

## **Improved Bias Correction Method using Wavelet Decomposition**

**Dhanya C.T. and Abhishek Gupta**

*Civil Engineering Department, IIT Delhi, Hauz Khas, New Delhi, INDIA*

### **Abstract**

Bias correction methods need to be applied to General Climate Models (GCMs) before they give us the correct value of climatic parameters. Out of the several bias correction techniques available, this paper considers the delta change approach and tries to improve upon it by applying wavelet transformation. The difference from traditional delta change approach is that it breaks up the future GCM data into constituent wavelets and then adds to it the difference between the wavelets of baseline data and observed data. Using wavelets is advantageous because it can analyse frequency data over longer time periods, and help us deal with dominating frequencies over non stationary data. The proposed approach has been tested for baseline MIROC (Model for Interdisciplinary Research on Climate) precipitation data (1950-1970) and applied to MIROC future values (1980-2000). The results have been compared with actually observed precipitation during 1980-2000, obtained from IMD (India Meteorological Department). The comparison is just preliminary as of now, using mean and standard deviation. It shows a fair amount of similarity with the observed values. And keeping these results in mind, it has been used to project rainfall conditions for 2080-2100, which show a small spike in the average rainfall and in the number of extreme events in eastern parts of the country. The study thus opens up the possibility of wavelets being incorporated in bias correction methods and obtaining more realistic future projections.

**Keywords:** Bias correction, delta change, wavelet.

## **1. Introduction**

### **1.1 Global Climate Models and Bias correction**

Global Climate Models (GCMs) are models that simulate the climate of the earth using our extant scientific knowledge. They divide the earth into various grids and thus try to predict the climate at different places. The two main types of General Circulation Models are Atmospheric and Ocean models. Separately they account for the changes within the atmosphere and the ocean respectively. Put together they make up a complete climate model. When two systems are connected the resulting more complete system is called a 'coupled' model.

Coupled atmosphere-ocean General Circulation Models (AOGCMs) are the models most often used to make the predictions of future climate. These predictions are sometimes called scenarios Based on different scenarios the IPCC can recommend measures to mitigate global warming and can predict what results a specific reduction in greenhouse gas emission are likely to have advanced coupled atmosphere-ocean General Circulation Model can also to some extent predict regional climate changes. These predictions are used to determine what actions nations and regions should take to prepare for the coming changes.

However, the prediction of GCMs is not up to the mark because of our poor understanding of many physical processes. This results in errors which we call bias and need to account for in our final predictions using past deviations. Some factors responsible for these errors are clouds, pollution, scale and resolution.

Clouds, and their particle sizes, make analysis and projection of climate more complex for a number of reasons. Daytime clouds block some of the sunlight from reaching the planet's surface, while clouds throughout the day block some of the heat from escaping from the surface. The latter effect is most detectable at night. The location of the cloud in the atmosphere affects whether they are more effective at blocking sunlight or trapping heat. So does their color and the size of particles and water droplets they contain smaller droplets in clouds are less likely to rain out than bigger ones. Along with precipitation rates, this can affect temperature because clouds containing smaller particles are more likely to persist, and thus continue to block some sunlight. Droplet size is affected by pollution particles, among other things.

In case of pollution particles, it can affect both temperature and precipitation in ways that are difficult to model, increasing the uncertainty in climate projections. Also known as aerosols, particles created from combustion mask the ongoing global warming to an unknown degree because of their effect on temperature and clouds. Aerosols have a cooling effect on temperature, too. They promote formation of smaller water droplets, which helps clouds persist while simultaneously making them better reflectors of sunlight. Aerosols themselves can also reflect sunlight. As a result, the recent warming has been less severe than it would have been without the shielding presence of pollution particles. The resulting "dimming" of the sun from the ongoing influx of pollution particles is similar to the type that occurs when volcanic aerosols reach above the atmosphere's weather layer, where they can remain airborne for a year or more. Aerosol pollution has a shorter-lived effect on temperature than greenhouse

gases. These particles typically rain out within a week or so, whereas greenhouse gases can remain airborne for many decades.

Different climate processes operate on different scales in time and space. Spatial scales are often called “resolution.” Loosely comparable to the number of pixels in a digital photo, resolution relates to the number of grid cells dividing the globe for modeling purposes. These scale issues affect the processes models can capture. With their relatively coarse resolution, GCMs tend to reproduce larger scale phenomenon, which include temperature fluctuations, better than smaller-scale phenomenon, such as extreme precipitation events. For instance, while heat waves tend to affect regions across several grid cells of a climate model, the intense rainfall that leads to floods often occurs at scales smaller than a grid cell. This is one of the reason clouds and rainfall are challenging to model.

## **1.2 Objectives**

There are several bias correction methods that are used-delta change approach, multiple regression, analogue method, local intensity scaling, quantile mapping to name a few. They all have some advantages and disadvantages. Attempts continue to improve upon them or find better techniques altogether. Here we try to improve the simple delta change approach by applying wavelets and see if it gives better correlations.

This paper is arranged as follows: Section 2 deals with the data used and the methodology for the proposed bias correction approach. Section 3 discusses the results that we get from the above approach and the actually observed values for the same time period (1980-2000). It then goes on to project, albeit in a raw manner, the expected precipitation in 2080-2100. Section 4 sums up the study and points out areas for further improvements. Section 5 is the references, which though not explicitly used here, but have been crucial in the overall understanding.

## **2. Data and Methodology**

### **2.1. Model Data**

The GCM data used is the coupled atmosphere-ocean general circulation model, called MIROC (Model for Interdisciplinary Research on Climate). It has been developed by Centre for Climate System Research, University of Tokyo; National Institute for Environmental Studies; and the Frontier Research Centre for Global Change. MIROC uses two main setups for climate studies-one with higher resolution and another with a low resolution. The higher resolution setup is the main setup, and is used for phenomena with relatively small spatial scales like extreme climate events or regional climate change. But it is expensive so the cheaper lower resolution is helpful.

The future MIROC precipitation data (for 2040-2100) was prepared using the RCP4.5 scenario. RCP (Representative Concentration Pathway) 4.5 is a scenario that stabilizes radiative forcing at  $4.5 \text{ W/m}^2$  in 2100 (without exceeding it). It includes long term global emissions of GHGs (Greenhouse gases), short lived species, and land use cover. While there are many other ways to achieve a  $4.5 \text{ W/m}^2$  limit, the RCP4.5 gives

a common platform for climate models to explore the climate system response to stabilizing the anthropogenic components of radiative forcing.

## 2.2. Observed data

It is  $1^\circ \times 1^\circ$  spatial resolution data for 1950-2004, obtained from IMD [Raajevan *et. al.* 2006]. It was prepared using the daily rainfall data archived at the National Data Centre, IMD Pune, which has the rainfall records of 6329 stations.

## 2.3. Methodology

The GCM (MIROC) data is first collected for 1950-2012 and 2040-2100. The observations (by IMD) correspond to 1950-2004. All these are broken down into constituent wavelets, using MATLAB. The wavelet filter used is ‘haar’ and the level of decomposition is 3. There are 4 resulting wavelets for each data set. The corresponding wavelets of baseline GCM and actual data are subtracted to get the respective delta (difference). This same difference is then added to the future GCM data to get predicted actual data. A flow chart for the same is shown in figure 3.1.

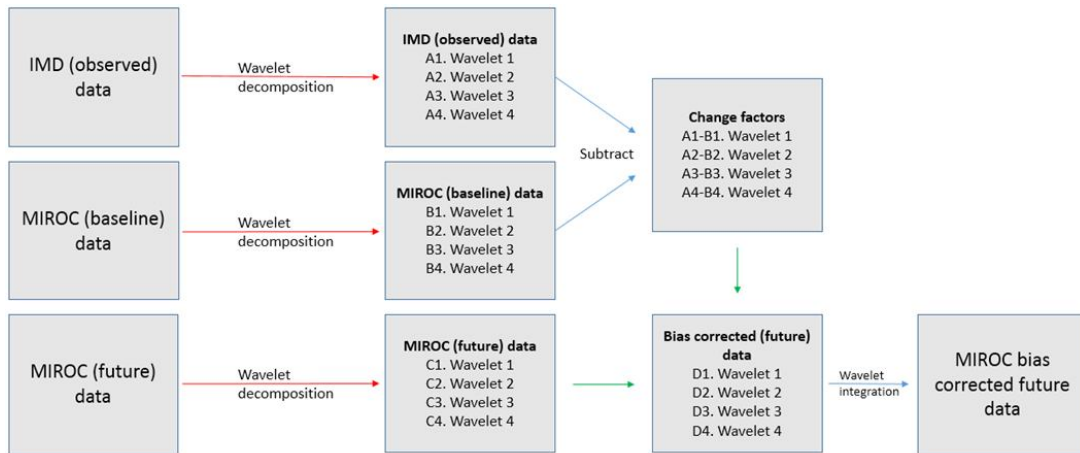


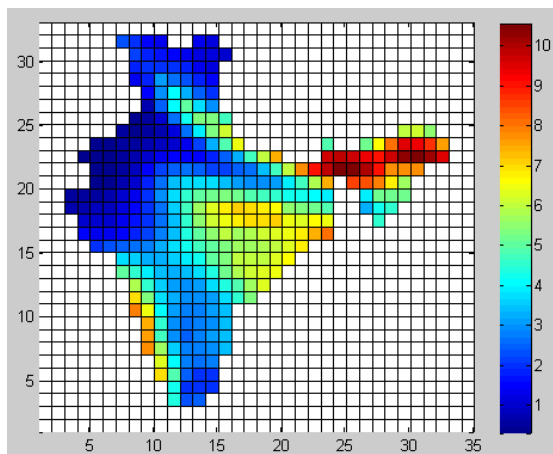
Figure 3.1: Methodology for the proposed bias correction method.

## 3. Results and Discussion

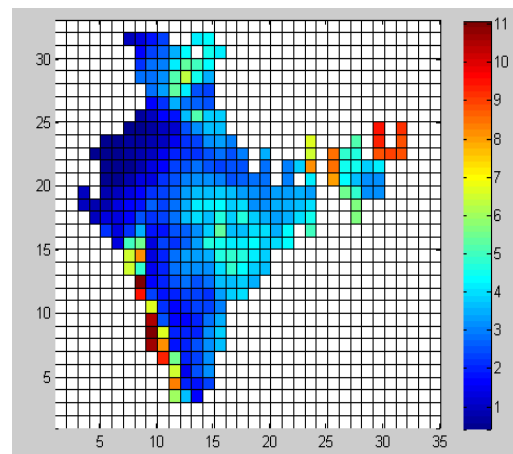
### 3.1 Methodology Validation

The method was first used to predict the precipitation for 1980-2000, for which we have the actual data as well. First we discuss the observations using the haar wavelet. The mean value comparison for all the grids shows a decent level of similarities between the two (see figure 3.2 and 3.3). The model predicts a low level of precipitation (in the range of 0-1mm) for the North Western part of India (mainly Rajasthan). The observed data shows that this region received around 0-2mm daily precipitation on an average, which is fair enough. The eastern part of top north, i.e. Jammu and Kashmir is predicted to have rainfall in the range of 4-5mm and some shades of 5-6mm. Observations show almost a similar variation though with very

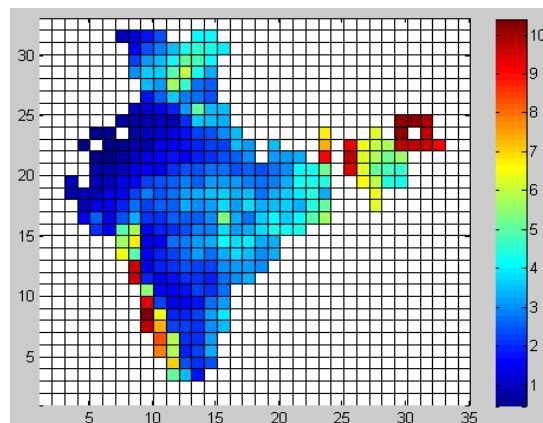
marginal reduction in mean values of the region. The western coastal part is expected to have varying rainfall—mostly in the higher range. This is also very similar to the varying pattern that we see in the observations. The central east parts that form a congregated area of higher rainfall (as compared to its surrounding areas), received somewhat different precipitation actually. The region on a whole received more rainfall than predicted, especially the areas around Kolkata and eastern Bihar. The North East region was supposed to get 3-5.5mm mean daily rainfall, with parts of eastern Arunachal Pradesh getting very high rainfall (approximately 9-10mm). The results again showed that the precipitation was more than predicted, in the range 5-6.5mm, and eastern parts getting even higher rainfall, in the range 9.5-11mm.



**Fig. 3.1:** Original MIROC mean values for 1980-2000. Values in mm.



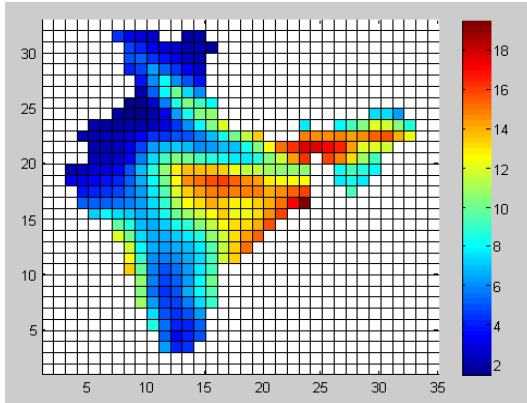
**Fig. 3.2:** Bias corrected MIROC mean values for 1980-2000. Values in mm.



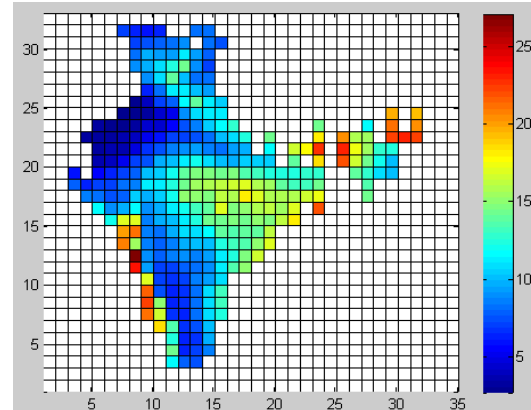
**Fig 3.3:** IMD (observed) mean values for 1980-2000. Values in mm.

On the whole the model data still shows a very marginal bias for lesser precipitation, though the difference is not so high. Also, the regional variation has been

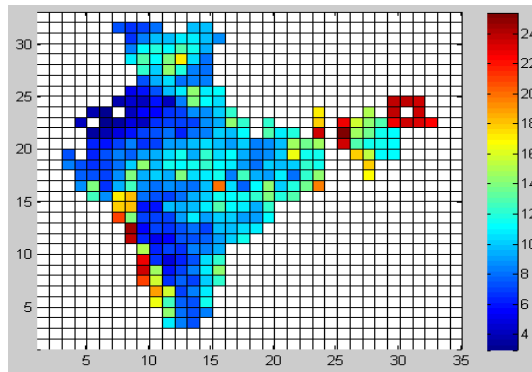
accurately captured. If we see the original MIROC mean values (Figure 3.1), it shows a much simplified version and lacks regional variations that have been corrected to a great extent after bias correction. However it should be remembered that these comparisons are for the mean values for daily data of 20 years. This is a very gross assumption, and a detailed and accurate comparison should be season wise and for shorter time durations.



**Fig. 3.4:** Original MIROC values for standard deviation (1980-2000)  
Values in mm.



**Fig. 3.5:** Bias corrected MIROC values for standard deviation (1980-2000).  
Values in mm.



**Fig 3.6** IMD (observed) standard deviation values (1980-2000). Values in mm.

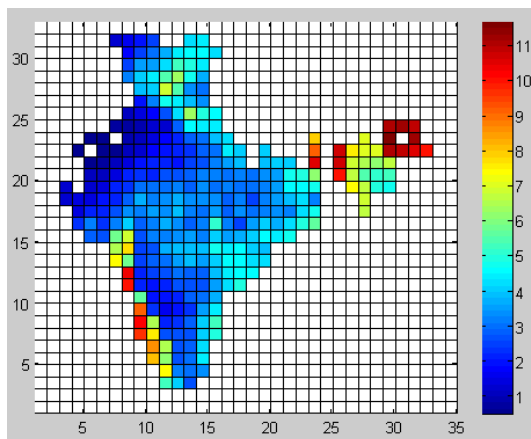
The standard deviation, which helps us decide the crowding of values around the mean, gives us an idea about the of daily rainfall data over 20 years. Though a strict comparison should have been between seasons, but still we know that it would be higher for regions that receive most of their rainfall from the South West monsoon. The Western coastal regions and the extreme North East parts have high standard deviations (see figures 3.5 and 3.6), which has a tad lower value than the actual data. The discrepancy is in the central and eastern parts where the model predicted standard deviation in the range of 12-17, whereas it is actually lower. It means that the model

predicts higher incidences of extreme climate events, whereas it is not so high in reality.

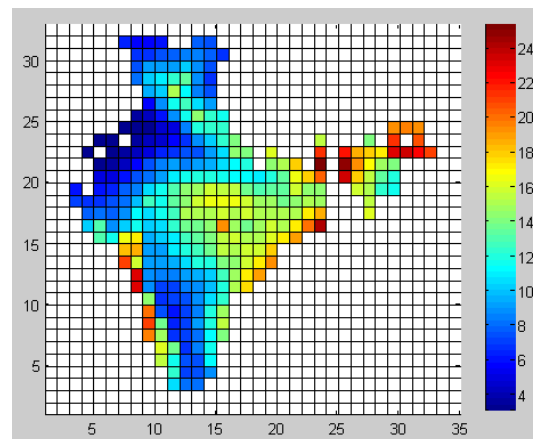
But the bias corrected result is surely better than the original standard deviation values (in figure 3.4), which lacked the regional variation that is observed in reality. It seems to follow the broad physical processes but is ignorant about the regional specifications, and that is why bias correction is applied. Thus bias correction using wavelets seems to work fine than uncorrected data. However its comparison with other techniques needs to be worked out, but that is outside the scope of this study.

### 3.2. Projections for the future

Considering the above correlations in mind, we now look at what the bias corrected predictions for 2080-2100 look like. The mean values for daily rainfall in 2080-2100 (figure 3.7) look a lot like the observed IMD data for 1980-2000. And so, it can be concluded that the rainfall patterns in the future will remain almost the same, albeit with a slight higher intensity. Thus, the north western parts should get 0-2mm rainfall on an average. The eastern areas of Jammu and Kashmir (the northern most part of India) are expected to receive 3-4.5 mm. The western coastal part will get a rainfall on the higher side as compared to other regions, but not higher in comparison to previous years. The eastern and central parts will get moderate rainfall, with coastal areas getting more rain as compared to the interiors. However some parts of central India may receive a bit less rainfall than predicted. The North East is again supposed to have high rainfall, with the expected range of 4.5-7.5mm for middle sub-region and 10-11.5mm for the north east and western sub regions. However, it should again be borne in mind that these are the mean expected values for daily rainfall for 20 years, which deserves a detailed attention rather than a bird's eye view as has been done here for simplification purpose.



**Fig. 3.7:** Bias corrected 'mean' MIROC values for 2080-2100. Values in mm.



**Fig. 3.8:** Bias corrected MIROC standard deviation values for 2080-2100. Values in mm.

The standard deviation results show a broadly similar result (in figure 3.8) as seen for the method validation period (1980-2000). But this case has more intense values of standard deviation on a whole than before, which implies that there would be higher cases of extreme events than seen in the past-especially in the eastern, central eastern and north eastern regions. But if we take a cue from the previous case, the values may not exactly be true for all places in these regions, though they may have higher extreme events on the whole.

#### **4. Conclusion**

The objective of our study was to produce a better bias correction method than the simplistic delta change method. The validation result shows that the method indeed comes close to a realistic picture for precipitation, though it might have been over the long term period of 20 years. This study is just a preliminary investigation into the potential of wavelets being applied in this area, and is by no means a final proof of correctness of the method. It needs to be checked by using various other wavelet filters and different decomposition levels. The observations in this study have been tallied using only two methods-mean and standard deviation. This needs a lot of improvement using mathematical tools like correlation and regression. The time period of the study also needs detailed attention as pointed out before, with a seasonal comparison if we have to get specific daily information for the future. Nevertheless, the application of wavelets definitely looks promising enough.

#### **References**

- [1] Eisner S, F. Voss, and E. Kynast (2012): Statistical bias correction of global climate projections – consequences for large scale modeling of flood flows, *Adv. Geosci.*, 31, 75–82
- [2] Lenderink G, A. Buishand and W. Deursen (2007): Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, *Hydrol. Earth Syst. Sci*, 11(3), 1145-1159
- [3] Li H, J. Sheffield, and E.F. Wood (2010): Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching; *Journal of Geophysical Research*, Vol. 115, D10101, doi:10.1029/2009JD012882.
- [4] Rajeevan ,M., J. Bhate, J.D.Kale and B. Lal, 2006: High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells, *Current Science*, 91(3), 296-306.
- [5] Teutschbein C, J. Seibert (2012): Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, *Journal of Hydrology* 456–457 (2012) 12–29.