Project: **1245**

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

Principal investigator: **Elena Xoplaki, Justus Liebig University Giessen**

Report period: **2023-11-01 to 2024-10-31**

1. Brief overview of the project results acquired until now

Based on the seasonal forecast with a 7-month lead time for 7 meteorological variables, composed of 51 ensemble members over a grid of 600 x 1008 points and a 5-year period at daily time resolution, an AI downscaling and bias correction model was developed and implemented. Using the downscaled and bias-corrected variables, a heat wave detection model was implemented to identify intensity, duration, and cumulative intensity characteristics. Subsequently, evapotranspiration and streamflow were computed using LISVAP and LISFLOOD. Finally, a neural network and random forest bias correction were applied to predict tick distribution. The computational resources used correspond to a total of 9 199 CPU node hours, and 9 258 GPU node hours (3 830 GPU DAKI and 5 428 GPU).

2. Downscaling Bias-correction of the seasonal forecast models

We performed extensive model testing with the objective of finding the best configuration for each variable (6 in total) and month (each month was studied separately). After obtaining the most suitable architecture for each variable, the models were run to obtain the inputs for the hydrological model. The outputs were evaluated against observations to account for errors and assess the performance of the approach (Figure 1a). The analog method was used to construct the training dataset to overcome the lack of synchronicity of long-range forecasts with observations, while also benefiting from the broad variability of the seasonal forecast. After comparison, we verified that applying the analog search on multiple GPH levels increased the robustness of the downscaling system by providing more coherent input features, thus finally enhancing the accuracy of the downscaling system.

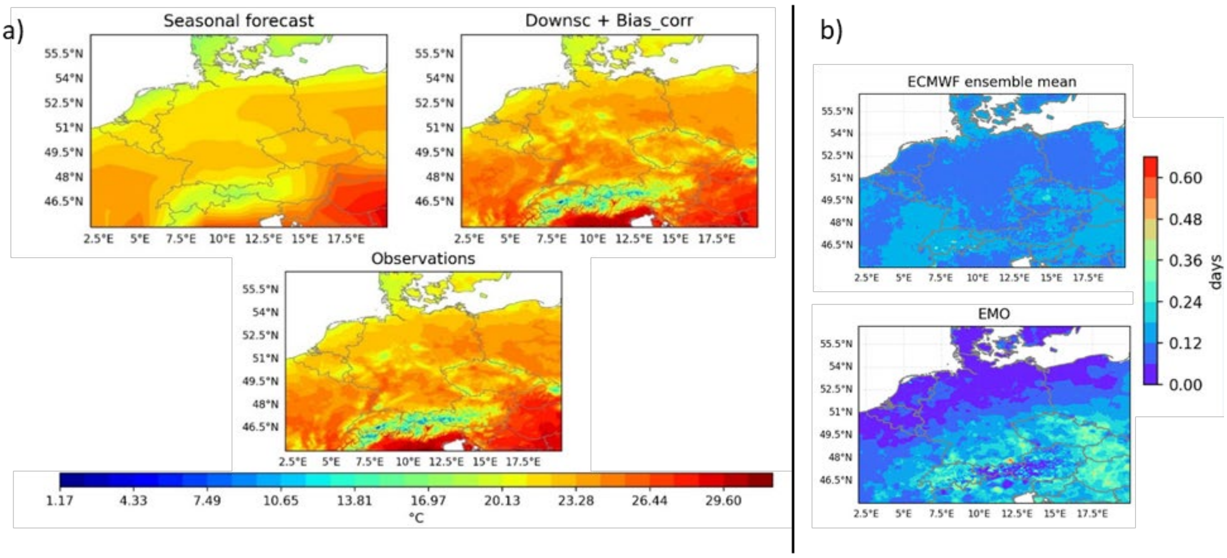

Figure 1: a) Maximum temperature for June-July-August (2012-2015): raw seasonal forecast (left), downscaled to 1 arcminute and bias corrected (centre), and observations (right). b) Heat wave duration averaged over the test period (2012-2015): model output (upper panel), and observations (bottom panel).

3. Extreme events detection – Heat waves and droughts

A heat wave detection model was applied to reforecast (1993-2015) and forecast (2017-2023) data, defining heat waves as over 3 consecutive days with maximum temperature above the 90th percentile, based on 1993–2015 data. Intensity, duration, and cumulative intensity were calculated for each heat wave. Heat wave magnitude was assessed using the HWMId index, with the sum of magnitudes for consecutive days. Detection was performed on all ensemble members, providing a probability range. Event duration was averaged for the 2012–2015 test period, and the results were compared with observations (Figure 1b).

4. Hydrological modelling

Meteorological inputs were prepared, computing the potential bare-soil evaporation (es), potential open water evaporation (e0), and potential reference evapotranspiration (et). We then run the hydrological model LISFLOOD to compute the streamflow (Q) for 51 ensemble members and 7 lead times of the ECMWF model for the reforecast period (1993 - 2017) based on bilinear re-grid, and for the forecast period (2017 – 2023) based on the AI-downscaled meteorological forcings (Figure 2a and b). Finally, we prepared the data set, and we are developing a demonstrator for exploring and visualizing the downscaled meteorological variables, heat waves, simulated streamflow, and the climatic water balance (Figure 2c).

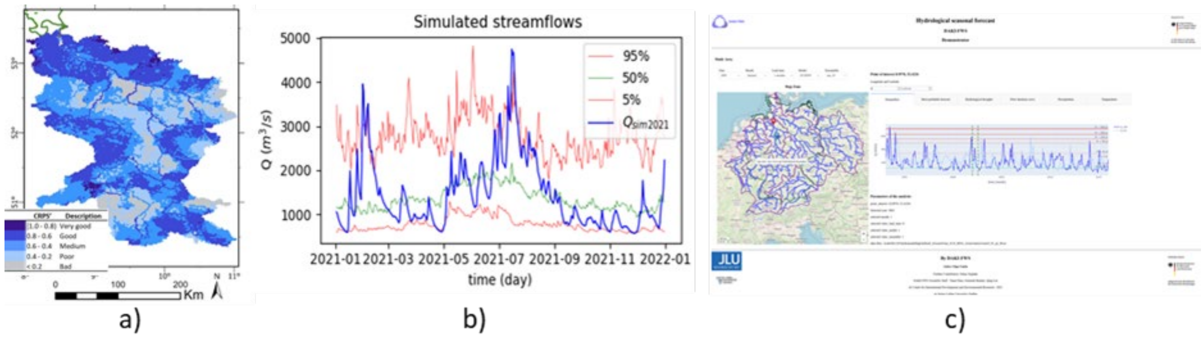

Figure 2: a) Performance of the seasonal forecast for the reforecast period. b) Example of the simulated streamflow for the Weser catchment. c) Overview of the demonstrator.

5. Ticks

We developed a model based on a dense Neural Network (NN) and a Random Forest bias correction (RFBC), which takes meteorological variables, elevation, and land use as the features to predict the tick's distribution on a logarithmic scale. The NN consists of 7 dense layers, with an L2 regularization term in layers one, three, and five. The NN model was trained with 603 samples. Since the number of training samples is small, the NN required cross-validation method. The performance of the NN ticks' model has been conducted using the metrics RMSE, ROC curve and ROI AUC value, which indicate that the model is able to reproduce the general trend and the tendency of small values but cannot reproduce high values. An example of the predictions is shown in the Figure 3.

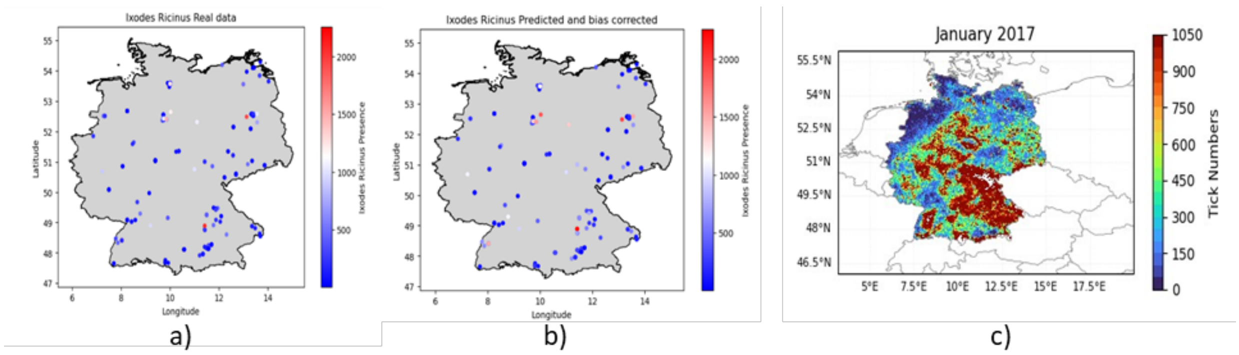

Figure 3: Ixodes Ricinus: a) observed data vs b) predictions for the testing period, and c) prediction for the forecast period, ensemble member 1, lead time 0.

6. Summary and outlook

In summary, the AI downscaling, AI bias correction, hydrological and NN tick's models were implemented for the ECMWF seasonal forecast over the river catchments in Germany and neighboring countries. Currently, the project continues with the evaluation of the performance of all models in the forecast period (i.e. with data that the models have not seen before), the further development and improvement of the demonstrator, and the application of the approach for two use cases in the agricultural sector.