# IMPROVING THE ACCURACY OF SATELLITE-BASED NEAR SURFACE AIR TEMPERATURE AND PRECIPITATION PRODUCTS

Ç.H. Karaman [a, *], Z. Akyurek [a,b]

[a] METU, Geodetic and Geographic Information Technologies, Natural and Applied Sciences, 06800 Ankara, Turkey
[b] METU, Civil Eng. Dept., 06800 Ankara, Turkey

cagri.karaman@metu.edu.tr
zakyurek@metu.edu.tr

**KEY WORDS:** Near-Surface Air Temperature, Precipitation, Satellite, Downscaling, Machine Learning

**ABSTRACT:**

In this study, we evaluate the performance of several reanalyses and satellite-based products of near-surface air temperature and precipitation to determine the best product in estimating daily and monthly variables across the complex terrain of Turkey. Each product's performance was evaluated using 1120 ground-based gauge stations from 2015 to 2019, covering a range of complex topography with different climate classes according to the Köppen-Geiger classification scheme and land surface types according to the Moderate Resolution Imaging Spectroradiometer (MODIS). Furthermore, various traditional and more advanced machine learning downscaling algorithms were applied to improve the spatial resolution of the products. We used distance-based interpolation, classical Random Forest, and more innovative Random Forest Spatial Interpolation (RFSI) algorithms. We also investigated several satellite-based covariates as a proxy to downscale the precipitation and near-surface air temperature, including MODIS Land Surface Temperature, Vegetation Index (NDVI and EVI), Cloud Properties (Cloud Optical Properties, Cloud Effective Radius, Cloud Water Path), and topography-related features. The agreement between the ground observations and the different products, as well as the downscaled temperature products, was examined using a range of commonly employed measures. The results showed that AgERA5 was the best-performing product for air temperature estimation, while MSWEP V2.2 was superior for precipitation estimation. Spatial downscaling using bicubic interpolation improved air temperature product performance, and the Random Forest (RF) machine learning algorithm outperformed all other methods in certain seasons. The study suggests that combining ground-based measurements, precipitation products, and features related to topography can substantially improve the representation of spatiotemporal precipitation distribution in data-scarce regions.

## 1. INTRODUCTION

Accurate depiction of the near-surface air temperature and precipitation in both time and space is crucial for various fields, including hydrometeorology and climatology, as they are critical components in the water cycle. Such accurate representation is necessary for predicting extreme weather events, monitoring droughts, flood forecasting, in shaping the ecology of the different regions and ecosystems, in remote sensing, in hydrological and environmental modelling, in understanding the impacts of the climate change, and for agriculture (Arismendi et al., 2014; Arnaud et al., 2011; Colombi et al., 2007; Graae et al., 2012; Liou & Kar, 2014; Mu et al., 2021; Terando et al., 2018). Precise assessments of spatial and temporal variations in air temperature and precipitation across the Earth's land surface are vital for comprehending the climate and the mechanisms that support life on land (Prihodko & Goward, 1997). These variables have been obtained through ground-based measuring instruments with a long history of providing in-person observations (Timmermans et al., 2019).

Ground-based stations offer precise measurements of these variables, but only over a limited region and time frame. When these point measurements are extended or estimated to cover larger areas, a significant amount of uncertainty is introduced. Interpolation methods can result in substantial spatial sampling errors, particularly in areas with few ground stations and uneven distribution (Elgamal et al., 2017).

Furthermore, hydro-climatological research is hampered in many semi-arid and arid regions due to the rough terrain, harsh climate conditions, and absence of monitoring stations. Several factors, such as limited observation period, inadequate spatial coverage, missing data, and biases from various sources, can impact the accuracy of measurements for these variables. These limitations hinder accurately depicting the temporal and spatial variability of hydrological processes at the watershed scale (Chen & Brissette, 2017).

State Meteorological Service (TSMS) is responsible institution for measuring climate variables throughout Turkey. However, considering Turkey's rough and vast topography, the data required for hydrological analysis are only available in some locations. Some spatial interpolation methods and statistical analysis exist to predict climate data. Several methods are available to interpolate point measurements of climate data over larger areas, such that Geographically Weighted Regression (GWR) (Brunsdon et al., 2010), Kriging interpolation (Hengl et al., 2007), Inverse Distance Weighting (IDW) (Shepard D, 1968). However, these methods introduce a considerable amount of uncertainty. Therefore, continuous and spatially distributed climate data are needed.

* Corresponding author. Çağrı Karaman

Satellites can measure various environmental factors using remote sensing techniques, including temperature, moisture, cloud cover, precipitation, vegetation, and soil indices. These measurements can obtain information about the Earth's surface and environmental conditions (Condom et al., 2020).

There are mainly three data sources spatially varying climate data (precipitation, temperature, etc.):

• Satellite-based remote sensing data
• Gridded data of point measurements by spatial interpolation methods
• Blending ground observations with satellite-based products

There is the lack of accurate and downscaled remote sensed climate variables (near-surface temperature and precipitation) datasets in areas having complex topography, like Turkey. The aims of this study can be listed as follows:
- To assess the effectiveness and accuracy of satellite-based products (near-surface air temperature, precipitation) in Turkey in diverse climate using ground-based stations.
- To improve the spatial resolution of the selected products representing the fine-scale variability of temperature and precipitation distribution in Turkey through spatial downscaling to a specific grid utilizing various interpolation techniques.
- To implement algorithms for blending ground observations with satellite-based precipitation products.

## 2. MATERIALS AND METHODS

### 2.1 Ground-based Observation Data and Datasets

In this study, ground-based observations from the Turkish State Meteorological Service (TSMS) are utilized. There are two types of meteorological stations: Automated Weather Observing Systems (AWOS) and pluviometer-type stations. AWOS stations provide hourly data, while pluviometer-type stations record data three times a day. To ensure data quality control, collocation between AWOS stations and pluviometer stations is implemented (Derin & Yilmaz, 2014). The historical data used in this study encompasses daily mean air temperature and daily total precipitation measurements taken at 1120 meteorological stations from January 2010 to December 2019. The locations of these stations, depicted in Figure 1, are displayed on a Digital Elevation Map (DEM) with a spatial resolution of 30 meters, which can be retrieved from https://earthdata.nasa.gov/ (Farr et al., 2007). Turkey's average altitude is 1130 meters, and it rises progressively from the central part of Anatolia towards the east. Despite Turkey's primary location in the Mediterranean region, its diverse topography, including parallel coastal mountains, leads to a range of climatic conditions.

In this study, the stations are grouped into classes based on the Köppen-Geiger climate classification system developed by German botanist-climatologist Wladimir Köppen that takes into account local vegetation and seasonality of monthly air temperature and precipitation (Köppen, Wladimir; Geiger, 1936). Climate classes of stations and their mean elevations are presented in Table 1.

In the context of this study, we utilized the MODIS MCD12Q1 Version 6 data product for land cover classification, which employs land cover classes derived from supervised classifications of MODIS Terra and Aqua reflectance data. This data product provides global land cover types yearly from 2001 to 2019 with a spatial resolution of 500 meters. It comprises five land cover classification schemes, with the primary scheme identifying 17 land cover classes (Friedl et al., 2010) (Figure 2).

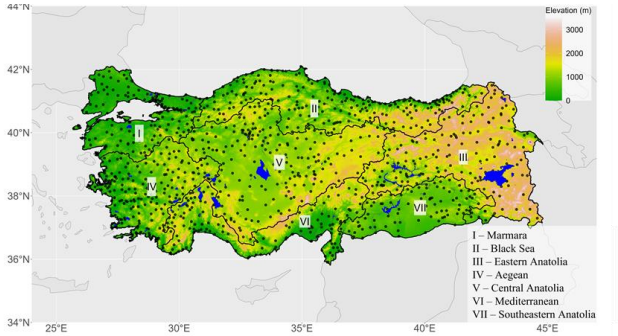


**Figure 1.** Distribution of Meteorological Stations over Digital Elevation Model (DEM) in Turkey

| Symbol | Description | No. of Stations | Mean Elevation (m) |
|---|---|---|---|
| BSh | Arid steppe hot | 2 | 370 |
| BSk | Arid steppe cold | 268 | 1003 |
| BWk | Arid desert cold | 3 | 843 |
| Cfa | Temperate no dry season hot summer | 92 | 145 |
| Cfb | Temperate no dry season warm summer | 32 | 442 |
| Csa | Temperate dry summer hot summer | 476 | 362 |
| Csb | Temperate dry summer warm summer | 27 | 896 |
| Dfb | Cold no dry season warm summer | 48 | 1452 |
| Dfc | Cold no dry season cold summer | 1 | 1114 |
| Dsa | Cold dry summer hot summer | 71 | 1266 |
| Dsb | Cold dry summer warm summer | 98 | 1531 |
| Dsc | Cold dry summer cold summer | 1 | 2052 |

**Table 1.** Climate classes of stations and their mean elevations

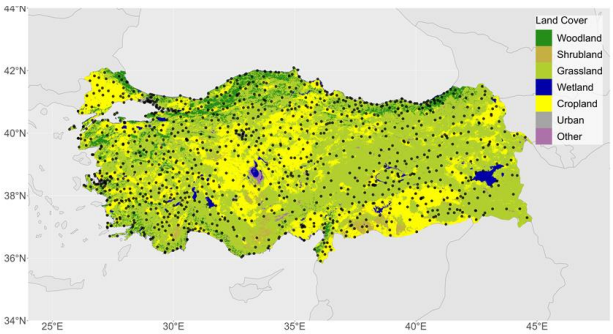


**Figure 2.** Distribution of Meteorological Stations over land cover in Turkey

The near-surface air temperature products used in this study are given in Table 2 and the precipitation products are given in Table 3. All products are arithmetically averaged into daily and monthly time series. Then, performance analysis is conducted on a daily and monthly time scale separately.

| | Grid Size (Lat x Lon) | Temporal Resolution | Data format | Time Period | Type | Reference |
|---|---|---|---|---|---|---|
| Era5 | $0.25^0$ x $0.25^0$ | Hourly | Netcdf | 2015-2020 | Reanalysis Observational | (Hersbach et al., 2019) |
| Era5 Land | $0.1^0$ x $0.1^0$ | Hourly | Netcdf | 2015-2020 | Reanalysis Observational | (Hersbach et al., 2019) |
| AgEra5 | $0.1^0$ x $0.1^0$ | Daily | Netcdf | 2015-2020 | Reanalysis Observational | (Hersbach et al., 2019) |
| Merra2 | $0.5^0$ x $0.625^0$ | Daily | H5 | 2015-2020 | Reanalysis | (Gelaro et al., 2017; Rienecker et al., 2011) |
| JRA-55 | $1.5^0$ x $1.5^0$ | 6 Hourly | Grib | 2015-2020 | Reanalysis | (Kobayashi et al., 2015) |

**Table 2.** Characteristics of the near-surface temperature products

| Product | Grid Size (Lat x Lon) | Temporal Resolution | Type | Latency | Reference |
|---|---|---|---|---|---|
| **ERA5** | $0.25^0$x$0.25^0$ | Hourly | Reanalysis | 5 days | (Hersbach et al., 2019b) |
| **ERA5 Land** | $0.1^0$x$0.1^0$ | Hourly | Reanalysis | 10 days | (Hersbach et al., 2019b) |
| **IMERG V06 Early** | $0.1^0$x$0.1^0$ | Hourly | Satellite | 4 hr | (Hou et al., 2014; Huffman et al., 2007) |
| **IMERG V06 Late** | $0.1^0$x$0.1^0$ | Hourly | Satellite | 12 hr | (Hou et al., 2014; Huffman et al., 2007) |
| **IMERG V06 Final** | $0.1^0$x$0.1^0$ | Hourly | Satellite Observational | 3 months | (Hou et al., 2014; Huffman et al., 2007) |
| **TMPA 3B42** | $0.25^0$x$0.25^0$ | 3 Hourly | Satellite | 10 days | (Huffman et al., 2007) |
| **TMPA 3B42RT** | $0.25^0$x$0.25^0$ | 3 Hourly | Satellite | 7 hrs | (Huffman et al., 2007) |
| **CHIRPS 2.0** | $0.05^0$x$0.05^0$ | Daily | Satellite Observational | 3 weeks | (Funk et al., 2015) |
| **CMORPH CDR V1.0** | $0.08^0$x$0.08^0$ | Half Hourly | Satellite, Observational | 3 months | (Joyce et al., 2004b) |
| **PERSIANN** | $0.25^0$x$0.25^0$ | Hourly | Satellite | 2 days | (Nguyen et al., 2019) |
| **PERSIANN CCS** | $0.04^0$x$0.04^0$ | Hourly | Satellite | 1 hr | (Khan et al., 2011) |
| **PERSIANN CDR** | $0.25^0$x$0.25^0$ | Daily | Satellite Observational | 3 months | (Ashouri et al., 2015) |
| **MSWEP V2.2** | $0.1^0$x$0.1^0$ | 3 Hourly | Reanalysis Gauge Satellite | 3 hr - years | (Beck et al., 2019) |
| **SM2RAIN-ASCAT V1.5** | $0.11^0$x$0.11^0$ | Daily | Satellite | 5 months | (Brocca et al., 2014; Camici et al., 2019) |
| **JRA-55** | $1.5^0$x$1.5^0$ | 6 Hourly | Reanalysis | 4 days | (S. Kobayashi et al., 2015)(Kobayashi et al., 2015) |
| **MERRA-2** | 0.50 x 0.6250 | Daily | Reanalysis | 2-3 months | (Gelaro et al., 2017; Rienecker et al., 2011) |

**Table 3.** Characteristics of the precipitation products

Standard statistical estimators are utilized to assess accuracy by comparing the agreement between ground observations, different products, and the downscaled temperature product. The station-based dataset serves as the ground truth, and performance metrics are calculated for evaluation. Daily, monthly, seasonal, and annual statistical indices such as mean absolute error (MAE), root-mean-square error (RMSE), correlation coefficient (CC), and bias (Eq 1- Eq 4) are used to assess accuracy of the near-surface air temperature products. Additionaly, Eq 5-Eq 6 are used to assess accuracy of the precipitation products.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left(\left|T_{o,i} - T_{p,i}\right|\right) \quad (1)$$

$$CC = \frac{\sum_{i=1}^{N}\left(T_{o,i} - \bar{T}_o\right)\left(T_{p,i} - \bar{T}_p\right)}{\sqrt{\sum_{i=1}^{N}\left(T_{o,i} - \bar{T}_o\right)^2}\sqrt{\sum_{i=1}^{N}\left(T_{p,i} - \bar{T}_p\right)^2}} \quad (2)$$

$$RMSE = 1 - \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(T_{o,i} - T_{p,i}\right)^2} \quad (3)$$

$$Bias = \frac{1}{N}\sum_{i=1}^{N}\left(T_{o,i} - T_{p,i}\right) \quad (4)$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (5)$$

$$\beta = \frac{\mu_s}{\mu_o} \quad (6)$$

$$\gamma = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} \quad (7)$$

In the above equations, $T_{o,i}$, $T_{p,i}$, $\bar{T}_o$, $\bar{T}_p$ refer to the observed temperature, product temperature estimate, the average observed temperature, and the average product temperature estimates, respectively. The letter N indicates the number of stations, and o and p represent the observed and predicted data values, respectively. The r value represents the correlation coefficient, $\mu_s$ and $\mu_o$ represent the mean of the prediction and observation values, respectively, $\sigma_s$ and $\sigma_o$ represent the variability of the prediction and observation values.

### 2.2 Downscaling Methods

The study employs a range of conventional interpolation techniques, including bilinear, bicubic (Keys, 1981), distance-weighted average remapping (Eum et al., 2012), largest area fraction remapping (Jones, 1997), nearest neighbour (NN) (Thévenaz et al., 2009), first-order conservative remapping (Jones, 1999) methods, along with the random forest algorithm (Breiman, 2001) to downscale the near-surface air temperature data obtained from the reanalysis product. As distance-based interpolation methods Nearest Neighbourhood interpolation, Bilinear interpolation, Bicubic interpolation, Distance-weighted average remapping, Largest area fraction remapping, First-order conservative remapping methods are used. As machine learning methods Random Forest and Random Forest Spatial Interpolation methods are used. The choice of predictors plays a crucial role in building a downscaling model. Hence, it is essential for the predictors to be informative, and the relationship between them and their response should remain constant or stationary. (Pang et al., 2017). Several predictor variables are selected for developing the RF model. The availability of high spatial resolution is considered as a significant factor in selecting these predictors. Table 4 provides information on the spatial and temporal resolutions of the predictors used. Furthermore, auxiliary data was also taken into account by calculating the distance of each station to the coast. For downscaling the precipitation products cloud properties (Cloud Optical Thickness, Cloud Effective Radius, Cloud Water Path) derived over the 2015-2020 period using MOD06_L2 product are used as additional predictor data.

| Abbreviation | Variable | Source | Explanation |
|---|---|---|---|
| LST_Day, | Land Surface Temperature | MODIS | Daily average Lanc Surface Temperatur derived over the 2015-2020 period usi MOD11A1 produc |
| LST_Night | | | |
| LST_DN | Day-Night Land Surface Temperature Difference | | |
| NDVI | Normalized difference vegetation index | MODIS | Normalized differen vegetation index deri from 16-day composite MOD13A2 produc |
| Lon | Longitude | Meteorological data | The geographical loca |
| Lat | Latitude | | |
| El | Elevation | SRTM | Extracted from 30 r resolution Digital Elevation Mo (SRTM-DEM) |
| Sl | Slope | | |
| As | Aspect | | |
| Dis | Distance to coast | Meteorological data | Based on the GSHH coastline |

**Table 4.** Summary of predictor data used in downscaling

## 3. RESULTS

### 3.1 Near-surface Air Temperature Products Evaluation

Several statistical metrics are employed to assess the performance of reanalysis products against ground-based observations. The initial evaluation covers monthly data from 2015 to 2019, focusing on the mean monthly temperature. Validation results for all seasons are summarized in Table 5, comparing observed mean temperature values with data from various products across the study domain. The findings show that reanalysis-based products effectively capture the mean monthly temperature in Turkey throughout all seasons (Karaman & Akyürek, 2023).

Several statistical metrics are employed to evaluate the performance of multiple reanalysis products compared to ground-based observations for daily time series. The initial evaluation covers daily data from 2015 to 2019. Table 6 presents a summary of the validation results for all seasons by comparing daily observed mean temperature values with the corresponding data from the reanalysis products.

| Season | Metric | AgEra5 | Era5 | Era5 Land | Merra2 | JRA-55 |
|---|---|---|---|---|---|---|
| Complete Time Series | MAE | 1.38 | 1.42 | 1.47 | 1.81 | 1.78 |
| | CC | 0.98 | 0.98 | 0.98 | 0.97 | 0.96 |
| | RMSE | 1.87 | 1.92 | 1.95 | 2.40 | 2.40 |
| | Bias | 0.90 | 0.30 | 0.89 | 0.94 | 0.31 |
| Winter | MAE | 1.47 | 1.50 | 1.76 | 2.22 | 1.87 |
| | CC | 0.93 | 0.91 | 0.92 | 0.88 | 0.85 |
| | RMSE | 2.01 | 2.07 | 2.29 | 2.88 | 2.55 |
| | Bias | 1.05 | 0.44 | 1.34 | 1.65 | 0.27 |
| Spring | MAE | 1.46 | 1.35 | 1.47 | 1.82 | 1.69 |
| | CC | 0.95 | 0.93 | 0.95 | 0.92 | 0.90 |
| | RMSE | 1.97 | 1.86 | 1.96 | 2.38 | 2.26 |
| | Bias | 1.14 | 0.30 | 1.07 | 1.25 | 0.41 |
| Summer | MAE | 1.36 | 1.46 | 1.38 | 1.63 | 1.83 |
| | CC | 0.92 | 0.88 | 0.91 | 0.85 | 0.80 |
| | RMSE | 1.83 | 1.95 | 1.81 | 2.17 | 2.45 |
| | Bias | 0.73 | 0.13 | 0.65 | 0.36 | 0.23 |
| Autumn | MAE | 1.23 | 1.35 | 1.28 | 1.60 | 1.73 |
| | CC | 0.97 | 0.96 | 0.96 | 0.94 | 0.93 |
| | RMSE | 1.67 | 1.81 | 1.71 | 2.15 | 2.34 |
| | Bias | 0.68 | 0.31 | 0.55 | 0.57 | 0.31 |
| Annual | MAE | 1.26 | 1.30 | 1.32 | 1.60 | 1.64 |
| | CC | 0.91 | 0.87 | 0.90 | 0.84 | 0.77 |
| | RMSE | 1.72 | 1.76 | 1.75 | 2.16 | 2.24 |
| | Bias | 0.90 | 0.30 | 0.89 | 0.95 | 0.31 |

**Table 5.** Seasonal and annual statistics for the entire study area

| Season | Metric | AgEra5 | Era5 | Era5 Land | Merra2 | JRA-55 |
|---|---|---|---|---|---|---|
| Complete Time Series | MAE | 1.63 | 1.68 | 1.86 | 2.15 | 2.01 |
| | CC | 0.98 | 0.97 | 0.97 | 0.96 | 0.96 |
| | RMSE | 2.18 | 2.24 | 2.46 | 2.84 | 2.69 |
| | Bias | 0.86 | 0.32 | 0.89 | 0.94 | 0.38 |
| Winter | MAE | 1.85 | 1.89 | 2.27 | 2.69 | 2.22 |
| | CC | 0.93 | 0.91 | 0.90 | 0.87 | 0.87 |
| | RMSE | 2.48 | 2.55 | 2.95 | 3.48 | 3.00 |
| | Bias | 0.97 | 0.46 | 1.33 | 1.63 | 0.39 |
| Spring | MAE | 1.69 | 1.62 | 1.84 | 2.13 | 1.92 |
| | CC | 0.94 | 0.93 | 0.93 | 0.91 | 0.90 |
| | RMSE | 2.25 | 2.17 | 2.44 | 2.78 | 2.54 |
| | Bias | 1.10 | 0.33 | 1.06 | 1.22 | 0.45 |
| Summer | MAE | 1.49 | 1.59 | 1.64 | 1.80 | 1.94 |
| | CC | 0.92 | 0.89 | 0.89 | 0.86 | 0.82 |
| | RMSE | 1.98 | 2.10 | 2.19 | 2.37 | 2.58 |
| | Bias | 0.72 | 0.18 | 0.65 | 0.37 | 0.27 |
| Autumn | MAE | 1.50 | 1.61 | 1.71 | 1.98 | 1.97 |
| | CC | 0.96 | 0.95 | 0.95 | 0.93 | 0.92 |
| | RMSE | 1.99 | 2.13 | 2.25 | 2.61 | 2.63 |
| | Bias | 0.64 | 0.33 | 0.55 | 0.57 | 0.41 |
| Annual | MAE | 1.21 | 1.27 | 1.31 | 1.57 | 1.62 |
| | CC | 0.92 | 0.88 | 0.91 | 0.86 | 0.79 |
| | RMSE | 1.67 | 1.73 | 1.74 | 2.14 | 2.20 |
| | Bias | 0.86 | 0.32 | 0.89 | 0.95 | 0.38 |

**Table 6.** Seasonal and annual mean statistics for the entire study area (daily)

Figure 3 illustrates MAE (°C) for monthly time series of Köppen-Geiger climate classes. Across all classes, AgERA5, ERA5, and ERA5-Land products exhibit a similar pattern. However, AgERA5 consistently outperforms the other products. Among the compared products, JRA-55 demonstrates the lowest performance, followed by the MERRA-2 product. The maximum mean absolute error of 2°C is observed for the (Dfb) class, characterized by wet seasons, cold winters, and hot summers, as well as the (Dsa) class, characterized by severe winters and dry, very hot summers.

In Figure 4, MAE (°C) is displayed for daily time series across different climate classes. In the analysis, the arid steppe hot (Bsh), arid desert cold (Bwk), cold no dry season cold summer (Dfc), and cold dry summer cold summer (Dsc) classes are excluded due to an insufficient number of stations, making it difficult to derive statistically meaningful results. However, considering the daily time series, the AgERA5 product exhibits superior performance across all climate classes. On the other hand, the AgERA5 product performed best in all climate classes, followed by the JRA-55 product and the MERRA-2 product.

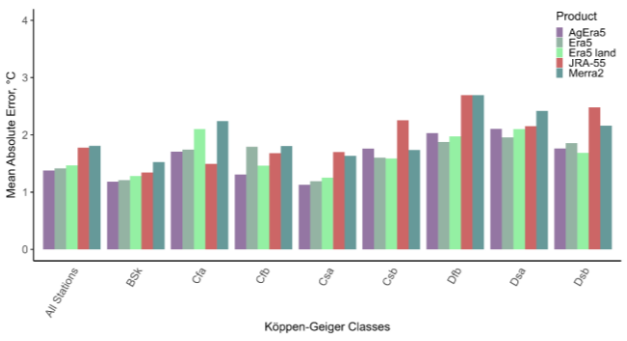


**Figure 3.** MAE (ºC) for monthly time series on varying climatic conditions

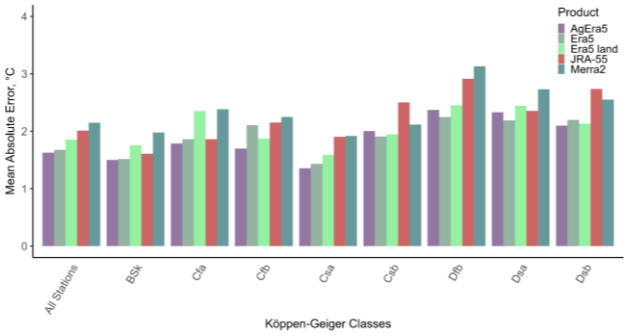


**Figure 4.** MAE (ºC) for daily time series on varying climatic conditions

### 3.2 Precipitation Products Evaluation

Figure 5 and Figure 6 present box-and-whisker plots showing the KGE scores for 16 precipitation datasets for monthly and daily data series, respectively. The mean KGE of all datasets is 0.47. Datasets performed better on a monthly scale with improved KGE scores compared to daily time series.

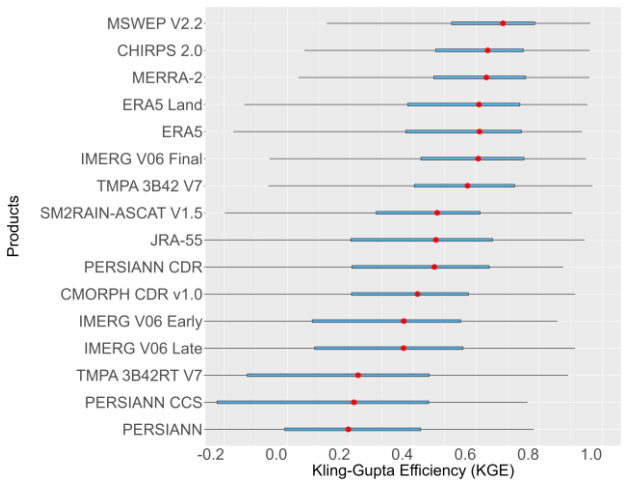


**Figure 5.** Box-and-whisker plots of KGE score of precipitation products in monthly time series

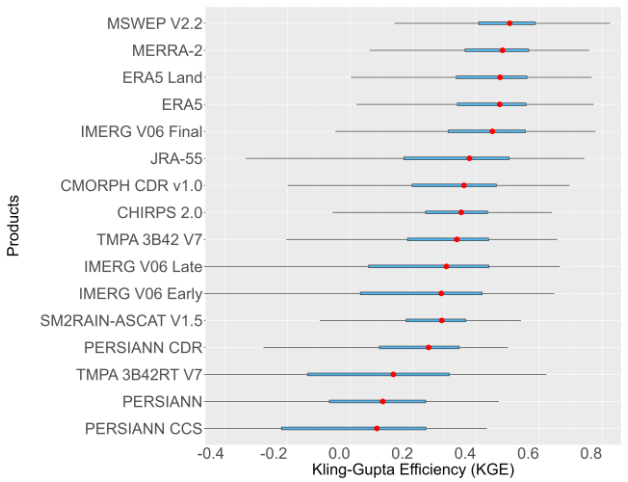


**Figure 6.** Box-and-whisker plots of KGE score of precipitation products in daily time series

### 3.3 Downscaling of Temperature Product

From the evaluation of the near-surface temperature products AgERA5 is found superior among other products. Therefore, this product is downscaled to 5 km for monthly time series and daily time series. The results are given in Table 7 for monthly and Table 8 for daily time series, respectively. Results indicate that different interpolation methods result in similar results with minor differences.

| Season | Metric | AgEra5 | Nearest Neighbor | Bilinear | Bicubic | Distance-Weighted Average | Largest Area Fraction | First-Order Convervative | Random Forest |
|---|---|---|---|---|---|---|---|---|---|
| Complete Time Series | MAE | 1.50 | 1.52 | 1.48 | 1.47 | 1.50 | 1.55 | 1.48 | 1.22 |
| | CC | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.99 |
| | RMSE | 2.04 | 2.05 | 2.00 | 1.99 | 2.02 | 2.11 | 2.00 | 1.65 |
| | Bias | 0.95 | 0.93 | 0.97 | 0.95 | 0.99 | 1.00 | 0.97 | 0.66 |
| Winter | MAE | 1.82 | 1.82 | 1.81 | 1.80 | 1.82 | 1.89 | 1.81 | 1.51 |
| | CC | 0.92 | 0.92 | 0.92 | 0.92 | 0.93 | 0.92 | 0.92 | 0.94 |
| | RMSE | 2.45 | 2.45 | 2.43 | 2.43 | 2.43 | 2.57 | 2.43 | 2.03 |
| | Bias | 1.23 | 1.21 | 1.27 | 1.25 | 1.28 | 1.29 | 1.27 | 0.98 |
| Spring | MAE | 1.53 | 1.53 | 1.52 | 1.49 | 1.53 | 1.58 | 1.52 | 1.14 |
| | CC | 0.95 | 0.95 | 0.95 | 0.95 | 0.95 | 0.94 | 0.95 | 0.96 |
| | RMSE | 2.06 | 2.05 | 2.02 | 2.00 | 2.04 | 2.13 | 2.02 | 1.54 |
| | Bias | 1.16 | 1.14 | 1.18 | 1.15 | 1.20 | 1.20 | 1.18 | 0.54 |
| Summer | MAE | 1.36 | 1.38 | 1.33 | 1.31 | 1.35 | 1.40 | 1.33 | 1.12 |
| | CC | 0.91 | 0.90 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.93 |
| | RMSE | 1.84 | 1.86 | 1.78 | 1.76 | 1.81 | 1.87 | 1.78 | 1.52 |
| | Bias | 0.74 | 0.72 | 0.76 | 0.72 | 0.78 | 0.78 | 0.76 | 0.53 |
| Autumn | MAE | 1.32 | 1.34 | 1.29 | 1.28 | 1.30 | 1.36 | 1.29 | 1.10 |
| | CC | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.98 |
| | RMSE | 1.77 | 1.80 | 1.73 | 1.73 | 1.74 | 1.83 | 1.73 | 1.47 |
| | Bias | 0.67 | 0.66 | 0.70 | 0.68 | 0.72 | 0.73 | 0.70 | 0.60 |
| Annual | MAE | 1.31 | 1.31 | 1.30 | 1.28 | 1.31 | 1.37 | 1.30 | 0.97 |
| | CC | 0.91 | 0.90 | 0.91 | 0.91 | 0.91 | 0.90 | 0.91 | 0.94 |
| | RMSE | 1.79 | 1.80 | 1.75 | 1.74 | 1.77 | 1.87 | 1.75 | 1.34 |
| | Bias | 0.95 | 0.93 | 0.97 | 0.95 | 0.99 | 1.00 | 0.97 | 0.66 |

**Table 7.** Performance metrics for different interpolation algorithms for monthly time series

| Season | Metric | AgEra5 | Nearest Neighbor | Bilinear | Bicubic | Distance-Weighted Average | Largest Area Fraction | First-Order Convervative | Random Forest |
|---|---|---|---|---|---|---|---|---|---|
| Complete Time Series | MAE | 1.62 | 1.65 | 1.58 | 1.57 | 1.6 | 1.63 | 1.58 | 1.59 |
| | CC | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.97 | 0.98 | 0.97 |
| | RMSE | 2.17 | 2.2 | 2.12 | 2.11 | 2.14 | 2.19 | 2.12 | 2.09 |
| | Bias | 0.82 | 0.82 | 0.83 | 0.8 | 0.86 | 0.82 | 0.83 | 0.58 |
| Winter | MAE | 2.06 | 2.1 | 2.04 | 2.03 | 2.06 | 2.1 | 2.04 | 1.85 |
| | CC | 0.92 | 0.91 | 0.92 | 0.92 | 0.92 | 0.91 | 0.92 | 0.93 |
| | RMSE | 2.75 | 2.79 | 2.72 | 2.72 | 2.73 | 2.82 | 2.72 | 2.45 |
| | Bias | 1.2 | 1.22 | 1.23 | 1.21 | 1.26 | 1.22 | 1.23 | 0.92 |
| Spring | MAE | 1.7 | 1.71 | 1.67 | 1.65 | 1.69 | 1.71 | 1.67 | 1.66 |
| | CC | 0.94 | 0.94 | 0.94 | 0.94 | 0.94 | 0.94 | 0.94 | 0.93 |
| | RMSE | 2.26 | 2.27 | 2.21 | 2.2 | 2.24 | 2.29 | 2.21 | 2.18 |
| | Bias | 1.12 | 1.12 | 1.13 | 1.1 | 1.16 | 1.13 | 1.13 | 0.53 |
| Summer | MAE | 1.51 | 1.54 | 1.46 | 1.45 | 1.49 | 1.5 | 1.46 | 1.54 |
| | CC | 0.9 | 0.9 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.89 |
| | RMSE | 2.01 | 2.05 | 1.95 | 1.93 | 1.98 | 2 | 1.95 | 2.01 |
| | Bias | 0.7 | 0.71 | 0.71 | 0.67 | 0.74 | 0.69 | 0.71 | 0.51 |
| Autumn | MAE | 1.54 | 1.58 | 1.51 | 1.51 | 1.52 | 1.56 | 1.51 | 1.53 |
| | CC | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 | 0.96 |
| | RMSE | 2.05 | 2.09 | 2.01 | 2.01 | 2.02 | 2.08 | 2.01 | 2.01 |
| | Bias | 0.65 | 0.66 | 0.67 | 0.65 | 0.7 | 0.67 | 0.67 | 0.56 |
| Annual | MAE | 1.27 | 1.28 | 1.25 | 1.24 | 1.26 | 1.32 | 1.25 | 0.96 |
| | CC | 0.91 | 0.9 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.94 |
| | RMSE | 1.74 | 1.75 | 1.7 | 1.68 | 1.71 | 1.8 | 1.7 | 1.32 |
| | Bias | 0.84 | 0.82 | 0.87 | 0.84 | 0.88 | 0.89 | 0.87 | 0.61 |

**Table 8.** Performance metrics for different interpolation algorithms for daily time series

### 3.4 Downscaling of Precipitation Product

From the evaluation of the precipitation products MSWEP V2.2 is found superior among other products. Therefore, this product is downscaled to 5 km for monthly time series and daily time series. The results presented in Table 9 and Table 10 allow for a comparison of the performance of the original MSWEP V2.2 products with that of the downscaled product for monthly and daily time series, respectively. The various statistical measures, such as KGE, MAE, CC, RMSE, and bias, provide information on the accuracy and reliability of the downscaled product for

different seasons and for the entire study domain. This comparison can help to identify any potential improvements or limitations of the downscaling method.

| Season | Metric | MSWEP V2.2 | Randım Forest Cloud | Random Forest LST |
|---|---|---|---|---|
| Complete Time Series | KGE | 0.55 | 0.18 | 0.19 |
| | MAE | 1.62 | 4.88 | 0.59 |
| | CC | 0.62 | 0.24 | 0.31 |
| | RMSE | 4.91 | 9.36 | 2.49 |
| | Bias | -0.01 | -0.36 | -0.10 |
| Winter | KGE | 0.59 | 0.16 | 0.18 |
| | MAE | 2.25 | 6.01 | 0.76 |
| | CC | 0.64 | 0.20 | 0.33 |
| | RMSE | 5.99 | 10.94 | 3.38 |
| | Bias | -0.07 | -0.87 | -0.14 |
| Spring | KGE | 0.56 | 0.21 | 0.19 |
| | MAE | 1.74 | 4.35 | 0.66 |
| | CC | 0.61 | 0.27 | 0.37 |
| | RMSE | 4.32 | 7.69 | 2.14 |
| | Bias | -0.10 | -0.69 | -0.19 |
| Summer | KGE | 0.35 | -0.04 | 0.16 |
| | MAE | 1.13 | 3.70 | 0.59 |
| | CC | 0.45 | 0.19 | 0.26 |
| | RMSE | 4.30 | 8.39 | 2.63 |
| | Bias | 0.08 | 0.80 | -0.04 |
| Autumn | KGE | 0.57 | 0.20 | 0.19 |
| | MAE | 1.36 | 4.75 | 0.44 |
| | CC | 0.64 | 0.30 | 0.31 |
| | RMSE | 4.87 | 9.49 | 1.96 |
| | Bias | 0.04 | 0.01 | -0.09 |
| Annual | KGE | 0.63 | 0.65 | 0.34 |
| | MAE | 145.30 | 136.53 | 27.30 |
| | CC | 0.80 | 0.71 | 0.47 |
| | RMSE | 218.94 | 202.16 | 43.08 |
| | Bias | -3.36 | -28.37 | -12.70 |

**Table 9.** Seasonal and annual mean statistics for the entire study area (Monthly time series)

Downscaling MSWEP with selected predictors doesn't perform well in general. Therefore, Random Forest Spatial Interpolation method is used through blending the product with in situ measurements. The RFSI model was used to generate a regression model for each day in the dataset. The stations were separated into two groups, training (%80) and testing (%20). These stations were selected randomly with consideration for their uniform spatial distribution. A 10-fold cross-validation was also used during the training process of each model. The daily RFSI models were trained with the training dataset, which allowed for the analysis of the sensitivity of the results to spatial information. The models were built using 500 trees, and five neighbouring stations were selected to calculate spatial autocorrelation. To calculate spatial autocorrelation, five adjacent stations were chosen.

Box and whisker plots for the complete time series of MSWEP V2.2, IMERG V06 Final and merged product are also presented in Figure 7. The spatial distribution of KGE indicates a trend towards higher values (Figure 8). There is a noticeable improvement in KGE scores across all regions of Turkey. Despite the challenging topography in the Eastern Black Sea region, with its high elevation differences affecting the accuracy of precipitation products, the merged product still demonstrates much better performance with high accuracy. The merging method has played a role in addressing these difficulties. In addition, MSWEP V2.2 and merged products on estimating precipitation over varying climatic conditions are investigated. Figure 9 presents MAE (mm) for complete-time series on varying climatic conditions for MSWEP V2.2 and Merged products.

| Season | Metric | MSWEP V2.2 | Random Forest Cloud | Random Forest LST |
|---|---|---|---|---|
| Complete Time Series | KGE | 0.74 | 0.65 | 0.27 |
| | MAE | 20.07 | 19.38 | 4.21 |
| | CC | 0.80 | 0.68 | 0.40 |
| | RMSE | 35.49 | 35.85 | 9.62 |
| | Bias | -0.61 | -2.73 | -1.06 |
| Winter | KGE | 0.72 | 0.61 | 0.24 |
| | MAE | 27.74 | 30.10 | 3.29 |
| | CC | 0.80 | 0.68 | 0.41 |
| | RMSE | 46.38 | 51.93 | 9.67 |
| | Bias | -2.38 | -8.96 | -0.83 |
| Spring | KGE | 0.71 | 0.58 | 0.27 |
| | MAE | 19.13 | 21.54 | 4.15 |
| | CC | 0.76 | 0.63 | 0.46 |
| | RMSE | 29.70 | 32.22 | 7.39 |
| | Bias | -3.10 | -7.04 | -1.77 |
| Summer | KGE | 0.63 | 0.36 | 0.24 |
| | MAE | 15.31 | 10.51 | 5.83 |
| | CC | 0.71 | 0.63 | 0.34 |
| | RMSE | 30.95 | 22.99 | 13.25 |
| | Bias | 2.08 | 3.70 | -0.58 |
| Autumn | KGE | 0.70 | 0.58 | 0.34 |
| | MAE | 18.09 | 16.57 | 3.58 |
| | CC | 0.81 | 0.65 | 0.46 |
| | RMSE | 32.35 | 30.94 | 6.80 |
| | Bias | 0.95 | 0.07 | -1.06 |
| Annual | KGE | 0.64 | 0.65 | 0.34 |
| | MAE | 144.90 | 136.53 | 27.29 |
| | CC | 0.80 | 0.71 | 0.47 |
| | RMSE | 217.09 | 202.16 | 43.07 |
| | Bias | -7.33 | -28.37 | -12.70 |

**Table 10.** Seasonal and annual mean statistics for the entire study area (Daily time series)

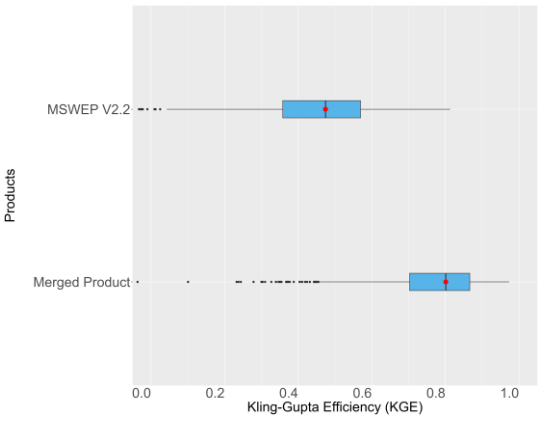


**Figure 7.** KGE values for MSWEP V2.2 and Merged Products

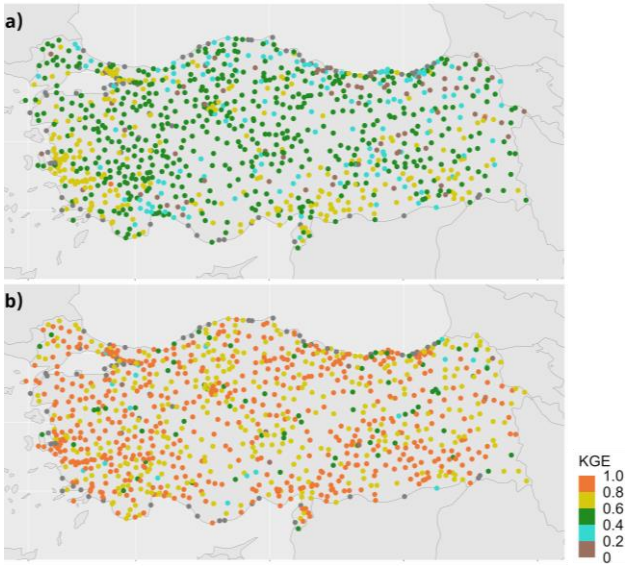


**Figure 8.** Spatial distribution KGE scores of stations:
a) MSWEP V2.2 product, b) Merged MSWEB V2.2 product

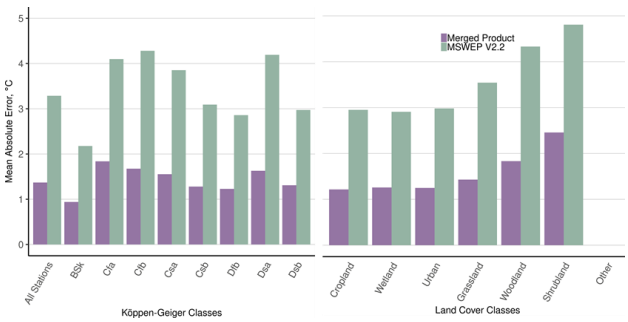


**Figure 9.** Comparison of MSWEP V2.2 and merged products on climatic conditions and Land Cover Classes

## 4. DISCUSSIONS

To evaluate the accuracy of reanalysis products (AgERA5, ERA5, ERA5-Land, MERRA-2, and JRA-55) in estimating near-surface air temperature in Turkey, a comprehensive assessment is conducted by comparing these products with ground-based measurements. The comparison is performed at both daily and monthly time scales to assess their performance across different temporal resolutions. Furthermore, the performance of each product is analyzed across various climatic conditions and land cover types, utilizing Köppen-Geiger classifications and land cover classes retrieved from the MODIS MCD12Q1 Version 6 data product.

It is evident that the AgERA5 product demonstrates remarkable accuracy in estimating near-surface air temperature across Turkey, both on a seasonal and annual basis. Subsequently, to further improve its performance, various interpolation algorithms are utilized to enhance the spatial resolution of the AgERA5 product. To achieve this, a combination of traditional distance-based methods and a machine learning-based approach using the RF algorithm is implemented to downscale the AgERA5 product. Random forest performs best among all interpolation methods. However, it has disadvantage of loss of data due to cloud contamination. It definitely provides much higher accuracy in clear-sky conditions (Karaman & Akyürek, 2023).

Performance evaluation of reanalysis, gauge, and satellite-based products of ERA5, ERA5 Land, IMERG V06, TMPA 3B42, TMPA 3B42-RT, SM2RAIN-ASCAT V1.0, MSWEP V2.2, CHIRPS 2.0, CMORPH, PERSIANN, PERSIANN CCS, PERSIANN CDR, and JRA-55 products on estimation of daily and monthly total precipitation over Turkey has been performed by comparing ground-based measurements. The effectiveness of each product was evaluated by comparing them on daily and monthly time scales, as well as in different climatic conditions using the Köppen-Geiger classification system. Initial analysis revealed that the MSWEP V2.2 product outperforms others in estimating daily and monthly total precipitation across Turkey both seasonally and annually. To enhance its performance, the spatial resolution of the MSWEP V2.2 product was improved through the use of various interpolation algorithms. Traditional downscaling methods were combined with a machine learning approach using the Random Forest algorithm to improve the MSWEP V2.2 products spatial resolution. The effectiveness of each interpolation algorithm in enhancing the product's estimate was investigated. Despite the use of various interpolation algorithms, the results showed that the performance of the product was not improved significantly. In addition, different predictors were used in estimating daily and monthly precipitation values by applying the Random Forest machine learning algorithm. It was found that incorporating cloud properties such as cloud effective radius (CER), cloud optical thickness (COT), and cloud water path (CWP) led to higher performance in downscaling precipitation, compared to using land surface temperature (LST), normalized difference vegetation index (NDVI) parameters.

It was found that the KGE scores of the merged product are higher in all seven regions of Turkey when compared to the MSWEP V2.2 products. The distribution of the KGE scores of the stations in each region was analysed using box and whisker plots presented in Figure 7. The results indicate a clear improvement in the performance of the merged product in all regions, with higher KGE scores observed compared to MSWEP V2.2. In order to enhance the precision of precipitation products, we utilized a merging algorithm that incorporates ground-based observations. The results indicate a substantial improvement in the daily performance of the precipitation product. The performance of the merged product was also analysed spatially in different climatic conditions and land cover classes.

The results highlight the significance of incorporating ground observations in developing satellite-based products and combining remote sensing data with observations in the RFSI model, which results in the most favourable outcomes.

The results of this study can provide useful information for decision-makers and researchers who use near-surface air temperature and precipitation data in various applications, such as hydrological modelling, weather forecasting, and climate change studies. Additionally, the results can be used to improve the quality of satellite and reanalysis data products by highlighting their strengths and weaknesses and suggesting areas for future improvement. While this study concentrates on climate variables such as precipitation and temperature, the proposed methods are applicable to any type of climate or land surface variables, such as soil moisture, snow cover, snow water equivalent, etc.

## ACKNOWLEDGEMENTS

The Turkish State Meteorological Service is greatly appreciated for providing the ground-based data.

## REFERENCES

Arismendi, I., Safeeq, M., Dunham, J. B., & Johnson, S. L., 2014. Can air temperature be used to project influences of climate change on stream temperature? *Environmental Research Letters*, 9(8). https://doi.org/10.1088/1748-9326/9/8/084015

Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., & Javelle, P., 2011. Sensitivity of hydrological models to uncertainty in rainfall input. *Hydrological Sciences Journal*, 56(3). https://doi.org/10.1080/02626667.2011.563742

Breiman, L., 2001. Random forests. Machine Learning. https://doi.org/10.1023/A:1010933404324

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E., 2010. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Condom, T., Martínez, R., Pabón, J. D., Costa, F., Pineda, L., Nieto, J. J., López, F., & Villacis, M., 2020. Climatological and Hydrological Observations for the South American Andes: In situ Stations, Satellite, and Reanalysis Data Sets. *Frontiers in Earth Science*. https://doi.org/10.3389/feart.2020.00092

Chen, J., & Brissette, F. P., 2017. Hydrological modelling using proxies for gauged precipitation and temperature. *Hydrological Processes*, 31(22), 3881–3897. https://doi.org/10.1002/hyp.11304

Colombi, A., Pepe, M., Rampini, A., & Michele, C. De., 2007. Estimation of daily mean air temperature from MODIS LST in Alpine areas. *EARSeL EProceedings*, 6(January 2007), 563–573.

Derin, Y., & Yilmaz, K. K., 2014. Evaluation of multiple satellite-based precipitation products over complex topography. *Journal of Hydrometeorology*. https://doi.org/10.1175/JHM-D-13-0191.1

Elgamal, A., Reggiani, P., & Jonoski, A., 2017. Impact analysis of satellite rainfall products on flow simulations in the Magdalena River Basin, Colombia. *Journal of Hydrology: Regional Studies*, 9, 85–103. https://doi.org/10.1016/j.ejrh.2016.09.001

Eum, H. Il, Gachon, P., Laprise, R., & Ouarda, T., 2012. Evaluation of regional climate model simulations versus gridded observed and regional reanalysis products using a combined weighting scheme. *Climate Dynamics*, 38(7–8), 1433–1457. https://doi.org/10.1007/s00382-011-1149-3

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. E., 2007. The shuttle radar topography mission. *Reviews of Geophysics*. https://doi.org/10.1029/2005RG000183

Graae, B. J., de Frenne, P., Kolb, A., Brunet, J., Chabrerie, O., Verheyen, K., Pepin, N., Heinken, T., Zobel, M., Shevtsova, A., Nijs, I., & Milbau, A., 2012. On the use of weather data in ecological studies along altitudinal and latitudinal gradients. Oikos, 121(1). https://doi.org/10.1111/j.1600-0706.2011.19694.x

Hengl, T., Heuvelink, G. B. M., & Rossiter, D. G., 2007. About regression-kriging: From equations to case studies. *Computers and Geosciences*. https://doi.org/10.1016/j.cageo.2007.05.001

Jones, P., 1997. A user’s guide for SCRIP: A spherical coordinate remapping and interpolation package. Los Alamos National Laboratory. http://oceans11.lanl.gov/trac/SCRIP/export/21/trunk/SCRIP/doc/SCRIPusers.pdf

Jones, P. W., 1999. First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*. https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2

Karaman, C., & Akyürek, Z. (2023). Evaluation of Near-surface Air Temperature Reanalysis Datasets and Downscaling with Machine Learning based Random Forest Method for Complex Terrain of Turkey. *Advances in Space Research*. https://doi.org/https://doi.org/10.1016/j.asr.2023.02.006

Keys, R. G., 1981. Cubic Convolution Interpolation for Digital Image Processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6). https://doi.org/10.1109/TASSP.1981.1163711

Köppen, Wladimir; Geiger, R., 1936. Handbuch der Klimatologie: *Das geographische System der Klimate. In Verlag von Gebrüder Borntraeger*.

Liou, Y. A., & Kar, S. K., 2014. Evapotranspiration estimation with remote sensing and various surface energy balance algorithms-a review. *In Energies* (Vol. 7, Issue 5). https://doi.org/10.3390/en7052821

Mu, Y., Biggs, T., & Shen, S. S. P., 2021. Satellite-based precipitation estimates using a dense rain gauge network over the Southwestern Brazilian Amazon: Implication for identifying trends in dry season rainfall. *Atmospheric Research*, 261. https://doi.org/10.1016/j.atmosres.2021.105741

Pang, B., Yue, J., Zhao, G., & Xu, Z., 2017. Statistical Downscaling of Temperature with the Random Forest Model. *Advances in Meteorology*. https://doi.org/10.1155/2017/726517

Prihodko, L., & Goward, S. N., 1997. Estimation of air temperature from remotely sensed surface observations. *Remote Sensing of Environment*, 60(3), 335–346. https://doi.org/10.1016/S0034-4257(96)00216-7

Shepard D., 1968. Two- dimensional interpolation function for irregularly- spaced data. *Proc 23rd Nat Conf*. https://doi.org/10.1145/800186.810616

Terando, A., Youngsteadt, E., Meineke, E., & Prado, S., 2018. Accurate near surface air temperature measurements are necessary to gauge large-scale ecological responses to global climate change. *In Ecology and Evolution* (Vol. 8, Issue 11). https://doi.org/10.1002/ece3.3972

Thévenaz, P., Blu, T., & Unser, M., 2009. Image interpolation and resampling. *Handbook of Medical Image Processing and Analysis*. https://doi.org/10.1016/B978-012373904-9.50037-4

Timmermans, B., Wehner, M., Cooley, D., O’Brien, T., & Krishnan, H., 2019. An evaluation of the consistency of extremes in gridded precipitation data sets. *Climate Dynamics*. https://doi.org/10.1007/s00382-018-4537-0