BINGO
a better future under
CLIMATE CHANGE

BRINGING INNOVATION TO ONGOING WATER MANAGEMENT

D2.8

Stochastic weather generator producing large ensembles for decadal prediction

June 2017

www.projectbingo.eu

The BINGO project has received funding from the European Union's Horizon 2020 Research and Innovation programme, under the Grant Agreement number 641739.

,

Horizon 2020 Societal challenge 5:
Climate action, environment, resource
efficiency and raw materials

BINGO

Bringing INnovation to onGOing water management –

a better future under climate change

Grant Agreement n° 641739, Research and Innovation Action

Deliverable number:	**D2.8**
Deliverable name:	**Stochastic weather generator producing large ensembles for decadal prediction**
WP / WP number:	WP2
Delivery due date:	Project month x (xx/yy/201z)
Actual date of submission:	26/06/2017
Dissemination level:	Public
Lead beneficiary:	FUB
Responsible scientist/administrator:	Uwe Ulbrich, Henning Rust
Estimated effort (PM):	2
Contributor(s):	Komlan Kpogo-Nuwoklo, Edmund Meredith, Christos Vagenas (all FUB)
Estimated effort contributor(s) (PM):	10
Internal reviewer:	Not yet determined

Changes with respect to the DoW

A stochastic weather generator can only be built if sufficient data is made available for training. Even if our general concept is transferable to all the Research Sites, not all sites will be able to provide sufficient data to make use of the weather generator. The RS Wupper, Badalona, Bergen and partially Targus declared interest in providing sufficient data for its implementation. However, Bergen requested a modified version based on a single station. RS Cyprus has its own stochastic generator; here we aim for a comparison of the precipitation generator developed here with the Cyprus generator.

Dissemination and uptake

Data from the weather generator will be used by the Research Sites’ modelling groups.

Short Summary of results (<250 words)

This deliverable extends the stochastic precipitation generator built in D2.6 to a precipitation guided conditional stochastic weather generator. Given a set of driving variables (the conditions) from a large scale simulation and the precipitation fields (from D2.6), this generator is capable of simulating large ensembles of 2-dimensional fields of various climate variables requested by the research sites’ modelling groups at 1-km spatial resolution.

Evidence of accomplishment

This report and the simulations made with this conditional stochastic weather generator; the latter will be available for download via DECO (the Data Extraction and Conversion platform of FUBs FREVA Web-System).

TABLE OF CONTENTS

1. Introduction

Deliverable 2.8 consists of the development of a precipitation guided conditional stochastic weather generator for the various BINGO research sites. In the Deliverable 2.6, we presented a stochastic precipitation generator that is able to produce large ensembles of precipitation fields over the catchments. Here, we augment the latter by various other climate variables requested by the hydrological modeling groups at the BINGO research sites. Our goal is to build a conditional stochastic weather generator that is able to simulate large ensembles of fields of those variables at 1-km spatial resolution, conditioned on large scale drivers and the simulated precipitation fields resulting from D2.6.

The central idea of the weather generator built here is based on an Analogue approach. This approach has been very successful in a conditional stochastic weather generator setting used for various downscaling studies [Zorita et al. 1999; Vautard et al. 2009; Yiou et al. 2013]. The first requirement in building such a stochastic weather generator is a reference database that contains all climates variables which we want to simulate. This can be a set of observational data, reanalyses or also other simulations, e.g. from a dynamical climate model. In the absence of observational data, the COSMO-CLM simulations, whose spatial resolution and temporal resolution are respectively 12-km and 3 hours, are used as reference data here. Thus, to simulate 3-hourly data at 1-km spatial resolution, we adopt a two-step method (see flowchart in Figure 1):

1. We first use the analogue approach to simulate 3-hourly climate variables at 12-km spatial resolution which are consistent with the large scale drivers and the simulated precipitation fields. As all fields (e.g. temperature, humidity, radiation, etc.) stem from a spatially and physically consistent reference data set, the resulting fields are also spatially and physically consistent.[1]

2. In a second step, we interpolate the 12-km simulations to 1-km using kriging.

Please note, that precipitation as the variable showing the largest spatio-temporal variability is NOT interpolated but simulated using a proper conditional stochastic precipitation generator for spatial fields. The remaining variables, such as temperature, humidity, or radiation, are comparably smooths and a higher resolution simulation of these variables will not add more information.

1 Physically consistent means that a day with a large fraction of cloud cover will not show large values for radiation. Or, as another example, days with no cloud cover will not show precipitation.

This document is organized as follows: the data used in the study are presented in section 2. In section 3, the conditional stochastic weather simulator based on the analogue approach is described; this approach is used to simulate spatially and physically consistent fields of various climate variables at a 12-km-resolution. The spatial interpolation to 1-km resolution is presented in section 4, followed by summary in section 5.

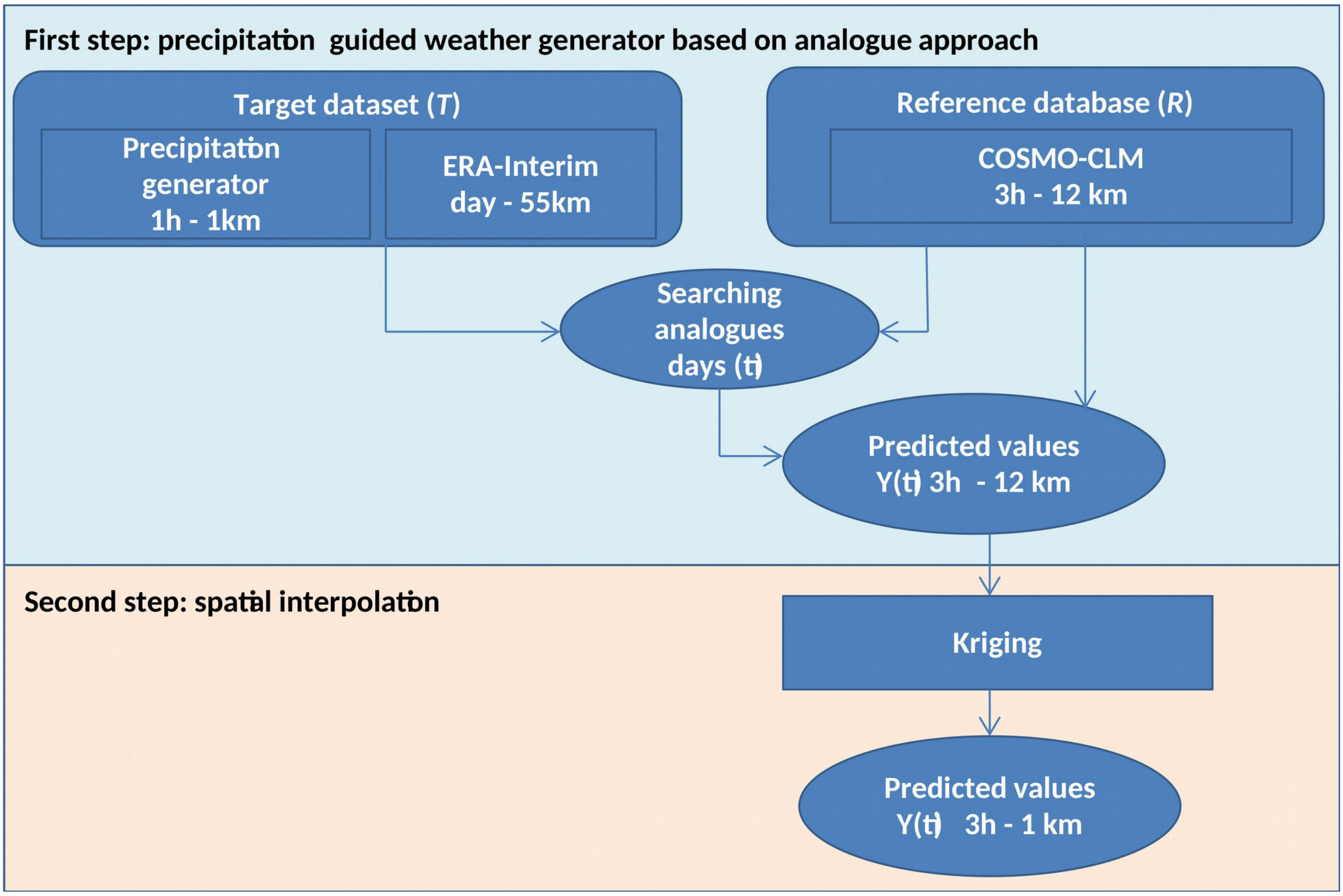

Figure 1: Flowchart presenting the main steps of the method used to generate 1-km data

2. Data description

The predictors that have been used to condition the weather generator come from the hourly precipitation fields generated by the stochastic precipitation generator (described in BINGO deliverable D2.6) and from ERA-interim reanalysis on a 0.5 × 0.5 degree grid, for the time period 1980-2015. The list of predictors is detailed in the Table 1. The ERA-interim predictors are daily-based and are the same variables which have been used to condition the precipitation generator (D2.6). Additionally, we generate 3 more predictors at the daily scale from the simulated precipitation fields: 1) daily precipitation sum integrated over the research sites' catchment, 2) daily maximum 3-hourly precipitation over the catchment, 3) 3-hourly precipitation profile; the latter consists of 8 entries with 3-hourly precipitation sum over the research sites' catchment, see also Table 1.

Predictors (X)	Database (*T*)
sum of precipitation over the area three hourly values maximum precipitation three hourly precipitation profile	Ensemble of precipitation fields simulated by the spatio-temporal stochastic precipitation generator
mean sea level pressure 10 metre U wind component 10 metre V wind component 2 metre temperature 2 metre dewpoint temperature	ERA-Interim reanalysis grid 0.5 × 0.5 degree

Table 1: List of predictors used to condition the weather generator

As a reference database, we use 3-hourly COSMO-CLM simulations (1979-2015) at 12-km spatial resolution. In addition to precipitation, this database contains the variables (denoted as **Y**), described in Table 2, which we want to reproduce with the weather generator. These variables are requested by the research sites' modeling gropus.

Variable description	Variable ID
Total Cloud Fraction	Clt
Near-Surface Relative Humidity	Hurs
Near-Surface Specific Humidity	Huss
Surface Air Pressure	Ps
Sea Level Pressure Surface	Psl
Downwelling Longwave Radiation	Rlds
Surface Downwelling Shortwave Radiation	Rsds
Near-Surface Air Temperature	Tas
Near-Surface Dew Point Temperature	Tdps
Eastward Near-Surface Wind	Uas
Northward Near-Surface Wind	Vas

Table 2: List of variables that are produced by the weather generator

3. A precipitation guided weather generator based on analogue method

The analogue method is probably the most intuitive weather generation approach which – in its conditional variant – is used also for statistical downscaling. It is based on the idea that a situation (or a day) simulated by a large-scale atmospheric flow model, for instance a GCM, is compared to available historical observations at the same scale, and the most similar situation (or day) – in a sense that has to be defined – is chosen as its *analogue* [Zorita et al. 1999]. The simultaneously observed local weather is then associated to the simulated large-scale pattern.

Our goal is to simulated consistent fields of climate variables **Y** (described in Table 2). We have a “target” dataset *T* of predictors **X** (see Table1 for the list of predictors). We also have a reference database *R* of climate variables (including **Y** and **X**). Here our reference database is the COSMO-CLM simulations. For a given field $X_i^{(T)}$ in the target dataset *T* we find a day t_i in the reference database *R* that optimizes a distance:

$$t_i = argmin_{j \in R} distance\left(X_i^{(T)}, X_j^{(R)}\right) \text{ , with } j = d_i - \tau, \ldots, d_i, \ldots, d_i + \tau$$

Where $d_i = 1, \ldots, 366$ is the day of the year which correspond to $X_i^{(T)}$ and τ a time lag to be defined (e.g. $\tau = 15$). The Euclidian distance is a canonical choice and thus selected for this study. Predictors are used as anomalies with respect to their seasonal cycle. For each day in the target dataset, we determine an analogue date t_i in the reference database *R.* Finally, the predicted values of **Y** for this target day are $Y(t_i)$.

We should point out the advantage of the analogue method regarding multi-variate simulations compared to other weather generators: For a target day, the most critical step is to find an analogue day t_i . Once this analogue day is found, all climate variables can be easily produced at the same time by taking their values at t_i . This also helps to keep the correlation (spatial and physical consistency) between variables and in space.

This approach can be used for simulating data sets in the frame of the historical period as well as for the decadal predictions with the reference database coming from the downscaled COSMO-CLM decadal predictions. Yet, for predictors, in particular sea level pressure, temperature, dewpoint temperature and wind, the global decadal prediction system (based on MPI-ESM-LR) is be used to condition the stochastic weather generator.

4. Spatial interpolation using kriging

The spatial resolution of the data simjulated by the analogue method described above is 12-km. We use ordinary kriging [e.g., Cressie, 1993] to interpolate this data to 1-km. Note that this procedure does not involve precipitation! The spatial preciptation fields have been simulated directly at 1-km resolution by means of a separate conditional stochastic precipitation generator. The latter is based on station data (described in detail in D2.6). The variables involved in the interpolation procedure (see Table 2) can be considered as smooth enough in space which justifies our choice to interpolate them to the desired resolution. A direct simulation of the high resolution will not introduce significantly more information.

The ordinary kriging is one of the most commonly used kriging techniques, a geostatistical interpolation technique. The principle of kriging is to estimate values of a regionalized variable at a selected location (Y_k), based on the surrounding existing values (Y_i). Selected locations are assigned a relevant weighting coefficient (λ_i) which represents the influence of particular data on the value of the final estimation at the selected grid node. The weights are estimated from the variogram or the covariance of the data.

We show in Figure 1 an example of three consecutive days simulations of temperature at 12-km and the interpolation to 1-km.

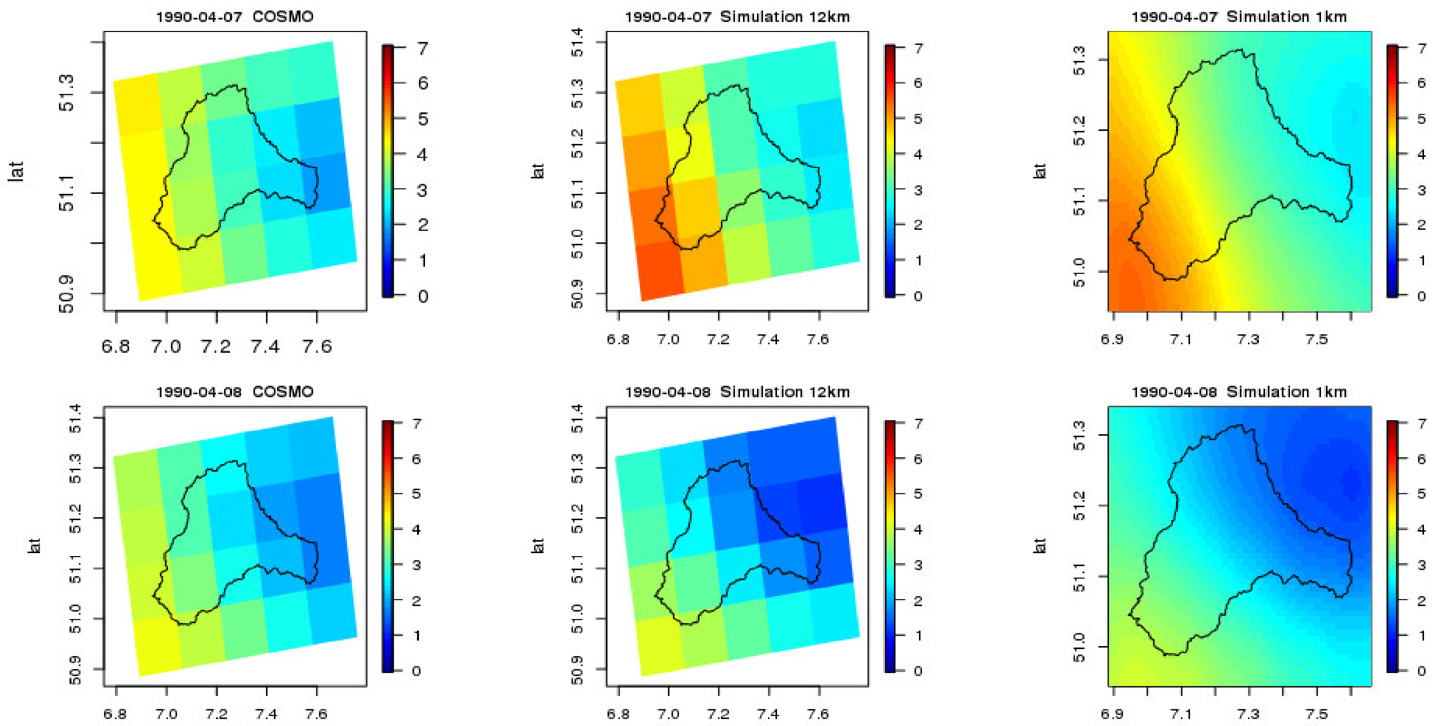

Figure 1: Temperature map for three consecutive days: each row corresponds to a one day, the first column correspond to the COSMO data (12 km), the second to the simulation from the weather generator at 12 km and the last to the simulation interpolated to 1km

5. Summary

A contional precipitation guided stochastic weather generator capable to simulate climate variables, other than precipitation, at 1-km spatial resolution has been developed. It consists of two-steps: First, an analogue approach is used to simulate 3-hourly, 12-km data. For a given target day, its analogue is searched in the reference database using predictors based on daily statistics obtained from a high resolution precipitation fields simulated with a conditional stochastic precipitation generator (see D2.6) and various other predictors obatin from a large scale driving field (ERA-Interim or MiKlip global decadal prediction system, respectively). These are sea level pressure, temperature, dewpoint temperature and wind. The first step leads to the simulation of 3-hourly, 12-km spatial resolution data. Then, the 12-km data is then spatially interpolated to 1-km using ordinary kriging. Note that the spatial interpolation is carried out for fields which can be assumed to relatively smooth in space and direct simulation of these fields at a 1-km resolution will not introduce significantly more information. The only field with a high spatio-temporal variability is precipitation; this field is, however, simulated with a particular dedicated conditional stochastic precipitation generator developed in D2.6.

6. References

Noel AC Cressie. Statistics for spatial data: Wiley series in probability and mathematical statistics. Find this article online, 1993.

Vautard, Robert, and Pascal Yiou, "Control of recent European surface climate change by atmospheric flow." *Geophysical Research Letters* 36.22 (2009).

Yiou, Pascal, et al. "Ensemble reconstruction of the atmospheric column from surface pressure using analogues." *Climate dynamics* 41.5-6 (2013): 1333-1344.

Zorita, Eduardo, and Hans Von Storch. "The analog method as a simple statistical downscaling technique: comparison with more complicated methods." *Journal of climate* 12.8 (1999): 2474-2489.