Performance of an Interpolated Stochastic Weather Generator in Czechia and Nebraska

M. Dubrovsky1,2, M. Trnka3, M.J. Hayes2, M.D. Svoboda2, D. Semeradova1, L. Metelka4, P. Hlavinka3

1. Abstract

Met&Roll is a WGEN-like parametric four-variate daily weather generator (WG), with an optional extension allowing the user to generate additional variables (i.e. wind speed and water vapor pressure). It is designed to produce synthetic weather data by interpolating present and future climate conditions to be used as an input into various models (e.g. crop growth and rainfall-runoff models). This contribution summarizes recent experiments, in which we tested the performance of the interpolated WG, with the aim to examine whether the WG may be used to produce synthetic weather series even for sites having no meteorological observations. The experiments discussed here include: (1) the comparison of 3 interpolation methods where their performance is compared in terms of the accuracy of the interpolation for selected WG parameters; (2) assessing the ability of the interpolated WG in the territories of Czechia and Nebraska to reproduce extreme temperature and precipitation characteristics; (3) indirect validation of the interpolated WG in terms of the modeled crop yields simulated by the STICS crop growth model (in Czechia).

The experiments are based on observed daily weather series from two regions: Czechia and Nebraska. Even though Nebraska exhibits a much lower density of stations, this is offset by the state’s relatively flat topography, which is an advantage in using the interpolated WG.

2. Methods

2.1 Stochastic weather generator Met&Roll (basic version − Richardson’s WGEN (1993))

- Daily variables: PREC (precip.), RAD (global solar radiation), TMAX and TMIN (daily max/min temperatures)
- Model: PREC: - occurrence ~ Gamma distribution (parameters: α, β / shape, scale)
  - all parameters are assumed to vary during the year
  - daily WG is optionally linked to monthly WG, which is based on AR(1) model and improves interannual variability
- Additional variables (e.g. water vapor pressure and/or wind speed) may be optionally added by nearest neighbors resampling

2.2 Nearest neighbors interpolation

The interpolated value Z_i,j is defined as a weighted average of observed values of parameter Z_i,j restricted to the nearest stations by applying a correction for between-station differences in latitude, longitude and altitude:

\[ Z_i,j = \sum_k w_k(Z_k,j) \]

\[ w_k = \left( \frac{1}{d_k} \right)^2 \]

where
- \( Z_k,j \) is a weight accounting for the distance \( d_k \) between the two stations
- \( w_k = \left( \frac{1}{d_k} \right)^2 \) for \( d_k \geq 1 \)
- \( w_k = 0 \) for \( d_k < 1 \)
- \( d_k \) is the distance between the location of which we interpolate, its value is a subject of optimisation

3. Regions & Data

Czechia
- area: 78,804 km²
- 125 stations available (subarea: CZ-25 subset)
- 1961-1990

Nebraska
- area: 205,530 km²
- 25 stations
- 1988-2006

3.1 Comparison of 3 methods for interpolation of WG parameters (only selected WG parameters are shown)

<table>
<thead>
<tr>
<th>Interpolation Method</th>
<th>Comparison</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Box 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Box 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Box 3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Experiment

1. Calibration of WG from the observed weather data (OBS) for each station in the set [WGsite = site-calibrated WG] with a synthetic series (WGint = interpolated WG)

2. Interpolation of the WG for each station in the verification set and using learning set of stations of the interpolated station is also in the learning set, the station is omitted from the learning set [WGint = interpolated WG] with a synthetic series (WGint = interpolated WG)

3. Implementation of WG parameters interpolated by three methods into GTOPO30 grid (Box 1 and Box 2)

Figure 1.1 WG parameters interpolated by three methods into GTOPO30 grid (0.5°-0.3°)

Box 1: Comparison of 3 methods for interpolation of WG parameters (only selected WG parameters are shown)

Table. Correlation between site-calibrated and interpolated WG parameters in terms of RV = Reduction of Variance

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily temperature maximum</td>
<td>0.72</td>
<td>0.69</td>
<td>0.67</td>
</tr>
<tr>
<td>Daily temperature minimum</td>
<td>0.74</td>
<td>0.68</td>
<td>0.69</td>
</tr>
<tr>
<td>Daily temperature range</td>
<td>0.68</td>
<td>0.65</td>
<td>0.67</td>
</tr>
<tr>
<td>Precipitation daily occurrence</td>
<td>0.70</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>Precipitation monthly occurrence</td>
<td>0.72</td>
<td>0.67</td>
<td>0.68</td>
</tr>
<tr>
<td>Precipitation annual occurrence</td>
<td>0.73</td>
<td>0.69</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Box 2: Comparisons of interest

- Comparison of climatic characteristics derived from WGsite vs. OBS: testing WG ability to reproduce climatic characteristics, misfits are due to WGEN imperfections
- Comparison of climatic characteristics derived from synthetic series produced by site-calibrated (WGsite) and interpolated (WGint) generators (imperfections are due to interpolation errors)
- Comparison of crop yield characteristics modelled with OBS, WGsite and WGint weather series. Misfits are due to WG (C1), interpolation errors (C2), both (C3)
- Assessing accuracy of interpolated crop yields
Validation of interpolated WG in terms of TMAX increase

Box 1: choice of the interpolation method: The Nearest Neighbours interpolator was found to be the most effective of the three candidate interpolation methods. This method was therefore selected for the additional experiments presented here (Boxes 2 and 3).

Box 2: reproduction of annual and 30-year extreme temperature and precipitation characteristics by WG

Box 3: Indirect validation of interpolated weather generator through crop model: mean wheat yields simulated by STICS model

Q: When we want to estimate PDF of crop yields (or other crop model output statistics) in the weather-ungauged site, which of the two approaches provides us with a better estimate:

A. 1. crop-simulation in sites with available weather, 
2. interpolation of model yields (*)

B. 1. calibration of WG in sites with available weather (WGsites) 
2. interpolation of WG into ungauged site (WGint)
3. generation of synthetic series
4. crop-simulation with the synthetic series (MYint)

(*) interpolation method is the same as the one used for interpolation of WG

Comparisons of interest:

WGsite vs OBS (testing ability of WG to reproduce climatic characteristics; misfits are due to WG imperfections)

WGsite vs WGint (imperfections are due to interpolation errors)

where the fit is quantified by:

- RV (Reduction of Variance): 1 indicates perfect fit, below zero values indicate that the reference value is a worse estimate than the sample mean
- Bias*: RV-bias/STD: 0 indicates no systematic deviation, above/below zero values indicate that the values are larger than the reference values
- FS (%) = % of significantly different: FS<100 indicates that the values are significantly different at all stations in a verification set

The experiment was made for 3 sets of stations:

- CZ-25 vs CZ-125 (Czechia): learning set = verification set = 25 stations
- NE (Nebraska): learning set = verification set = 25 stations

Validation is made in terms of annual and 30-year maxima of:

a) temperature characteristics
TMAX = daily max. temperature
TMIN = daily min. temperature

b) precipitation characteristics
PREC = daily precipitation
L(Heat) = interdiurnal TMAX drop
L(Cold) = interdiurnal TMIN drop
L(Wet) = length of wet spell*
L(Dry) = length of cold spell

Validation is made in terms of annual and 30-year maxima of:

TMAX = daily max. temperature
TMIN = daily min. temperature

Precipitation:
PREC = daily precipitation
L(Heat) = interdiurnal TMAX drop
L(Cold) = interdiurnal TMIN drop
L(Wet) = length of wet spell*
L(Dry) = length of dry spell*

The present tests revealed that the best-calibrated interpolated raw WG and then crop-simulated provides us with a better estimate of crop yields in the weather-ungauged site compared to site simulated and then interpolated.

The present tests revealed that the best calibrated interpolated raw WG and then crop-simulated provides us with a better estimate of crop yields in the weather-ungauged site compared to site simulated and then interpolated.

The present results indicate that the interpolated weather generator is applicable if we have enough stations in the area of interest (the required density of stations may depend on the orography) the weather series produced by the interpolated weather generator may provide a representative weather series for a weather-ungauged location.

The employment of the interpolated weather generator and subsequent crop model simulation at a site of interest results in an estimate of crop model characteristics that are superior to the interpolation of the crop model output from the surrounding stations where the crop model was run with the observed series.