LONG-TERM PREDICTION OF CASPIAN SEA LEVEL UNDER CMIP6 SCENARIOS USING ARTIFICIAL NEURAL NETWORKS

S. Mahya Hoseini[1] and Mohsen Soltanpour[1]

Artificial Neural Network (ANN) is employed to predict the long-term Caspian Sea level (CSL). 114-year observed CSL data (1900–2014) and the precipitation and temperature of historical and future scenarios of Coupled Model Intercomparison Phase 6 (CMIP6) are used to predict the future fluctuations of CSL (2015–2050). The values of the statistical indices in training, validating and testing periods (1900-2014) indicate the efficiency of the ANN in reconstruction of the CSL. Considering the outputs of different climate change projections (CMIP6) and excluding the human interventions, the study predicts the CSL fluctuation range of -28 m to -26m until 2050.

Keywords: Caspian Sea level; Artificial Neural Networks; climate change; CMIP6

INTRODUCTION

The Caspian Sea (CS) is the world's largest inland water. CS locates in a semi-arid area between latitudes 36°-47°N and longitudes 47°-54°E. This endorheic basin, situated between Europe and Asia, is surrounded by five countries, i.e. Russia, Kazakhstan, Turkmenistan, Iran, and Azerbaijan (see Fig. 1). Excluding the Kara-Bogaz-Gol Bay, CS area is approximately ~371,000 km^2 with a length from north to south of about 1200 km (Rodionov, 1994).
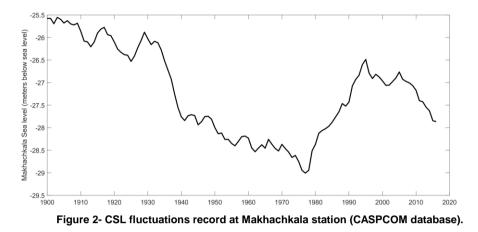
The CSL is below the mean sea level of the ocean. The catchment area of the rivers discharging into the CS is approximately 3.5 million km^2 (close to 2% of the global land), with a length of 2500 km (35°N to 60°N) and an average width of 1000 km (40°E to 60°E) (Arpe et al., 2000a; Rodionov, 1994). Approximately 80% of the river discharge into the CS belongs to the Volga River, supplying about 240 km^3 of water annually, which corresponds to 60 cm of the Caspian Sea level (CSL). 15 cm of CSL is the inputs from other rivers (e.g., Ural, Sulak, Samur, Kura and Terek) and 20 cm is attributed to the precipitation over the sea. The level is balanced by evaporation and, to a small extent, by the outflow to Kara-Bogaz-Gol Bay in the east.

Figure 1- The Caspian Sea.

Historical data and water level indicators reveal that the CSL has been strongly fluctuating in the past. In contrast to global sea level changes, CSL fluctuations have been occurred 100 times faster over the last century (Giralt et al., 2003). Fig. 2 illustrates historical observations of CSL in Makhachkala station for the period 1900-2016 (CASPCOM Working Group, 2011). During 1930 to 1977, sea water level sharply fell by more than 3 meters to its lowest level in the last 400 years, i.e., 29 meters below sea level (Baltic datum). During 1977 to 1995, an unexpected rise of about 2.5 meters (up to -26.5 m) resulted in enormous flooding and various problems in neighboring countries. After this sharp rise, the

[1] Faculty of Civil Engineering, K. N. Toosi University of Technology, No. 1346, Valiasr Street, Mirdamad Intersection, Tehran, Iran

CSL was gradually decreasing with the present value of about -27.8 m. This fall of water level has resulted in different kinds of problems, such as navigability at ports.

Figure 2- CSL fluctuations record at Makhachkala station (CASPCOM database).

The overwhelming majority of scientists who studied the cause of CSL fluctuations' state that the CSL is controlled by climate-induced changes of the sea's water budget. They also reported that tectonic factors have a very insignificant effect on CSL fluctuations; inversely, climatic components are of primary importance in CSL variations (Berg, 1934; Kalinin, 1968; Varuschenko et al., 1978; Golitsyn and Panin, 1989; Panin and Divakov, 1991) Additionally, during the 20th century, anthropogenic activities such as land-use changes and reservoir improvements influenced the CSL (Renssen et al., 2007). The main water budget of the CS comprises river discharges into the CS dominated by Volga River discharge, precipitation and evaporation over the CS and discharge to Kara-Bogaz-Gol (Rodionov, 1994). Considering the insignificant role of groundwater influx/outflux in CSL fluctuations (Nazarali et al., 2016), the annual variation of CSL can be formulated by Eq. 1:

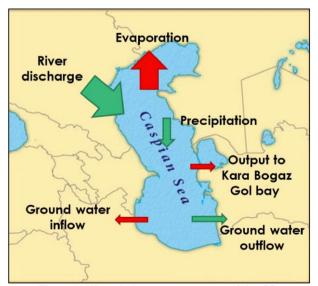$$\Delta L = \frac{Q_{Volga}}{A_{CS}} + \sum_{i=1}^{N} P_{CS,i} - \sum_{i=1}^{N} E_{CS,i} - KBG \qquad (1)$$

Figure 3- Water balance components of the CS.

where ΔL is change in CSL (cm/yr), Q_{Volga} is river discharge of Volga (cm^3/yr), A_{CS} is the CS surface area (cm^2), P_{CS} is precipitation over the CS (cm/yr), E_{CS} is evaporation over the CS (cm/yr) and KBG is the sea-level loss due to outflow to the Kara-Bogaz-Gol Bay (cm/yr) (see Fig. 3) (Rodionov, 1994).

Because of the broad socio-economic impacts of CSL changes, several attempts have been carried out to predict and discover the phenomena. K. Arpe et al. (2000a) found a high degree of correlation between the CSL and the Southern Oscillation Index, the driver of El Niño events. Overeem et al. (2003) revealed the very high correlation between the rate of CSL variations and discharge of the Volga River that was later supported by the findings of Arpe et al. (2000b).

Existing literature shows that the conducted studies have been focused on the reconstruction of the past CSL fluctuations through building regional climate models (RCMs), using the outputs of Atmosphere-Ocean General Circulation Models (AOGCMs) (Elguindi and Giorgi, 2006a; Chen et al., 2017). Future CSL has also been predicted using AOGCM and RCM outputs (Elguindi and Giorgi, 2006a; 2007; Renssen et al., 2007). Moreover, CSL was predicted by applying a simple statistical model to AOGCMs outputs (Roshan et al., 2012). A number of scientists who tried to forecast CSL fluctuations using stochastic time series conclude that stochastic models of time series, particularly Integrated Moving Average and Multiplicative (ARIMA), are good tools for simulating and prediction of short-term CSL variations. However, these models are not recommended for the forecasts longer than one year (Vaziri, 1997; Imani et al., 2014; Dehbashi et al., 2017).

In spite of all past efforts, most of the long-term predictions of CSL are still not satisfactory, even not be able to present the general trend of increasing/decreasing of water level. Moreover, the derived trends of long-term CSL predictions from different climate models indicate considerable differences and they are often inconsistent with each other. Prediction of future levels of the CS has remained a major challenge so far. In order to meet this challenge, the present study offers a different innovative method for the forecast of future water level of the CS. Past studies revealed that the changes of the CS water balance components, including precipitation and evaporation over the CS as well as the discharge of entering rivers, are the main cause of water level changes. Thus, a concise and accurate model, which establishes the relationship between the precipitation and temperature over the CS and the catchment area of the Volga River (80% of the riverine inlet) with historical records of CSL can be applied to precipitation and temperature forecasts scenarios of AOGCMs in order to predict future CSL fluctuations. In addition to considering the effects of climatic variables and climate change on CSL fluctuations, here the prediction is performed by Artificial Neural Network (ANN), as a well-known model to establish nonlinear relationships between variables. The main path of this study includes: (1) to test ANN architecture and analyze its performance in establishing a relationship between climatic variables and historical records of CSL, (2) to evaluate CMIP6 model outputs in the CS region, and (3) to assess which SSPs scenarios capture well present state of the CSL.

CSL MODELING

ANN is an established technique with a flexible mathematical structure capable of identifying complex nonlinear relationships between input and output data based on the brain's physiology. ANNs are helpful computational strategies for sea level research and forecasting. Because of their independence with regard to the assumptions of functional models, probability distribution, and smoothness, the applications of ANNs are very common (Imani et al., 2014). ANNs has also been frequently used for the precipitation and temperature analyses (Goyal and Ojha, 2012; Okkan and Inan, 2015; Okkan and Kirdemir, 2016; Vu et al., 2016; Alotaibi et al., 2018). ANNs have also been applied to satellite altimetry data of sea level to develop the CSL forecast model (Imani et al., 2014). They concluded that ANN models can successfully predict short-term sea level anomalies, even more precise in comparison with routine autoregressive moving average (ARMA) models.

An ANN model was trained in order to acquire a higher correlation between CSL with temperature and precipitation outputs of an AOGCM in representative stations over the Volga River basin and the CS. The red dots in Fig. 4 present the locations of mentioned representative stations. AOGCM output data was adopted from the latest Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2015). The CMIP6 data is obtained from the Canadian Earth System Model version 5 named CanESM5, which consists of the three-dimensional atmosphere (T63 spectral resolution equivalent roughly to 2.8∘) and ocean (nominally 1∘) general circulation models, a sea-ice model, a land surface scheme, and explicit land and ocean carbon cycle models (Swart et al., 2019). The CanESM5 data are available for various emission scenarios, which comprise historical simulations (available for 1850 to 2014), and future simulations including Shared Socioeconomic Pathway (SSP) experiments (available for 2015 to 2100).
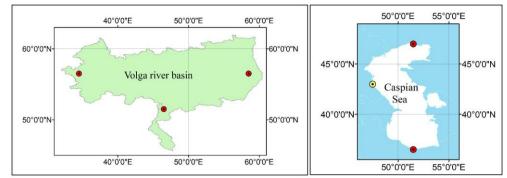
Figure 4- Locations of derived climate outputs (red dots) and Makhachkala sea level data (yellow dot).

Historical simulations are forced with estimates of natural climate forcings (e.g., volcanic eruptions, solar and orbital variability) and anthropogenic climate forcings (e.g., Greenhouse gas emissions, aerosols and land use change) in order to simulate climate change and variability over the observational record time period. The SSPs reflect various future greenhouse gas emissions and land-use change scenarios, estimated from Integrated Assessment Models (IAMs) and based on different assumptions relating to economic growth, climate mitigation efforts, and global governance. Using these assumptions, the SSPs are utilized to produce diverse radiative forcing pathways, and related warming, up to the end of the 21st century (O'Neill et al., 2016).

The entire data comprising both historical (1850–2014) simulations conducted as part of the core DECK experiments (Eyring et al., 2015) and seven SSPs (2015–2100) from ScenarioMIP (O'Neill et al., 2016) were downloaded from the earth system grid federation (EGSF) website (https://esgf-node.llnl.gov/search/cmip6) (Cinquini et al., 2014). Seven SSP's simulations were considered to simulate the ranges of probable future fluctuations of the CSL:

SSP1-2.6 (2.6 Wm^{-2}; low forcing, sustainable development),
SSP2-4.5 (4.5 Wm^{-2}; medium forcing, middle-of-the-road development),
SSP3-7.0 (7.0 Wm^{-2}; medium-to high-end forcing pathway, regional rivalry),
SSP5-8.5 (8.5Wm^{-2}; high-end forcing pathway fossil fuel-driven development),
SSP1-1.9 (1.9 Wm^{-2}; relatively low forcing, inequality),
SSP4-3.4 (3.4 Wm^{-2}; relatively low forcing, inequality),
SSP4-6.0 (6.0 Wm^{-2}; relatively high forcing, inequality),
and SSP5-3.4-OS (3.4 Wm^{-2}; relatively low forcing, fossil fuel-driven development, an overshoot pathway).

The ANN usually comprises the hidden and output layers. A feed forward multi-layer perceptron (MLP), i.e. the type of ANN broadly used in forecasting models, was applied here (Patil et al., 2015). The input layer consists of the precipitation flux and average temperatures of historical outputs from the CanESM5 model in 5 locations over the Volga River basin and the Caspian Sea. Yearly CSL fluctuations, recorded at Makhachkala station over 114 years (1900-2014), was selected as the output layer. As *z*-score standardization makes the training of a downscaling model capable, the standardization process was exerted on all data before being inputted into ANN. The standardization procedure with the *z*-score method was also applied to future scenarios of the CanESM5 model (Eq. 2), where x, μ, σ and z represent raw value, dataset mean, dataset standard deviation and normalized value (z-score).

$$z = \frac{x - \mu}{\sigma} \qquad (2)$$

Before applying ANN to derive the nonlinear relationship for future CSL forecast, the whole data was divided into three groups, i.e. 70% for model training, 20% for model validating and 10% were for model testing. A MATLAB code was developed for the proposed training and prediction strategies (Fig. 5). The ANN training and validation were repeated 20 times, and the best trained ANN was selected for the prediction, based on statistical parameters of Correlation Coefficient (CC) and Root Mean Square Error (RMSE), formulated in Eq. 3 and Eq. 4, respectively. Applying precipitation and temperature of CMIP6 scenarios, including SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0, SSP4-3.4, SSP4-6.0 and SSP5-8.5 to the trained ANN, the prediction was performed for the period of 2015-2050.
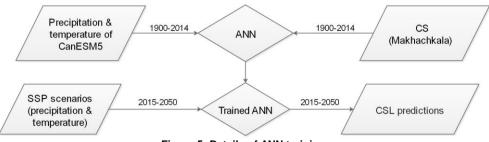
Figure 5- Details of ANN training.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - S_i)^2} \qquad (3)$$

$$CC = \frac{\sum_{i=1}^{N}(O_i - \overline{O_i})(S_i - \overline{S_i})}{\sqrt{\sum_{i=1}^{N}(O_i - \overline{O_i})^2 \frac{1}{N}\sum_{i=1}^{N}(S_i - \overline{S_i})^2}} \qquad (4)$$

RESULTS

The performance of ANN in CSL reconstruction was evaluated using CC and RMSE. Table. 1 presents the values of these statistical indices for training, validating and testing periods, indicating the efficiency of the ANN in the reconstruction of the CSL. Fig. 6 shows the observed CSL against modeled CSL, which confirms the good performance of ANN for simulating the past fluctuations (1900-2014). Comparing the predicted and recorded CSLs, it is observed that all scenarios of CanESM5 models are higher than the observed values of CSL in 2015 and 2016. It might be related to the overall performance of CMIP6 models that differs across different climatic zones (Yazdandoost et al., 2020). The outputs of different scenarios reveal that CSL predictions by SSP1-1-2.6 and SSP4-60 scenarios have minimum mean errors (see Table 2). Moreover, it is observed that considering the outputs of different climate change projections (CMIP6), CSL will probably fluctuate between -28 m and -26 m until 2050. However, it should be added that dramatic human interventions, particularly the programs of dam constructions on Volga River and vast irrigation projects can highly affect this prediction.

Table 1. ANN performance.

Result	Training	Validating	Testing
CC	0.967	0.928	0.926
RMSE	0.25	0.55	0.39

Table 2. Prediction's relative error.

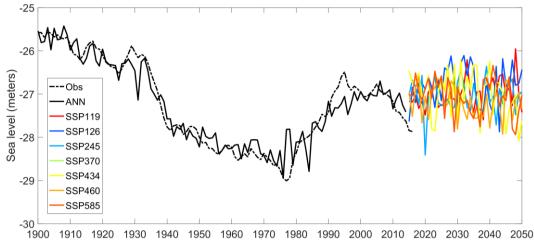| SSP scenarios / Year	1-1.9	1-2.6	2-4.5	3-7.0	4-3.4	4-6.0	5-8.5
2015	0.024	0.007	0.029	0.05	0.05	0.032	0.03
2016	0.025	0.037	0.043	0.04	0.04	0.013	0.03
Average	0.024	0.022	0.036	0.045	0.045	0.022	0.03

Figure 6- Observed and modeled past and future CSL under different CMIP6 scenarios.

SUMMARY AND CONCLUSION

The CSL has high environmental effects and huge economic impacts on the coastal infrastructures of neighboring countries and planned future developments. The CS is experiencing substantial fluctuations, which is mainly due to the net balance between inflows and over-sea precipitation versus outflows and over-sea evaporation. The CS and Volga River basin in north-south alignment extends over a distance of approximately 10 degrees, which leads to crossing several different climatic zones. In order to predict CSL in future, an ANN approach was presented establishing a non-linear relationship between AOGCM's temperature, precipitation over the CS and Volga River basin, and historical records of CSL. The effectiveness of ANN in CS simulations of CanESM5 for the latest CMIP6 scenarios was assessed and trained ANN was used for CSL forecast. Seven different shared socioeconomic pathways were considered and the CanESM5 outputs were selected as it covers all scenarios in the study area. Various climate change scenarios lead to fluctuating CSL between -28 m and -26 m until 2050. In future, it is recommended to utilize various AOGCMs across different parts of the CS and Volga River basin, based on their different climatic zone to increase the accuracy of CSL forecast and limit the overestimate predictions.

REFERENCES

Alotaibi, K., Ghumman, A.R., Haider, H., Ghazaw, Y.M., Shafiquzzaman, M., 2018. Future predictions of rainfall and temperature using GCM and ANN for arid regions: A case study for the Qassim region, Saudi Arabia. Water (Switzerland) 10. https://doi.org/10.3390/w10091260

Arpe, K., Bengtsson, L., Golitsyn, G.S., Mokhov, I.I., Semenov, V.A., Sporyshev, P. V., 2000. Connection between Caspian Sea level variability and ENSO. Geophys. Res. Lett. 27, 2693–2696. https://doi.org/10.1029/1999GL002374

Arpe, K, Bengtsson, L., Golitsyn, G.S., Mokhov, I.I., Semenov, V.A., Sporyshev, P. V, 2000. Connection between Caspian Sea level variability the rivers flowing into the Sea is about mean to December of next year) and SSTs (mean mean). Values. Geophys. Res. Lett. 27, 2693–2696.

Berg, L.S., 1934. The Caspian Sea level for the historical period. Probl. Fiz. Geogr. (Problems Phys. Geogr. 11-64 (In Russian).

CASPCOM Working Group, 2011. General Catalogue of the Caspian Sea Level. Coord. Comm. Hydrometeorol. Pollut. Monit. Casp. Sea. URL http://www.caspcom.com/ (accessed 9.5.20).

Chen, J.L., Pekker, T., Wilson, C.R., Tapley, B.D., Kostianoy, A.G., Cretaux, J.F., Safarov, E.S., 2017. Long-term Caspian Sea level change. Geophys. Res. Lett. 44, 6993–7001. https://doi.org/10.1002/2017GL073958

Cinquini, L., Crichton, D., Mattmann, C., Harney, J., Shipman, G., Wang, F., ... & Pobre, Z., 2014. The Earth System Grid Federation: An open infrastructure for access to distributed geospatial data. Futur. Gener. Comput. Syst. 36, 400–417. https://doi.org/https://doi.org/10.1016/j.future.2013.07.002

Dehbashi M, Azarmsa S A, Vafakhah M, 2017. Water Level Fluctuation Analysis and Forecast in the

Caspian Sea Using Stochastic Time Series models. marine-engineering 13, 23–33.

Elguindi, N., Giorgi, F., 2007. Simulating future Caspian sea level changes using regional climate model outputs. Clim. Dyn. 28, 365–379. https://doi.org/10.1007/s00382-006-0185-x

Elguindi, N., Giorgi, F., 2006a. Simulating multi-decadal variability of Caspian Sea level changes using regional climate model outputs. Clim. Dyn. 26, 167–181. https://doi.org/10.1007/s00382-005-0077-5

Elguindi, N., Giorgi, F., 2006b. Projected changes in the Caspian Sea level for the 21st century based on the latest AOGCM simulations. Geophys. Res. Lett. 33, 4–7. https://doi.org/10.1029/2006GL025943

Eyring, V., Bony, S., Meehl, G.A., Senior, C., Stevens, B., Stouffer, R.J., Taylor, K.E., 2015. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organisation. Geosci. Model Dev. Discuss. 8, 10539–10583. https://doi.org/10.5194/gmdd-8-10539-2015

Giralt, S., Julià, R., Leroy, S., Gasse, F., 2003. Cyclic water level oscillations of the KaraBogazGol - Caspian Sea system. Earth Planet. Sci. Lett. 212, 225–239. https://doi.org/10.1016/S0012-821X(03)00259-0

Golitsyn, G.S., and G.N. Panin, 1989. The water balance and modern variations of the level of the Caspian Sea. Sov. Meteorol. Hydrol. 46–52.

Goyal, M.K., and C. S.P. Ojha, 2012. Downscaling of surface temperature for lake catchment in an arid region in India using linear multiple regression and neural networks. Int. J. Climatol. 32, 552–566. https://doi.org/10.1002/joc.2286

Imani, M., You, R.J., Kuo, C.Y., 2014. Caspian Sea level prediction using satellite altimetry by artificial neural networks. Int. J. Environ. Sci. Technol. 11, 1035–1042. https://doi.org/10.1007/s13762-013-0287-z

Kalinin, G.P., 1968. The Global Hydrology Problems. Gidrometeoizdat 378 (in Russian).

O'Neill, B.C., Tebaldi, C., Van Vuuren, D.P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.F., Lowe, J., Meehl, G.A., Moss, R., Riahi, K., Sanderson, B.M., 2016. The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geosci. Model Dev. 9, 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Okkan, U., Inan, G., 2015. Statistical downscaling of monthly reservoir inflows for Kemer watershed in Turkey: Use of machine learning methods, multiple GCMs and emission scenarios. Int. J. Climatol. 35, 3274–3295. https://doi.org/10.1002/joc.4206

Okkan, U., Kirdemir, U., 2016. Downscaling of monthly precipitation using CMIP5 climate models operated under RCPs. Meteorol. Appl. 23, 514–528. https://doi.org/10.1002/met.1575

Overeem, I., Veldkamp, A., Tebbens, L., Kroonenberg, S.B., 2003. Modelling Holocene stratigraphy and depocentre migration of the Volga delta due to Caspian Sea-level change. Sediment. Geol. 159, 159–175. https://doi.org/10.1016/S0037-0738(02)00257-9

Panin, G.N., and I.V. Divakov, 1991. Estimation of possible variations of evaporation from the North Caspian Sea under the conditions of sea level rising. Sov. Meteorol. Hydrol. 51–57.

Patil, N.S., Laddimath, R.S., Hooli, S., 2015. Downscaling of Precipitation Data from GCM outputs using Artificial Neural Network for Bhima basin 10, 1493–1508.

Renssen, H., Lougheed, B.C., Aerts, J.C.J.H., de Moel, H., Ward, P.J., Kwadijk, J.C.J., 2007. Simulating long-term Caspian Sea level changes: The impact of Holocene and future climate conditions. Earth Planet. Sci. Lett. 261, 685–693. https://doi.org/10.1016/j.epsl.2007.07.037

Rodionov, S.N., 1994. Global and regional climate interaction: the Caspian Sea experience, Global and regional climate interaction: the Caspian Sea experience.

Roshan, G., Moghbel, M., Grab, S., 2012. Modeling Caspian Sea water level oscillations under different scenarios of increasing atmospheric carbon dioxide concentrations. J. Environ. Heal. Sci. Eng. 9, 1–10.

Swart, N.C., Cole, J.N.S., Kharin, V. V., Lazare, M., Scinocca, J.F., Gillett, N.P., Anstey, J., Arora, V., Christian, J.R., Hanna, S., Jiao, Y., Lee, W.G., Majaess, F., Saenko, O.A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., Von Salzen, K., Yang, D., Winter, B., 2019. The Canadian Earth System Model version 5 (CanESM5.0.3). Geosci. Model Dev. 12, 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Varuschenko, S.I., Varuschenko, A.H., and Klige, R.K., 1978. Changes in Regime of the Caspian Sea and Closed Water Bodies in Paleotime (In Russian).

Vaziri, B.M., 1997. Predicting Caspian Sea Surface Water Level 158–162.

Vu, M.T., Aribarg, T., Supratid, S., Raghavan, S. V., Liong, S.Y., 2016. Statistical downscaling rainfall using artificial neural network: significantly wetter Bangkok? Theor. Appl. Climatol. 126, 453–467. https://doi.org/10.1007/s00704-015-1580-1

Yazdandoost, F., Moradian, S., Izadi, A., Aghakouchak, A., 2020. Evaluation of CMIP6 precipitation simulations across different climatic zones: Uncertainty and model intercomparison. Atmos. Res. 105369. https://doi.org/10.1016/j.atmosres.2020.105369