Development and Evaluation the Performance of Ann Based Statistical Downscaling Models For Daily and Monthly Precipitation.

Khaled Alramlawi (✉ khaled.aramlawi@ogr.deu.edu.tr)

Dokuz Eylul Universitesi Muhendislik Fakultesi <https://orcid.org/0000-0002-3432-2660>

Okan Fistikoglu

Dokuz Eylul University: Dokuz Eylul Universitesi

Mohamed Najjar

Dokuz Eylül Üniversitesi: Dokuz Eylul Universitesi

Research Article

Keywords: Statistical downscaling model, daily and monthly precipitation, Artificial Neural Network) ANN), large-scale atmospheric variables

Posted Date: February 22nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1363323/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

DEVELOPMENT AND EVALUATION THE PERFORMANCE OF ANN BASED STATISTICAL DOWNSCALING MODELS FOR DAILY AND MONTHLY PRECIPITATION.

Khaled Alramlawi 1
Okan Fistikoglu 2
Mohamed Najar 3

1 Dokuz Eylul University, The Graduate School of Natural And Applied Sciences,
Civil Engineering Department, Izmir, Turkey - khaled.alramlawi@ogr.deu.edu.tr -
<https://orcid.org/0000-0002-3432-2660>

2 Dokuz Eylul University, The Graduate School of Natural And Applied Sciences,
Civil Engineering Department, Izmir, Turkey - okan.fistikoglu@deu.edu.tr -
<https://orcid.org/0000-0002-9483-1563>

3 Dokuz Eylül University, Graduate School of Natural and Applied Sciences, GIS
Dept., Izmir, Turkey eng.mhmd.najar@gmail.com - <https://orcid.org/0000-0002-9107-961X>

Abstract

Statistical downscaling techniques represent a quantitative relationship between large-scale atmospheric variables (predictors) and local-scale meteorological variables (predictand) such as precipitation. This study uses large-scale atmospheric as predictor variables derived from the National Centre for Environmental Prediction and National Centre for Atmospheric Research (NCEP/NCAR) reanalysis data set and precipitation in Izmir city stations as predictant.

The purpose of this study is to develop statistical downscaling models for daily and monthly precipitation over Izmir city by using Artificial Neural Network (ANN) methods and comparison of the performance of those models. The results revealed that the performance of the daily model improves with the aggregated of daily results. Although the performance of the daily model gives fair (not high) results (e.g., R^2 ranging from 0.362 to 0.331), the aggregated model gives very good results at monthly

level (e.g., R^2 is more than 0.73). The monthly statistical downscaling model gives very good results for the study area (e.g. R^2 ranging between 0.65 to 0.73).

It was founded that a slight variation from the performance of aggregated monthly model and the monthly model. However, the aggregated monthly provided better results. Although the accuracy of the aggregated monthly model is higher, it required a significant amount of time and effort.

ANN-based downscaling model could be used to determine the inputs of water potential studies, but the results of the model have to be improved by Statistical methods such as Bias correction in order to use it in floods study.

Keywords

Statistical downscaling model, daily and monthly precipitation, Artificial Neural Network (ANN), large-scale atmospheric variables.

INTRODUCTION

Global climate changes can lead to hydrological changes that will affect almost all aspects of humans, e.g, regional water availability, agricultural productivity, flood control. So it will be necessitates understanding how a change in global climate could affect regional water availability (Kusangaya et al. 2014).

General circulation models (GCMs) are one of the most important tools for studying climate change, GCMs are numerical models and describe the atmospheric processes through mathematical equations. GCMs which have been developing over several decades are able to simulate features of the global climate, represent various earth systems including the atmosphere, oceans, land surface, and sea ice, and offer considerable potential for studying climate change (Liu et al. 2016).

The performance of GCMs are coarse at the smaller spatial and temporal scales relevant to regional climate, because the spatial resolution grids are too coarse to resolve many important sub grid-scale processes, therefore GCMs outputs are often unreliable at regional scales (Xu 1999).

One possible solution to overcome this problem is using downscaling techniques to downscale the output from GCMs or reanalysis datasets to a higher resolution in space and time. The basic idea of downscaling techniques is transferring large-scale changes in atmospheric variables to local weather (Hanssen-Bauer et al. 2005).

Downscaling methodologies are classified into two main groups as dynamic downscaling and statistical downscaling (Fistikoglu and Okkan 2011). Dynamic downscaling, which use of regional climate models (RCMs) to produce higher resolution outputs, is very computationally intensive, very complex, and requires substantial computational resources. Therefore, it is not the best options for the studies

that require a quick response. The statistical downscaling methodology requires less computational effort than dynamical downscaling applications and it is sound statistically. Statistical downscaling consists of a varied group of approaches that are often almost simple to implement, but requires an adequate amount of quality observed data (Trzaska and Schnarr 2014). Statistical downscaling can be divided into three main categories, which are Weather classification, Weather generators, and Transfer functions.

Weather classification methods use synoptic patterns to classify weather patterns based on their synoptic similarity. After that, they establish quantitative relationships to assign Predictants (Bardossy et al. 2005; Fistikoglu and Okkan 2011). Although these methods are appropriate for downscaling nonnormal distributions, they require a large amount of observed data (more than 30 years) to evaluate all probable weather conditions. These approaches are more computationally intensive than linear approaches (Trzaska and Schnarr 2014).

Weather generators methods are replicators of local weather statistical features under the large-scale variables conditions (Fistikoglu and Okkan 2011; Kilsby et al. 2007). These methods are widely used in temporal downscaling.

Transfer functions methods build the relationships between large-scale atmospheric variables and local surface variables. These methods are considered the most common methods in statistical downscaling. Transfer functions use various applications including linear and non-linear regression, artificial neural networks, redundancy analysis, and canonical correlation analysis (Benestad et al. 2007; Wilby et al. 2004).

The applications of transfer functions are range from linear and nonlinear regression types to artificial neural networks (ANNs), principal component analysis (PCA),

canonical correlation, and redundancy analysis (Schoof et al. 2007),(Fistikoglu and Okkan 2011).

ANN is one of the most popular approach of transfer function for downscaling studies. There are several successful studies on the downscaling of large-scale atmospheric variables by the ANN based transfer functions (Goyal & Ojha, 2012; Laddimath & Patil, 2019; Nourani et al., 2019; Okkan & Kirdemir, 2016; Salimi et al., 2019; Vu et al., 2016; Wang et al., 2020).

Studies comparing different statistical downscaling methods are now relatively common (Singh and Kumar 2020; Liu et al. 2016). The results of these studies have shown that different methods have different performance in a certain area, and a certain method has different performance in different study areas.

Many studies were implemented SDM using monthly precipitation as predictands (Okkan and Kirdemir 2016) and also relatively few previous studies focused on daily precipitation (Singh and Kumar 2020). This paper take in consideration the daily and monthly precipitation.

In this study, a statistical downscaling method based on ANN is applied to predict precipitation over the city of Izmir. The purpose of the study is to develop two different models, the first one is a model that downscale daily precipitation, and through this model we calculate a cumulative rainfall for different periods for two days until we reach the monthly precipitation by aggregate the rainfall day by day, while the second model is to downscale monthly rainfall directly by using monthly data. At the end of the paper, we compare the results of the monthly model and monthly cumulative results of the daily model.

STUDY REGION AND DATASETS

The study region is Izmir city, which is the third largest city In Turkey, it is located at the Aegean coast of Turkey (see Fig. 1). Izmir city and its environs reflect typical Mediterranean climate characteristics. Precipitation amounts in study area is around 610 mm per year, and it is mostly abundant in winter, whereas up to 80% of the total annual precipitation falls between November to March.

Predictands

There are many stations are available within the region, 6 stations were selected in this study. The available data and the location of the stations are listed in table 1. Daily and monthly precipitation data from meteorological stations in Izmir city used as predictand. The other neighbour stations around Izmir station were used in order to test the results of downscaling models regarding to the spatial consistency of the daily and monthly rainfall estimations. The required data were collected and extracted mainly from the Turkish state Meteorological Service in Turkey

Fig. 1: Location of selected meteorological stations within study area.

Table 1: Selected meteorological stations.

No	No of stations	Name of stations	Latitude	longitude	Elevation	available data	
1	17220	Izmir	38.3949	27.0819	29	1948	2018
2	17219	Adnan menderes	38.295	27.1479	119	2009	2018
3	17820	Seferihisar	38.199	26.835	22	1980	2011
4	17221	Cesme	38.3036	26.3724	5	1965	2018
5	17822	Odemis	38.2157	27.9642	111	1960	2011
6	17854	Selcuk	37.9423	27.3669	18	1965	2011

It's worth mention that, the stations (Adnan menderes, Selcuk, Seferihisar Cesme, and Odemis) were selected as the stations are geographically closed to each other, they are sharing similar meteorological and atmospherically conditions. Besides, the stations shows a strong +ve correlation (coefficient of determination R^2) with the main station used in this paper is Izmir station. As its clearly shown in Table 2, at the daily level, the obtained correlation between the main station (Izmir) and A.M. havalimani and Seferihisar stations are (0.72 and 0.65, receptivity), while odemis and cesme selcuk stations recorded relatively low R^2 regards to the other stations (havalimani and sefirhisar). On the other hand, at the monthly level, a strong correlation was observed between all stations ranging between (0.75 and 0.91), see Table 2.

Table 2: Summary of the R^2 between the precipitation in Izmir station and other neighbor stations

Stations	R^2 at daily level	R^2 at monthly level
Adnan menders	0.721	0.91
Seferehisar	0.654	0.842
Selcuk	0.506	0.829
Cesme	0.373	0.745
Odemis	0.403	0.793

Predictors:

The large-scale atmospheric variables from the National Center for Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis data set were used as predictors. The NCEP/NCAR data set is considered to represent the atmospheric conditions of the study area considerably well (Fistikoglu and Okkan, 2011). The large-scale atmospheric variables of NCEP/NCAR reanalysis dataset are

selected as the predictors for the period from 1948 - 2018. The spatial resolution of $2.5^{\circ} \times 2.5^{\circ}$. The latitudes range from 36.25°N to 38.75°N , and the longitudes range from 26.25°E to 28.75°E , see Fig. 1.

The first step in statistical downscaling is to select the appropriate predictors which can change from one location to another. Generally, predictors are chosen by a combination of experts' judgments and some statistical measures, such as correlation, and partial correlation (Anandhi et al. 2009; Wilby and Wigley 2000). According to the previous studies (Chen et al. 2010; Wilby et al. 1999), the predictors can be selected using stepwise regression and analysing correlation.

In this study, the main criteria for choosing the relevant predictors are the predictors should be physically and conceptually sensible for the predictand (Wilby et al. 1999). In addition, it has an acceptable correlation with the predictant which is daily precipitation in this study (Fistikoglu and Okkan 2011). Therefore, 12 large scale atmospheric variables were selected as predictors, which taking into consideration precipitation generation mechanism. The selected NCEP/ NCAR reanalysis variables are listed in Table 3. The data were downloaded from the NCEP/NCAR reanalysis project web site <http://www.esrl.noaa.gov/psd/data/grided/data.ncep.reanalysis.html>.

Table 3 : Selected Atmospheric predictors from NCEP/NCAR Reanalysis data set

Atmospheric level	Selected variables	Unit	Acronym
200 hpa	Mean geopotential Height	m	hgt200
500 hpa	Mean air temperature	degK	air500
	mean geopotential height	m	hgt500
	mean relative humidity	%	rhum500
850 hpa	mean air temperature	degK	air850
	mean geopotential height	m	hgt850
	mean relative humidity	%	rhum850
Surface	mean air temperature at surface	degK	Air
	mean relative humidity at surface	%	Rhum
	mean sea level pressure	Pascals	Slp
	mean v-wind at 10 m	m/s	Vwnd
	mean daily precipitation rate	kg/m ² /s	Prate

METHODOLOGY

The overall proposed methodology is displayed in Fig. 2. The first step in this study is the collection and preparing the required data from local scale and large scale. In the presented paper, two models were developed, the daily model refers to the original data extracted from daily Atmospheric (NCEP/NCAR) reanalysis data set, while the monthly model refers to the model that use the original data driven from monthly large scale Atmospheric (NCEP/NCAR) reanalysis data set.

Fig. 2: flow chart illustrates the proposed Methodology

Statistical Downscaling techniques

Statistical downscaling techniques were developed to establish quantitative relationships between large-scale atmospheric variables (predictors) and local surface variables (predictants) (Goly et al. 2014; Sennikovs and Bethers, 2009).

Statistical downscaling models can be defined as $Y = F(X)$, where Y is the predictand, which represent the daily/monthly precipitation, and X is the predictors. The predictand is the target data, which is the small-scale variable representing in this study precipitation from meteorological stations in Izmir city (see Table 1). Predictors are the input that represent the large-scale atmospheric variables from NCEP/NCAR reanalysis data set (see Table 3).

Statistical Downscaling using ANN

ANN was used as a well-known machine learning technique to estimate the observed daily rainfall from the large-scale atmospheric reanalysis parameters (Wilby et al. 1998).

ANN refers to computing systems whose central theme is borrowed from the analogy of biological neural networks. ANN are able to learn and generalize from examples to produce meaningful solutions to problems (Jiang and Cotton 2004). The ANN is utilized as a practical black-box tool for developing a non-linear regression between the large-scale atmospheric dataset (predictors) and observed daily rainfall (predictands) (Harpham and Wilby 2005; Khan et al. 2006).

The ANN structure designed consists of three types of layers; the first one is an input layer, which has n neurons for each variable; here, the input layer is the one where the predictors (large-scale atmospheric variables) were defined. The second layer is the hidden layer which may have several neurons, and all hidden neurons

transform the inputs nonlinearly into another dimension through weight and a bias term shown in Eq. (1). In this study, the assumption that the number of neurons is double the number of predictors was accepted (by applying the trial and error technique). The last layer is the output layer (predictand) which in this study contains daily/monthly precipitation variables from station records. See Fig. 3.

Fig. 3 : Structure of the downscale model

$$y_k = f_2 \left(\sum_{j=1}^k \left[f_1 \left(\sum_{i=1}^n x_i w_{ij} + w_j \right) \right] w_{jk} + w_k \right) \quad (1)$$

Where : y_k : output ($j=1,2, \dots,k$), x_i : input ($i=1,2,\dots,n$), w : bias value, $f_{1,2}$ are activation function at the hidden and output layer.

Three popular transfer functions are tried out in ANN construction trials: tangent sigmoid, linear, and log-sigmoid. In this study, the tangent sigmoid function was found suitable. See Eq. (2)

$$f(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}} \quad (2)$$

The search for the best value of weights and biases is referred to as ANN learning phase, which is carried out with a known input and output set. Each training phase

involves a set of inputs passed forward through the network to generate trial outputs, and then compare the observed outputs. When the difference (residual) exceeds the desired value, the error is passed back to the network.

The training algorithm adjusts the connection weights based on the error. This process is referred to as back-propagation. Once, the comparison error has been reduced to an acceptable level for the entire training set, the training phase is complete. After training, the network is evaluated using a set of cases withheld from it during the training session. After the model was completed, it was ready to be applied to any other situation.

The network is trained using the Levenberg-Marquardt feed forward-back-propagation algorithm, as shown in Eq. (3). The Levenberg-Marquardt feed forward-back-propagation algorithm is a second-order non-linear optimization technique, that is typically faster and more reliable than any other back-propagation methodology. (Fistikoglu and Okkan, 2011).

$$\Delta \mathbf{w} = - [\mathbf{J}^T(\mathbf{w})\mathbf{J}(\mathbf{w}) + \mu \mathbf{I}]^{-1} \mathbf{J}^T(\mathbf{w})\mathbf{e}(\mathbf{w}) \quad (3)$$

$\mathbf{J}$ =Jacobian matrix, $\mathbf{w}$ is parameter vector, μ is Marquardt parameter.

RESULTS AND DISCUSSION

The performance of the model is evaluated by comparing the results of statistical downscaled model **P(t)_down.** and precipitation observed **P(t)_obs.** from stations data. To examine and evaluate the performance of models, the coefficient of determination (R^2), and root mean square error (RMSE) indices were used. See Eq. 4, 5,

$$R^2 = 1 - \frac{\sum_{i=1}^n (\text{observed}(i) - \text{predicted}(i))^2}{\sum_{i=1}^n (\text{observed}(i) - \text{Observed}(\text{mean}))^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\text{predicted}(i) - \text{observed}(i))^2}{N}} \quad (5)$$

Daily SDM

To calibrate the model, the time series of Izmir station was divide for training and test datasets. The precipitation data of the years between 1948 and 1990 were served as training period, while rain-data recorded for the years of 1991 to 2018 employed as testing period.

The cumulative precipitation was obtained through aggregating the daily unit to larger units (i.e 2, 3, 4, 5, 6, 7, 15, 30-days). Eq. 6 mathematically expresses the aggregation process.

$$P(n)_{cum} = \sum_n^1 P(n) \quad (6)$$

Where $P(n)_{cum}$ is the cumulative precipitation aggregated from daily precipitation, and n is the number of days.

The performance of the model for all the aggregated units (from the daily unit) combined with the descriptive statistics of the training time series are presented in Table 4. Similarly, Table 5 shows the same statistics for the tested time series.

The results show that the mean values for both observed and downscaled precipitation are very close to each other in both training and testing periods, regardless aggregated units. However, in terms of standard deviation, as the aggregated unit becomes larger, the gap between the observed and downscaled values decreases, giving a strong indicator that the performance of the model is enhanced as the

aggregated unit increases. Besides, when comparing the coefficient of variation between the observed and downscaled precipitation, the abovementioned conclusion can be drawn as Fig. 4 illustrate.

Fig. 7 shows that the differences of the box plots of observed and downscaled depresses in long duration like 15 days and 30 days.

For the aims of spatially validate the neighbour stations, the concept of the aggregated unit was also applied. The model behaviour maintains valid (R^2 increase with the aggregated unit increase) for these stations, as Fig. 6 and Table 6 help indicate. To sum up, it was clear that the performance of the model improves as the aggregated unit increases. Although the performance of the model gives fair results at daily level, a significant improvement has been recorded at 30 days level (e.g., R^2 is more than 0.73).

Table 4: Summary of the results of daily model for the training periods for Izmir station (1948-1990)

	precipitation for																	
	1 day		2 days		3 days		4 days		5 days		6 days		7 days		15 days		30days	
type	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.
Mean	1.84	1.83	3.68	3.66	5.52	5.50	7.36	7.33	9.21	9.16	11.04	10.99	12.88	12.82	27.62	27.48	55.18	54.93
Std. dev.	6.59	3.89	10.38	6.91	13.44	9.38	16.24	11.76	18.74	13.80	21.13	15.69	23.50	17.67	39.85	31.89	66.08	55.44
Skew.	6.06	3.43	4.67	2.97	3.93	2.68	3.47	2.50	3.29	2.26	2.95	2.13	2.80	1.94	2.02	1.61	1.55	1.31
Cv	3.57	2.13	2.82	1.89	2.43	1.71	2.21	1.61	2.04	1.51	1.91	1.43	1.83	1.38	1.44	1.16	1.20	1.01
RMSE	5.26		7.54		9.31		11		12.35		13.72		15.06		22.25		34.15	
R ²	0.3627		0.4733		0.5207		0.5412		0.5663		0.5787		0.5899		0.6892		0.733	

Table 5: Summary of the results daily model for the testing periods for Izmir station (1991-2018)

	precipitation for																	
	1 day		2 days		3 days		4 days		5 days		6 days		7 days		15 days		30days	
type	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.	obs.	down.
Mean	1.94	1.51	3.88	3.03	5.81	4.54	7.75	6.06	9.69	7.57	11.63	9.09	13.57	10.60	29.09	22.72	57.94	45.34
Std. dev.	7.15	3.37	11.13	5.89	14.69	8.02	17.32	9.92	20.17	11.67	22.45	13.34	24.73	14.95	40.80	25.57	65.81	43.79
Skew.	6.77	3.71	4.84	3.16	4.10	2.85	3.48	2.66	3.15	2.40	2.90	2.24	2.63	2.20	1.79	1.45	1.44	1.23
Cv	3.69	2.23	2.87	1.94	2.53	1.77	2.23	1.64	2.08	1.54	1.93	1.47	1.82	1.41	1.40	1.13	1.14	0.97
RMSE	5.9132		8.413		10.72		12.19		14.033		15.147		16.559		25.746		38.419	
R ²	0.3308		0.4552		0.5014		0.5399		0.5543		0.5872		0.5936		0.6601		0.7308

Fig. 4: Coefficient of variation in training and test periods

Fig. 5: Coefficient of variation in training and test periods

Table 6: Summary of the value of R^2 of precipitation for 1 day and cumulative days

Station	Izmir	Izmir	Adnan menders	Selcuk	Seferihisar	Cesme	Odemis
Type of period	Training period	Testing period	Testing period	Testing period	Testing period	Testing period	Testing period
Period	1948-1990	1991-2018	2009 - 2018	1965 - 2011	1980 - 2011	1965 - 2018	1960 - 2011
Rainfall for	R^2	R^2	R^2	R^2	R^2	R^2	R^2
1 day	0.36	0.34	0.37	0.32	0.26	0.27	0.28
2 days	0.48	0.45	0.50	0.46	0.36	0.35	0.41
3 days	0.52	0.50	0.58	0.50	0.41	0.40	0.46
4 days	0.55	0.53	0.61	0.55	0.46	0.44	0.50
5 days	0.56	0.55	0.62	0.56	0.49	0.44	0.53
6 days	0.58	0.59	0.66	0.59	0.52	0.48	0.55
7 days	0.59	0.59	0.67	0.59	0.52	0.48	0.56
15 days	0.69	0.66	0.76	0.69	0.61	0.58	0.66
30 days	0.74	0.72	0.83	0.76	0.69	0.67	0.74

Fig. 6: The performance of the model (R^2) improves with an increase in the duration.

Monthly SDM

Monthly model uses the original data driven from monthly large scale (NCEP/NCAR) reanalysis variables. The performance of the monthly model combined with the descriptive statistics of the training and testing time series are presented in Table 7. The model gives very good results, i.e. R^2 is 0.73 and 0.63 for training and testing period. In addition, the monthly model provides high R^2 records for the neighbour stations (see Table 8).

Table 7: Summary of the results of monthly model for Izmir station

	Training 1948-1990		Testing 1991-2018	
Type	obs.	down.	obs.	down.
Mean	55.45	59.13	58.99	57.68
Std. dev.	56.66	57.17	66.73	55.46
Skew.	1.48	1.23	1.38	1.34
Cv	1.02	0.97	1.13	0.96
RMSE	35.01		39.32	
R^2	0.735		0.653	

Monthly model vs. aggregated monthly model

As explained earlier in the study, the term “aggregated monthly model” refers to a monthly model that aggregated from the daily model with daily input units (e.g., 30 or 31 days), while the term “monthly model” pointing out the direct use of monthly NCEP/NCER data.

When comparing the results of aggregated monthly model with the results of monthly model, similar outputs were obtained. However, for all investigated stations, the aggregated monthly model results provided better accuracy (see Table 8), e.g., for Izmir station, when applying the aggregated monthly model, the results of training period show that R^2 and RMSE equal to 0.75 and 33.78 respectively, while it is equal to 0.73, and 35.11 respectively for monthly model.

Table 8: The performance of cumulative monthly results and monthly model.

Stations	Observation period		Monthly cumulative results of daily model		Monthly model	
			R ²	RMSE	R ²	RMSE
Izmir	training	1948-1990	0.7515	33.784	0.7345	35.016
Izmir	test period	1991-2018	0.7433	38.708	0.6536	39.32
Adnan menders	test period	2009-2018	0.7935	36.12	0.7505	38.45
Selcuk	test period	1965-2011	0.7621	35.438	0.709	36.75
Odemis	test period	1980-2008	0.731	30.04	0.688	32.37
Cesme	test period	1965-2018	0.691	32.844	0.657	35.077
Seferehisar	test period	1960-2011	0.696	31.9	0.682	31.196

CONCLUSION

The ANN technique was used to establish statistical downscaling models for estimating station-based precipitation from the large-scale NCEP/NCAR atmospheric variables. The downscaling models were established using the records in Izmir station, and for spatial validation of the model, the nearby stations were utilized. Here it was downscaled daily and monthly precipitation in separated models, and then the performance of these models was compared.

In the daily model, although the performance of the model gives fair (not high) results at daily level, a significant improvement has been recorded as the aggregated day increases. That means, the accuracy at weekly precipitation level is significantly better than at daily level, as well as at monthly precipitation better than at weekly (e.g., R² are 0.36, 0.59 and 0.73 for daily, weekly, and monthly model respectively for Izmir station in training period). Also, the model behaviour maintains valid for the neighbour stations.

In the Monthly downscaled model, where mean monthly large scale atmospheric variables were used, it is observed that the model gives very good results at training and test period, (e.g. R^2 ranging 0.65 to 0.73).

In general, performance of statistical downscaling model shows that ANN model produces very good results for monthly model, while daily model produces fair (not high) results. This is explained by the presence of uncertainty in daily predictors and predictants, which is much higher than that in monthly data. (Khan et al., 2006). In addition to that, the ANN downscaling model simulates the mean values perfectly, because ANN uses MSE for obtaining the best results, so it produces small amount of precipitation in each dry days, but the mean values are nearly the same with observed values.

So, the aggregated daily sums approaches to the observed cumulative as the aggregate day increases. For this reasons, the monthly downscaling models have relatively strong performance comparing with the daily downscaling. It is found that there are similar findings of the aggregated monthly model and the monthly model. However, the aggregated monthly provided a little better result. Although the accuracy of the aggregated monthly model is higher, it required a significant amount of time and effort.

This study recommends that if monthly precipitation is required for climate studies, the monthly downscaling approach could be preferred conveniently rather than daily downscaling due to the time of the process.

Overall, these results indicate that ANN method proved to be a good tool for producing local surface variables like precipitation, especially at the monthly level, from large-scale atmospheric variables.

REFERENCES

- Anandhi A, Srinivas V, Kumar DN, Nanjundiah RS (2009) Role of predictors in downscaling surface temperature to river basin in India for IPCC SRES scenarios using support vector machine. *International Journal of Climatology* 29:583–603. <https://doi.org/10.1002/joc.1719>
- Bardossy A, Bogardi I, Matyasovszky I (2005) Fuzzy rule-based downscaling of precipitation. *Theoretical and Applied Climatology* 82:119–129. <https://doi.org/10.1007/s00704-004-0121-0>
- Benestad R E, Hanssen-Bauer I, Førland E J (2007) An evaluation of statistical models for downscaling precipitation and their ability to capture long-term trends. *International Journal of Climatology* 27:649–665. <https://doi.org/10.1002/joc.1421>
- Chen H, Guo J, Xiong W, Guo S, Xu CY (2010) Downscaling GCMs using the Smooth Support Vector Machine method to predict daily precipitation in the Hanjiang Basin. *Advances in Atmospheric Sciences* 27:274–284. <https://doi.org/10.1007/s00376-009-8071-1>
- Fistikoglu O, Okkan U (2011) Statistical downscaling of monthly precipitation using NCEP/NCAR reanalysis data for Tahtali river basin in Turkey. *Journal of Hydrologic Engineering* 16:157–164. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000300](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000300)
- Goly A, Teegavarapu RV, Mondal A (2014) Development and evaluation of statistical downscaling models for monthly precipitation. *Earth Interactions* 18:1–28. <https://doi.org/10.1175/EI-D-14-0024.1>
- Goyal MK, Ojha, CP (2012) Downscaling of surface temperature for lake catchment

- in an arid region in India using linear multiple regression and neural networks. *International Journal of Climatology* 32:552–566. <https://doi.org/10.1002/joc.2286>
- Hanssen-Bauer I, Achberger C, Benestad RE, Chen D, Forland EJ (2005) Statistical downscaling of climate scenarios over Scandinavia. *Climate Research* 29:255–268. <https://doi.org/10.3354/cr029255>
- Harpham, C., & Wilby, R. L. (2005). Multi-site downscaling of heavy daily precipitation occurrence and amounts. *Journal of Hydrology*, 312:235–255. <https://doi.org/10.1016/j.jhydrol.2005.02.020>
- Jiang H, Cotton WR (2004) Soil moisture estimation using an artificial neural network: A feasibility study. *Canadian Journal of Remote Sensing* 30:827–839. <https://doi.org/10.5589/m04-041>
- Khan MS, Coulibaly P, Dibike Y (2006) Uncertainty analysis of statistical downscaling methods. *Journal of Hydrology* 319:357–382. <https://doi.org/10.1016/j.jhydrol.2005.06.035>
- Kilsby CG, Jones PD, Burton A, Ford AC, Fowler HJ, Harpham C, James P, Smith A, Wilby RL (2007) A daily weather generator for use in climate change studies. *Environmental Modelling and Software* 22:1705–1719. <https://doi.org/10.1016/j.envsoft.2007.02.005>
- Kusangaya S, Warburton ML, Archer E, Garderen V, Jewitt GW (2014) Impacts of climate change on water resources in southern africa: A review. *Physics and Chemistry of the Earth* 67:47–54.
- Laddimath RS, Patil NS, (2019) Artificial Neural Network Technique for Statistical Downscaling of Global Climate Model. *Mapan - Journal of Metrology Society of*

- India 34:121–127. <https://doi.org/10.1007/s12647-018-00299-0>
- Liu J, Yuan D, Zhang L, Zou X, Song X (2016) Comparison of three statistical downscaling methods and ensemble downscaling method based on Bayesian model averaging in upper Hanjiang River Basin. *Advances in Meteorology*, 1–12. <https://doi.org/10.1155/2016/7463963>
- Nourani V, Razzaghzadeh Z, Baghanam AH, Molajou A (2019) ANN-based statistical downscaling of climatic parameters using decision tree predictor screening method. *Theoretical and Applied Climatology*, 137:1729–1746. <https://doi.org/10.1007/s00704-018-2686-z>
- Okkan U, Kirdemir U (2016) Downscaling of monthly precipitation using CMIP5 climate models operated under RCPs. *Meteorological Applications*, 23: 514–528. <https://doi.org/10.1002/met.1575>
- Salimi AH, Masoompour Samakosh J, Sharifi E, Hassanvand MR, Noori A, von Rautenkranz H (2019) Optimized Artificial Neural Networks-Based Methods for Statistical Downscaling of Gridded Precipitation Data. *Water* 11:1653. <https://doi.org/10.3390/w11081653>
- Schoof JT, Pryor SC, Robeson SM (2007) Downscaling daily maximum and minimum temperatures in the midwestern USA: a hybrid empirical approach. *International Journal of Climatology* 27:439–454. <https://doi.org/10.1002/joc.1412>
- Sennikovs J, Bethers U (2009) Statistical downscaling method of regional climate model results for hydrological modelling. In 18th World IMACS / MODSIM Congress. <http://mssanz.org.au/modsim09>
- Singh VK, Kumar D (2020) Downscaling daily precipitation over the Upper Shivenath basin: A comparison of three statistical downscaling techniques. *International*

- Journal of Current Microbiology and Applied Sciences 9:1676–1688.
<https://doi.org/10.20546/ijcmas.2020.901.185>
- Trzaska Y, Schnarr E (2014) A Review of downscaling methods for climate change projections (USAID, 2014).
- Vu MT, Aribarg T, Supratid S, Raghavan SV, Liong SY (2016) Statistical downscaling rainfall using artificial neural network: significantly wetter Bangkok? Theoretical and Applied Climatology, 126:453–467. <https://doi.org/10.1007/s00704-015-1580-1>
- Wang J, Hu L, Li D, Ren M (2020) Potential Impacts of Projected Climate Change under CMIP5 RCP Scenarios on Streamflow in the Wabash River Basin. Advances in Meteorology. <https://www.hindawi.com/journals/amete/2020/9698423/>
- Wilby R, Charles S, Zorita E, Timbal B, Whetton P, Mearns L (2004) Guidelines for use of climate scenarios developed from statistical downscaling methods. <https://doi.org/10.5281/ZENODO.1438320>
- Wilby RL, Hay LE, Leavesley GH (1999) A comparison of downscaled and raw GCM output: Implications for climate change scenarios in the San Juan River Basin, Colorado. Journal of Hydrology, 225:67–91. [https://doi.org/10.1016/S0022-1694\(99\)00136-5](https://doi.org/10.1016/S0022-1694(99)00136-5)
- Wilby RL, Wigley TL (2000) Precipitation predictors for downscaling: Observed and general circulation model relationships. International Journal Of Climatology. A Journal of the Royal Meteorological Society, 20:641-661.
- Wilby RL, Wigley TL, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: A comparison of

methods. Water Resources Research, 34:2995–3008.
<https://doi.org/10.1029/98WR02577>

Xu CY (1999) Climate change and hydrologic models: A review of existing gaps and recent research developments. Water Resources Management, 13:369–382.
<https://doi.org/10.1023/A:1008190900459>

Statements & Declarations

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. The authors have no relevant financial or non-financial interests to disclose.

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Khaled Alramlawi, Okan fistikoglu, Mohamed Najar. The first draft of the manuscript was written by khaled Alramlawi and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.