Bias correction techniques for meteorological data of A2 scenario climate model output in Chao Phraya River Basin of Thailand

Somchai Baimoung¹, Taikan Oki², Boonlert Archevarahuprok³, Aphantree Yuttaphan³ and Manoon Pangpom³

¹Ministry of Information and Communication Technology, Thailand
²The University of Tokyo, Japan
³Thai Meteorological Department, Thailand

Abstract:

Statistical and dynamic methods were used in the downscaling process from Global Climate Model (GCM) to Regional Climate Model (RCM). We selected the European Centre for Medium-Range Weather Forecasts model, Hamburg version 4 (ECHAM4) with 300 × 300 km resolution for A2 scenario. We focused on SE Asia domain located between 20°S to 30°N and 80°E to 135°E for 1960–2099 with wind components, temperature, geo-potential height, and specific humidity as data input in Providing Regional Climates for Impacts Studies (PRECIS) RCM analysis. The downscaling process output was 50 km resolution for 1971–2010 and precipitation, temperature, wind, relative humidity, radiation from 8 meteorological stations in Chao Phraya River Basin; Lampang, Suphanburi, Nan, Sisamonrong, Takfa, Chainat, Uthong and Bangna selected and used for bias correction. Three methods, namely 1) adjusting the mean based on RCM, 2) adjusting the mean based on observation, and 3) quantile-based mapping were used. Methods were compared using observed climatic data, RCM outputs of calibration period, and RCM outputs from the validation period. RMSE was found to be lower for method 2 compared to other methods implying a relatively superior technique for improving the model. As such method 2 was used to correct the PRECIS products during 2001–2009. These products are useful in the studies of impact of climate change and for early warning systems in Thailand.

KEYWORDS A2 scenario climate model; PRECIS; bias correction techniques

INTRODUCTION

Climate is usually described in terms of the mean and variability of temperature, precipitation and wind over a period of time, ranging from months to millions of years (the classical period is 30 years). A major limiting factor to the predictability of weather beyond several days is a fundamental dynamical property of the atmosphere. In the 1960s, meteorologist Edward Lorenz discovered that very slight differences in initial conditions can produce very different forecast results. A significant problem facing weather forecasting is the need to know all the conditions at the start of the forecast period, so it can be useful to think of climate as dealing with the background conditions for weather. More precisely, climate can be viewed as concerning the status of the entire Earth system, including the atmosphere, land, oceans, snow, ice and living things that serve as the global background conditions for determining weather patterns.

Projections of future climate are shaped by fundamental changes in heat energy balance in the Earth system, in particular the increasing intensity of the greenhouse effect that traps heat near the Earth’s surface, determined by the amount of carbon dioxide and other greenhouse gases in the atmosphere. Projecting changes in climate due to changes in the concentration of greenhouse gases 50 years from now is a very different and theoretically solvable problem in contrast with forecasting weather patterns several weeks ahead. To put it another way, long-term variations brought by changes in the composition of the atmosphere are much more predictable than individual weather events (IPCC, 2007).

IPCC Special Report on Emissions Scenarios (SRES) were constructed to explore pathways of future greenhouse gas and aerosol precursor emissions, derived from self-consistent sets of assumptions about energy use, population growth, economic development, and other factors. The SRES team defined four scenario families, labeled A1, A2, B1 and B2, describing the relationships between greenhouse gas and aerosol emissions as the driving forces and their evolution during the 21st century for large world regions and globally as presented in Table I.

Several human activities are the main sources of the increase in the concentration of greenhouse gases. The increased trend may lead to change in global and regional climate features, such as variation of average temperature and precipitation etc.

The predicted output by Global Climate Model (GCM) and Regional Climate Model (RCM) of the 21st century driven by greenhouse gases and aerosol emission scenarios such as B1, A1B and A2 are mostly used to study climate change and their effects. One factor that has limited the use of climate forecast information from GCM in climate change impact prediction such as hydrological prediction is the scale mismatch between climate model output and the spatial scale at which the hydrological models are applied (e.g., Wilby et al., 2000; Wood et al., 2002).

Although RCMs are powerful tools for describing regional and even smaller scale climate conditions they still feature severe systematic errors (Themell et al., 2011), particularly the small-scale patterns of daily precipitation.
which are highly dependent on model resolution and some parameterization thus often making them unsuitable for
direct use in climate change impact assessment studies
(Fowler et al., 2007). This problem has been solved by
applying the various bias-correction techniques to RCM
outputs to make them suitable for the studies.

RESEARCH OBJECTIVES

This study was part of a project which aimed to predict
the yield of sugarcane under climate change condition. The
project links the climate model with crop simulation model
for predicting sugarcane yield. The daily meteorological
outputs of the climate model were used as input data in the
crop simulation model. The crop simulation model was not
a simplified function of mean seasonal climatic conditions,
but a dynamic function and nonlinear interactions between
weather, soil, water and plant. Most crop simulation models
typically simulate these interactions on a daily time step.
Therefore, this study attempted to derive climatic parameters
which would be used for crop simulation model with the
objective as follows:

- To use Regional Climate Model (RCM) so-called
  PRECIS which stands for Providing Regional Climates for
  Impacts Studies (Jones et al., 2004) for projecting the
  climatic parameters from 1960 to 2099 under A2 Scenario
  by using the data from meteorological stations along with
  the Chao Phraya River Basins.

- To use the appropriate bias correction techniques for
  filtering the climate model data set output and generate new
daily climate data set which can be used for impact and
adaptation studies of climate change on several human
activities.

STUDY AREA

The study area is the Chao Phraya River Basin located
in the northern and central parts of Thailand. The basin
can be divided into eight sub-basins based on the natural
distribution of its river system. The sub-basins consist of
Ping, Wang, Yom, Nan, Chao Phraya, Sakae Krang, Tha
Chin and Pasak as shown in Figure 1. The data from
meteorological stations, namely; Lampang, Suphanburi,
Nan, Sisamrong, Takfa, Chainat, Uthong and Bangna were
used in this study.

MATERIALS AND METHODS

All data and processes were collected and carried out step by step as shown in Figure 2.

Daily meteorological parameters collection from
ECHAM4 products

In order to investigate climatic changes and their impact
in the study area which needed to be performed at a much
finer spatial scale, one of the techniques for adding small-
scale detailed information to the large-scale global projection
is through the use of RCMs. Therefore, the RCM-PRECIS
which was developed by the Hadley Centre at the UK Met
Offices, has been implemented on the open SUSE 10.3
LINUX operating system.

The PRECIS climate model is an atmospheric and land
surface model of limited area and high resolution which is
locatable over any part of the globe. Dynamical flow, the

![Figure 1. Study area](image-url)
atmospheric sulphur cycle, clouds and precipitation, radiative processes, the land surface and the deep soil are all formulated. The atmospheric component of the PRECIS model is a hydrostatic version of the full primitive equations. There are 19 vertical levels in the atmosphere, the lowest at ~50 m and the highest at 0.5 hPa. The horizontal resolution is 0.44° × 0.44° and a time step of 5 minutes is used to maintain numerical stability (Jones et al., 2004).

The climate simulation in this research used initial and boundary conditions of the European Centre for Medium-Range Weather Forecasts model, Hamburg version 4 (ECHAM4) GCM data for scenario A2 with a low resolution of 2.8° × 2.8° or 300 × 300 kilo-meters for the period 1960–2099, and coverage spanning 20°S to 30°N and 80°E to 135°E in Southeast Asia. The lateral and surface boundary conditions are updated every six hours. The GCM data consisted of six meteorological parameters namely; wind in horizontal and vertical components (u, v, w), temperature, height and specific humidity. Downscaling was performed to a higher 50 kilometers resolution for period 1960–2099. Given the performance of the computer system used, it was possible to run the analysis at 50 kilometers grid resolution. The processing of PRECISE on the computer took six months.

In this study, the output of PRECIS would be a higher 50 kilometers resolution for the period 1960–2099, and meteorological parameters such as precipitation, temperature, wind, relative humidity and radiation during 1971–2010 from eight meteorological stations in northern and central parts beyond the Chao Phaya River Basin of Thailand namely; Lampang, Suphanburi, Nan, Sisamrong, Takfa, Chainat, Uthong and Bangna meteorological stations respectively have been selected and used for bias correction procedures.

Bias correction procedures

This study has performed the bias correction of daily meteorological data such as temperature, rainfall, wind speed, relative humidity and radiation from the PRESICS model for the Chao Phraya River Basin. Three statistical bias correction methods were applied for each meteorological station separately, and then the method giving the output of the simulated current climates closest to the observation data was selected. The validation of the method was carried out using six daily meteorological parameters from the PRESICS model for the eight meteorological stations in the study area.

In order to perform and compare the bias correction methods, three datasets in two defined periods were needed namely the observed climatic data, given RCM outputs for the same time of calibration period and the RCM outputs for the validation period. Further, due to the difference of time at the beginning of meteorological parameter observation and collection in Thailand, precipitation and air temperature data were available for 40 years but the wind speed, relative humidity and radiation data were available only for 30 years. Therefore, the calibration and validation periods for all bias corrections were divided into half of the period for each data. Subsequently, the temperature and rainfall data were calibrated against the corresponding observed data from 1971 to 1990 and then were validated from 1991 to 2010, and the other data were calibrated from 1981–1995 and validated from 1996–2010. These three methods have been used for the bias correction techniques as follows:

Method 1 – adjusting the mean based on RCM (Weedon et al., 2010)

This technique has corrected the RCM which is based on the mean difference between RCM and observation data during the calibration period. The future corrected RCM could be constructed as follow:

\[ x_{corr} = x_{RCM,future} + (\bar{x}_{obs,calib} - \bar{x}_{RCM,calib}) \]  

(1)

Where, \( x_{corr} \) and \( x_{RCM,future} \) are the bias corrected data and RCM output in future period, and \( \bar{x}_{obs,calib} \) and \( \bar{x}_{RCM,calib} \) are the mean for the calibration period of observation data and RCM output respectively. The purpose of this function is to remove the difference between the means of the RCM and observations. Applying this method, the daily variation of corrected data becomes identical to that of the original RCM output.

Method 2 – adjusting the mean based on observation (Lehner et al., 2006)

This technique corrects the RCM based on the mean difference of RCM between future and calibration periods. The future corrected RCM can be constructed as follows:

\[ x_{corr} = x_{obs,calib} + (\bar{x}_{RCM,future} - \bar{x}_{RCM,calib}) \]  

(2)

Where \( x_{obs,calib} \) is observation data in the calibration period. \( \bar{x}_{RCM,future} \) and \( \bar{x}_{RCM,calib} \) are the mean of RCM output in future and calibration periods respectively. Applying this method, the daily variation of corrected data becomes identical to that of the observation data in the calibration period.

Method 3 – quantile-based mapping (Panofsky and Brier, 1968)

In using this method, the statistical distribution of data must be specified. The normal distribution was assumed for maximum and minimum temperatures, relative humidity and radiation. The gamma distribution was assumed for rainfall, while the weibull distribution was assumed for wind speed. The correction has been done by generating the Cumulative Distribution Function (CDF) for the observed data and RCM output. The CDFs for both RCM and observed data have then been related through the probability threshold to define the quantile map which would be used for bias removal from the RCM output in future period as shown in Figure 3.
Root mean square error (RMSE) comparison of bias correction

The comparison of all bias correction results has been performed by comparing the corrected and observed data at the validation period using RMSE technique as follows:

\[
RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{corr} - x_{obs})^2}
\]  

(3)

Where \(x_{corr}\), \(x_{obs}\) and \(n\) are the bias corrected data, observation data and number of data, respectively.

RESULTS AND DISCUSSION

The daily meteorological parameters output at the 50 kilometers resolution

The daily meteorological parameters output with 50 kilometers resolution for the period 1960–2099 has been derived from the PRECIS model. This output is grid data at the meteorological station, and an example is shown in Table II which was the output of maximum temperature in 1960 at Suphanburi meteorological station in central part of Thailand.

Bias correction results

The results of each bias correction technique has been derived and shown in Table III which was an example of the comparison between a daily maximum temperature output of RCM, observed data and the results after three bias correction methods in 2000 of Suphanburi meteorological station in the validation period.

Root mean square error (RMSE) results

From the comparison between all daily corrected and observed data in the validation period for six meteorological parameters in eight meteorological stations by using root mean square error (RMSE), it was found that method 2 which involved adjusting the mean based on observation data has generally given lower RMSE than other methods as shown in Table IV. Note that due to the formulation of method 2, the bias-corrected products in the validation period becomes close to those in the calibration period since the two periods are close and the term delta (\(\bar{x}_{RCM, future} - \bar{x}_{RCM, calib}\)) in equation 2 is negligible. It is also likely that the observed data of these two periods are similar. This might explain why method 2 gives the lowest RMSE of the three methods. Figure 4 presents the comparison between an observed daily maximum temperature, the RCM output and the result after the bias correction with method 2 at Suphanburi meteorological station during January–March 2010. Note that the fluctuation pattern difference between the observed data and the bias corrected data could be explained from the meaning of delta terms (\(\bar{x}_{RCM, future} - \bar{x}_{RCM, calib}\)) in equation 2 which were the mean difference for the same day of 20 years of data. Whenever, these delta

| Table II. Daily maximum temperature in 1960 of Suphanburi meteorological station from PRECIS model |
| --- |
| Year | Month | Date | Max. Temp. |
| --- | --- | --- | --- |
| 1960 | 1 | 1 | 33.6 |
| 1960 | 1 | 2 | 27.9 |
| 1960 | 1 | 3 | 29.9 |
| 1960 | 1 | 4 | 30.1 |
| 1960 | 1 | 5 | 28.6 |
| 1960 | 1 | 6 | 27.4 |
| 1960 | 1 | 7 | 27.2 |
| 1960 | 1 | 8 | 27.0 |
| 1960 | 1 | 9 | 24.4 |
| 1960 | 1 | 10 | 25.7 |
| 1960 | 1 | 11 | 26.0 |

Table III. Comparison between daily maximum temperature output of RCM, observed data and the results after three bias correction methods in 2000 for Suphanburi meteorological station |

| Year | Month | Date | Observed | RCM | Method 1 | Method 2 | Method 3 |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 2000 | 1 | 1 | 33.1 | 39.3 | 35.5 | 31.5 | 35.3 |
| 2000 | 1 | 2 | 33 | 36.5 | 32.3 | 31.3 | 31.9 |
| 2000 | 1 | 3 | 32.6 | 39.2 | 34.7 | 31.2 | 34.3 |
| 2000 | 1 | 4 | 34 | 39.4 | 34.1 | 32.2 | 36.1 |
| 2000 | 1 | 5 | 34.3 | 39.3 | 34.1 | 30.7 | 36.0 |
| 2000 | 1 | 6 | 33.5 | 35.5 | 30.5 | 29.5 | 30.6 |
| 2000 | 1 | 7 | 33.5 | 27.3 | 23 | 29 | 25.6 |
| 2000 | 1 | 8 | 33.2 | 34.2 | 30 | 30.1 | 30.4 |
| 2000 | 1 | 9 | 33 | 32.3 | 27.4 | 31.1 | 28.1 |
| 2000 | 1 | 10 | 33.8 | 34.8 | 30.5 | 30.9 | 31 |
| 2000 | 1 | 11 | 34.1 | 36.1 | 31.1 | 30.8 | 31.2 |

Figure 3. Quantile-based mapping method
terms were used for bias correction then the fluctuation patterns were different also. Evaluation of the performance of the bias correction methods was done not only using daily data but also computing RMSE of monthly data. The results of monthly data RMSE were similar to daily data. Method 2 gave lower RMSE than other methods. Therefore, for all bias correction processes, the climate output and observation data in the calibration period of 1981–2000 have been used for the climate output in the period of 2001–2099.

**Table IV. Daily RMSE from eight meteorological stations using three bias correction methods for six meteorological parameters**

| Parameter               | Station          | Lampang | Nan  | Sisamrong | Takfa | Chainat | Suphanburi | Uthong | Bangna |
|-------------------------|------------------|---------|------|-----------|-------|---------|------------|--------|--------|
| Maximum Temp. (°C)      | Method 1         | 4.8     | 4.6  | 4.3       | 4.4   | 4.3     | 4.0        | 4.1    | 3.6    |
|                         | Method 2         | 3.5     | 3.4  | 3.2       | 3.0   | 3.0     | 3.0        | 3.2    | 2.5    |
|                         | Method 3         | 3.8     | 3.6  | 3.4       | 3.1   | 3.0     | 3.1        | 3.4    | 2.5    |
| Minimum Temp. (°C)      | Method 1         | 2.9     | 3.1  | 2.8       | 2.8   | 2.7     | 2.6        | 2.6    | 2.5    |
|                         | Method 2         | 2.7     | 2.9  | 2.7       | 2.5   | 2.4     | 2.4        | 2.6    | 2.5    |
|                         | Method 3         | 3.0     | 3.2  | 2.9       | 2.8   | 2.7     | 2.6        | 2.9    | 2.7    |
| Rainfall (mm)           | Method 1         | 9.6     | 10.8 | 10.7      | 11.1  | 9.9     | 9.7        | 9.3    | 12.7   |
|                         | Method 2         | 11.6    | 13.6 | 13.9      | 13.7  | 12.6    | 13.0       | 13.3   | 16.4   |
|                         | Method 3         | 9.5     | 12.6 | 14.3      | 12.1  | 12.4    | 11.4       | 12.3   | 15.4   |
| Relative humidity (%)   | Method 1         | 17.0    | 17.2 | 14.7      | 18.3  | 16.0    | 14.0       | 14.3   | 14.4   |
|                         | Method 2         | 10.7    | 10.9 | 10.4      | 13.6  | 10.6    | 9.8        | 11.6   | 10.7   |
|                         | Method 3         | 10.4    | 9.9  | 10.3      | 12.0  | 10.2    | 9.4        | 11.3   | 10.4   |
| Wind speed (m/s)        | Method 1         | 0.8     | 0.9  | 1.4       | 1.2   | 1.7     | 1.3        | 1.5    | 1.3    |
|                         | Method 2         | 0.7     | 1.2  | 1.7       | 1.3   | 1.8     | 1.2        | 1.6    | 1.3    |
|                         | Method 3         | 1.9     | 1.9  | 1.7       | 2.5   | 2.4     | 2.5        | 2.6    | 2.9    |
| Radiation (kJ/m²/d)     | Method 1         | 6883.7  | 6419.2 | 6190.7  | 6118.6 | 5975.0  | 5756.2     |
|                         | Method 2         | 5738.4  | 5816.8 | 5599.5  | 5618.4 | 5890.7  | 5735.1     |
|                         | Method 3         | 5649.7  | 5884.4 | 5611.5  | 5714.3 | 5941.8  | 5816.6     |

**Figure 4.** Comparison between an observed daily maximum temperature, the RCM output and the result after the bias correction with method 2 at Suphanburi meteorological station during January–March 2010

CONCLUSIONS AND RECOMMENDATIONS

According to the aforementioned results, it has been clearly found that the bias correction method involving adjustment of the mean based on observation data was an appropriate technique for data assimilation of meteorological parameters in Thailand. This method can be used to improve the simulation output of any model with high correlation to the observed data.

In conclusion, the climate prediction based on RCM will give a finer scale and is thus suitable for climate impact study on a regional scale in the future. However, the results
of RCM still feature severe systematic errors which need bias correction prior to use and application for several activities in Thailand, such as the study of the impact of climate change in the near future, meteorological early warning system and forecasting major crop yield for advance planning.

ACKNOWLEDGEMENTS

ECHAM4 and PRECIS data used in this study were provided by the Hadley Centre, U.K. This study was fully supported by the IMPAC-T project of the Japanese government programs SATREPS, JICA and JST. Also, we highly appreciate the help of the Thai Meteorological Department for their support with all related facilities.

REFERENCES

Fowler HJ, Blenkinsop S, Tebaldi C. 2007. Linking climate change modeling to impact studies: recent advances in downscaling techniques for hydrological modeling. *International Journal of Climatology* 27: 1547–1578. DOI: 10.1002/joc.1556.

Intergovernmental Panel on Climate Change (IPCC). 2007. Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA; 966.

Jones RG, Noguer M, Hassell DC, Hudson D, Wilson SS, Jenkins GI, Mitchell JFB. 2004. Generating high resolution climate change scenarios using PRECIS, Met Office Hadley Centre, Exeter, UK; 40.

Lehner B, Düll P, Alcamo J, Henrichs T, Kaspar F. 2006. Estimating the impact of global change on flood and drought risks in Europe: A continental, integrated analysis. *Climatic Change* 75: 273–299. DOI: 10.1007/s10584-006-6338-4.

Panofsky HA, Brier GW. 1968. *Some Applications of Statistics in Meteorology*. Pennsylvania State University Press, State College, PA; 224.

Themeßl MJ, Gobiet A, Leuprecht A. 2011. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology* 31: 1530–1544. DOI: 10.1002/joc.2168.

Wood AW, Maurer EP, Kumar A, Lettenmaier DP. 2002. Long Range Experimental Hydrologic Forecasting for the Eastern U.S. *Journal of Geophysical Research* 107: 4429. DOI: 10.1029/2001JD000659.