SELECTING OF REGIONAL CLIMATE MODEL SIMULATIONS FOR
MODELING CLIMATE CHANGE IMPACTS ON THE WATER QUALITY
STATUS OF LAKE BURULLUS, EGYPT

A. Shalby¹, M. Elshemy² and B. A. Zeidan³

¹Faculty of Engineering, Tanta University, Egypt, E-mail: ahmed.shalby@f-eng.tanta.edu.eg
²Faculty of Engineering, Tanta University, Egypt, E-mail: m.elshemy@f-eng.tanta.edu.eg
³Faculty of Engineering, Tanta University, Egypt, E-mail: b.zeidan@f-eng.tanta.edu.eg

ABSTRACT

Future climatic changes have significant impacts on the quantity and the quality of the water resources. Till present, limited water quality impacts studies are addressed in the literature. This may return to the uncertainties in future climate change projections, which are considered as a key challenge for climate change studies. Using of Regional Climate Models (RCMs) simulations, which present appropriate spatial and temporal resolutions for the climatic projections, is essential to minimize these uncertainties. In this work, the most appropriate RCM simulations for Lake Burullus was selected out of six different RCMs simulations. Six different statistical measurements were used to examine the comparisons between collected meteorological records at the study area and the six RCMs simulations for the study area. The RCMs simulations were dynamically downscaled from a combination of Global Circulation Models (GCMs) in space to 0.44° latitude-longitude resolution (about 50 km) and in time to a daily average. The simulations cover two periods; a historical period (up to 2005) and a future period (2006-2100). The comparisons were done for two periods; a historical period (1994-2005) and a future, according to the RCMs simulations, period (2006-2015). Three meteorological characteristics were chosen for the comparisons; mean air temperature, maximum air temperature and minimum air temperature. The above methodology is used to narrow uncertainty in regional climate predictions, which will be used to modify a calibrated and validated water quality model of Lake Burullus to investigate the future climatic impacts.

Keywords: Climate change, Lake Burullus, RCMs, RCMs simulations selection, Uncertainty, Water quality model.

1 INTRODUCTION

Certainly, climate change (CC) is a very serious problem which is affecting the global environment. CC is taking place as a result of a growing increase in levels of Greenhouse Gases (GHGs) emissions and working against the promotion of sustainable development. It is urgently recommended to mitigate its dramatic impacts on the environment. Mitigation and adaptation studies are essential and impact studies are considered as the guidance for it. Impacts modeling studies show how CC is affecting us and the environment (Wilcke and Barring, 2016). Impacts models can simulate the impacts in local and global scale, e.g. crop yields, hydrological, and water quality models (Pulatov et al., 2016). The crucial step in impacts modeling studies is to obtain and carefully select the input future projections of meteorological conditions (Mendlik and Gobiet, 2016).

Impacts modeling studies use climate model output data for the simulation of future climate change projections. These projections are not “predictions” but they represent designedly examples of many climates that may occur in the twenty-first century (Cayan et al., 2008). Future climatic data are available via different Global Climate Models (GCMs) which have been developed by a number of research groups. The World Climate Research Programme (WCRP) has established the framework Coupled Modelling Intercomparison Programme which has now become its fifth phase, CMIP5. It includes more than 50 models from 24 modeling groups with a higher resolution and more ensemble members for individual experiments (Taylor et al., 2012). It meant to provide a framework for
coordinated climate change experiments and thus includes simulations for the fifth assessment reports (AR5) report of the Intergovernmental Panel on Climate Change's (IPCC) (CMIP5 website). The IPCC from its first to fifth assessment reports discussed several families of future scenarios. Future scenarios are uncertain because they depend on future social, political, and technological decisions which control GHGs emissions. AR5, the most recent one, introduced representative concentration pathway (RCP) scenarios (i.e. RCP26, RCP45, RCP60, and RCP85) as a new family of future scenarios. The numbers indicating ten times the assumed radiative forcing at the end of the twenty-first century (Rogelj et al., 2012).

Although, GCM is a very useful tool to project future climate, its spatial resolution is too coarse to use its outputs in impact studies (Themeßl et al., 2011). Downscaling approaches are used to downscale GCM outputs to get finer spatial and temporal resolutions (Themeßl et al., 2011). Regional Climate Models (RCMs) can be used for dynamically downscaling the global resolution onto a regional scale and then, it can present future climate information on an appropriate scale for most impacts modeling studies. The RCMs output simulations are coordinated in frameworks such as CORDEX, ENSEMBLES and PRUDENCE which produce growing output for ensembles of GCM and RCM combinations with the aim to provide input for a variety of climate impact studies.

The results of any impacts models are not “exact predictions,” but rather an approach of indication of the effects of climate change. Uncertainties in climate change modeling is a key challenge in climate change impacts and adaptation studies. Different sources of uncertainty in climate change impact studies have been identified (Gaur and Simonovic, 2015). Uncertainty in modeling the impacts can be reduced by using multiple models (Deser et al., 2012). Using of RCMs simulation can reduce the uncertainty of future climatic projection because of its finer resolution than GCMs outputs. Uncertainty of RCMs simulation selection is due to its growing and non-limitation simulations. To cover this source of uncertainty, Tebaldi and Knutti (2007) recommended to take all available climate model data into consideration rather than using only a single model. The ensembles of GCM-RCM simulations are very big to be handled in the impact study (Wilcke and Barring, 2016), as taking all of them is not practical in research projects due to time restrictions and limited computer resources (Pulatov et al., 2016). It's more recommend to develop a selection method. Whetton et al. (2012) showed that the “model selection” issue arises under a growing number of GCM and downscaled results of different techniques available for application.

Selection criteria depends on the assessment of model performance to represent some relevant climate variables in the recent and historical periods with the aim to get the best simulations (e.g. Deidda et al., 2013) or to eliminate poor performing RCM simulations from consideration (e.g. Evans et al., 2014). In 2007, the performance of nine different RCMs in reproducing the present-day climate of a validation period (1961-1990) over the European region is investigated (Jacob et al.). The evaluation of the models were based on the comparison of simulated seasonal and annual means against observations as well as comparison of observed and simulated inter-annual variability for temperature. Anagnostopoulos et al. (2010) examine how well several model outputs fit measured data of different variables, such as mean temperature and precipitation, at 55 points around the globe at several temporal scales using two main statistical indices: the correlation coefficient and the coefficient of efficiency as well as the average, the standard deviation, the first-order autocorrelation coefficient and the Hurst coefficient. Perkins et al. (2007) evaluated 14 RCMs based on their skill in simulation of observed precipitation, and maximum and minimum temperatures over 12 regions in Australia using probability density functions and a simple quantitative measure. They also ranked the model using a skill score based on the overlap between the observed and modeled probability density functions. Deidda et al. (2013) ranked 14 RCMs for six representative catchments in the Mediterranean region, according to their relative performance in providing daily precipitation and surface temperatures during a validation period (1951 to 2010), the ranking of the model's performance were applied using different dimensions normalized metrics. Ouda et al. (2016) developed an ensemble model to estimate potential evapotranspiration under climate change estimations, which are composed of three different models simulations, based on the goodness of fit between measured and projected relevant climate variables.
Following the selection studies, no particular RCM is the best for all climatic variables. Knutti et al. (2010) stated that, good agreement of the model simulation with observations for one variable does not guarantee good performance in other variables. Selection of the study variables is vital and depends on the subsequent impact study and its relevant climate variables. Gleckler et al. (2008) concluded that the model which well represents the basic variables, such as temperature and precipitation, often also perform well in other variables.

In this work, a selection procedure for the most appropriate RCM simulations for Lake Burullus was applied. It depended on a comparison process to get the simulation which gives the highest closeness with measured meteorological records of the study area. The best RCM simulations will be used in a subsequent impact study to investigate the future climatic impacts on the water quality status of Lake Burullus. Three meteorological characteristics of air temperature were chosen for the comparisons. The choice of these three variables was based on their major effects on the case study characteristics. The most direct impact of climate change on water resources characteristics is the water temperature rising. Sea surface temperature (SST) has risen by 0.5° C globally since the 1980s and is predicted to continue increasing throughout the 21st century (IPCC, 2007a). The contribution of Global warming, as a new additional hazard, to degrade the water quality status of surface watercourses is confirmed (UNEP, 2007), however, limited water quality impact studies are addressed in the literature.

2 STUDY AREA

The Egyptian coast along the Mediterranean is considered as one of the regions which is expected to experience the worst effects of climatic changes, in particular sea level rise. It includes five lagoons, Lake Burullus represents the second largest one. It extends in the central part of the northern shoreline of the Nile Delta between the two Nile River branches: Damietta and Rosetta. Its surface area is about 410 Km² (Hossen and Negm, 2016). Lake Burullus is a shallow brackish lake and connected to the Mediterranean by a small outlet (Boughaz), which is about 44 m width near El Burg village. It extends between longitudes 30° 31′ E and 31° 05′ E and latitudes 31° 25′ N and 31° 35′ N. It collects agricultural drainage water from about 4000 km² of cultivated land in the catchment through eight drains. It is also connected to the Nile River by Brimbal Canal, Figure 1. The climate conditions of Burullus area are arid Mediterranean, with mean annual rainfall less than 200 mm and the temperature ranges from 9 °C in winter to 30 °C during summer (El-Adawy et al., 2013).
3 METHODOLOGY

In this work, an evaluation of the relative performance of several RCMs simulations, for representing the air temperature projections at the lake region, was investigated as a selection procedure. To select the most appropriate RCM simulations for the subsequent proposed water quality impact study for the lake, a selection procedure based on the comparison between simulated and measured meteorological variables using different statistical metrics was developed. Six different RCMs simulations for air temperature from the CORDEX project (Cordex Project) and the Canadian Centre for Climate Modelling and Analysis (CCCMA) were selected according to appropriate spatial and temporal conditions, see section 3.1. Recorded air temperature series, for the study area, were obtained from a Land-Based meteorological station, see section 3.2. The used six different statistical metrics for the selection procedure are presented in section 3.3.

3.1 RCMs simulations

The ensemble of climate simulations used in this study consists of six different GCM-RCM combinations, Table 1. Five simulations were selected from the CORDEX project (Cordex Project), for three different representative concentration pathway (RCP) scenarios; RCP26, RCP45, and RCP85. The sixth simulation is downloaded from the Canadian Centre for Climate Modelling and Analysis (CCCMA) for only two RCP scenarios; RCP45, and RCP85. The six simulations cover two domains [Middle East - North Africa domain (MENA) and Africa domain (AFR)] with a grid spacing of 0.44° × 0.44° (approx. 50 km × 50 km). Daily near surface air temperature in its three components [average (tas), maximum (tasmax), and minimum (tasmin)] were selected to be extracted at the grid points for the study region.

Table 1. GCM-RCM combinations of the simulations

<table>
<thead>
<tr>
<th>Domain</th>
<th>GCM</th>
<th>RCM</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>MENA-44</td>
<td>ICHEC-EC-EARTH</td>
<td>SMHI- RCA4</td>
<td>S1</td>
</tr>
<tr>
<td>AFR-44</td>
<td>ICHEC-EC-EARTH</td>
<td>MPI-CSC- REMO2009</td>
<td>S2</td>
</tr>
<tr>
<td>AFR-44</td>
<td>MPI-M-MPI-ESM-LR</td>
<td>MPI-CSC- REMO2009</td>
<td>S3</td>
</tr>
<tr>
<td>AFR-44</td>
<td>MOHC-HadGEM2-ES</td>
<td>SMHI- RCA4</td>
<td>S4</td>
</tr>
<tr>
<td>AFR-44</td>
<td>MPI-M-MPI-ESM-LR</td>
<td>SMHI- RCA4</td>
<td>S5</td>
</tr>
<tr>
<td>AFR-44</td>
<td>CanESM2</td>
<td>CCCma- CanRCCM4</td>
<td>S6</td>
</tr>
</tbody>
</table>

3.2 Recorded meteorological data

The meteorological data of air temperature were obtained for the nearby local meteorological station (Baltim station), see Figure 1, for the period [1994-2015] (websites of “Weather Underground” and “NOAA”). The recorded meteorological data of maximum and minimum air temperature were available as once daily. While the mean air temperature was based on the average of about eight times observations uniformly distributed along the day. Annual averages of mean, maximum and minimum air temperatures are presented from Figure 2 –to Figure 4, respectively. Also, Figure 5 shows the time series distribution of the mean, maximum and minimum air temperatures in year of 2010.
Figure 2. Annual averages of measured mean air temperature records at Baltim station

3.3 Selection procedure

According to Chai and Draxler (2014), a combination of metrics is required to assess the models performance (Chai and Draxler, 2014). Six different statistical procedures were used to examine the goodness of fit between the corresponding observed and simulated (projected) air temperature for two study periods; a historical period (1994-2005) and a future period (2006-2015), according to the RCMs simulations. Absolute Mean Error (AME), Root Mean Square Error (RMSE) and Percent Bias (PBIAS) were used as standard metrics for simulations errors (i.e. lower scores indicate better performance). The coefficient of Efficiency (E), index of agreement (d) and the correlation coefficient (r) were used as acceptance indicators (i.e. higher scores indicate better performance). The coefficient of efficiency (E) and the index of agreement (d) represent a decided improvement over the coefficient of determination ($R^2$) (Legates and McCabe, 1999).

Both AME and RMSE are regularly employed in evaluate models performance studies (Chai and Draxler, 2014). RMSE gives the general standard deviation of simulation error (Willmott and Matsuura, 2005) and computed as in Equation (1).

Figure 4. Annual averages of measured minimum air temperature records at Baltim station
\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (S_i - O_i)^2} \]  

Where, \(S_i\) and \(O_i\) are the simulated and observed values at the \(i\)th time step, respectively, \(n\) represents the number of corresponding observed and simulated values used in comparison.

\[ \text{AME} = \frac{\sum_{i=1}^{n} (S_i - O_i)}{n} \]  

AME, in contrast with RMSE, gives the same weight to all errors and often used as a good indicator of average model performance (Chai and Draxler, 2014). AME computed as in Equation (2) with variables have similar meanings to those in Equation (1).

\[ \text{PBIAS} = \frac{\sum_{i=1}^{n} (O_i - S_i) \times 100}{\sum_{i=1}^{n} O_i} \]  

PBIAS compares the average tendency of the simulated data to the corresponding observed data (Gupta et al., 1999). The optimal value of PBIAS is zero. A positive value indicates that the model has underestimated and a negative value indicates a bias toward overestimation (Gupta et al., 1999). PBIAS is computed as in Equation (3) with variables have similar meanings to those in Equation (1).

\[ E = 1 - \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} (O_i - \bar{O})^2} \]  

\(E\) shows how much better the model can simulate/predict observation and computing the simulation by just using the mean of observation as a predictor. It determines the relative magnitude of the residual variance compared to the measured data variance (Legates and McCabe, 1999) and computed as in Equation (4).

Where, \(\bar{O}\) is the average of the observed data and other variables have similar meanings to those in Equation (1).

\(E\) values can range from \(-\infty\) to 1. An \(E\) value of 1 corresponds to a perfect match of observed to simulated data. An \(E\) value between 0 and 1 is considered an acceptable level of performance, whereas an \(E\) value \(\leq 0\) suggests that the observed average is a better predictor than the model (Dile and Srinivasan, 2014; Legates and McCabe, 1999).

\[ d = 1 - \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} [(S_i - \bar{O})^2 + (O_i - \bar{O})^2]} \]  

d is the standardized measure of the degree of model simulation error (Willmott, 1981) and computed as in Equation (5) with variables have similar meanings to those in Equations (1) and (4).

\(d\) values can range from 0 to 1. A value of 1 indicates a perfect match of observed to simulated data, and value of 0 indicates no agreement at all (Willmott, 1981).
r is a number that quantifies a type of correlation and dependence, meaning statistical relationships between two or more values and computed as in Equation (6) with variables have similar meanings to those in Equation (1)

\[
r = \frac{\left( \sum_{i=1}^{n} (O_i \times S_i) \right) - \left( \sum_{i=1}^{n} O_i \right) \times \left( \sum_{i=1}^{n} S_i \right)}{\sqrt{\left[ \sum_{i=1}^{n} (O_i)^2 - \left( \sum_{i=1}^{n} O_i \right)^2 \right] \left[ \sum_{i=1}^{n} (S_i)^2 - \left( \sum_{i=1}^{n} S_i \right)^2 \right]}} \quad (6)
\]

r values can range from -1 to 1. An r value between 0 and 1 means a positive linear relationship, whereas an r value ≤ 0 means a negative linear relationship. An r value closer to 1 suggests the stronger relationship.

4 RESULTS AND DISCUSSIONS

The selection procedure was applied to check the goodness of fit test between observed and simulated values for the six selected RCMs simulations for the two studied periods. The most appropriate simulation can be evidenced by the lowest values of AME and RMSE, the lowest absolute values of PBIASE and the highest values of E and d. For the selected simulation, values of r should indicate stronger positive relationship between observed and simulated data.

4.1 Historical Period

Figures 6 – 7 present the averages of the test statistical models used for selected procedure for the historical period (1994-2005). Mean air temperature results are presented in Figure 6, max. air temperature results are presented in Figure 7, while min. air temperature results are presented in Figure 8. As can be seen, for each presented figure, the figure is composed from 2 separated charts, the simulation errors and the result of the acceptance measures.

![Figure 6. Averages of comparison metrics for mean air temperature - historical period (1994-2005)](image)

In Figure 6, it can be clearly noticed that the models simulations S4 represents the worst case, the values of AME, RMSE, PBIASE, E, d and r are 4.24, 3.48, 7.81, 0.66, 0.86 and 0.79, respectively. While S5 is considered as the best case, where the values of AME, RMSE, PBIASE, E, d and r are 2.47, 1.86, 0.42, 0.76, 0.93 and 0.88, respectively. All models are underestimated the simulations of air temperature, except S3 which gives a negative value of PBIASE indicating an overestimating bias. For
Figures 7 and 8, S5 is still representing the best case, although its correlation coefficient is lower than that of both S2 and S3. The worst models simulations are S4 and S6, according to the used metrics.

Figure 7. Averages of comparison metrics for max. air temperature - historical period (1994-2005)

Figure 8. Averages of comparison metrics for min. air temperature - historical period (1994-2005)

4.2 Future Projected Period

This study and the subsequent impact studies focus on the future climate, therefore the model performance in projection of future climate is crucial. Figures 9-11 present the averages of the test statistical models used for selected procedure for the future projected period (2006-2015), according to the models simulations categories. Mean air temperature results are presented in Figure 9, max. air temperature results are presented in Figure 10, while min. air temperature results are presented in Figure 11. The models simulations for RCP85 scenario were selected as an example for this work.

As can be seen in Figures 9-11, S5 models simulations represents the best simulation case, as previously in the historical period study. For mean temperature simulations, the values of AME, RMSE, PBIASE, E, d and r are 2.57, 2.02, 1.31, 0.74, 0.94 and 0.89, respectively. Figure 9 shows that the models simulations are underestimated the projected mean air temperature simulations. Although S4 has the last rank in all metrics, it gives a relative small average of P BIAS. While S1 has the relative maximum P BIAS. In Figures 10 and 11, S6 has a negative value of E indicating a poorly performance model.
5 CONCLUSIONS

An important source of the uncertainty in future climatic impact studies is related to the RCMs simulations selection. A set of six regional climate models used to simulate air temperature in the study area were purposely selected. Their simulation of air temperature (mean, maximum and minimum) have been validated against observed records using a selection procedure, to judge the quality of model performance. The results showed substantial differences between the
simulations. Generally, the models simulations of mean air temperature was better than the simulations of maximum and minimum air temperature. Based on the presented results, the model simulation of SMHI- RCA4 (for AFR-44 domain and [MPI-M-MPI-ESM-LR] GCMs) of CORDEX Project was the most accurate one. This selected simulation showed the best agreements with observed values in the two studied periods and for the three studied variables. Overall, this methodology is successfully used to narrow uncertainty in selecting regional climate models and thus may reduce the uncertainty of the subsequent climate change impact studies. Moreover, the results reveal the ability of the RCMs to simulate the past daily variability of air temperature, realistically. Examining the models performances for other meteoro logical variables and under more scenarios may be recommended, depending on the subsequent impact study.

ACKNOWLEDGMENTS

This article was originated as a part of the MSc thesis of the corresponding author at the Faculty of Engineering, Tanta University, Egypt. The authors extremely thank the World Climate Research Programme's Working Group on Regional Climate, and the Working Group on Coupled Modelling, former coordinating body of CORDEX and responsible panel for CMIP5 for making the simulations available through the ESGF web services. Also, sincere thanks must go to the reviewers who participated in enhancing the quality of this article.

ABBREVIATIONS

AR5 assessment report no. Fife
CC Climate Change
CMIP5 Coupled Modelling Intercomparison Programme, the fifth phase
GCM Global Circulation Model
GHG Greenhouse Gas
IPCC Intergovernmental Panel on Climate Change
RCM Regional Climate Model
RCP Representative Concentration Pathway
SST Sea surface Temperature
UNEP United nations environement programme
WCRP World Climate Research Programme

SYMBOLS

AME Absolute Mean Error
d index of agreement
E Coefficient of Efficiency
PBIAS Percent bias
RMSE Root Mean Square error
r correlation coefficient
$R^2$ coefficient of determination
$t_{an}$ near surface mean air temperature
$t_{asmax}$ near surface maximum air temperature
$t_{asmin}$ near surface minimum air temperature

REFERENCES


NOAA websites,(Accessed on 13th March 2017)
https://www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD&countryabbv=&georegiona bbv


