

SPATIAL DOWNSCALING OF GPM IMERG V06 GRIDDED PRECIPITATION USING MACHINE LEARNING ALGORITHMS

Manikandan Sathianarayanan¹, Pai-Hui Hsu² *

¹ Department of Civil Engineering, National Taiwan University, Taiwan-10617-d06521026@ntu.edu.tw,

² Department of Civil Engineering, National Taiwan University, Taiwan-10617-hsuph@ntu.edu.tw.

Commission IV, WG 7

Abstract

According to recent studies, Remote sensing data plays a significant role in filling gaps in the poor gauge station, particularly at high elevations and with complex underlying surface features. In order to provide high-resolution precipitation estimates over the poor gauge and with complex terrain areas, downscaling low-resolution satellite precipitation estimates using various environmental variables. In this paper, we tried to downscale the GPM IMERG V06 with a resolution of $(0.1^\circ \times 0.1^\circ)$ nearly 10km to $(1\text{km} \times 1\text{ km})$ using four machine learning algorithms namely, Decision Trees, Multiple Linear Regression, Support Vector Regressor and random forest. Vegetation indices Normalized difference in vegetation index (NDVI), Topography, Land Surface Temperature (LST), and latitude and longitude. This framework can downscale the 0.1° resolution of the GPM IMERG precipitation product to 1 km, by determining the importance of features, and automatically optimizing the model parameters. Additionally, ground recorded data from rain gauge stations have validated downscaled precipitation products. Spatial downscaling can generally increase the accuracy of GPM IMERG gridded precipitation data and results reveal that spatial downscaling is an acceptable way of investigating the precipitation over Taiwan.

KEY WORDS: Spatial downscaling, Machine learning algorithms, Satellite Precipitation estimation.

1. INTRODUCTION

The global hydrological cycle cannot function without precipitation, which is crucial for maintaining the hydro-climate balance and ecosystem activities. Even while the data from rain gauges can yield precise point-based measurements, it is difficult to extrapolate the data to create precise maps at the basin scale, particularly when the distribution of rain gauges is uneven or for ungauged basins (Luo et al., 2019; Zhang et al., 2019) and rain gauges serves as an evaluation tool for various precipitation products. Over the last few decades, rapid developments in remote sensing techniques provide an opportunity to estimate spatial continuous precipitation on a global scale. Satellite-based precipitation products (SPPs) are openly available to the public to understand the precipitation characteristics for hydrometeorological applications, especially over sparse rain gauge areas (Belabid et al., 2019; Pellet et al., 2019). Various SPPs are available in recent decades namely, Tropical Rainfall Measuring Mission (TRMM) Multi-satellite precipitation analysis (TMPA), Climate Prediction Center Morphing technique (CMORPH), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN), and Integrated Multi-satellitE Retrieval for Global Precipitation Measurement (IMERG). However, these SPPs are relatively low resolution varying from 0.1° to 0.5° , and these products are too coarse for analysis. An efficient strategy required to close the spatial scale gap between low/coarse resolution and high/fine resolution is to use spatial downscaling techniques. There are two major downscaling techniques available for SPPs, statistical downscaling (construct the empirical relationship between object and auxiliary variable) and dynamical downscaling (mathematical representation of the complex physical phenomenon of atmosphere, ocean, and land) (Sachindra & Perera, 2016). A critical step in the downscaling method is to select appropriate environmental variables and methods for downscaling precipitation, variables can be divided

between dynamical (variable can change over spatially and temporally) and static variables (variables remain constant). Numerous downscaling models has been developed recent decades using Univariate Regression (UR), multivariate regression (MR), and Geographic Weighted Regression (GWR) and it fails to reflect spatial heterogeneity between precipitation and land surface characteristics. In this study we employed non-parametric machine learning algorithms (Decision Trees, Multiple Linear Regression, Support Vector Regressor and random forest) to downscale the IMERG satellite estimation using various environmental variables.

2. STUDY AREA & DATASET

2.1 Study Area

The study area located in East Asia, on the western side of the Pacific Ocean and it consists of main land of Taiwan as well as small distant islands. The Pacific Ocean, Bashi channel, Taiwan Strait and East China sea are all located to east, south, west and north of Taiwan mainland respectively. Study area stretching from 120E to 122E and from 22N to 25N that covers 36197 km² which mainland dominated by mountain ranges in East and gently sloping plains in the West, approximately 394km long from North to South and 144 km broad from West to East.

2.2 Datasets

This study utilized satellite-based precipitation estimates (GPM IMERG V06), vegetation indices (NDVI, EVI), digital elevation model (SRTM), landcover (MCD12Q1), and Land Surface Temperature (LST) with their spatial and temporal resolution characteristics are given below in the Table.1.

2.2.1 Integrated Multi-satellitE Retrievals for GPM (IMERG)

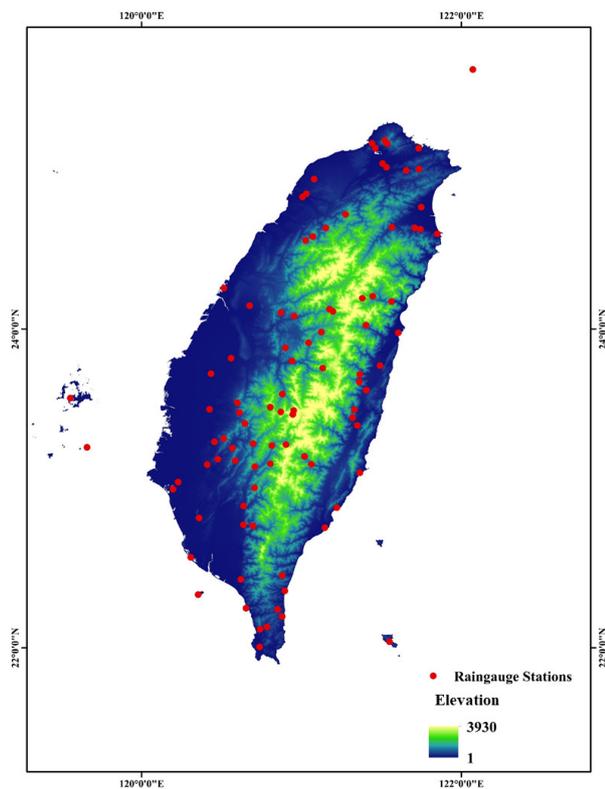


Figure 1. Study area with Elevation and Rain gauge station locations

Data	Spatial & Temporal Resolution
IMERG	0.1°×0.1°/daily
NDVI/MOD13A2	1 km/16 days
DEM/SRTM	90 meters
Land surface Temperature/MOD11A1	1 km/daily
Landcover/MCD12Q1	500 meter/ yearly

Table 1. Dataset required for this study

GPM mission, an international constellation of satellites consisting of one major observatory satellite and ten partner satellites that offers the next-generation global precipitation measurement, which was started by Japan Aerospace Exploration Agency (JAXA) and United States National Aeronautics and Space Administration (NASA) (Lu et al., 2018). IMERG has a multi-channel GPM Microwave Imager (GMI) that is combined with first space-borne Ku (13GHz) and Ka (35GHz) bands which can able to detect light rain (<0.5 mm/hr) (Hou et al., 2014).

Final Run IMERG products were utilised in this study out of other two products (early run, late run) as suggested by data provider and it can be downloaded approximately four months after real time observation from (<https://gpm.nasa.gov/data/directory>). Guo et al. (2016) suggested that PrecipitationCal performs better than PrecipitationUnCal over Taiwan since it has very less bias with Rain gauge station observations.

2.2.2 MODIS Land Surface Temperature (LST)

MOD11A1 version 6 daily product from Terra MODIS-LST product were used in this study at 1km spatial resolution. Daily

LST data can be downloaded from USGS website for about one complete year from 1 st January 2018 to 31 st December 2018 (<https://e4ftl01.cr.usgs.gov/MOLT/MOD11A1.006/>). Annual mean LST data were calculated by averaging daily LST data in a given month/year.

2.2.3. NDVI and Landcover

In this study, 16-day NDVI product (MOD13A2) from Moderate Resolution Imaging Spectroradiometer (MODIS) aboard Terra sensor with a spatial resolution of 1km by 1km which can be downloaded from NASA Land Processes Distributed Active Center (<https://lpdaac.usgs.gov/products/mod13a2v006/>). Annual mean NDVI were calculated by averaging 16-day NDVI data in a given year. Anomalous pixels (snow cover, urban areas and water bodies) should be removed from NDVI pixels which was influenced by vegetation growth (Wang et al., 2019). Anomalous pixels were removed from original NDVI data based on Land cover data (MCD12Q1) from MODIS onboard Terra sensor at 500m by 500m spatial resolution.

2.2.4 Elevation and Geographic locations

Digital elevation model with a spatial resolution of 90m from the Shuttle Radar Topography Mission provided the elevation data which was used in this study and it can be downloaded from USGS website ([EarthExplorer \(usgs.gov\)](https://earthexplorer.usgs.gov/)). 90 m elevation data were resampled to 1km and 10 km spatial resolution by using pixel averaging methods. Latitude and Longitude data utilised in this study were derived using elevation data as well and it serves as a common feature to combine all data for further processing and analysis.

3. METHODOLOGY

3.1 Machine learning algorithms

In this study, we used Four machine learning algorithms of the scikit-learn python (Pedregosa et al., 2011) that contains Random Forest, Decision Tree, Support Vector Machine and Multiple Linear Regression, along with other model called Adaptive Network-based Fuzzy Inference System (ANFIS) utilised to model the complex relationship between IMERG precipitation and environmental parameters for downscaling the precipitation products. In order to categorize the input variables into an m-dimensional feature space with a maximal margin, which may be determined by solving a quadratic problems depends upon on an optimization theory that employs a hyperplane (Smola & Schölkopf, 2004). Most widely used SVM tools called libsvm that developed by (Chang & Lin, 2011) and it is available freely in online that is being adopted for this study. libsvm includes all important parameters such as kernel function, capacity parameter cost. Random Forest is an improved technique which is integrating a large number of Classification and Regression Tree (CART) methods in to an ensemble. It has been widely utilized for numerous remote sensing applications, including regression and classification (Rhee et al., 2014). RF employs a bootstrap aggregating method in contrast to CART algorithm to enhance model performance. The aggregate output from numerous trees may smooth the variance between trees and produce more accurate prediction results since each tree is constructed using a random subset of training data and a random subset of predictor variables (Breiman, 2001). According to Zhang et al. (2017), conventional method of Multiple Linear Regression

relies on regression coefficient matrix that may depict the correlation between dependent and independent variables. Multiple Linear Regression has been used frequently to predict the values of dependent variable from a collection of predictor variables. In this study. We tried to build a Multiple Linear Regression function between satellite precipitation products and the numerous environmental variables.

3.2 Adaptive Network-Fuzzy Inference System

Zadeh (1965) made first attempt to prepare fuzzy set approach which is basically a combination of neural network (NN) and the concept of fuzzy logic (FL). Sugeno's system is one of the most widely used in fuzzy systems to create model with scarce or ambiguous data. Fuzzy Logic has a stronger capacity for condition adjustment during learning process. Thus, by using NN, it is possible to reduce the error rate according to FL rules and have membership function (MFs) vary naturally. ANFIS model developments has two main components, (i) during learning process, membership function (MF) is used to convert input values to fuzzy values from 0 to 1 and, (ii) using some IF-THEN rules in ANFIS model to describe the non-linear relationship between input and output space.

Membership function is the primary factor influencing the ANFIS model's predictability performance, hence choosing the right MF is one of the predefined modelling steps. Fig.2 shows

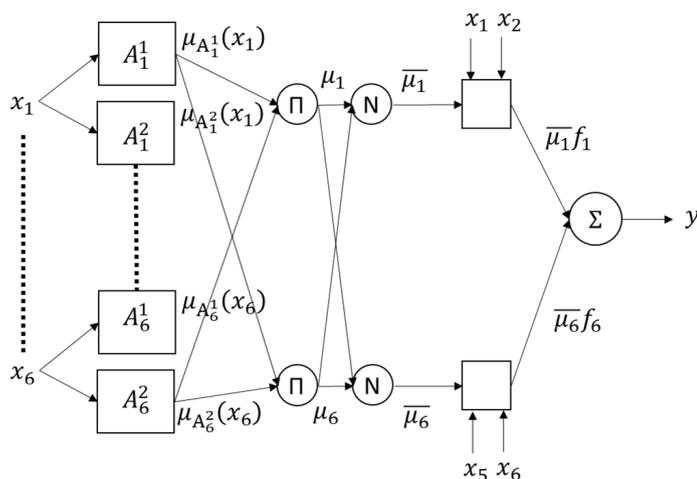


Figure 2. Architecture of ANFIS model equivalent to the inference system.

about structure of ANFIS model with 6 input variables(x_1 to x_6) and five different layers: fuzzification, product, normalization, de-fuzzification and output(Moosavi et al., 2013). The most common MFs are triangular, trapezoidal, generalized bell-shaped and Gaussian, Gaussian MF is used in this study. Jang et al. (1997) developed a hybrid teaching for the neuro-fuzzy model to calculate the model parameters more quickly and accurately than back-propagation method which is based on gradient reduction.

3.3 Downscaling

The main goal of statistical spatial downscaling is to establish, an empirical statistical relationship between the object variable and the relevant auxiliary variables to low/coarse spatial resolution, High/fine spatial resolution is thought to be appropriate to the empirical statistical relationship. The downscaled object variable at fine spatial resolution is then produced by applying the established empirical statistical relationship to the auxiliary variables at fine spatial resolution.

The detailed flow of work involved in downscaling satellite precipitation in the Fig.3. In our study IMERG annual satellite precipitation data serve as an object variable that has to be downscaled from 0.1° (nearly 10 km) resolution to high resolution of 1km. six auxiliary variables are NDVI, Latitude, Longitude, Elevation, LST, Landcover selected from the past literatures. Initially monthly based auxiliary variables have been prepared with a spatial resolution of 10 km and 1km and then it converts in to yearly composition by averaging the monthly values and auxiliary variables have resampled from 1 km and 10km by the method of nearest neighbour method. Waterbody and urban built-up areas to be masked out from vegetation indices and landcover dataset due to negative impact on the downscaling model. To get rid of the impacts of different scale, standardize the variables by use of their means and standard deviation at scikit-learn library. Non-parametric machine learning regression models (SVR, Decision Trees, MLR, Random forest and ANFIS) has created between IMERG precipitation and six variables with 10km resolution. Residual errors can be calculated from predicted 10km precipitation with IMERG precipitation products. Resampled the residual error from 10 km to 1 km using spline interpolation and findings that have been downscaled using residual correction and are 1 km resolution might be attained by combining the residual errors from the interpolated the results after scaling without residual rectification.

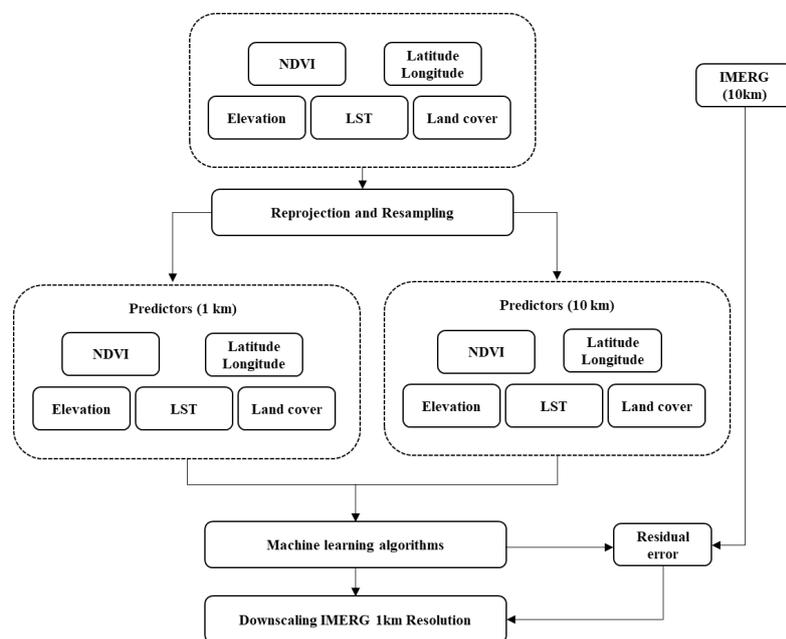


Figure 3. The process involved in Spatial downscaling of IMERG precipitation products.

3.3.1 Data Preparation and Pre-Processing

In this study, six environmental variables that include three static variables (elevation, longitude and latitude) and three dynamical variables (either NDVI or EVI, LST and Land cover) with a spatial resolution of 10km and 1km are created. Initially annual composite values were calculated for all the environmental variables from averaging monthly values. By employing the nearest neighbour method to resample the environmental variables at 10km and 1km spatial resolution. To prevent a detrimental impact on the downscaling model, Negative values in the vegetation indices as well as urban, built-

up areas and water bodies should be eliminated through Landcover dataset.

3.3.2 Data Normalization

All input environmental variables are normalized and scaled in the same order, in order to improve the evaluation accuracy and speed when working with raw input data and it has been collected from different sources. We adopted min-max normalization in this study to linearly rescale all the environmental variables in to (0 to 1) interval.

$$x_n = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

Where x_i indicates that normalized data and x_{min} and x_{max} are minimum and maximum value of that data.

3.3.3 Hyper-Parameter Optimization

Hyper-Parameter optimization plays an important role in machine learning algorithms to improve its performance. We used scikit-learn GridSearchCV algorithm with 10-fold cross-validation to identify the best hyper-parameters to construct optimal prediction model (Table.2).

Downscaling algorithms	Parameters	Parameter set
SVM	Kernel Type	Radial basis function
	Cost	20
	Gamma	0.0003
	Epsilon	0.008
Random Forest	Number of estimators	100
	Max depth	20
	Min sample leaf	1
	Min sample split	2
Decision Tree	Max depth	10
	Min sample leaf	5
ANFIS	Epoch	100
	Membership function	Gaussian
	FIS type	sugeno

Table.2 Parameters setting for machine learning methods-based downscaling

3.3.4 Statistical indices

Two assessment indices were adopted for this study to compare the performances of downscaling machine learning models that includes correlation of determination(R^2) and mean squared error(MSE)(Shi et al., 2015).

$$R^2 = \frac{\sum_{i=1}^n [(S_i - \bar{S})(P_i - \bar{P})]}{\sqrt{\sum_{i=1}^n (S_i - \bar{S})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (P_i - \bar{P}_i)^2$$

Where, S represents original IMERG and P predicted IMERG precipitation.

4. RESULT AND DISCUSSION

4.1 Downscaling model evaluation

Table 3 shows about the results of various downscaling models that adopted for this study to simulate IMERG precipitation data at 1km spatial resolution. Table.3 provides the information about annual IMERG precipitation data predicted by use of five machine learning models with different input variables combinations and shows better consistency with original IMERG data. Validation results revealed that ANFIS model performed well with highest R^2 (0.9736) with combination of all input variables and MSE value of 0.05244 mm/year followed by Random forest, SVR, Multiple Linear Regression and Decision Tree. Hence, ANFIS model was used for the downscaling analysis of IMERG precipitation data for the year 2018.

Predictors	Downscaling methods	R-Square	MSE
LC+ NDVI+DEM	Random Forest	0.147779	0.379764
	SVR	0.058532	0.419544
	Decision Tree	0.2180377	0.34848
	Multiple Linear Regression	0.10983	11.9502
	ANFIS	0.2881	0.1701998
NDVI+DEM	Random Forest	0.1477797	0.3797892
	SVR	0.058532	0.4194
	Decision Tree	0.218037	0.3456
	Multiple Linear Regression	0.109836	11.9502
	ANFIS	0.27	0.2295648
NDVI+DEM+LC+ Longitude + Latitude	Random Forest	0.931656	0.0304571
	SVR	0.61831	0.170064
	Decision Tree	0.5774924	0.18828
	Multiple Linear Regression	0.61984	7.8084
	ANFIS	0.9584	0.2382057
NDVI+DEM+LC+LST+ Longitude + Latitude	Random Forest	0.89956	0.0447588
	SVR	0.650343	0.15552
	Decision Tree	0.38498	0.274068
	Multiple Linear Regression	0.60618	7.948692
	ANFIS	0.9736	0.0524488

Table.3 Downscaling model validation results for year 2018

4.2 Feature Importance

As ANFIS model has been selected for further downscaling analysis of IMERG precipitation and it is hard to identify the rate of influencing individual variable on prediction result in conventional regression model. However this can accomplish by analysing SHapely Additive exPlanations(SHAP) values. According to game theory-based presumptions, SHAP enables evaluation of the extent and nature of each explanatory variables impact on the outcome of any machine learning or deep learning model(Lundberg & Lee, 2017).Figure.4 depicts the summary plot of SHAP values on downscaling model which can explain the connection between input variables and prediction results. Positive or negative SHAP values indicates the positive or negative correlation and the higher values tends to affect the accuracy of prediction. Fig.4 shows that higher

latitude and longitude affects the prediction result negatively, while lower values affect it favourably.

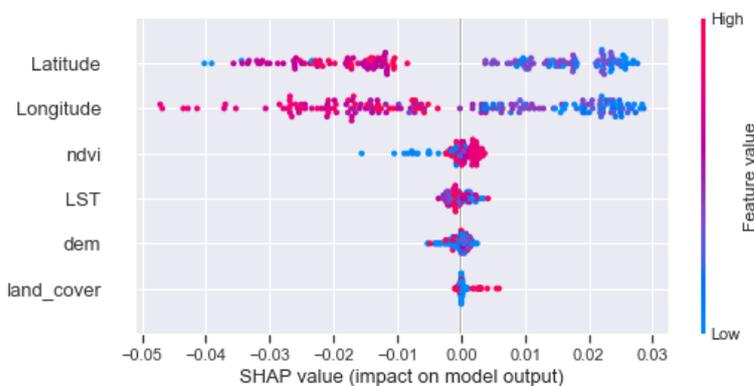


Fig.4 Summary plot of the SHAP values on downscaling model

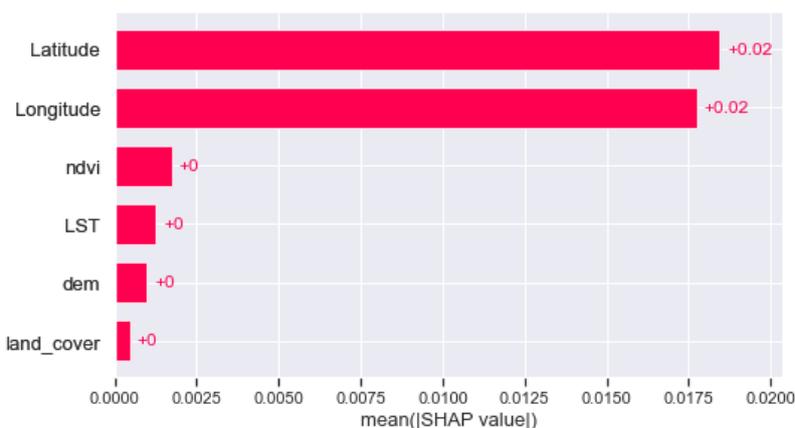


Fig.5 Feature importance plot of mean SHAP values on downscaling model

Fig.5 display the feature importance plot of the mean SHAP values on the ANFIS downscaling model, bar graph has been generated by averaging absolute value of SHAP value of each feature. Latitude and longitude have a most influencing factor followed by NDVI, LST and elevation, while land cover has least one in the list that impact the downscaling model.

Figure.6 explains about waterfall plot of predicted results of features with largest contribution in SHAP value. X axis represents value of target variable (ie., Precipitation) and $E[f(x)]$ provides the expected value of precipitation prediction. Longitude and latitude have the highest impact followed NDVI and LST and landcover was least one.

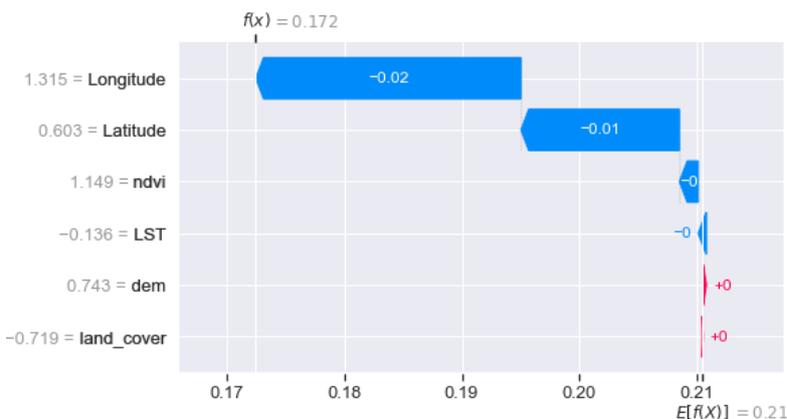


Fig.6 Waterfall plot for predicted results

4.3 Annual downscaled precipitation products

Figure 7 represents the original IMERG precipitation products and the predicted precipitation based on six input variables using ANFIS machine learning algorithm. Predicted results revealed that spatial pattern of downscaled IMERG as same as the original IMERG precipitation products. Annual mean of downscaled products are very close to original IMERG while slight difference in spatial distribution. Results indirectly proves that ANFIS algorithm performs well for downscaling the satellite precipitation.

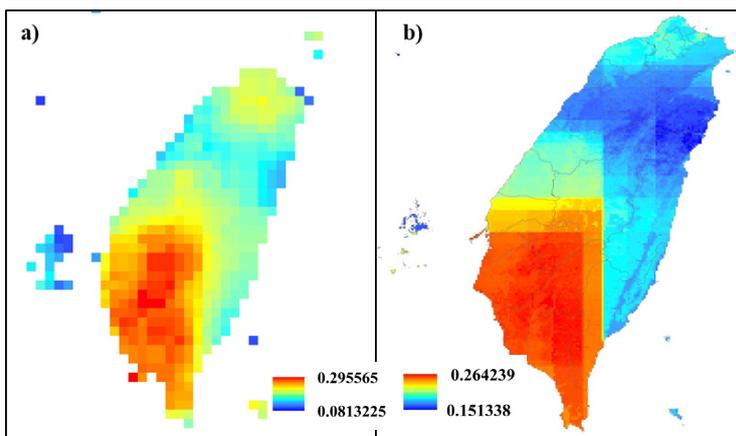


Fig.7 Annual mean original IMERG precipitation product (a), downscaled annual mean IMERG precipitation (b).

5. CONCLUSION

The main objective of this framework is to correct way for downscaling the IMERG precipitation products along with various static and dynamical variables (Elevation, Land cover, LST, NDVI, latitude and longitude) using suitable machine learning algorithms (SVR, MLR, Decision Tree, Random forest and ANFIS). The machine learning based algorithms provides smallest residual errors than conventional parametric models on downscaling with sufficient spatial details for meteorological analysis. Our results revealed that Adaptive Network-based Fuzzy Inference system (ANFIS) regression algorithm performed better and significant statistical metrics in downscaling the precipitation. Future studies will involve validation of downscaled precipitation against rain gauge observation values.

REFERENCES

Belabid, N., Zhao, F., Brocca, L., Huang, Y., & Tan, Y. (2019). Near-Real-Time Flood Forecasting Based on Satellite Precipitation Products. *Remote Sensing*, 11(3), 252. <https://www.mdpi.com/2072-4292/11/3/252>

Breiman, L. (2001). Random Forests. *Machine learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3), Article 27. <https://doi.org/10.1145/1961189.1961199>

- Guo, H., Chen, S., Bao, A., Behrangi, A., Hong, Y., Ndayisaba, F., Hu, J., & Stepanian, P. M. (2016). Early assessment of integrated multi-satellite retrievals for global precipitation measurement over China. *Atmospheric Research*, 176, 121-133.
- Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., & Iguchi, T. (2014). The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society*, 95(5), 701-722. <https://doi.org/10.1175/bams-d-13-00164.1>
- Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). Neuro-Fuzzy and Soft Computing-A Computational Approach to Learning and Machine Intelligence [Book Review]. *IEEE Transactions on Automatic Control*, 42(10), 1482-1484. <https://doi.org/10.1109/TAC.1997.633847>
- Lu, X., Wei, M., Tang, G., & Zhang, Y. (2018). Evaluation and correction of the TRMM 3B43V7 and GPM 3IMERGM satellite precipitation products by use of ground-based data over Xinjiang, China. *Environmental Earth Sciences*, 77(5), 1-18.
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.
- Luo, X., Wu, W., He, D., Li, Y., & Ji, X. (2019). Hydrological Simulation Using TRMM and CHIRPS Precipitation Estimates in the Lower Lancang-Mekong River Basin. *Chinese Geographical Science*, 29(1), 13-25. <https://doi.org/10.1007/s11769-019-1014-6>
- Moosavi, V., Vafakhah, M., Shirmohammadi, B., & Behnia, N. (2013). A Wavelet-ANFIS Hybrid Model for Groundwater Level Forecasting for Different Prediction Periods. *Water Resources Management*, 27(5), 1301-1321. <https://doi.org/10.1007/s11269-012-0239-2>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- Pellet, V., Aires, F., Munier, S., Fernández Prieto, D., Jordá, G., Dorigo, W. A., Polcher, J., & Brocca, L. (2019). Integrating multiple satellite observations into a coherent dataset to monitor the full water cycle – application to the Mediterranean region. *Hydrol. Earth Syst. Sci.*, 23(1), 465-491. <https://doi.org/10.5194/hess-23-465-2019>
- Rhee, J., Park, S., & Lu, Z. (2014). Relationship between land cover patterns and surface temperature in urban areas. *GIScience & Remote Sensing*, 51(5), 521-536. <https://doi.org/10.1080/15481603.2014.964455>
- Sachindra, D. A., & Perera, B. J. C. (2016). Statistical Downscaling of General Circulation Model Outputs to Precipitation Accounting for Non-Stationarities in Predictor-Predictand Relationships. *Plos One*, 11(12), e0168701. <https://doi.org/10.1371/journal.pone.0168701>
- Shi, Y., Song, L., Xia, Z., Lin, Y., Myneni, R. B., Choi, S., Wang, L., Ni, X., Lao, C., & Yang, F. (2015). Mapping Annual Precipitation across Mainland China in the Period 2001–2010 from TRMM3B43 Product Using Spatial Downscaling Approach. *Remote Sensing*, 7(5), 5849-5878. <https://www.mdpi.com/2072-4292/7/5/5849>
- Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199-222. <https://doi.org/10.1023/B:STCO.0000035301.49549.8>
- Wang, L., Chen, R., Han, C., Yang, Y., Liu, J., Liu, Z., Wang, X., Liu, G., & Guo, S. (2019). An improved spatial-temporal downscaling method for TRMM precipitation datasets in Alpine regions: A case study in northwestern China's Qilian Mountains. *Remote Sensing*, 11(7), 870.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353. [https://doi.org/https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/https://doi.org/10.1016/S0019-9958(65)90241-X)
- Zhang, H., Wu, P., Yin, A., Yang, X., Zhang, M., & Gao, C. (2017). Prediction of soil organic carbon in an intensively managed reclamation zone of eastern China: A comparison of multiple linear regressions and the random forest model. *Science of The Total Environment*, 592, 704-713.
- Zhang, Z., Tian, J., Huang, Y., Chen, X., Chen, S., & Duan, Z. (2019). Hydrologic Evaluation of TRMM and GPM IMERG Satellite-Based Precipitation in a Humid Basin of China. *Remote Sensing*, 11(4), 431. <https://www.mdpi.com/2072-4292/11/4/431>