Weather Generator and Climate Change Scenario Generator for Climate Risk Assessment

(version 0.1.0 BETA)
August,2019

Technical Documentation

Authors:

Dibesh Shrestha, Email: dibeshshrestha@live.com
Divas B. Basnyat, Email: divas@ndri.org.np
Water and Climate Team
Nepal Development Research Institute

Nepal Development Research Institute

Weather Generator and Climate Change Scenario Generator for Climate Risk Assessment

Technical Documentation, 2019

Table of Contents

PREFACE

This technical documentation will provide the methodology embedded in the main engine of this tool - 'Weather Generator and Climate Change Scenario Generator (version 0.1.0 Beta)'. It will focus brief description of the processes involved in generation weather generation and apply any shifts / changes to the weather variables in order to generate climate change scenarios.

This tool is developed in Python 3.7.

Please feel free to use the tool and send us email for queries, bugs or issues related to this tool.

Regards,

Dibesh Shrestha, Email: dibeshshrestha@live.com
Divas B. Basnyat, Ph.D, Email: divas@ndri.org.np
Water and Climate Team
Nepal Development Research Institute
August, 2019

Introduction

‘Weather Generator and Climate Change Scenario Generator (version 0.1.0 Beta)’ is a tool aiming to support climate risk assessments of water resources system. It is mainly designed to produce inputs for climate stress test and it provides an interfaces for weather generating process and enforcing changes in climatic means to produce climate change scenarios.

The tool is developed based on research papers by Apipattanavis et al (2007) and Steinschneider and Brown (2013). This documentation will provide a brief overview of steps followed in order to develop the main engine of weather generation process and apply the climate changes or shifts. Users are requested to read those papers for details about the weather generation processes used in this tool.

This tool provides five major interfaces. Interface ‘Annual Series Simulator’ provides tools to generate annual precipitation series based on historic precipitation series by using ARMA method. It is diversion from wavelet based approach as described by Steinschneider and Brown (2013). Interfaces ‘WG-CRA’ and ‘k-NN WG’ are for weather generation. The former one is conditioned on annual precipitation series as described in Steinschneider and Brown (2013) whereas latter is not conditioned but simply weather generator as formulated in Apipattanavis et al (2007). Interface ‘CC Scenario Generator’ allows to enforce shifts or changes in distributional properties of weather variables by quantile mapping approach for precipitation and simple shifting approach for other variables and it is described in Steinschneider and Brown (2013). Finally, interface ‘Result viewer’ is for graphically viewing the results generated by mentioned interfaces.

References

Apipattanavis, S., G. Podesta´, B. Rajagopalan, and R. W. Katz (2007), A semiparametric multivariate and multisite weather generator, Water Resour. Res., 43, W11401, doi:10.1029/2006WR005714

Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, Water Resour. Res., 49, 7205–7220, doi:10.1002/wrcr.20528.

Overview

This chapter describes basic information of the 'pages' (tabs) in the tool. Out of seven pages, two are information related to this tool. Each of remaining pages has its own independent functions (weather generation, climate change scenario generation, result viewing). They are briefly described as follows:

Pages	Brief description
(with blank tab)	It gives information on name, version and authors of the tool.
Annual Series Simulator	It is for simulating the annual precipitation series using ARMA method. Users can fit ARMA model with (p, q) parameters, simultaneously viewing autocorrelation and partial autocorrelation plots. Users can carry residual analysis and simulate annual series using the fitted ARMA model.
WG-CRA	Users can generate daily weather variables conditioned on annual precipitation values based on provided daily observed weather variables. Users have to supply observed and simulated annual precipitation series, observed areal averaged daily weather variables (can be single site or multisite) and declare which of the variable is precipitation. Users should provide the model parameters and initial conditions. Users can change the transition probability in case of enforcing climate change condition. Please refer to Steinschneider and Brown (2013) for technical details. Users can save the results.
k-NN WG	It is for simulating the weather based on provided daily observed weather variables but without conditioned on annual precipitation series. Users have to observed areal averaged daily weather variables (can be single site or multisite) and declare which of the variable is precipitation. Users should provide the model parameters and initial conditions. Users can change the transition probability in case of enforcing climate change condition. Results will be save in the output directory. Please refer to Apipattanavis et al (2007) for technical details.
CC Scenario Generator	It is for enforcing the long-term changes/ shifts in the simulated or observed climatic variables. For precipitation, users can enforce changes in mean and coefficient of variation of its distribution (which is modelled as gamma distribution). For other variables like temperature, users can shift mean. Users can save the output in desired location.
Result viewer	It is for graphically viewing the results and it is intended for basic viewing only. Users can view the generated daily series, annual / monthly sums and averages of generated series. Users can further use other advanced data analysis tools for further analysis.
Brief Description	It provides brief description of the tool including contact details of the developers.

Methodology

This chapter describes brief methodology on generating weather series and climate change scenarios.

Simulation of annual precipitation series

(Corresponding to Page – 'Annual Series Simulator' in the tool)

In order to model the observed precipitation series and generate new annual precipitation series, Autoregressive Moving Average Method ARMA of order (p, q) is used in the tool. 'Statsmodels' package (https://www.statsmodels.org/stable/index.html) is used for this purpose. Residual analysis can be done using tools as Autocorrelation test (Durbin-Watson statistic) and Normality Test (D'Agostino and Pearson's test).

Simulation of weather variables conditioned on annual precipitation series

(Corresponding to Page – 'WG-CRA' in the tool)

Simulated annual precipitation time series was then disaggregated into daily and at station scale using k-NN weather generator following the approach used by Steinschneider and Brown (2013). Brief description to this method is as follows:

1. For an annual precipitation value (say P) for given year in time series generated using ARMA process, the closest k nearest neighbour is searched from historical annual time series using Euclidean distance metric. Value of k is $\sqrt{N}$ where N is the number of years in annual series.

2. k-NN are sorted in ascending order based on distance metrics and they are given weights such probability of sampling them with replacement is given by the weights. Weights are assigned using discrete kernel function as:

$$K(j) = \frac{\frac{1}{j}}{\sum_{i=1}^{k} \frac{1}{i}}$$

   where $K(j)$ is the probability with which jth closest neighbor to P is resampled.

3. Using weights in step (2), user defined (n) numbers of resamples were selected for P. Daily data of weather variables (Eg: precipitation, maximum temperature, minimum temperature) were arranged year wise corresponding to each of the samples in order.

4. Daily time-series formed in step (3) were used to estimate parameters for daily weather generator. Using weather generator, new daily time series for period of one year for given P is formed.

5. Steps 1-4 are repeated for each of years simulated using ARMA process.

Once the new daily time series of weather variables are obtained from steps 1 to 5, it is used to simulate spatially correlated multivariate and multisite daily weather series. For this, method by Apipattanavis et al. (2007) and Steinschneider and Brown (2013) uses coupled Markov chain (Richardson 1981) and k-NN resampling approach (Rajagopalan and Lall 1999) is used. Brief description of approach for generating daily weather (Apipattanavis et al. 2007, Steinschneider and Brown 2013) is described as follows:

6. First order Markov chain for precipitation states, i.e. the precipitation state in given day only depends on precipitation state on its previous day, is used in this study. The precipitation states are defined as dry (S = 0), wet (S = 1) and extremely wet (S = 2). In this tool, user defined (Eg 80$^{th)}$ percentile of observed daily precipitation series is used as threshold to define wet and extremely wet states. Likewise, user defined precipitation value (Eg. 0.1 mm) is used to separate wet and dry days. Transition probabilities from one precipitation state to another state are computed from the daily time-series compiled in step (1 - 5). As three precipitation states are used, there are nine transition probabilities. These are computed for each of the month. We take data from all the years for month under consideration to compute transition probabilities. In this study, like Steinschneider and Brown (2013), we choose user-defined window in order to select k- nearest neighbors as described in following steps.

7. For a user defined window of size N days (eg- 15 days – if the day is 8th Jan , then it consists all the days from 1st Jan to 15th Jan for all the years in the series), assuming that the precipitation state for previous day, say t-1, is known, the precipitation state for given day, t, is given by the transition probability. Let us assume precipitation state on day t-1 is S_{t-1} (S_{t-1}= 0 or 1 or 2) (known), and three transition probabilities are $[\, p_{S_{t-1}|0}\ p_{S_{t-1}|1}\ p_{S_{t-1}|2}\,]$. Random number, u, from uniform distribution is generated and the state on day t, S_t , is decided as follows:
   (a) if u <=$p_{S_{t-1}|0}$ then state S_t = 0
   (b) if $p_{S_{t-1}|0}$< u <= $p_{S_{t-1}|0} + p_{S_{t-1}|1}$ then state S_t = 1 (c) if $p_{S_{t-1}|0} + p_{S_{t-1}|1}$ < u <= $p_{S_{t-1}|0} + p_{S_{t-1}|1} + p_{S_{t-1}|2}$ then state S_t = 2.

8. Once the states for day t-1 and t, [$S_{t-1}\, S_t$], are known, we search and arrange all the days in moving window with states [$S_{t-1}\, S_t$] in the dataset from all the years in series. Let the arranged data be $\left[X_{S_{t-1}}^{obs}\ , X_{S_t}^{obs}\right]$. Here, for example,

$$X_{S_{t-1}}^{obs} = \begin{bmatrix} P_{S_{t-1}}^{obs} & Tmax_{S_{t-1}}^{obs} & Tmin_{S_{t-1}}^{obs} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}_{nt\ x\ 3} \quad \&\ X_{S_t}^{obs} = \begin{bmatrix} P_{S_t}^{obs} & Tmax_{S_t}^{obs} & Tmin_{S_t}^{obs} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}_{nt\ x\ 3}$$

9. Superscript 'obs' indicate that the variables are from the time series from step (b - 3). $X_{S_{t-1}}^{obs}$ is set of variables corresponding to state S_{t-1} and $X_{S_t}^{obs}$ to state S_t . Example variables - P, Tmax and Tmin represent precipitation, maximum temperature and minimum temperature. It is to be noted that the indices (corresponding dates) are also to be stored accordingly. 'nt' represents the number of observations in dataset.

10. Let assume that weather variables in day t-1 are X_{t-1}= $[P_{t-1}, Tmax_{t-1}, Tmin_{t-1}]$ (example). We search the k-nearest neighbour to X_{t-1} in $X^{obs}_{S_{t-1}}$ using Euclidean distance metrics. Equal weights (=1) are given to each of the variables. Here, we set $k = \sqrt{nt}$ following Rajagopalan and Lall (1999).

11. k- nearest neighbours are arranged in ascending order according to values of distance metrics as:

$$X^{obs,k-NN}_{S_{t-1}} = \begin{bmatrix} \cdots & \cdots & \cdots \\ P^{obs\,(j)}_{S_{t-1}} & Tmax^{obs(j)}_{S_{t-1}} & Tmin^{obs(j)}_{S_{t-1}} \\ \cdots & \cdots & \cdots \end{bmatrix}_{k\,x\,3}$$

They are given weights such probability of sampling them with replacement is given by the weights. Weights are assigned using discrete kernel function as:

$$K(j) = \frac{\frac{1}{j}}{\sum_{i=1}^{k} \frac{1}{i}}$$

where $K(j)$ is the probability with which jth closest neighbour X_{t-1} to is resampled.

12. Let the resampled weather variables set in $X^{obs,k-NN}_{S_{t-1}}$ be $\left[P^{obs\,(l)}_{S_{t-1}} \quad Tmax^{obs(l)}_{S_{t-1}} \quad Tmin^{obs(l)}_{S_{t-1}}\right]$ with 'l^{th}'position. The weather variables in the successive day, i.e. $\left[P^{obs\,(l)}_{S_{t}} \quad Tmax^{obs(l)}_{S_{t}} \quad Tmin^{obs(l)}_{S_{t}}\right]$, is the simulated weather variables for day t.

13. Steps 6-12 are repeated to generate the daily weather series.

Simulation of weather variables **without** conditioning on annual precipitation series

(Corresponding to Page 'k-NN WG' in the tool)

Using observed daily time series of weather variables (single site or multisite), spatially correlated multivariate and multisite daily weather series are simulated by method described in Apipattanavis et al. (2007). Step-wise description is as follows:

1. First order Markov chain for precipitation states, i.e. the precipitation state in given day only depends on precipitation state on its previous day, is used in this study. The precipitation states are defined as dry (S = 0), wet (S = 1) and extremely wet (S = 2). In this tool, user defined (Eg 80$^{th)}$ percentile of observed daily precipitation series is used as threshold to define wet and extremely wet states. Likewise, user defined precipitation value (Eg. 0.1 mm) is used to separate wet and dry days. Transition probabilities from one precipitation state to another state are computed from the daily time-series compiled in step (1 - 5). As three precipitation states are used, there are nine transition probabilities. These are computed for each of the month. We take data from all the years for month under consideration to compute transition probabilities. In this study, like Steinschneider and Brown (2013), we choose user-defined window in order to select k- nearest neighbors as described in following steps.

2. For a user defined window of size N days (eg- 15 days – if the day is 8th January, then it consists all the days from 1st Jan to 15th Jan for all the years in the series), assuming that the precipitation state for previous day, say t-1, is known, the precipitation state for given day, t, is given by the transition probability. Let us assume precipitation state on day t-1 is S_{t-1} (S_{t-1}= 0 or 1 or 2) (known), and three transition probabilities are [$p_{S_{t-1}|0}$ $p_{S_{t-1}|1}$ $p_{S_{t-1}|2}$]. Random number, u, from uniform distribution is generated and the state on day t, S_t , is decided as follows:
   (a) if u <=$p_{S_{t-1}|0}$ then state S_t = 0
   (b) if $p_{S_{t-1}|0}$< u <= $p_{S_{t-1}|0} + p_{S_{t-1}|1}$ then state S_t = 1 (c) if $p_{S_{t-1}|0} + p_{S_{t-1}|1}$ < u <= $p_{S_{t-1}|0} + p_{S_{t-1}|1} + p_{S_{t-1}|2}$ then state S_t = 2.

3. Once the states for day t-1 and t, [S_{t-1} S_t], are known, we search and arrange all the days in moving window with states [S_{t-1} S_t] in the dataset from all the years in series. Let the arranged data be $\left[X_{S_{t-1}}^{obs}, X_{S_t}^{obs}\right]$. Here, for example,

$$X_{S_{t-1}}^{obs} = \begin{bmatrix} P_{S_{t-1}}^{obs} & Tmax_{S_{t-1}}^{obs} & Tmin_{S_{t-1}}^{obs} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}_{nt\ x\ 3} \quad \& \ X_{S_t}^{obs} = \begin{bmatrix} P_{S_t}^{obs} & Tmax_{S_t}^{obs} & Tmin_{S_t}^{obs} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}_{nt\ x\ 3}$$

4. Superscript 'obs' indicate that the variables are from the time series from step (b - 3). $X_{S_{t-1}}^{obs}$ is set of variables corresponding to state S_{t-1} and $X_{S_t}^{obs}$ to state S_t . Example variables - P, Tmax and Tmin represent precipitation, maximum temperature and minimum temperature. It is to be noted that the indices (corresponding dates) are also to be stored accordingly. 'nt' represents the number of observations in dataset.

5. Let assume that weather variables in day t-1 are X_{t-1}= [$P_{t-1}, Tmax_{t-1}, Tmin_{t-1}$] (example). We search the k-nearest neighbour to X_{t-1} in $X_{S_{t-1}}^{obs}$ using Euclidean distance metrics. Equal weights (=1) are given to each of the variables. Here, we set $k = \sqrt{nt}$ following Rajagopalan and Lall (1999).

6. k- nearest neighbours are arranged in ascending order according to values of distance metrics as:

$$X_{S_{t-1}}^{obs,k-NN} = \begin{bmatrix} \dots & \dots & \dots \\ P_{S_{t-1}}^{obs\,(j)} & Tmax_{S_{t-1}}^{obs(j)} & Tmin_{S_{t-1}}^{obs(j)} \\ \dots & \dots & \dots \end{bmatrix}_{k\ x\ 3}$$

   They are given weights such probability of sampling them with replacement is given by the weights. Weights are assigned using discrete kernel function as:

$$K(j) = \frac{\frac{1}{j}}{\sum_{i=1}^{k} \frac{1}{i}}$$

   where $K(j)$ is the probability with which jth closest neighbour X_{t-1} to is resampled.

7. Let the resampled weather variables set in $X_{S_{t-1}}^{obs,k-NN}$ be $\left[P_{S_{t-1}}^{obs\,(l)} \quad Tmax_{S_{t-1}}^{obs(l)} \quad Tmin_{S_{t-1}}^{obs(l)}\right]$ with 'l^{th}'position. The weather variables in the successive day, i.e. $\left[P_{S_t}^{obs\,(l)} \quad Tmax_{S_t}^{obs(l)} \quad Tmin_{S_t}^{obs(l)}\right]$, is the simulated weather variables for day t.

8. Steps 6-12 are repeated to generate the daily weather series.

Enforcement of climate change scenario

(Corresponding to Page 'k-NN WG' in the tool)

For Precipitation: Climate change scenario was enforced into this newly simulated daily series using quantile mapping for precipitation as in Steinschneider and Brown (2013). Basis of quantile mapping of a variate from reference distribution to a variate with target distribution is that the probability of being less or equal to given value of a variate in reference distribution is equal to the probability of being less or equal to corresponding value of the mapped variate in target distribution (Panofsky and Brier 1968). In this study, for precipitation, quantile mapping is carried out as follows:

1. For given month m, the precipitation series is fitted by gamma distribution with shape parameter (α) and scale parameter (β). This is reference distribution. The relation between the mean (μ) and standard deviation (σ) between these parameters are $\mu = \alpha\beta$ and $\sigma = \alpha\beta^2$.
2. Now we introduce the change in mean or Coefficient of Variation or both as climate change scenario. Then, new parameters are computed using the relationship in 2 defining new gamma distribution. This new distribution is distribution for climate change scenario. This is target distribution.
3. Now precipitation value from reference distribution is mapped to target distribution such that they have equi-probability in both distributions. The corresponding value from target distribution is the new value in climate change scenario.

For other variables (like Temperature) - For enforcing climate change scenario for other variables than precipitation, additive factor was used as in Steinschneider and Brown (2013). For a given month m, we add constant value of temperature to all data in that month.

References

Apipattanavis, S., G. Podesta´, B. Rajagopalan, and R. W. Katz (2007), A semiparametric multivariate and multisite weather generator, Water Resour. Res., 43, W11401, doi:10.1029/2006WR005714

Panofsky, H.A. and Brier, G.W. (1968), Some Applications of Statistics to Meteorology. Earth and Mineral Sciences Continuing Education, College of Earth and Mineral Sciences.

Rajagopalan, B., and Lall, U. (1999), A k-nearest-neighbor simulator for daily precipitation and other weather variables, *Water Resour. Res.*, 35(10), 3089– 3101, doi:10.1029/1999WR900028.

Richardson, C. W. (1981), Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resour. Res.*, 17(1), 182– 190, doi:10.1029/WR017i001p00182.

Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, Water Resour. Res., 49, 7205–7220, doi:10.1002/wrcr.20528.

Codes

This chapter provides the code written in Python 3.7.3 developed in order to run this tool.

'WGEN' package

Package developed was named as 'WGEN' with a single module called 'wg.py'. The codes are provided in the installation folder name CODES. Besides, it is provided in Annex – 1. details documentation of this package will be later released as open source.

GUI interface

GUI interface was developed using 'wxpython' and codes are provided in the installation folder named CODES under 'WG_CRA.py'. Likewise, the codes are provided in Annex – 2.

Annex – 1

CODES – ‘wg.py’

Annex - 2

CODES for GUI development – ‘WG_CRA.py’