Often, pillow lavas and sheeted dykes do not present a net geological limit, but the oldest lavas host the youngest dykes (Cleintaur et al., 1977). This transitional zone between pillow lavas and dykes takes the name of basal group. Throughout the ophiolitic complex, bedrock is usually found at shallow depths. According to the digital soil map of Cyprus (Camera et al. 2017), most of the soils over

Troodos are Lithic Leptosols with a stony gravelly texture and a predominant very shallow depth (0-10 cm), which can sometimes reach up to 100 cm. These characteristics highlight why rock fractures can be considered the main controlling factor

100   for the region's subsurface hydrology.

Due to its characteristic Mediterranean climate, more than 90% of a hydrologic year's (October-September) runoff from Troodos is produced between December and April. During the summer months, most rivers are completely dry (Le Coz et al., 2016). Due to their small areas and steep slopes, all watersheds have quite short times of concentration. Therefore, intense rainfall events lasting few hours can easily cause floods in the downstream plains.

105   Table 1 lists the 22 watersheds and summarizes their geology, as obtained from the geological map of Cyprus (Cyprus Geological Survey Department, 1995). Agios Nikolaos and Platania are sub-watersheds of Kargiotis; Lagoudera is a sub-watershed of Vyzakia; Kotsiati is a sub-watershed of Nisou.

110   **Figure 1. Geographical setting of the island of Cyprus and WRF-Hydro study area with the 22 target watersheds. For watershed short names refer to Table 1.**

4

Table 1. Geological characteristics of the studied watersheds.

Watershed	Watershed short name	Ultramafic complex [%]	Gabbro [%]	Sheeted Dikes [%]	Basal group [%]	Pillow Lavas [%]	Sedimentary formations [%]
Xeros	Xe	0	0	100	0	0	0
Agia Forest	Af	0	0	100	0	0	0
Stavros	St	0	0	42	13	0.17	26
Argaka	Ar	0	0	72	24	0.04	0
Pano Gialia	Pg	0	0	100	0	0	0
Leivadi	Le	0	4	96	0	0	0
Mavros Kremnos	Mk	0	8	92	0	0	0
Pyrgos	Py	0	0	100	0	0	0
Limnitis	Li	0	0	100	0	0	0
Marathasa	Ma	15	65	20	0	0	0
Agios Nikolaos	An	95	5	0	0	0	0
Platania	Pl	33	67	0	0	0	0
Kargiotis	Ka	30	41	25	3	1	1
Atsas	At	0	47	42	8	3	0
Lagoudera	La	0	12	76	11	0	0
Vyzakia	Vy	0	11	36	38	14	0
Peristerona	Pe	1	11	69	20	0	0
Akaki	Ak	0	2	37	47	11	2
Agios Onoufrios	Ao	0	0	33	57	9	0
Pedieos	Pd	0	0	52	35	11	1
Kotsiatis	Ko	0	1	11	28	59	1
Nisou	Ni	0	0	9	22	50	18

3 Data

3.1 Streamflow data

For the 22 watersheds, daily discharge data (m^3/s) from the Cyprus Water Development Department for the period 1980-2010 were analyzed. In addition, the original continuous hydrograph charts (water levels) of 16 streamflow stations from the Water Development Department, for the for the Jan-1989 and Nov-1994 events, were scanned and manually digitized through the GetData Graph Digitizer software (http://getdata-graph-digitizer.com). The digitized water levels were interpolated to obtain values precisely every 15 minutes (00.00, 00.15, 00.30, 00.45, 01.00....) and converted to discharge with the appropriate rating curve. Both interpolation and conversion were carried out by R scripts (https://www.r-project.org/). The 15-minute data were aggregated into hourly discharge values. Both hourly and daily values were used for model performance analysis.

3.2 Meteorological data

An hourly gridded dataset with a resolution of 1×1 km^2 was developed using hourly and daily rainfall data from the Cyprus Department of Meteorology stations and the daily gridded rainfall dataset of Camera et al. (2014). Data were extracted for two extreme events, with 42 rain gauges available over the island for Jan 1989 and 37 rain gauges available for Nov 1994. The

5

temporal disaggregation from daily to hourly gridded rainfall was developed through a FORTRAN code based on the method of hourly fractions (Di Luzio et al., 2008), which preserves the original daily values. The main steps of the disaggregation method are:

130
  a. The hourly rainfall observations (*ph*) are summed in 24-hour totals (*phs*). The 24-hour period ranges from 8.00 AM of the previous day until 8.00 AM of the attribution day, coherently with the daily gridded dataset.
  b. The fractions of the hourly rainfall data to the daily total rainfall are calculated as:

$$hfrac = ph/phs. \tag{1}$$

  c. The nearest gauge to each rainfall gridded dataset cell (*ng*) is found.
135
  d. The hourly rainfall at each grid cell (*phc*) is calculated by multiplying each gridded daily (*d*) rainfall value (*pdc*) with the hourly (*h*) fraction (*hfrac*) of the nearest valid gauge (*ng*).

$$phc(h,ng) = pdc(d,c) \cdot hfrac(h,ng). \tag{2}$$

4 Modelling setup

4.1 WRF-Hydro model description

140
The WRF-Hydro model is an extension package of the 1-D Noah LSM and Noah-MP LSMs, which are commonly coupled to WRF. In this study, the Noah LSM 2.7.1 version and the WRF-Hydro 3.0 version, as modified by Rummler et al. (2019), were used. WRF-Hydro, in comparison to the traditional 1-D LSM, enhances the physical description and mathematical resolution of surface and near surface hydrologic processes. It includes physics options for quasi 3-D saturated subsurface flow, 1-D or 2-D surface overland flow, 1-D channel routing, lake/reservoir routing, and baseflow processes. WRF-Hydro uses a
145
disaggregation-aggregation procedure to resolve the hydrologic processes at a finer resolution than the LSM. Below, a brief description of the main modeled processes and characteristics is presented. For a detailed description of the model components the reader can refer to Gochis et al. (2015). A schematic representation of the model structure, as used in this study, is presented in Fig. 2.
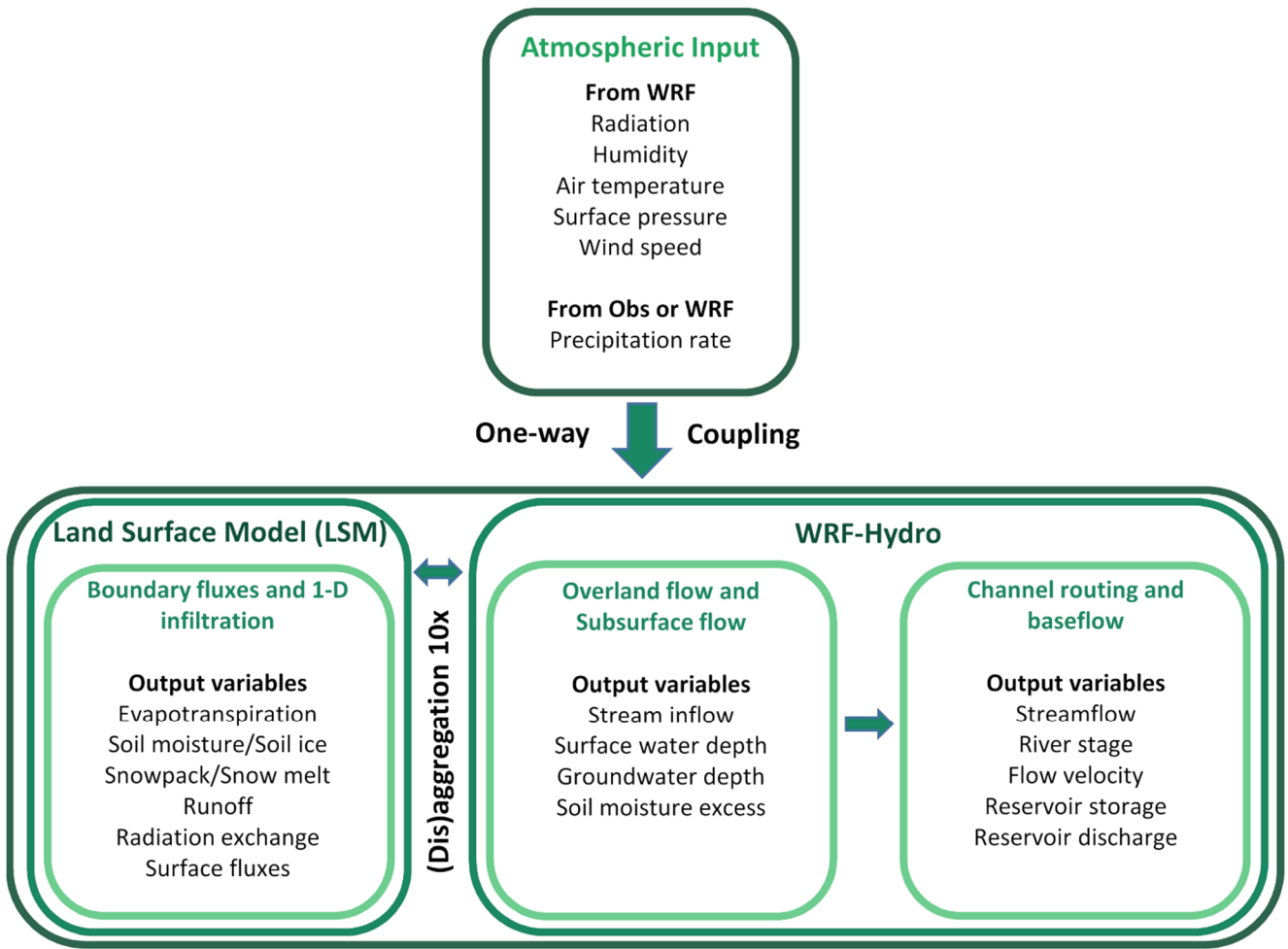
Figure 2. Schematic illustration of the model structure used in this study, including the coupling between WRF, the Noah Land Surface Model and WRF-Hydro routines (modified after Gochis et al., 2015).

One of the major advances of WRF-Hydro is the lateral subsurface flow component, which is calculated following the approach proposed by Wigmosta et al. (1994) and Wigmosta and Lettenmaier (1999). When precipitation reaches the surface, it can either infiltrate or run off. The partitioning between infiltration and runoff is controlled, besides the antecedent soil moisture conditions, by soil properties. In the Noah-LSM, the infiltration capacity (DDT) is defined as a function of the soil moisture deficit (DD) and an exponential scaled adjustment (VAL), which is a function of the parameter KDT. It follows the approach of Schaake et al. (1996), with the difference that KDT is not directly calibrated but is expressed as a function of the saturated hydraulic conductivity and two scaling coefficients:

$$DDT = DD \cdot VAL, \tag{3}$$

$$VAL = \left(1 - e^{(-KDT \cdot DT)}\right), \tag{4}$$

$$KDT = \frac{REFKDT \cdot K_{sat}}{REFDK}, \tag{5}$$

7

where DT is the time step duration [day]; K_{sat} [m s^{-1}] is the saturated hydraulic conductivity; REFDK is the reference (silty clay loam) saturated hydraulic conductivity (default 2E-06 m s^{-1}); and REFKDT is the infiltration partitioning scaling coefficient, which needs to be calibrated to empirically correct KDT for natural variability. As was demonstrated by previous studies (e.g., Naabil et al., 2017; Verri et al., 2017; Givati et al., 2016; Senatore et al., 2015), the model is sensitive to REFKDT. Once the water enters the soil, it moves vertically, through a four-layer soil column, until it reaches the saturated level and then laterally, according to the local gradient. In case the moisture content at the top of the soil column is larger than its water holding capacity (saturation), exfiltration occurs. The exfiltration amount is added to the infiltration excess and is routed over the surface. At the bottom of the soil column a vertical flux is calculated, using Richards equation (Richards, 1931). Drainage from the soil column is computed by multiplying the vertical flux with the SLOPE parameter, which can vary between 0-1, where 0 represents an impermeable boundary between the soil column and the underlying formations. The SLOPE parameter is assigned based on terrain slope classes through a table, however in an implicit way it expresses bedrock properties too (the higher the slope, the higher is the SLOPE coefficient in order to scale the projected map area over which deep drainage occurs). Drained water can be considered a loss or added to streamflow within the channel network through a conceptual baseflow module, if this is activated.

Regarding overland flow, WRF-Hydro allows water to pond on the earth's surface. A water retention depth is defined based on land use and vegetation cover. This parameter can be adjusted through a scaling factor (RETDEPRTFAC), which can be specified for each model cell. The fraction of ponded water exceeding the retention depth is available to overland flow routing. The routing is performed based on the diffusive wave formulation of Julien et al. (1995) and it can be resolved in both 1-D (Steepest Descent) or 2-D (x-y directions). Overland roughness is defined through the same tables as the retention depth and it can be adjusted through the overland-roughness routing factor (OVROUGHRTFAC). Overland flow can re-infiltrate, evaporate or enter the channel network.

Water entering the channel network, which the user defines through a Digital Elevation Model, is routed based on a streamflow algorithm that uses an implicit, one-dimensional, variable time stepping diffusive wave formulation. Such formulation is a simplification of the St. Venant equations for shallow water flow. The algorithm does not allow overbank flow and therefore the 2-D modelling of floods (Rummler et al., 2019). Channels are considered trapezoidal in section. Their geometrical properties, including roughness, are defined based on stream order. These model parameters are entered through a table and they can be set by expert knowledge or adjusted during calibration. Along the channel network, reservoirs can be added. Water can flow into reservoirs through the channel network or when overland flow intersects them. Water can flow out of the reservoir through weir overflow and gate-controlled flow. These fluxes are governed by the reservoir parametrization (reservoir area, maximum water level in the reservoir, weir length, gate area, gate elevation, gate aperture coefficient). No exchanges occur between the reservoir, the atmosphere, and the soil column around the reservoir (i.e., evaporation and subsurface lateral flow from the reservoir are not accounted for).

When deep drainage from the soil column is not considered as a loss, WRF-hydro allows two mathematical simple solutions to account for baseflow. For both solutions, baseflow is calculated within sub-watersheds. The first solution consists of a

simple pass-through model, meaning that the cumulated deep drainage occurring in a time step is equally redistributed to all channel segments within the watershed. The second solution consists of calculating a baseflow discharge [m^3 s^{-1}] (Q_{bf}) by means of an exponential bucket model, described by the following equation:

$$Q_{bf} = C \cdot \left(e^{a \cdot \frac{Z}{Z_{max}}} - 1 \right), \tag{6}$$

where C is the bucket coefficient [m^3/s], a is the bucket model exponent [-], Z_{max} is the maximum bucket level [m], and Z [m] is the bucket level at a certain time step. The user defines the C, a and Z_{max} parameters for each sub-watershed, together with a Z_{ini} [m] parameter to initialize the water storage in the bucket groundwater reservoir. At each time step the Z value is updated first adding the deep drainage contribution and subsequently subtracting Q_{bf}. Similar to the first solution, Q_{bf} is equally redistributed to channel segments. If Z equals Z_{max}, all deep drainage is transferred to the channel network.

4.2 WRF-Hydro model parameterization

The Noah LSM was parameterized over a 1×1 km^2 grid, while WRF-Hydro was run over a 100×100 m^2 grid. All simulations were performed in uncoupled mode, resolving the steepest descend formulation of the overland flow routine, with channel flow, baseflow and reservoir routines activated.

To run WRF-Hydro in uncoupled mode, the meteorological forcing needed are precipitation rate [mm s^{-1}], downward shortwave and longwave radiation [W m^{-2}], specific humidity [kg kg^{-1}], air temperature [K], surface pressure [Pa], near surface wind components [m s^{-1}]. For the calibration and validation runs, all variables except precipitation were taken from the WRF ERA-Interim downscaling experiments presented in (Zittis et al., 2017). These simulations incorporated the Grell-Freitas Ensemble Convection and the Ferrier Microphysics parameterization schemes, which were found to outperform the other tested configurations for the selected events. For precipitation, hourly observed gridded data were used (see Section 3.2 – Meteorological Data). For the simulation runs with WRF-modelled rainfall, all variables including precipitation were taken from the WRF experiments (Zittis et al., 2017). To derive soil moisture initial conditions, 15-day WRF spin-up runs were performed for both events. For Jan 89, the 15-day rainfall during spin-up was 99 mm and average soil moisture at the end of the simulation was 0.32 m^3 m^{-3}. The Nov-1994 event followed the dry summer and only a few scattered rain days occurred between the end of October and the beginning of November. The 15-day rainfall during spin-up was 18.4 mm and average soil moisture at the end of the simulation was 0.26 m^3 m^{-3}. Experimental data (Camera et al., 2018) show that in these conditions soil moisture can vary between 0.10 and 0.15 m^3 m^{-3}. Therefore, the WRF-derived initial soil moisture values for November were halved.

Soil, land use and vegetation cover data were derived from the MODIS dataset through the WRF Pre-Processing System. The related properties were attributed through the default table values implemented in WRF-Hydro (see Gochis et al., 2015). The hydrologic input layers (latitude, longitude, topography, flow direction, channel grid, lake grid, stream order, watersheds) were all calculated in ArcGIS® 10.2.2 starting from a 25×25 m^2 Digital Elevation Model (see Camera et al., 2017), resampled on the 100×100 m^2 grid, and the known locations of stream gauges and lakes. For the channel grid, a flow accumulation threshold

of 1500 cells was adopted. Nine slope terrain classes were derived following Silver et al. (2017) and SLOPE values were attributed accordingly. For each slope terrain class, the default SLOPE value listed in the WRF-hydro general parameters table was given. Other general parameters are REFKDT, and soil depth (SD), which were calibrated. REFDK, RETDEPRTFAC, and OVROUGHRTFAC were left to their default value.

Channel geometrical parameters were attributed based on the study area knowledge of the authors (Table 2). Default values were used for the Manning coefficients. The initial channel water depth was set to the default value for dry conditions. Six reservoirs were characterized in the model setup (Table 3) according to data from the Cyprus Water Development Department (2009). At all reservoirs, outflow occurs for overflow only; the structures do not have a gate. Vyzakia reservoir was completed in early 1994, therefore it was not included in the Jan-1989 simulation.

Regarding baseflow, the parameter C was set equal to the long-term baseflow index, calculated from the 1980-2010 data series with the program PART (Rutledge, 1988). The initial level of the conceptual reservoir (Z_{ini}) was set as a fraction of the maximum level (Z_{max}), based on the saturation degree of the deepest soil layer at time zero. The exponent a and Z_{max} were adjusted during calibration.

4.3 WRF-Hydro sensitivity analysis

A sensitivity analysis of the LSM parameters REFKDT and soil depth (SD), which have been identified as sensitive parameters in previous studies (e.g., Fersch et al., 2019; Senatore et al., 2015), was performed for the Jan-1989 event. For these simulations the baseflow routine was switched off. A reference scenario was set, with REFKDT equal to 1 and SD equal to 1.0 m. Parameters were changed one at a time. Eight values were tested for REFKDT (0.3, 0.5, 3.0, 5.0, 8.0, 10.0, 100.0, 1000.0) and two for SD (0.5 and 2.0 m). Also, to demonstrate the equifinality of calibrating REFDK and REFKDT, as suggested by eq. 5, two extra runs were performed for REFDK values of 4E-06 m s^{-1} and 6.67E-7 m s^{-1}. The relative sensitivity (S) was computed according to the following formula:

$$S = -\frac{(Vtot_i - Vtot_{ref})}{Vtot_{ref}}, \tag{7}$$

where $Vtot$ is the total volume discharged during the simulation period, ref refers to the reference scenario, and i to the perturbed value.

Table 2. WRF-Hydro channel parameter values used in this study (*Bw* is the channel bottom width, *HLINK* is the initial depth of water in the channel, *ChSSlp* is the channel side slope, and *MannN* is the Manning's roughness coefficient).

Stream Order	*Bw* [m]	*HLINK* [m]	*ChSSlp* [-]	*MannN* [-]
1	1.5	0.02	3.00	0.14
2	3.0	0.02	1.00	0.12
3	5.0	0.02	0.50	0.09
4	10.0	0.03	0.18	0.09

Table 3. Characteristics of the reservoirs included in the WRF-Hydro simulations; Long and Lat are longitude and latitude, respectively.

Watershed	Reservoir Name	Long [deg]	Lat [deg]	Reservoir Area [m^2]	Reservoir max elevation [m a.s.l.]	Reservoir ave elevation [m a.s.l.]	Weir length [m]
Vyzakia	Xyliatos	33.038	35.006	80000	537.5	529.9	15.0
Vyzakia	Vyzakia	33.029	33.029	160000	353.8	319.0	6.0
Akaki	Palaichori	33.130	34.928	110000	719.6	704.5	9.8
Akaki	Kalochorio	33.155	34.981	13000	533.5	528.5	22.5
Kotsiatis	Lythrodontas-1	33.274	34.944	10000	460.3	455.3	19.0
Kotsiatis	Lythrodontas-2	33.288	34.949	15000	422.5	413.5	33.8

4.4 WRF-Hydro calibration and validation with observed precipitation

Calibration runs were evaluated for each watershed against Jan 1989 daily observed streamflow, based on five performance indices. The selected set of indices contains both absolute error and goodness-of-fit measures, as suggested by Legates and McCabe (1999). They are *BIAS,* Mean Absolute Error (*MAE*), Nash-Sutcliffe Efficiency (*NSE* - Nash and Sutcliffe, 1970), modified Nash-Sutcliffe Efficiency (*mNSE*, Krause et al., 2005), and Kling-Gupta Efficiency (*KGE*, Kling et al., 2012).

Soil Depth is constant throughout the domain, therefore it was fixed at the value that returned the best performance indices in the majority of the watersheds, following an evaluation of the sensitivity analysis runs. Calibration focused on three parameters, REFKDT and two baseflow bucket parameters (α and Z_{max}). REFKDT was initialized, in each watershed, based on the evaluation of the sensitivity runs through performance indices, as for SD. For the baseflow bucket routine, initial parameters were set to the default. Next, REFKDT, α and Z_{max} were fine-tuned based on a trial and error procedure for all watersheds. Modifications were applied to a single parameter at the time and if changes could not improve the model performance according to three indices out of five after five attempts, the parameters were retained. Commonly applied changes were ±1 for REFKDT, ±0.5 for α, and ±10% of the actual value for Z_{max}. Smaller (larger) changes were applied only in watersheds where the response of streamflow was (not) particularly sensitive to specific parameters. The calibrated model was subsequently applied to the Nov-1994 event for validation. The same five model performance indices were used for the evaluation.

4.5 WRF-Hydro simulations with WRF-modelled precipitation

The WRF-modelled precipitation (Zittis et al., 2017) was averaged over each of the 22 watersheds and the daily values were compared to observed data by means of *BIAS*, *MAE* and *NSE*. To evaluate how deviations from the observed rainfall pattern affected the hydrologic model performance in these small mountain watersheds, the calibrated version of WRF-Hydro model was run with the WRF-modelled hourly precipitation forcing. Modelled streamflow was evaluated with observed data, similar as in the calibration phase.

4.6 WRF-Hydro evaluation with observed and modeled precipitation at hourly scale

280 For watersheds presenting daily *NSE* equal to or larger than 0.50 for both the calibration and the validation event, model performance was also investigated at hourly resolution. The *NSE*, *KGE* and *MAE* were computed for the hourly streamflow values simulated with both observed and modeled precipitation.

5 Results and discussion

5.1 Sensitivity analysis

285 The results of the sensitivity analysis are presented in Fig. 3 as boxplots. Each boxplot represents the sensitivity of the modelled total discharge volume, over the 22 watersheds, for the perturbation applied, in comparison to the reference simulation. The boxplots show that in the suggested calibration range (0.5-5.0, Gochis et al., 2015) REFKDT is very sensitive. Although the sensitivity decreases for REFKDT values larger than 5.0, variations in the discharged volume can be observed up to REFKDT values equal to 100.0. Further increases in REFKDT (see REFKDT 1000.0) do not cause any variations in discharge,

290 suggesting that the model already infiltrates at its maximum capacity. The variability over the watersheds is related to both infiltration processes (soil moisture conditions and texture controlling the hydraulic conductivity) and overland flow processes (controlled by local topography, land use and vegetation). Precipitation, which is not homogeneous throughout the study area, can play a role in causing different responses as well.

The two simulations ran with REFDK values of 4.00E-06 m s^{-1} and 6.67E-7 m s^{-1} returned discharged volumes equal to those

295 obtained with REFKDT values of 0.5 and 3.0, respectively. These results confirm the equifinality of the two parameters and make it clear that REFDK calibration should be avoided. As shown in Eq (3-5), REFDK automatically adjusts the infiltration capacity for the effect of soil texture, whereas any other effects on the partitioning of rainfall into surface runoff and infiltration can and should be calibrated through REFKDT.

The sensitivity analysis shows also an important role played by Soil Depth. Especially in mountainous areas, soils are usually

300 thin. This limited soil thickness affects the total amount of water retained by the soil, favoring a partitioning of the available water between infiltration and surface runoff towards the latter. Similar observations are reported by Fersch et al. (2019), while commenting the offset between modelled and observed soil moisture content in mountainous catchments in Bavaria (Germany). To overcome the issue, in other land surface models (e.g., Brunke et al., 2016) variable soil thickness has been implemented and tested.

305 **5.2 WRF-Hydro calibration and validation with observed precipitation**

The calibrated parameters are listed in Table 4. Soil depth was set equal to 1 m for all watersheds, because it was the value returning the best performance indices (Fig. 4) in 16 out of 22 catchments (average *NSE* improvement equal to 0.14). Fourteen watersheds have a REFKDT coefficient larger than 5.0, which is outside the 0.5-5.0 range suggested by Gochis et al. (2015), but none has a REFKDT lower than 0.5.

12
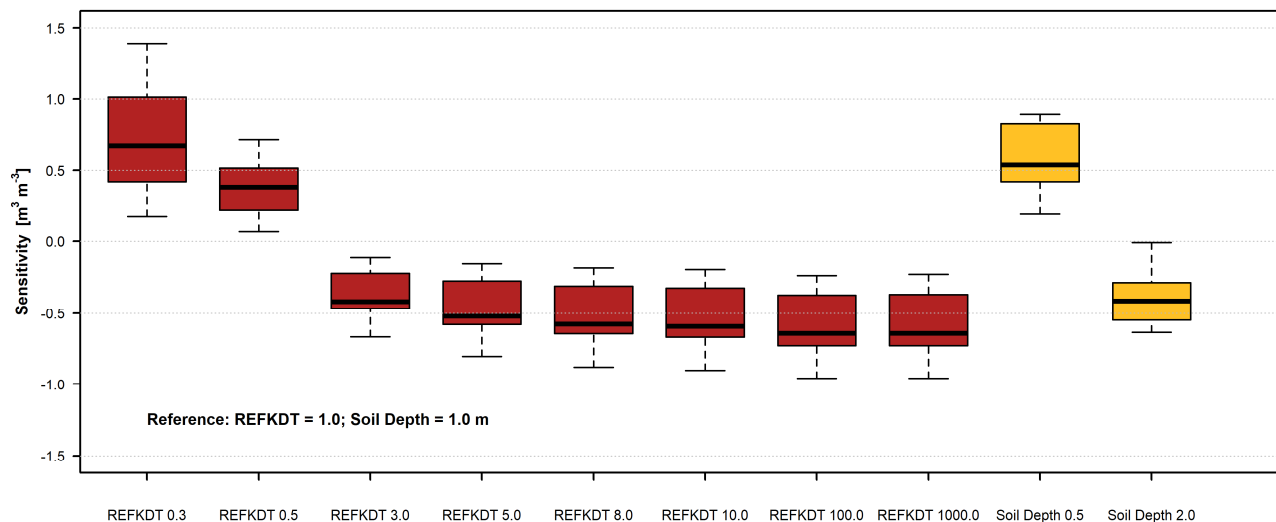
310

Figure 3. Boxplots of the sensitivity of the modelled streamflow to perturbations (x-axis) in REFKDT (infiltration partitioning scaling coefficient) and Soil Depth, relative to a defined reference scenario.

Table 4. Calibrated parameters (REFKDT, infiltration partitioning scaling coefficient; C, baseflow bucket coefficient; α, bucket exponent; Z_{max}, maximum bucket level) for the 22 watersheds with their maximum (MaxQ) and average (AveQ) discharges for the two analyzed events; for watershed short names refer to Table 1.

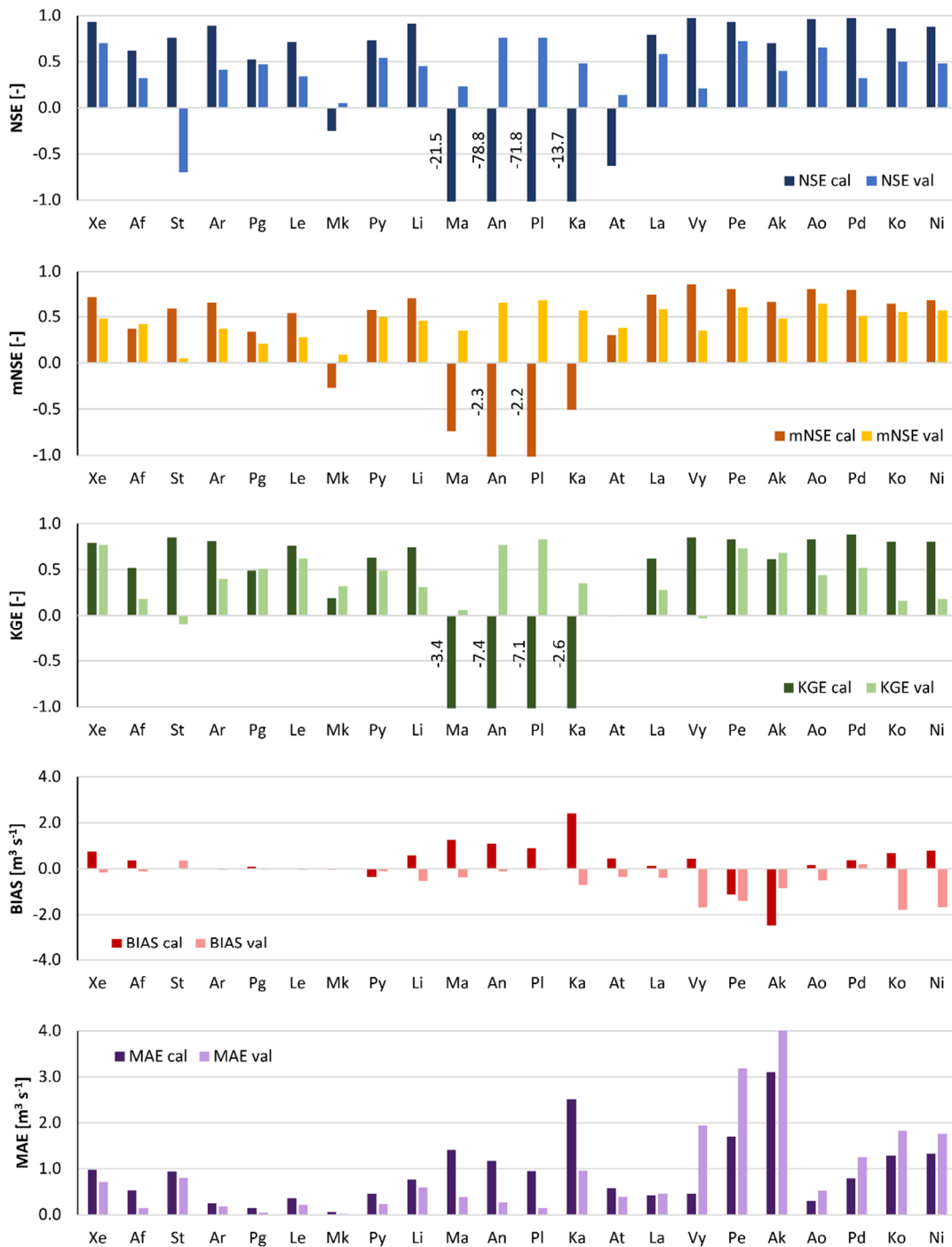| Watershed short name	Area [km²]	Max Q89 [m³ s⁻¹]	Ave Q89 [m³ s⁻¹]	Max Q94 [m³ s⁻¹]	Ave Q94 [m³ s⁻¹]	REFKDT [-]	C [m³ s⁻¹]	α [-]	Z_{max} [m]
Xe	67.5	19.7	3.7	8.2	1.1	5.0	0.30	2.0	30.0
Af	21.3	4.1	1.0	1.1	0.2	50.0	0.09	0.7	150.0
St	78.9	12.5	2.7	4.1	0.6	2.5	0.20	2.4	100.0
Ar	44.7	3.2	1.0	1.5	0.2	12.0	0.08	2.6	70.0
Pg	15.1	1.1	0.3	0.3	0.1	7.0	0.04	1.1	0.5
Le	27.9	3.3	0.9	1.4	0.3	2.0	0.07	2.2	1.4
Mk	5.2	0.3	0.1	0.1	0.0	8.0	0.01	3.2	10.0
Py	38.1	4.1	1.4	2.2	0.4	5.0	0.15	1.4	200.0
Li	48.0	12.3	3.0	5.4	1.0	7.0	0.27	1.2	70.0
Ma	22.6	3.7	1.2	2.9	0.6	100.0	0.19	1.2	100.0
An	15.7	1.8	0.7	3.3	0.8	100.0	0.24	1.6	100.0
Pl	10.2	1.3	0.4	1.9	0.3	100.0	0.05	0.8	130.0
Ka	64.6	9.2	2.6	10.5	1.9	100.0	0.30	1.6	130.0
At	32.7	2.9	1.0	1.9	0.5	100.0	0.04	0.8	60.0
La	14.5	8.8	1.6	6.4	0.9	8.0	0.05	2.3	15.0
Vy	81.0	15.9	3.5	12.0	2.4	6.0	0.09	3.0	12.0
Pe	78.2	58.0	7.5	35.0	5.9	1.0	0.29	2.5	8.0
Ak	96.7	49.0	7.7	28.0	5.7	0.8	0.20	3.0	5.0
Ao	14.2	9.9	1.4	5.3	1.1	3.0	0.03	2.1	10.0
Pd	29.8	26.0	3.2	8.9	1.9	2.0	0.07	2.2	10.0
Ko	74.1	18.0	3.6	15.0	3.0	6.0	0.05	3.5	2.0
Ni	95.6	18.2	4.1	16.5	3.0	7.0	0.05	2.6	3.0

13

Figure 4. Performance indices (NSE, Nash-Sutcliffe Efficiency; mNSE, modified Nash-Sutcliffe Efficiency; KGE, Kling-Gupta Efficiency; BIAS; MAE, Mean Absolute Error) calculated on daily streamflow resulting from observed rainfall for the 22 watersheds using the calibrated set of parameters for both Jan 1989 (cal) and Nov 1994 (val). For watershed short names refer to Table 1.

14

320   The parameterization of watersheds Ma, An, Pl, Ka, and At is peculiar, because even with very high REFKDT values no reasonable model fit indices could be obtained for the calibration period (Fig. 4). Conversely, the same watersheds show positive *NSE* values for the validation event. This is probably due to the fact that the upstream areas of these watersheds are located at the highest elevation and that part of the precipitation during the calibration was snow rather than rainfall, which is not explicitly considered in the model. In Fig. 5, the comparison between the observed and simulated daily hydrographs for
325   the Jan-1989 event is shown. The subdued response of the streamflow to the extreme precipitation is clear for watershed Pl, which is considered representative of the behavior of all five watersheds mentioned above, and it is clear that the simulated hydrograph overestimates the observed peak flow of the event. According to the MODIS soil map used, the soil type is uniform (clay loam) over the study area. However, geological differences between these five watersheds and the other watersheds are evident in Table 1. These five are the only watersheds with 70% or more of the surface bedrock made up of gabbro and
330   ultramafic rocks. The Troodos gabbro is very weathered and therefore permeable (Christofi et al., 2020). Improvements could have been obtained by using a more detailed soil map (e.g., Camera et al., 2017), possibly leading to an increase in the saturated hydraulic conductivity of the soil and an increase in drainage to the bedrock. Also, different bottom boundary conditions, as those implemented in the Noah Multi-Physics LSM, could improve the simulation results.

Overall, in all other watersheds the model behaves satisfactorily, with goodness-of-fit scores (*NSE*, *mNSE* and *KGE*, Fig. 4)
335   usually higher than 0.5 for the calibration run and larger than 0.0 for the validation event. Exceptions are watershed Mk for the calibration run and watershed St for the validation run. Looking at the hydrographs (Fig. 5 and Fig. 6), it is observed that Mk presents a very low discharge due to its limited area (Table 4). Therefore, small biases between observed and modelled streamflow produce poor goodness-of-fit indices. For St, the observed discharges show a two-day peak, while the model concentrates the discharge in the first day only, therefore affecting the performance scores.

340   In the eastern part of the modelling domain (La to Ni), for the calibration event both initial baseflow and the discharge peak are well modelled in all watersheds (Fig. 5). Differences between observed and simulated hydrographs can be observed in the post-peak, especially for watersheds Pe (not shown) and Ak. Both watersheds present a very high peak flow (> 50 m^3 s^{-1}) and an underestimation of the baseflow in the following days, which causes the high *BIAS* and *MAE* values visible in Fig 4. For the validation event (Fig. 6), the peak is well simulated in Ak and Ao (and in Pe and Pd, not shown) but the simulated
345   hydrographs show negative biases in comparison to the observed ones, in the post-peak phase, as for Jan 1989.

### 5.3 WRF-Hydro simulations with modeled precipitation

Figure 7 presents the performance indices of the WRF-modelled rainfall. The modelled rainfall is generally closer to observations for the Jan-1989 event than for the Nov-1994 event, as testified by the higher *NSE* (except for Le) and lower *MAE* values (Fig. 7). As can be seen in Fig. 5, the Jan-1989 event appears as a single day of intense precipitation, followed by
350   a few scattered low rainfall days that can show a moderate intensity towards the end of the simulation period. During Jan-1989, WRF-modelled rainfall is usually able to fit the observed daily precipitation trend over all watersheds, with slight

variations in the calculated daily amounts as suggested by the generally low bias (Fig. 7). In percentage, over the 22 watersheds *BIAS* vary between -35% and 50%, with an average of absolute values equal to 17%.
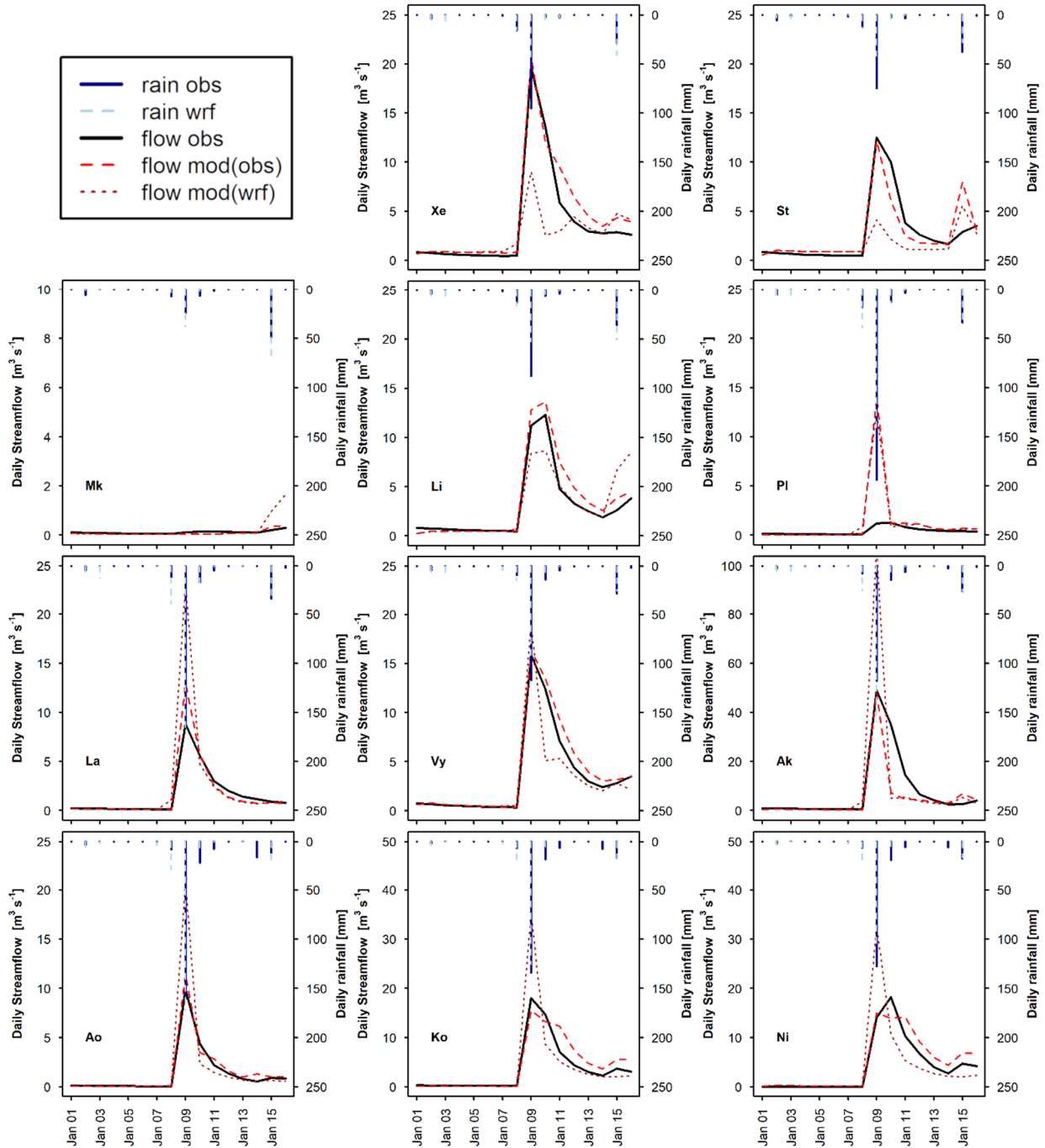
Figure 5. Observed daily hydrographs (flow obs) and hydrographs obtained with the calibrated WRF-Hydro model (flow mod) forced with observed rainfall (rain obs) and with WRF modelled rainfall (rain wrf) for the Jan-1989 calibration event, for 11 representative watersheds (see Table 1 for watershed short names).
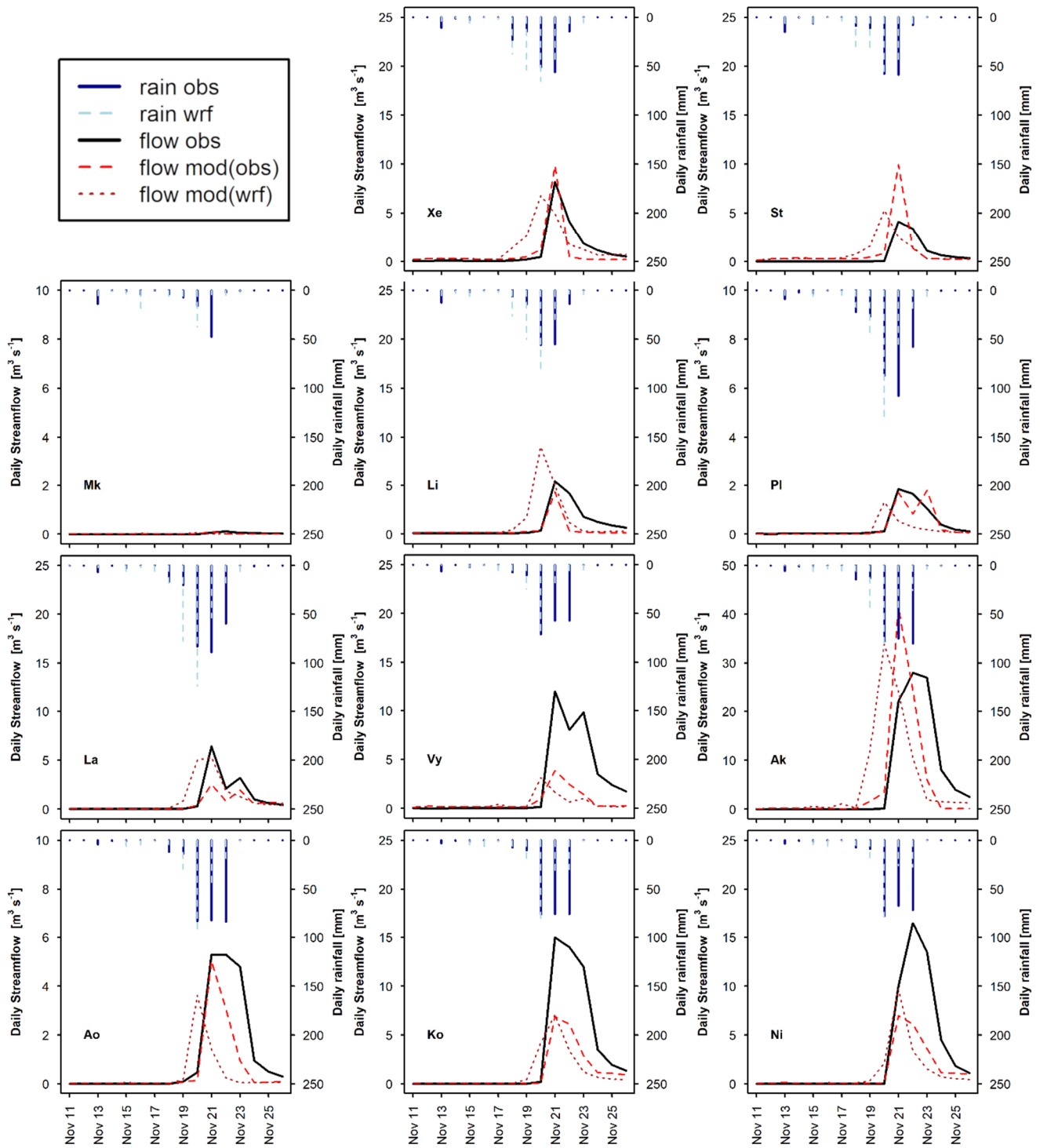
Figure 6. Observed daily hydrographs (flow obs) and hydrographs obtained with the calibrated WRF-Hydro model (flow mod) forced with observed rainfall (rain obs) and with WRF modelled rainfall (rain wrf) for the Nov-1994 validation event, for 11 representative watersheds (see Table 1 for watershed short names).

17
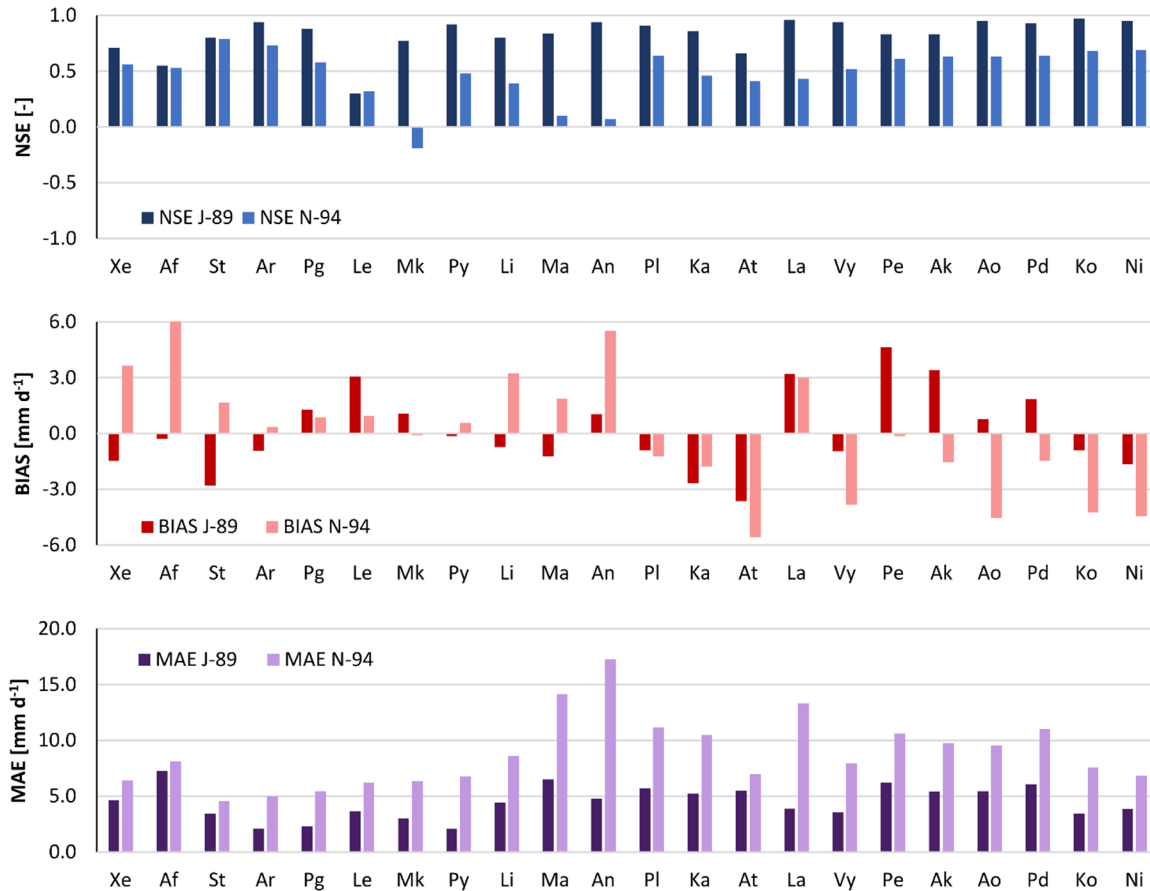
Figure 7. Performance indices (NSE, Nash-Sutcliffe Efficiency; BIAS; MAE, Mean Absolute Error) of daily WRF-modelled rainfall over the 22 watersheds both Jan 1989 (J-89) and Nov 1994 (N-94) events. For watershed short names refer to Table 1.

365　The modelled rainfall in Jan 1989 results in hydrograph shapes similar to the observed ones but still in goodness-of-fit indices that are often negative. With observed rainfall forcing, the simulated daily hydrograph returned negative *NSE, mNSE and KGE* values (Fig. 4) in six, five and five watersheds, respectively. With WRF-modelled rainfall forcing the number of watersheds with negative indices (Fig. 8) increases up to thirteen, eight, and ten, respectively. Passing from observed to WRF-modelled rainfall, both *NSE* and *MAE* indicate a loss in model performance in all watersheds except four (Ma, Pl, Ka, At), which are

370　those characterized by very negative goodness-of-fit indices in the calibration run. The average *MAE* increased by 1.09 m^3 s^{-1}, corresponding to 93%, and ranged between 0.12 m^3 s^{-1} in Py and 4.55 m^3 s^{-1} in Pe. These large errors were caused by relatively small errors in the WRF-modelled precipitation. Average *NSE* and *MAE* of the WRF-modelled rainfall, calculated over the daily averages of these 16 watersheds, are 0.83 and 4.2 mm d^{-1}, respectively. *BIAS* decreased in eight watersheds (Af, Li, Ma, Pl, Ka, Ak, Ko, Ni) but on average increased by 8.6 times (58 times in Le and 110 times in St).
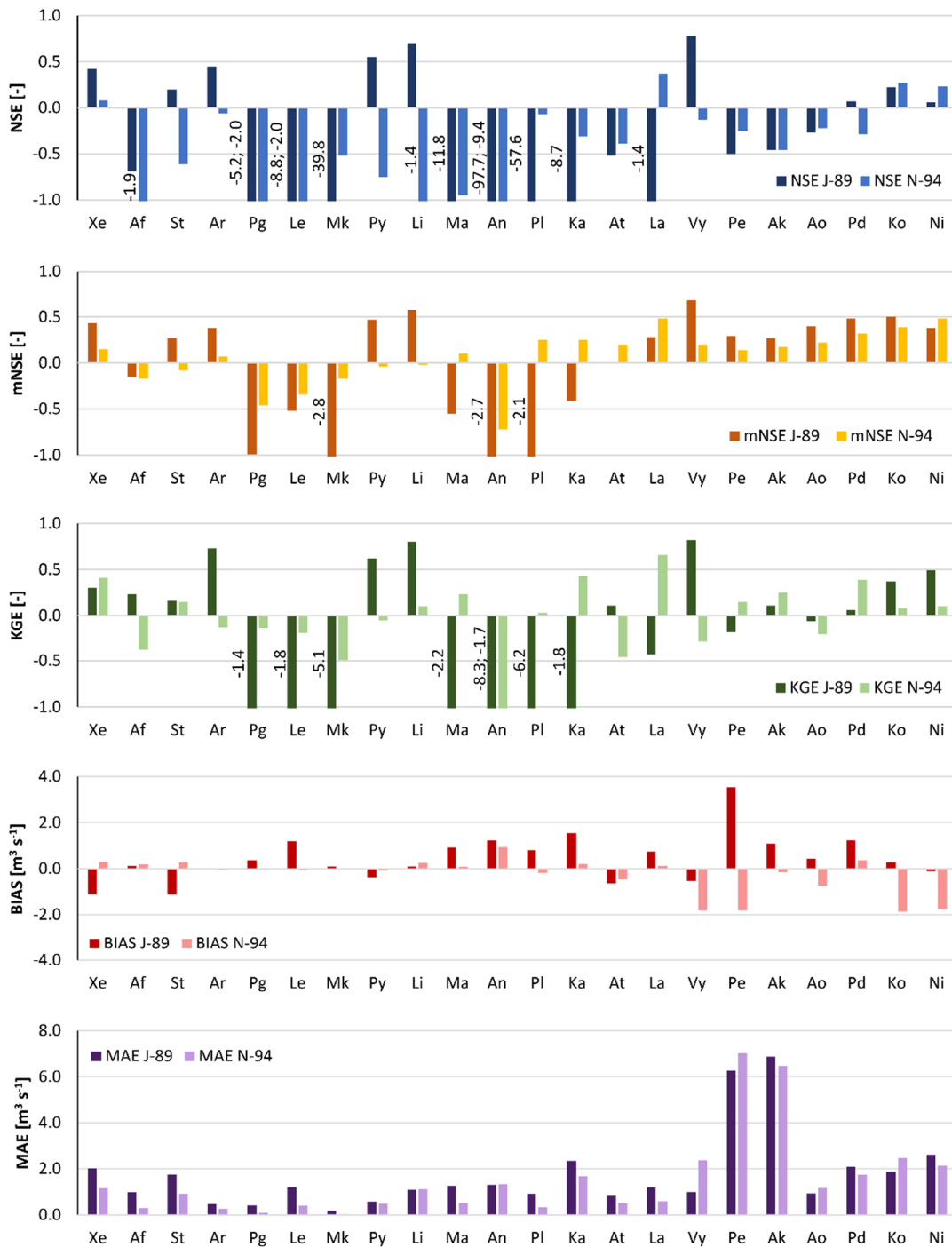
375

Figure 8. Performance indices (NSE, Nash-Sutcliffe Efficiency; mNSE, modified Nash-Sutcliffe Efficiency; KGE, Kling-Gupta Efficiency; BIAS; MAE, Mean Absolute Error) calculated on daily streamflow resulting from WRF-modelled rainfall for the 22 watersheds using the calibrated set of parameters for both Jan 1989 (J-89) and Nov 1994 (N-94) events. For watershed short names refer to Table 1.

Figure 6 shows that the Nov-1994 event is constituted of two days of moderately low precipitation, followed by three days of intense precipitation. The simulated event shows higher rainfall amounts in the preceding days and a loss of intensity after the first of the three high precipitation days. Consequently, the peak discharge is simulated to occur one day earlier than observed in most watersheds. This caused negative streamflow performance indices in eighteen watersheds for *NSE*, in eight watersheds for *mNSE*, and in ten watersheds for *KGE* (Fig. 8), while with the forcing of observed rainfall negative indices were found in one, zero, and two watersheds, respectively (Fig. 4).

These results indicate that a small shift in time or space of modelled rainfall, in comparison to observed precipitation, can strongly modify the hydrologic response of small watersheds to extreme events. The implementation of rainfall data correction or assimilation schemes could improve the forecasts of the atmospheric-hydrologic modelling chain, as demonstrated and discussed by previous studies (e.g., Avolio et al., 2019; Verri et al., 2017; Yucel et al., 2015). Recently, increasing efforts have been made to implement two-way coupled modelling systems, which were found to improve the overall skills of the modelling system (e.g., Senatore et al., 2015). However, the hydrologic component calibration is still usually performed based on observed precipitation data (e.g., Fersch et al., 2019; Givati et al., 2016).

5.4 WRF-Hydro with observed and modeled precipitation evaluation at hourly scale

Figure 9 shows the comparison between observed and modelled hourly hydrographs for three out of the six watersheds that had modelled daily streamflow *NSE* larger than 0.5 in both calibration and validation events. The two watersheds that are not shown are Pg (hourly streamflow data not available), Pe, and Ko (rating curve not available for peak flow). Looking at the streamflow modelled with observed rainfall as forcing, hourly peaks are generally overestimated and the modelled streamflow response to rainfall appears more immediate (pulse-like) than the observed streamflow. The overestimation is more evident for the Nov-1994 validation event than for the Jan-1989 calibration event. In addition, baseflow is well modelled for the calibration event but not so well for the validation event. The decent baseflow simulations lead to reasonable hourly performance indices for the Jan-1989 event. However, even with an *NSE* of 0.80 and a *KGE* of 0.72 for watershed Ao, the 17.9 m^3 s^{-1} hourly peak flow was overestimated by 18%.

The response of hourly streamflow to WRF-modelled rainfall shows similar behavior. The shape of the hydrographs is defined by rainfall pulses, in terms of both time of response and size of peaks. Even more than for daily outputs, it is evident that small differences in rainfall distribution and amount can cause large differences between observed and modelled streamflow (see performance indices).
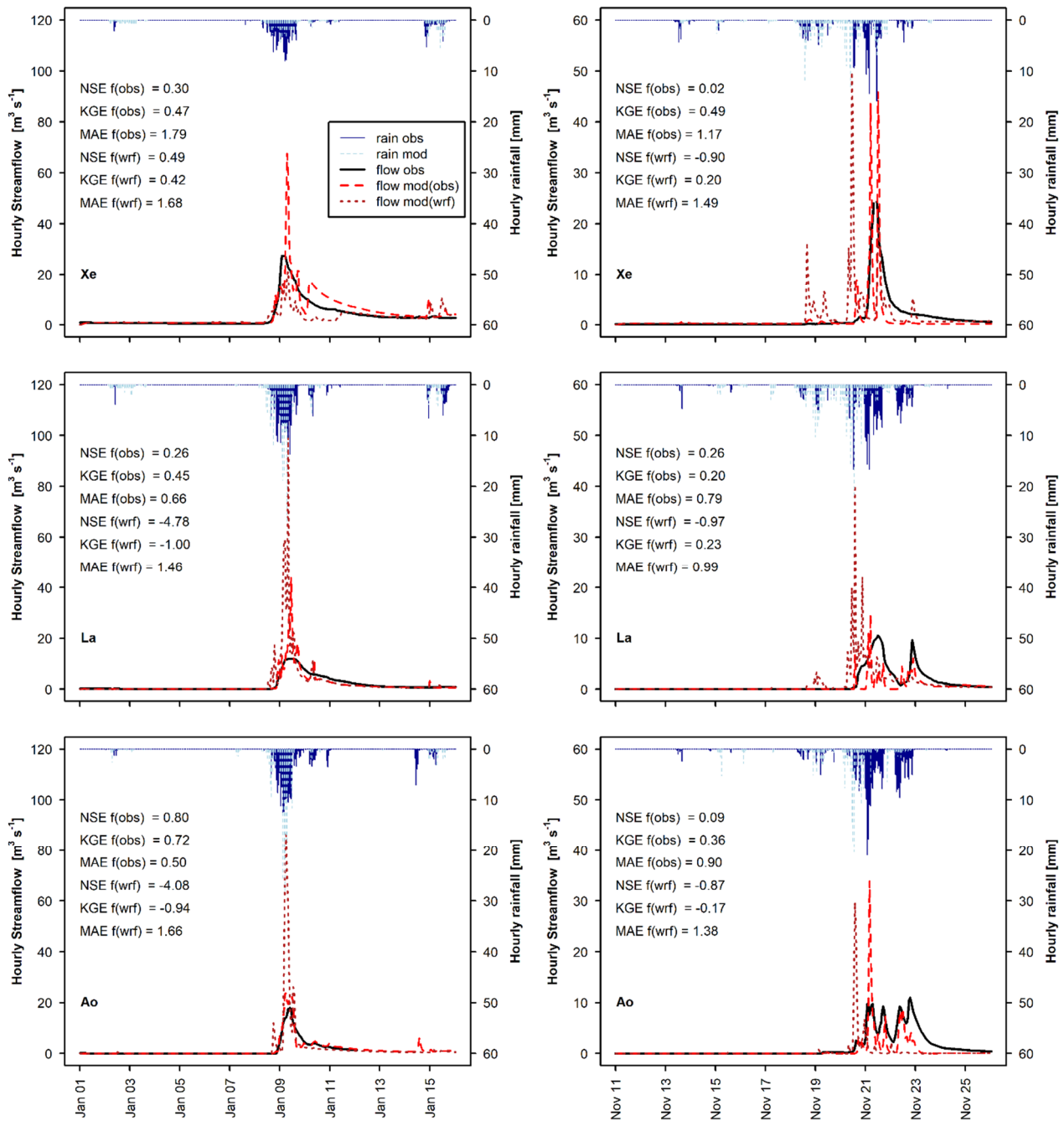
Figure 9. Observed hourly hydrographs (flow obs) and hydrographs obtained with the calibrated WRF-Hydro model (flow mod) forced with observed rainfall (rain obs) and with WRF-modelled rainfall (rain wrf) for both Jan 1989 (left) and Nov 1994 (right), for three watersheds (see Table 1 for watershed short names); modelled flow performance indices (NSE, Nash-Sutcliffe Efficiency; KGE, Kling-Gupta Efficiency; BIAS) are shown as well.

21

A possible improvement may be obtained by an increase in channel and overland roughness coefficients. This would allow slower flow, higher infiltration, and a smoothing of the peaks. Especially in dry Mediterranean areas, characterized by streams with seasonal flow, the vegetation (and consequently the roughness conditions) can be very different at the end of the dry

415 period (vegetation grown within the stream, dry understories and bushes and bare cropland overland) and in the middle of wet winter (water within the riverbed, green vegetation cover overland). This could be described with the inclusion of a seasonal variation of channel and overland roughness coefficients in the model. However, rainfall data with high spatial and temporal resolution would be essential to test this model modification.

6 Conclusion

420 This study evaluates streamflow simulations of the one-way coupled atmospheric-hydrologic model WRF-Hydro, forced with observed and WRF-modeled rainfall, during two extreme events, over 22 small mountain watersheds in Cyprus. Following model calibration and validation with observed rain, the model was run with WRF-downscaled (1×1 km^2) re-analysis precipitation data (ERA-Interim). These forcing data represent best-performing hindcasts of two extreme rainfall events, i.e. a model product that is as similar as possible to reality.

425 Overall, the selected four calibration parameters (REFKDT, Soil depth, the baseflow bucket exponent, and the maximum baseflow bucket capacity) were sufficient to obtain good model performance during model calibration in these steeply sloping and geologically complex watersheds. Sensitivity analysis showed that REFKDT can be calibrated beyond the suggested 0.5-5.0 range, having an effect on infiltration till a value of approximately 100.0. A Soil depth of 1.0 m, representative of the thin soils characterizing the study area, rather than the default value of 2.0 m, resulted in an average increase in *NSE* values of 0.14.

430 Calculated daily *NSE* values were higher than 0.5 in 16 out of the 22 modeled watersheds in Jan 1989 (calibration) and in eight watersheds in Nov 1994 (validation). Negative *NSE* values were found in the smallest watershed (5.2 km^2), caused by high relative errors for small differences; and in five watersheds with a high fraction of permeable bedrock, where drainage from the bottom of the soil column may have occurred faster than was simulated.

The comparison of modelled and observed hourly streamflow showed that almost all peak flows were overestimated by the

435 calibrated model. Modelled hourly streamflow fit the Jan 1989 hydrographs relatively well, but much less so the Nov 1994 discharges. This performance loss in Nov 1994 was due to a pulse-like behavior of the modeled streamflow related to an immediate response to rainfall, which could be attenuated by higher roughness coefficients.

Streamflow obtained with WRF-modelled rainfall forcing showed high discrepancies with observations, despite the good agreement between modelled and observed precipitation (average *NSE* of 0.83 and 0.49 for Jan 1989 and Nov 1994,

440 respectively). This suggests that model calibration with modelled rainfall forcing is not optimal for small mountain watersheds and therefore WRF rainfall forecasts may not be sufficiently accurate for predicting the location and size of the floods of such watersheds. However, modelled rainfall data could be suitable for investigating the effect of climate change on extreme rainfall and flood events. From the results presented and discussed, it emerges that future studies could focus on various aspects of the

22

modelling system to improve the simulation results of both precipitation and streamflow. The model could be improved by incorporating an option for time-dependent roughness coefficients to represent vegetation growth in ephemeral and intermittent streams in semi-arid environments. A model configuration with variable soil depths could also improve model performance, especially in mountain environments.

7 Acknowledgments

References

Arnault, J., Wagner, S., Rummler, T., Fersch, B., Bliefernicht, J., Andresen, S. and Kunstmann, H.: Role of runoff–infiltration partitioning and resolved overland flow on land–atmosphere feedbacks: A case study with the WRF-Hydro coupled modeling system for West Africa, J. Hydrometeorol., 17(5), 1489–1516, doi:10.1175/JHM-D-15-0089.1, 2016.

Arnault, J., Wei, J., Rummler, T., Fersch, B., Zhang, Z., Jung, G., Wagner, S. and Kunstmann, H.: A joint soil-vegetation-atmospheric water tagging procedure with WRF-Hydro: Implementation and application to the case of precipitation partitioning in the upper Danube river basin, Water Resour. Res., 55(7), 6217–6243, doi:10.1029/2019WR024780, 2019.

Avolio, E., Cavalcanti, O., Furnari, L., Senatore, A. and Mendicino, G.: Brief communication: Preliminary hydro-meteorological analysis of the flash flood of 20 August 2018 in Raganello Gorge, southern Italy, Nat. Hazards Earth Syst. Sci., 19(8), 1619–1627, doi:https://doi.org/10.5194/nhess-19-1619-2019, 2019.

Brunke, M. A., Broxton, P., Pelletier, J., Gochis, D., Hazenberg, P., Lawrence, D. M., Leung, L. R., Niu, G.-Y., Troch, P. A. and Zeng, X.: Implementing and Evaluating Variable Soil Thickness in the Community Land Model, Version 4.5 (CLM4.5), J. Clim., 29(9), 3441–3461, doi:10.1175/JCLI-D-15-0307.1, 2016.

Camera, C., Bruggeman, A., Hadjinicolaou, P., Pashiardis, S. and Lange, M. A.: Evaluation of interpolation techniques for the creation of gridded daily precipitation (1 × 1 km2); Cyprus, 1980–2010, J. Geophys. Res. Atmospheres, 119(2), 693–712, doi:10.1002/2013JD020611, 2014.

Camera, C., Zomeni, Z., Noller, J. S., Zissimos, A. M., Christoforou, I. C. and Bruggeman, A.: A high resolution map of soil types and physical properties for Cyprus: A digital soil mapping optimization, Geoderma, 285, 35–49, doi:10.1016/j.geoderma.2016.09.019, 2017.

Camera, C., Djuma, H., Bruggeman, A., Zoumides, C., Eliades, M., Charalambous, K., Abate, D. and Faka, M.: Quantifying the effectiveness of mountain terraces on soil erosion protection with sediment traps and dry-stone wall laser scans, CATENA, 171, 251–264, doi:10.1016/j.catena.2018.07.017, 2018.

Christofi, C., Bruggeman, A., Kuells, C. and Constantinou, C.: Hydrochemical evolution of groundwater in gabbro of the Troodos Fractured Aquifer. A comprehensive approach, Appl. Geochem., 114, 104524, doi:10.1016/j.apgeochem.2020.104524, 2020.

Cleintaur, M. R., Knox, G. J. and Ealey, P. J.: The geology of Cyprus and its place in the eastern Mediterranean framework, Geol. En Mijnb., 56(1), 66–82, 1977.

Constantinidou, K., Zittis, G. and Hadjinicolaou, P.: Variations in the Simulation of Climate Change Impact Indices due to Different Land Surface Schemes over the Mediterranean, Middle East and Northern Africa, Atmosphere, 10(1), 26, doi:10.3390/atmos10010026, 2019.

Cyprus Geological Survey Department: Geological map of Cyprus (1:250,000), [online] Available from: http://www.moa.gov.cy/moa/gsd/gsd.nsf/page32_en/page32_en?OpenDocument (Accessed 28 January 2020), 1995.

Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K. and Arnold, J. G.: Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States, J. Appl. Meteorol. Climatol., 47(2), 475–497, doi:10.1175/2007JAMC1356.1, 2008.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G. and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res. Atmospheres, 108(D22), doi:10.1029/2002JD003296, 2003.

Fersch, B., Senatore, A., Adler, B., Arnault, J., Mauder, M., Schneider, K., Völksch, I. and Kunstmann, H.: High-resolution fully-coupled atmospheric–hydrological modeling: a cross-compartment regional water and energy cycle evaluation, Hydrol. Earth Syst. Sci. Discuss., 1–37, doi:https://doi.org/10.5194/hess-2019-478, 2019.

Givati, A., Gochis, D., Rummler, T. and Kunstmann, H.: Comparing one-way and two-way coupled hydrometeorological forecasting systems for flood forecasting in the Mediterranean region, Hydrology, 3(2), 19, doi:10.3390/hydrology3020019, 2016.

Gochis, D. J., Yu, W. and Yates, D. N.: WRF-Hydro Technical Description and User's Guide, version 3.0, NCAR Tech. Doc., 123 pages, 2015.

Julien, P. Y., Saghafian, B. and Ogden, F. L.: Raster-Based Hydrologic Modeling of Spatially-Varied Surface Runoff1, JAWRA J. Am. Water Resour. Assoc., 31(3), 523–536, doi:10.1111/j.1752-1688.1995.tb04039.x, 1995.

Kling, H., Fuchs, M. and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol., 424–425, 264–277, doi:10.1016/j.jhydrol.2012.01.011, 2012.

Krause, P., Boyle, D. P. and Bäse, F.: Comparison of different efficiency criteria for hydrological model assessment, in Advances in Geosciences, vol. 5, pp. 89–97, Copernicus GmbH., 2005.

Lahmers, T. M., Gupta, H., Castro, C. L., Gochis, D. J., Yates, D., Dugger, A., Goodrich, D. and Hazenberg, P.: Enhancing the Structure of the WRF-Hydro Hydrologic Model for Semiarid Environments, J. Hydrometeor., 20(4), 691–714, doi:10.1175/JHM-D-18-0064.1, 2019.

Le Coz, M., Bruggeman, A., Camera, C. and Lange, M. A.: Impact of precipitation variability on the performance of a rainfall–runoff model in Mediterranean mountain catchments, Hydrol. Sci. J., 61(3), 507–518, doi:10.1080/02626667.2015.1051983, 2016.

Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233–241, doi:10.1029/1998WR900018, 1999.

Lin, T.-S. and Cheng, F.-Y.: Impact of soil moisture initialization and soil texture on simulated land–atmosphere interaction in Taiwan, J. Hydrometeorol., 17(5), 1337–1355, doi:10.1175/JHM-D-15-0024.1, 2016.

Maidment, D. R.: Conceptual Framework for the National Flood Interoperability Experiment, JAWRA J. Am. Water Resour. Assoc., 53(2), 245–257, doi:10.1111/1752-1688.12474, 2017.

Naabil, E., Lamptey, B. L., Arnault, J., Olufayo, A. and Kunstmann, H.: Water resources management using the WRF-Hydro modelling system: Case-study of the Tono dam in West Africa, J. Hydrol. Reg. Stud., 12, 196–209, doi:10.1016/j.ejrh.2017.05.010, 2017.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, J. Hydrol., 10(3), 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.

Ning, L., Zhan, C., Luo, Y., Wang, Y. and Liu, L.: A review of fully coupled atmosphere-hydrology simulations, J. Geogr. Sci., 29(3), 465–479, doi:10.1007/s11442-019-1610-5, 2019.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M. and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, J. Geophys. Res. Atmospheres, 116(D12), doi:10.1029/2010JD015139, 2011.

Richards, L. A.: Capillary conduction of liquids through porous mediums, Physics, 1(5), 318–333, doi:10.1063/1.1745010, 1931.

Rummler, T., Arnault, J., Gochis, D. and Kunstmann, H.: Role of Lateral Terrestrial Water Flow on the Regional Water Cycle in a Complex Terrain Region: Investigation With a Fully Coupled Model System, J. Geophys. Res. Atmospheres, 124(2), 507–529, doi:10.1029/2018JD029004, 2019.

Rutledge, A. T.: Computer Programs for Describing the Recession of Ground-Water Discharge and for Estimating Mean Ground-Water Recharge and Discharge from Streamflow Records--Update, [online] Available from: https://water.usgs.gov/ogw/part/ (Accessed 28 January 2020), 1988.

Schaake, J. C., Koren, V. I., Duan, Q.-Y., Mitchell, K. and Chen, F.: Simple water balance model for estimating runoff at different spatial and temporal scales, J. Geophys. Res. Atmospheres, 101(D3), 7461–7475, doi:10.1029/95JD02892, 1996.

Senatore, A., Mendicino, G., Gochis, D. J., Yu, W., Yates, D. N. and Kunstmann, H.: Fully coupled atmosphere-hydrology simulations for the central M editerranean: Impact of enhanced hydrological parameterization for short and long time scales, J. Adv. Model. Earth Syst., 7(4), 1693–1715, doi:10.1002/ 2015MS000510, 2015.

Silver, M., Karnieli, A., Ginat, H., Meiri, E. and Fredj, E.: An innovative method for determining hydrological calibration parameters for the WRF-Hydro model in arid regions, Environ. Model. Softw., 91, 47–69, doi:10.1016/j.envsoft.2017.01.010, 2017.

Skamarock, W. C. and Klemp, J. B.: A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, J. Comput. Phys., 227(7), 3465–3485, doi:10.1016/j.jcp.2007.01.037, 2008.

25

Verri, G., Pinardi, N., Gochis, D., Tribbia, J., Navarra, A., Coppini, G. and Vukicevic, T.: A meteo-hydrological modelling system for the reconstruction of river runoff: the case of the Ofanto river catchment, Nat. Hazards Earth Syst. Sci., 17(10), 1741, doi:10.5194/nhess-17-1741-2017, 2017.

550 Water Development Department: Dams of Cyprus, Ministry of Agriculture, Natural Resources and Environment, Nicosia, Cyprus., 2009.

Wehbe, Y., Temimi, M., Weston, M., Chaouch, N., Branch, O., Schwitalla, T., Wulfmeyer, V., Zhan, X., Liu, J. and Mandous, A. A.: Analysis of an extreme weather event in a hyper-arid region using WRF-Hydro coupling, station, and satellite data, Nat. Hazards Earth Syst. Sci., 19(6), 1129–1149, doi:10.5194/nhess-19-1129-2019, 2019.

555 Wigmosta, M. S. and Lettenmaier, D. P.: A comparison of simplified methods for routing topographically driven subsurface flow, Water Resour. Res., 35(1), 255–264, doi:10.1029/1998WR900017, 1999.

Wigmosta, M. S., Vail, L. W. and Lettenmaier, D. P.: A distributed hydrology-vegetation model for complex terrain, Water Resour. Res., 30(6), 1665–1679, doi:10.1029/94WR00436, 1994.

Yucel, I., Onen, A., Yilmaz, K. K. and Gochis, D. J.: Calibration and evaluation of a flood forecasting system: Utility of
560 numerical weather prediction model, data assimilation and satellite-based rainfall, J. Hydrol., 523, 49–66, doi:10.1016/j.jhydrol.2015.01.042, 2015.

Zhang, Z., Arnault, J., Wagner, S., Laux, P. and Kunstmann, H.: Impact of Lateral Terrestrial Water Flow on Land-Atmosphere Interactions in the Heihe River Basin in China: Fully Coupled Modeling and Precipitation Recycling Analysis, J. Geophys. Res. Atmospheres, 124(15), 8401–8423, doi:10.1029/2018JD030174, 2019.

565 Zittis, G., Hadjinicolaou, P. and Lelieveld, J.: Role of soil moisture in the amplification of climate warming in the eastern Mediterranean and the Middle East, Clim. Res., 59(1), 27–37, doi:10.3354/cr01205, 2014.

Zittis, G., Bruggeman, A., Camera, C., Hadjinicolaou, P. and Lelieveld, J.: The added value of convection permitting simulations of extreme precipitation events over the eastern Mediterranean, Atmospheric Res., 191, 20–33, doi:10.1016/j.atmosres.2017.03.002, 2017.

570