Evaluation of multisite synthetic data generated by spatial weather generator and long climate data series

Leszek Kuchar[1,*], and *Slawomir* Iwanski[1]

[1]Wroclaw University of Environmental and Life Sciences, Department of Mathematics, Grunwaldzka 55, 50357 Wroclaw, Poland

Abstract. In this paper a new validation test for the spatial weather generator SWGEN producing the multisite daily time series of solar radiation, temperature and precipitation is presented. The method was tested by comparing statistics of 1000 years of generated data with extra long series of 35 years of observed weather parameters and 24 sites of meteorological stations for south-west Poland. The method evaluation showed that the means (sums) and variances of generated data were comparable with observed climatic data aggregated for months, seasons and years.

1 Introduction

Compared with the well-known weather generators producing the data for the simple station, a spatial weather generator is much more sophisticated method from the mathematical point of view. The main idea behind constructing the spatial weather generator is an application for the complex processes evaluation in a given region including correlations among sites. The most well-known applications of spatial data simulated by spatial weather generators are used in hydrology and water management. Particularly, the forthcoming hydrology in the river catchments, runoff response, and predictions are strongly impacted by future climate and possible change [1-13]. The daily flow simulation, particularly seasonal extremes for future climate conditions given by different scenarios are important for water management and hydrology for several reasons [10-11, 14-18]. Above applications indicate that high-quality spatial weather generator is required. A high-quality spatial weather generator means that the generated data are same distributed as the climate data [10, 19-20]. This means that basic statistical measures such as averages, variances and others are very close.

The paper presents a new evaluation of the spatial weather generator SWGEN based on long climate data series and 24 stations from south-west Poland.

---

* Corresponding author: Leszek.Kuchar@gmail.com

2 Spatial weather generator − SWGEN

The idea of simulation of spatial data by the spatial weather generator is widely described in the literature [1, 6, 13, 16-17, 21-24]. For several applications in Poland the spatial weather generator SWGEN is used as the best downscaling method to produce *n* years of synthetic daily data on potentially possible weather course at *k* stations [9, 25-26]. The SWGEN model generates total precipitation by means of the first-order Markov chain to determine the occurrence of wet/dry days, and then for the amount of precipitation the multidimensional two-parameter gamma distribution is used [8, 11, 18]:

$$\left(\Gamma_m(\alpha_1, \beta_1), \dots, \Gamma_m(\alpha_k, \beta_k)\right) \tag{1}$$

where *m* is the month number (*m* = 1, …, 12, i.e. January = 1, Feburary = 2, ..., December = 12) and *k* is the location number. Daily values of solar radiation (SR), temperature maximum (T_{max}) and minimum (T_{min}) are treated as a multidimensional time series AR(1) in the following form:

$$\mathbf{X_t} = \mathbf{\Phi_m} \cdot \mathbf{X_{t-1}} + \mathbf{\varepsilon_t} \tag{2}$$

where $\mathbf{X_t}$ and $\mathbf{X_{t-1}}$ are vectors ($3k \times 1$) of standardized values for all three variables for day *t* and *t*-1, $\mathbf{\varepsilon_t}$ is a vector ($3k \times 1$) of independent random components normally distributed with vector of means equal to zero and matrix of covariance $\mathbf{\Sigma_m}$, and $\mathbf{\Phi_m}$ (for *m* = 1, …, 12) is a matrix of parameters [8, 11, 17].

3 Data generation

According to the model of spatial weather generator, characteristics of required meteorological parameters: solar radiation (SR), minimum and maximum temperature (T_{min} and T_{max}) and precipitation total (P) were estimated for each station of the area. In some cases of missing data, solar radiation values were obtained from sky cover measurements (IS) according to the Black formula [25, 27]. The characteristic of each analyzed meteorological parameter was represented by its monthly mean value and standard deviation. In addition, the spatial correlations among variables from all stations were added to the characteristics. For the stations with the lack of measurements of a given meteorological parameter, these characteristics were obtained with the use of interpolation techniques: ordinary kriging and inverse distance weighting method. The selection of better interpolation techniques was performed with the use of cross-validation method with the criteria based on the value of RMSE (root mean square error) [25, 27-28].

Then, the spatial weather generator SWGEN is used to produce new long series of 1000 years (comparing to previous study [25-27]) of synthetic data for 24 stations.

For the comparison of observed and generated data, the procedure shown in Fig. 1 is applied.
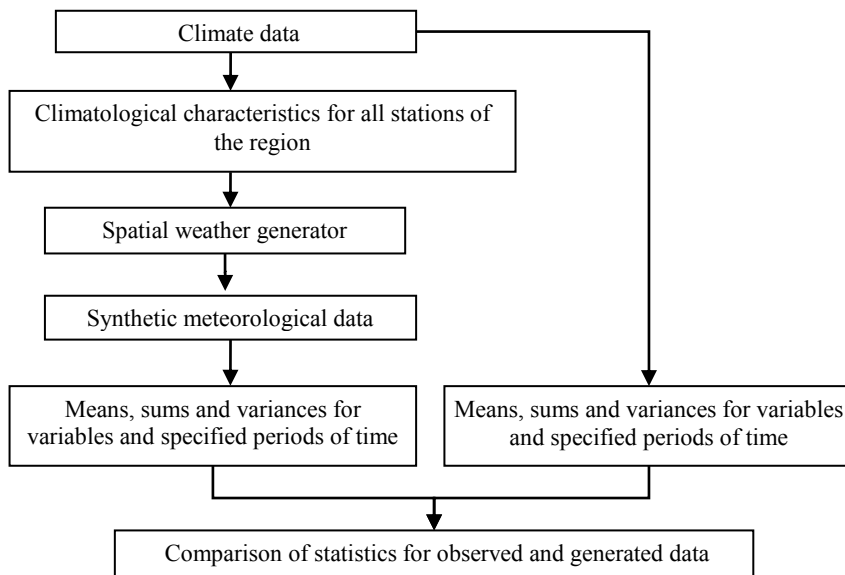
Fig. 1. Diagram of evaluation of generated data (see also [27]).

4 Meteorological data and study area

The data simulation using the scheme (Fig. 1) was applied to a south-west region of Poland in the basin of the Odra River accounting for the area of approximately 65000 km^2 (Fig. 2). Daily data of solar radiation, maximum and minimum air temperature, and total precipitation of a 35-year data series (1981−2015) of the meteorological network within the study area, were obtained for 24 stations from the Institute of Meteorology and Water Management National Research Institute (Table 1).

Fig. 2. Area of study – south-west of Poland, approximately 65000 km^2.

Table 1. List of weather stations used for SWGEN model test.

Weather stations	Longitude	Latitude	Elevation	Meteorological variable (P – total precipitation [mm], SR – solar radiation [MJ/m^2], IS – cloud cover in tenths, T – temperature [°C])
1. Bolkow	16°06'E	50°55'N	361 m	P
2. Chocianow	15°55'E	51°25'N	146 m	P
3. Chojnow	15°56'E	51°17'N	162 m	P
4. Dobromierz	16°15'E	50°55'N	278 m	P
5. Iwiny	15°42'E	51°12'N	211 m	P
6. Jawor	16°11'E	51°03'N	178 m	P
7. Kaczorow	15°58'E	50°55'N	487 m	P
8. Legnica	16°12'E	51°12'N	107 m	SR, IS, T, P
9. Lubin	16°12'E	51°24'N	127 m	P
10. Stanislawow	16°01'E	51°04'N	373 m	P
11. Strzegom	16°21'E	50°58'N	227 m	P
12. Tomaszow Boleslawiecki	15°41'E	51°17'N	186 m	T, P
13. Twardocice	15°45'E	51°06'N	258 m	P
14. Wojcieszow Dolny	15°55'E	50°59'N	397 m	P
15. Zagrodno	15°52'E	51°12'N	193 m	P
16. Zlotoryja	15°56'E	51°07'N	223 m	P
17. Chwalkowie	16°37'E	51°27'N	159 m	T
18. Jelenia Gora	15°48'E	50°54'N	345 m	SR, IS, T
19. Leszno	16°32'E	51°50'N	87 m	SR, IS, T
20. Polkowice Dolne	16°03'E	51°30'N	150 m	T
21. Pszenno	16°33'E	50°51'N	205 m	T
22. Wroclaw	16°53'E	51°06'N	110 m	SR, IS, T
23. Zgorzelec	15°02'E	51°08'N	198 m	T
24. Zielona Gora	15°32'E	51°56'N	169 m	SR, IS, T

4 Results

The SWGEN model was examined by comparison: observed data vs. generated data. Daily data of solar radiation (SR), minimum and maximum temperature (T_{min}, T_{max}) and precipitation total (P) were aggregated to annual, two hydrologic seasons (from April to September (H-S) and from October to March (H-W)) and 12 monthly periods (January, ..., December) for all 24 stations. For all synthetic data as well as for the data from climatology and for given periods, means (sum) and standard deviation (SD) were determined. For the above computations, 1000 years of generated data and series of 35 years of observed data were used. Next, absolute differences (abs) between observed and generated parameters (mean, SD), as well as relative absolute differences (Rel) in the form *Rel = abs(observed – generated)×100% /observed* were evaluated [13, 29].

Obtained results are presented in Table 2.

Table 2. Average absolute and relative error of means (sum) and standard deviations of solar radiation, temperature maximum and minimum, and precipitation for generated and observed data in different time periods (average for 24 stations).

Time period	Error	SR [%]		T_{max} [%]		T_{min} [%]		P [%]	
		Mean	SD	Mean	SD	Mean	SD	Sum	SD
Year	Abs.	0.02	0.02	0.01	0.02	0.01	0.04	3.30	1.05
	Rel.	0.19	0.24	0.05	0.23	0.18	0.65	0.54	17.53
H-S	Abs.	0.02	0.04	0.02	0.04	0.00	0.04	2.06	1.47
	Rel.	0.10	0.42	0.11	0.61	0.04	1.03	0.53	19.03
H-W	Abs.	0.03	0.01	0.01	0.03	0.01	0.10	1.48	0.54
	Rel.	0.42	0.19	0.15	0.51	3.48	2.03	0.68	15.60
Jan	Abs.	0.04	0.09	0.09	0.29	0.04	0.32	0.35	0.30
	Rel.	1.37	4.94	2.42	5.68	1.58	6.20	1.39	11.94
Feb	Abs.	0.01	0.04	0.09	0.18	0.14	0.29	1.04	0.18
	Rel.	0.23	1.19	1.73	3.30	6.85	5.31	3.25	7.03
Mar	Abs.	0.04	0.11	0.01	0.34	0.01	0.05	0.52	0.56
	Rel.	0.46	2.43	0.06	6.91	1.83	1.29	1.14	14.66
Apr	Abs.	0.07	0.14	0.05	0.15	0.08	0.09	0.76	0.97
	Rel.	0.45	1.94	0.32	2.82	2.49	2.31	1.77	19.43
May	Abs.	0.29	0.06	0.15	0.22	0.02	0.12	0.75	0.87
	Rel.	1.37	0.70	0.79	4.63	0.24	3.72	1.32	13.95
Jun	Abs.	0.21	0.36	0.16	0.21	0.08	0.05	1.04	0.90
	Rel.	0.99	4.31	0.75	4.80	0.71	1.74	1.39	12.08
Jul	Abs.	0.06	0.20	0.04	0.07	0.02	0.12	1.78	3.04
	Rel.	0.30	2.46	0.16	1.68	0.17	4.35	1.77	24.68
Aug	Abs.	0.11	0.17	0.14	0.10	0.03	0.10	1.43	1.52
	Rel.	0.55	2.27	0.55	2.09	0.25	3.32	2.41	20.03
Sep	Abs.	0.03	0.24	0.01	0.17	0.01	0.10	0.47	0.87
	Rel.	0.28	4.35	0.06	4.22	0.06	3.16	0.93	14.43
Oct	Abs.	0.06	0.18	0.05	0.10	0.06	0.05	0.84	0.60
	Rel.	0.79	4.66	0.33	1.99	1.26	1.14	2.35	15.82
Nov	Abs.	0.03	0.02	0.02	0.11	0.00	0.08	0.43	0.87
	Rel.	0.97	0.74	0.35	2.32	2.73	1.95	1.12	21.78
Dec	Abs.	0.03	0.03	0.08	0.20	0.08	0.09	0.48	0.27
	Rel.	1.31	2.09	2.50	4.09	3.52	1.78	1.41	10.51

The above study and results are comparable to earlier research [25-26]. However due to much longer data series used for the test, higher number of stations and larger area of study, new results are significant. Table 2 shows that the spatial weather generator reproduces satisfactorily daily values of meteorological variables with proper averages, sums and standard deviations.

The averages for generated solar radiation (SR) and both types of temperature (T_{min}, T_{max}) correspond to the observed ones. As in previous research, the higher values of temperature (T_{max} vs. T_{min}) for simulated data had a higher standard deviation than the observed ones and, as well, the standard deviation of simulated T_{max} was slightly higher than that of observed T_{max} values. The observed error defined as the relative absolute difference (Rel) does not exceed 7%.

Evaluation of the spatial weather generator was made by means of produced data also in terms of its ability to reproduce the precipitation distribution by comparison of generated and observed daily precipitation totals and standard deviation for all collected stations.

Although simulations show good compliance for the sum of precipitation totals, for observed and generated data, some uncertainties were identified for standard deviations.

The relative difference (Rel) of standard deviation for the annual period is about 18% while for monthly periods varies from 7% (February) to 25% (July).

5 Conclusion

The above validation test for the network of 24 stations, long data series (35 years) and a relatively large area (65000 km^2) leads to the following remarks.

Generated daily meteorological data, i.e. solar radiation, maximum and minimum temperatures as well as total precipitation, counted in periods from month to year are produced with low absolute and relative errors. Generated data errors are acceptable from the point of view of several applications such as hydrology prediction in river basins, evaluation of climate change for river hydrology, future water management in the river catchments up to 40−80 years, various possibilities for climate change study and others.

The presented test proved that the spatial weather generator producing daily meteorological data is an important tool for a decision support system, as the best downscaling technique.

References

1. G.A. Baigorria, J.W. Jones, J. Clim. **23**, 22, 5990 (2010)
2. J.J. Beersma, T.A. Buishand, Clim. Res. **25**, 121 (2003)
3. S. Bergström, B. Carlsson, M. Gardelin, G. Lindström, A. Pettersson, M. Rummukainen, Clim. Res. **16**, 101 (2001)
4. F. Brissette, M. Khalili, R. Leconte, J. Hydrol. **345**, 121 (2007)
5. H. Chen, C.Y. Xu, S. Guo, J. Hydrol. **434-435**, 36 (2012)
6. J. Chen, F.P. Brissette, X.C. Zhang, Trans. ASABE **57**, 1375 (2014)
7. IPCC WG I, II and III, *Climate Change, Synthesis Report* (IPCC, Geneva, 2014)
8. D.I. Jeong, A. St-Hilaire, T.B.M.J. Ouarda, P. Gachon, Clim. Res. **54**, 2, 129 (2012)
9. R. Srikanthan, G.G.S. Pegram, J. Hydrol. **371**, 142 (2009)
10. T. Tokarczyk, W. Szalinska, A. Tiukalo, J. Jelowicki, A. Chorazyczewski, Meteorol. Hydrol. Water Manage. **4**, 1, 65 (2016)
11. C.G. Kilsby, P.D. Jones, A. Burton, A.C. Ford, H.J. Fowler, P. James, A. Smith, R.L. Wilby, Environ. Mod. Soft. **22**, 12, 1705 (2007)
12. D.S. Wilks, WIREs Clim. Change **1**, 898-907 (2010) doi:10.1002/wcc.85
13. Y. Xu, Ch. Ma, S. Pan, Q. Zhu, Q. Ran, J. Zhejiang Univ. Sci. A **15**, 3, 219 (2014)
14. T. Barnett, T. Malone, W. Pennell, D. Stammer, B. Semtner, W. Washington, Clim. Change **62**, 1 (2004)
15. D.N. Graham, M.B. Butts, *Watershed Models* V.P. Singh, D.K. Frevert (Eds.), (CRC Press, Boca Raton, 2006)
16. M. Khalili, F. Brissette, R. Leconte, J. Am. Wat. Res. Ass. **47**, 2, 303 (2011)
17. M. Khalili, R. Leconte, F. Brissette, J. Hydrometeorol. **8**, 396 (2007)
18. M. Wdowikowski, B. Kazmierczak, O. Ledvinka, Meteorol. Hydrol. Water Manage. **4**, 1, 53 (2016)
19. M. Khalili, V-T-V. Nguyen, Clim. Dyn. **49**, 7-8, 2261 (2017)
20. A. Michalski, Meteorol. Hydrol. Water Manage. **4**, 1, 41 (2016)
21. J. Chen, F.P. Brissette, X.J. Zhang, J. Hydrol. Eng. **21**, 2 (2016)
22. D.I. Jeong, A. St-Hilaire, T.B.M.J. Ouarda, P. Gachon, Int. J. Clim. **33**, 10, 2431 (2013)
23. D.I. Müller-Wohlfeil, G. Bürger, W. Lahmer, Clim. Change **47**, 61 (2000)

24. D. Rayner, C. Achberger, D. Chen, Advances in Clim. Change Res. **7**, 70 (2016)
25. L. Kuchar, S. Iwanski, L. Jelonek, W. Szalinska, Geografie **119**, 1 (2014)
26. L. Kuchar, S. Iwanski, L. Jelonek, E3S Web of Conf. **17**, 00046 (2017) doi: 10.1051/e3sconf/20171700046
27. S. Iwanski, L. Kuchar, Acta Scientiarum Polonorum - Formatio Circumiectus **2**, 113 (2003)
28. R.E. Walpole, R.H. Myers, S.L. Myers, K. Ye, *Probability and statistics for engineers and scientists* (Prentice Hall, NJ 2002)
29. L. Kuchar, Math. Comp. Symul. **65**, 1-2, 69 (2004)