Project: **1245**

Project title: **Weather and climate modules of the AI-based early warning system DAKI-FWS**

Principal investigator: **Elena Xoplaki**

Report period: **2022-11-01 to 2023-10-31**

1. Brief overview of the project results acquired until now

The works performed during 2023 were: Analysis of the meteorological forcing datasets, (EMO 1arc min and HYRAS); Several machine learning approaches were tested and implemented to perform bias correction and downscaling of seasonal forecast simulations; Preparation and pre-processing of the heat wave detection datasets, ERA5 Tmax, Tmix, Tmean and 500 hPa geopotential height, NOAA SST; Calibration of the LISFLOOD hydrological model for all German and transboundary catchments; Collection and pre-processing of the tick occurrence dataset for Germany from various sources, including ECDC. A description of the resources used up to this stage of the project is given in Table 1.

Table 1. Overview of the performed calculations and resource usage

Calculations	Description	Node-hours
Bias-correction and downscaling	Data preprocessing, Nonlinear PCA, Analogs search, Implementation of ResNets for daily temperature and precipitation, Implementation of CNNs targeting architectural variations in the families of UNETs and ResNets	4852 GPU 284 CPU
	Convolutional Neural Networks targeting architectural variations in the families of UNETs and ResNets	229 GPU 565 CPU
Extreme events detection	AI module for drivers-based detection of heatwaves & droughts AI-automated non-stationary detection tool for concurrent extremes	1825 GPU 3876 CPU
Hydrological modelling	Pre-processing of meteorological forcings – Running Lisvap Calibration of LISFLOOD Test/benchmark LISFLOOD using the EMO 1arcmin meteorological forcings	4596 CPU
Ticks	Data pre-processing and data exploration	67 CPU

2. Bias-correction and downscaling of the seasonal forecast models

A large part of the works was devoted to data pre-processing, non-linear PCA, analogs for the training data set. Deep Res-CNN was implemented to upscaled (1°x1°) and non-biased observational temperature data. The model can successfully increase the resolution by 60 times to 1 arcmin. This is novel as such a ratio is not available in literature.

The outputs of several high-resolution NN architectures are shown in Figure 1a. The temperature downscaling results (Figure 1b) have achieved high accuracy. Overall, the downscaled temperature maps are very similar to the reference observational dataset. With monthly below 1 C, over most of the area.

a)

b)

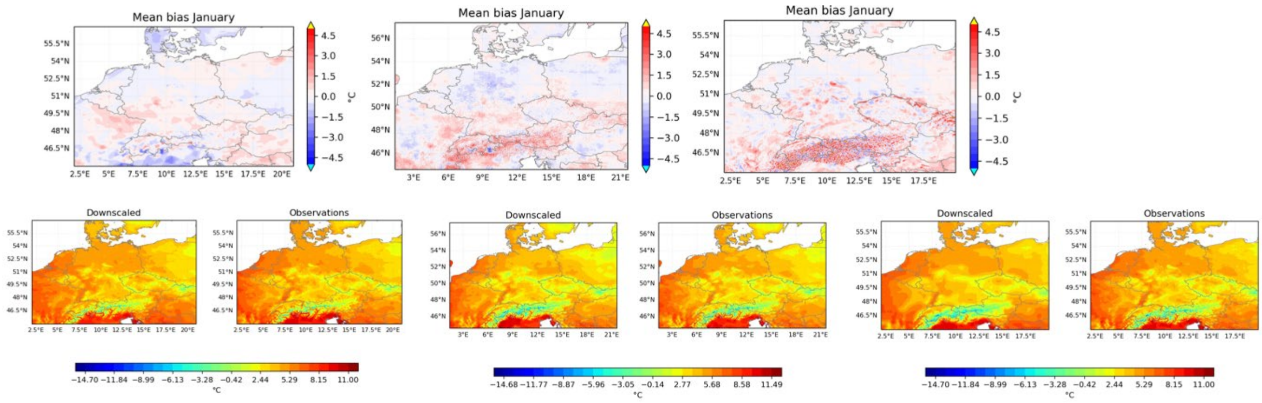

Figure 1. *(a) Monthly mean bias comparison for MLP + conv + dense (left), Conv-Dens-trConv (centre), ResNet (right). (b) comparison downscaled vs observed monthly (January) temperature (C) for MLP+conv+dense (left), Conv-Dens-trConv (centre), ResNet (right).*

3. Extreme events detection – Heat waves and droughts

The AI detection and prediction of heat waves and droughts based on drivers, teleconnections and atmospheric patterns considers the use of stations and gridded observations over a 30-year period, seasonal forecasts, and geographical, topographical information maps. The work carried out was 1) Calculation of heat waves (90th percentile based on moving windows) and their characteristics (duration and intensity) and indices at grid point level; 2) Calculation of heat wave intensity and duration over 36 smaller regions; 3) Preparation and stacking of datasets, NOAA SST, ERA5 500 hPa gph; 4) Reference data preparation heat wave magnitude (in progress). The output of the system is deviation of daily maximum temperature from the 90th percentile at grid point. An example is shown in Figure 2.

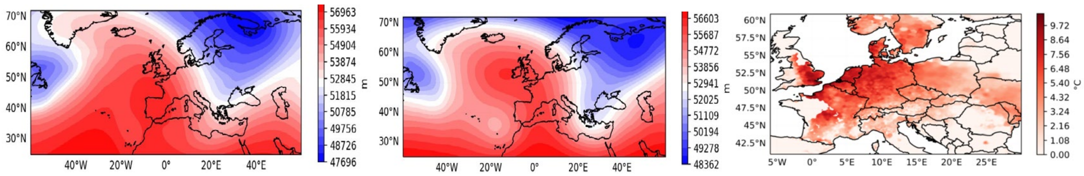

Figure 2. *(left, middle) 500 hPa GPH as input for heatwave detection. (right) AI-based HW-magnitude detection for June 15th 2003.*

4. Hydrological modelling

The implementation of the LISFLOOD hydrological model in the study area involved several steps: 1) preparation of meteorological forcings; 2) compilation of hydrological observations; 3) configuration of the LISVAP pre-processor to compute the reference evapotranspiration; 4) configuration of LISFLOOD; 5) model calibration; 6) mapping of calibrated parameters. We prepared a set of meteorological forcings to calculate potential evapotranspiration (Figure 3a), ran the calibration to obtain 14 parameter maps for LISFLOOD (Figure 3b), analysed the performance of the calibrated LISFLOOD using the KGE metric (Figure 3c), and finally ran LISFLOOD from 1990 to 2021 (Figure 3d).

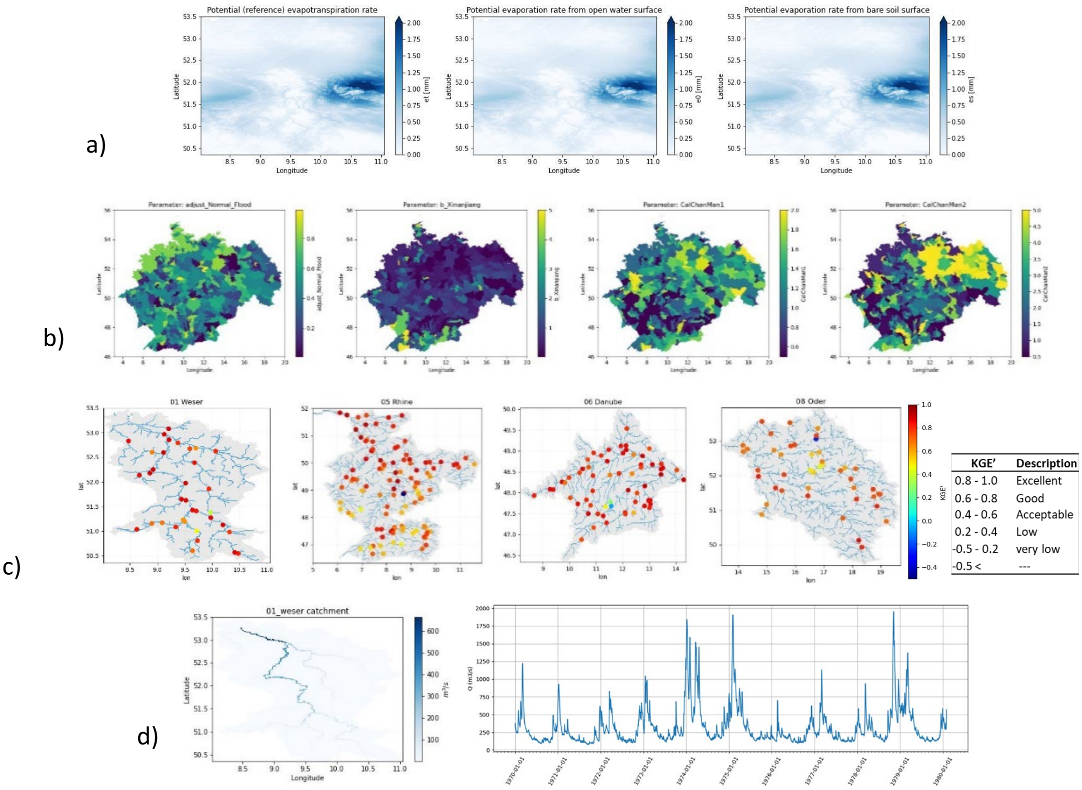

Figure 3. *Main results of the implementation of the hydrological model LISFLOOD in German catchments. (a) shows the computed potential evapotranspiration (ET0, EW0, ES0). (b) shows 4 of 14 parameters of LISFLOOD for the all catchments of Germany. (c) shows the evaluation of the calibration performance using the KGE' criteria for the Weser, Rhine, Danube and Oder catchments. (d) shows the spatial-temporal distribution of the streamflows of the Waser catchment computed for the period 1990-2021.*

5. Ticks

One approach to manage the health risk of tick-borne diseases is to implement a regional model that translates early warnings, based on seasonal forecasts, into public health alerts. We compiled tick presence data from the European Centre for Disease Prevention and Control and the Global Biodiversity Information Facility for 2004 to 2023 and 1905 to 2023. The data have been pre-processed and the distribution of ticks and diseases is identified. A random forest scheme is being developed to predict the presence of ticks based on rainfall, temperature and relative humidity.

6. Additional information

A total of 5449 CPU node hours expired due to: 1) Limitations in parallelising LISFLOOD thus increasing run time as different instances of LISFLOOD competed for resources, where the best approach was to run the calibration on 64 CPUs per node. 2) The sequential nature of the calibration process along the river network, from the highlands to the lowlands, contributed to the increased runtime of the calibration procedure. 3) In addition, we experienced delays in downscaling and bias correction, partly due to force majeure involving one of our team members.

Finally, Levante, as an HPC for Earth System Research, provides the necessary hardware and software configuration adapted to the requirements of the DAKI - FWS project.