Article

Not peer-reviewed version

Scale Dependence of Errors in Snow Water Equivalent Simulations using ERA5 Reanalysis over Alpine Basins

Susen Shrestha *, Mattia Zaramella, Mattia Callegari, Felix Greifeneder, Marco Borga

Posted Date: 10 May 2023

doi: 10.20944/preprints202305.0762.v1

Keywords: TOPMELT; ERA5; SWE; Drizzle; Bias Adjustment

Article

Scale dependence of errors in snow water equivalent simulations using ERA5 reanalysis over alpine basins

Susen Shrestha [1,*], Mattia Zaramella [3], Mattia Callegari [2], Felix Greifeneder [4], and Marco Borga [3]

[1] EURAC, Center for Climate Change and Transformation (CCT), Bolzano, Italy;
[2] EURAC, Institute for Earth Observation, Bolzano, Italy;
[3] Department of Land, Environment, Agriculture, and Forestry, University of Padova, Italy;
[4] Chloris Geospatial, Boston, USA;

* Correspondence: sshrestha@eurac.edu

Abstract: This study evaluated the potential of ERA5 reanalysis as a reference dataset for snow water equivalent (SWE) modeling in 16 Alpine basins of varying catchment sizes using the semi-distributed snow model (TOPMELT) in the Upper Adige river basin in the Eastern Italian Alps. The study aimed to identify errors in ERA5 meteorological variables and assess their impact on SWE computation from 1992 to 2019. The findings revealed that ERA5 precipitation overestimated low-intensity rainfall (drizzle problem) and underestimated high-intensity rainfall, while ERA5 temperature underestimated observations. The overestimation of low-intensity rainfall created fictitious low-intensity snowfall events, which, when combined with colder ERA5 temperature, resulted in delayed snowmelt and increased fictitious snow cover days over the study area. The Quantile Mapping (QM) technique was used to remove errors in ERA5 variables. For precipitation, a monthly correction factor accounting the calibration period (1992-2005) was considered to correct the entire period. Temperature errors were corrected individually for calibration and validation periods. The corrected ERA5 SWE simulation showed improved performance, with a decrease in fictitious snow cover days. The study also highlighted the importance of temperature correction over precipitation correction in SWE simulation, particularly for smaller basins, while considering the monthly moving mean to remove seasonality.

Keywords: TOPMELT; ERA5; SWE; Drizzle; Bias Adjustment

1. Introduction

Mountain snowpack and its melt dominates the surface hydrology of many regions, with implications for water supply, hydropower and ecological processes [1, 2]. Reliable snow cover (SC) delineation and snow water equivalent (SWE) assessment remains crucial for snowmelt-runoff prediction, operational flood control, water supply planning, and water resource management in snowmelt-dominated watersheds [3]. However, the proper assessment of SC and SWE in mountain watersheds still remains a challenge. In data scarce basins, the use of reanalysis data, such as ERA5 [4] as an input to snow pack models, has been shown to provide reasonable simulations of SWE for mountain basins [5, 6].

The study by [5] compared the performance of MERRA-2 and ERA-5 reanalysis meteorological forcing with the distributed SnowModel in the high Atlas region. Both the reanalysis products showed comparable performance for snow simulation. Nonetheless, the ERA5 simulation, due to its finer spatial resolution as compared to MERRA-2, showed better performance. Whereas, [6] considered ERA5 as the boundary and initial condition to force the weather and forecasting (WRF) model for two mountain ranges in Lebanon. The Intermediate Complexity Atmospheric Research Model (ICAR) was nested inside the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

WRF to develop a 1-km regional-scale snow reanalysis. The results showed a good temporal and spatial correlation of the snow variables with the MODIS fractionals snow-covered area and the ground observations of SWE.

Reanalysis data combine a wide array of measured and remotely sensed information within a dynamical–physical coupled numerical model. They use the analysis part of a weather forecasting model, in which data assimilation forces the model toward the closest possible current state of the atmosphere [7]. Of particular interest to the hydrological community are the largely improved (with respect to earlier reanalysis products) spatial (30 km) and temporal (1 h) resolutions of ERA5. The spatial resolution is now similar to or better than that of most observational networks in the world, with the exception of some parts of Europe and the United States. However, the ERA5 spatial resolution has been shown to be still too coarse to represent correctly the influence of topography on meteorological variables, which is crucial for snow modeling in mountainous regions [6, 8, 9].

Whereas a number of studies proposed downscaling techniques to approach this problem, the scale dependent structure of errors in SWE simulations using ERA5 reanalysis over alpine basins remains largely unexplored. This is in spite of the need to better understand at which spatial scales the ERA5 spatial resolution affects mostly the SWE simulations, and then under which circumstances the implementation of downscaling techniques is required.

Against this background, the aim of this study is to evaluate the quality of hourly temperature and precipitation data from ERA5 reanalysis to simulate SWE dynamics in a mountainous, snow-controlled river system, with respect to corresponding SWE simulations obtained from a relatively high-density and quality-controlled data set obtained from ground station, used as a reference. More specifically, this study aims i) to isolate the impact of the input spatial aggregation on the accuracy of SWE simulations, by quantifying the effects of aggregating the reference precipitation and temperature data at the ERA5 grid scale, and ii) to evaluate the scale-dependence of ERA5-based simulation errors when SWE is aggregated over a range of spatial scales.

The study is carried out in the Upper Adige river basin, a 6924 km²-wide basin located in the Eastern Italian Alps, over the 1992-2019 period. The study area is selected due to the dense network of meteorological stations used to provide input to a snowpack model and to validate the ERA-5 performance. The TOPMELT model [10] is used for snowpack dynamics and SWE simulation. TOPMELT is a semi-distributed snowpack model based on an extended temperature index approach capable to estimate the full spatial distribution of the SWE at each time step. TOPMELT exploits a statistical representation of the distribution of clear sky potential solar radiation to drive the snowpack model which drastically reduces the computational costs associated to the fully spatially distributed simulation of SWE over vast areas and extended period of time while preserving simulation accuracy [10]. The good accuracy of TOPMELT SWE and SC simulations over the study area has been tested with respect to available in situ data and MODIS observations by [11, 12].

2. Materials and Methods

2.1 *Study area and in-situ data*

The study area is located in the upper Adige river basin closed at Bronzolo, in the Eastern Italian Alps (Figure 1). This is an alpine catchment with a drainage area of approximately 6924 km². The elevation ranges from about 200 m a.s.l. at the southern valley bottoms, to around 3900 m a.s.l. in the western upper ranges, with a mean elevation of 1800 m a.s.l. The steep terrain and the high elevation gradients govern the spatial precipitation distribution [13] with the precipitation ranging from 500mm in the northwest region to 1600mm in the southern region [14]. During the winter season, the precipitation is stored as snow and the streamflow is minimum. The streamflow shows two maxima: the first due to snowmelt in the early summer and the second due to intense precipitation in autumn [15]. The main agricultural areas in the northern part of the catchment are located

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

in the Venosta valleys, and cultivation comprises mainly fruit trees and grapes. Land use at high elevations is dominated by grass, grazing, and forest.

The study area is characterized by a rather dense network of meteorological stations, with 88 rain gauges (1 per 72 km²) and 124 temperature gauges (1 per 55 km²) covering the study region. The Hydrographic Office of Bozen, Bolzano, has made hourly temperature, precipitation, and runoff data available from 1991 until 2019. To assess the performance of the model at varying basin sizes, sixteen watersheds were selected for analysis, with drainage areas ranging from 49 km² to 6,924 km² (Figure 1, Table 1). These watersheds were chosen for analysis because the relatively minor impact of operations from artificial reservoirs allowed for the use of a hydrological model, and the ensuing comparison of simulated versus observed discharges.

Table 1. Drainage area and elevation of the study basins

Sn	Name	Elevation Range (m)	Mean Elevation (m)	Area (Km2)
1	Rio Plan	1561-3445	2387	49
2	Rio Riva at Seghe	1523-3421	2386	76
3	Rio Anterselva at Bagni	1092-3421	2026	82
4	Rio Braies at Braies	1124-3074	1911	93
5	Rio Riva at Caminata	855-3421	2278	115
6	Rio Casies at Colle	1196-2815	1961	117
7	Rio Gadera at Pedraces	1318-3111	2027	125
8	Aurino at Cadipietra	811-3111	2162	150
9	Rio Ridanna at Vipiteno	1046-3417	1933	210
10	Gadera at Mantana	944-3441	1855	397
11	Rio Passirio at Merano	336-3445	1851	414
12	Aurino at Caminata	844-3421	2117	420
13	Aurino at S.Giorgio	817-3421	2036	608
14	Rienza at Vandoies	732-3421	1859	1919
15	Adige at Ponte Adige	236-3889	1895	2732
16	Adige at Bronzolo	236-3889	1805	6924

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

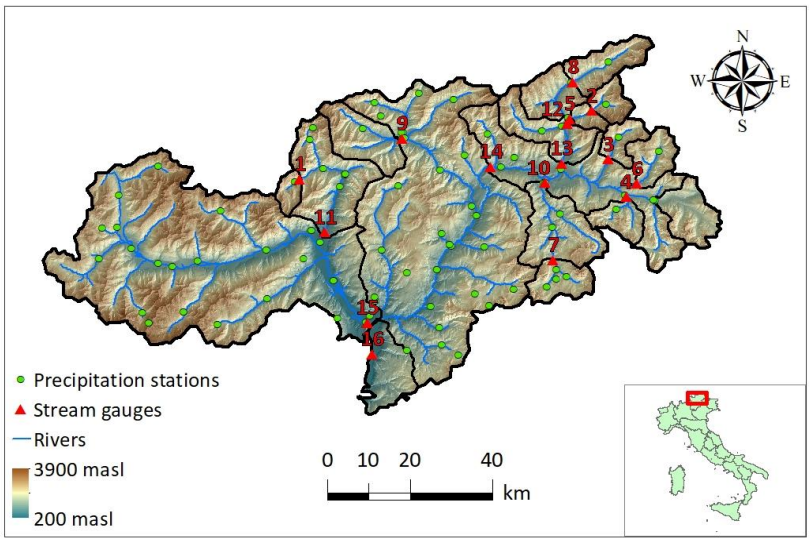


Figure 1. The upper Adige River basin closed at Bronzolo. The red triangles indicate the stream gauges, which are listed in Table 1.

2.2 *ERA5 reanalysis*

ERA5 reanalysis is a state-of-the-art fifth-generation ECMWF (European Centre for Medium-Range Weather Forecasts) atmospheric reanalysis of the global climate [4]. It is one of the fundamental elements of the Copernicus Climate Change Service (C3S), which is funded by the European Union. ERA5 provides multiple atmospheric, land, and oceanic climate data variables with availability spanning from 1959 to the present, at spatial resolution of 0.25 degrees and temporal resolution of 1 hour at the global scale. For this research, only temperature and precipitation will be considered from ERA5. Further information about ERA5 is available on the online data documentation (https://confluence.ecmwf.int/display/CKB). It provides a detailed description of the various products and a list of all available geophysical parameters, which can be freely downloaded.

Twenty-seven ERA5 grid cells that cover the Upper Adige River Basin at Bronzolo were considered for this study. Based on geometrical analysis, the ERA5 precipitation was partitioned over the 16 study basins. The ERA5 air temperature is scaled to the center of mass of each study basin based on the climatological monthly lapse rate valid for the region.

2.3 *The snowpack model: TOPMELT*

The snowpack model used in this work is TOPMELT [10]. TOPMELT is a semi-distributed snowpack model, which takes advantage of the extended temperature index approach to simulate SWE at full spatial distribution for each hourly time step [10, 12]. Clear sky shortwave solar radiation is computed at each element of the digital terrain model (DTM) by taking into account shadow and complex topography, calculating the apparent sun motion, and the intersection of radiation with topography. Diffuse radiation is computed by accounting for self-shading (by slope and aspect) and occlusions produced by the visible horizon. For model simulation, each basin is divided into elevation bands, and the full spatial distribution of clear sky potential solar radiation is discretized into a number of radiation classes for each elevation band. This is achieved by dividing each elevation band into a number n_c of equally distributed radiation classes, where the i^{th} class contains the band sub-area corresponding to the i^{th} percentile of the incident radiation energy. For each elevation band, the spatial extent of the model is represented by n_b elevation bands and by n_c radiation classes. TOPMELT accounts separately for snow and glacier melt and to consider the glacier area associated with each energy class, the model cells within $n_b \times n_c$ are provided with corresponding glacier area fraction.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

The calculation of the full spatial distribution of SWE reduces to each radiation class, rather than each element of the given elevation band, which significantly improves the computational efficiency. The analysis by [10] showed that the division of the elevation bands into ten radiation classes provides results comparable to a full spatially distributed model. Therefore, in this work, the elevation bands are subdivided into ten radiation classes based on the spatial distribution of the clear sky solar radiation on the pixels included in each elevation band.

The model uses the Thiessen polygon method to estimate the mean precipitation over the basin while the air temperature data is used to calculate the unique hourly vertical lapse rate for the entire basin. The precipitation data are corrected with the snow correction factor (SCF) to account for the inefficiency of the gauge during snow periods. The SCF is a multiplier applied to precipitation data when the station temperature goes below the threshold temperature T_c. Lastly, the basin precipitation, p_{basin}, is obtained by applying a precipitation correction factor (PCF), which is a non-dimensional constant used to take into account the poor spatial coverage of the rain-gauge stations. For a given i^{th} elevation band, the model computes the precipitation p_i (mm h^{-1}) by applying a vertical precipitation gradient G (km^{-1}), which considers increased precipitation over elevation as given in equation 1:

$$P_i = P_{basin}\left(1 + G\frac{h_i - h_{ref}}{1000}\right) \quad (1)$$

Where h_i, h_{ref} (m.a.s.l) are the mean altitude of the i^{th} elevation band and of the basin respectively. For the temperature, T_i(°C) is provided as input for each time step and elevation band. The use of a vertical temperature lapse rate helps to obtain the mean air temperature over the i^{th} elevation. The estimation of the precipitation phase (solid or liquid) is performed using the threshold temperature T_c. The snowmelt algorithm is applied to each radiation class of a given elevation band. For a given i^{th} elevation band and j^{th} radiation class, the snowmelt rate $F_{i,j}(t)[mm\ h^{-1}]$ at time t is calculated in equation 2:

$$F_{i,j} = CMF.RI_{i,j}(t).\left(1 - alb_i(t)\right).\max\{0; T_i(t) - T_b\} \quad (2)$$

Here, CMF is the combined melting factor considering both thermal and radiative effects, $RI_{i,j}(t)$ is the cell radiation index, $alb_i(t)$ is the snow albedo [-] at time t, T_i(t) is the air temperature at the i-th elevation band at time t while T_b=0°C is the temperature threshold above which snowmelt is assumed to occur, both in [°C]. The snow albedo at each elevation band is computed using the method described by [16] as given in equation 3:

$$alb(t) = ALBS - \beta_2.[\log_{10}\sum_k T_i(t_k)] \quad (3)$$

Where ALBS is the albedo of fresh snow, $\beta 2$ is a dimensionless parameter and $\sum_k T_i(t_k)$ (°C) is the summation of the positive hourly temperatures that are above the threshold base temperature (T_b) from the last snowfall until the current time t. The model accounts for rain-on-snow and melt during the night employing a temperature index approach, through two additional parameters: the rain melt factor (RMF) and the night melt factor (NMF), respectively.

The SWE ($we_{i,j}$; mm) for each model cell is updated considering snow accumulation, rain-on-snow, melt, and freezing water. After snowmelt or rainfall, the water is retained in the snowpack as interstitial water as liquid water $liqw_{i,j}$ (mm). When $liqw_{i,j}$ exceeds the water-holding capacity of the snowpack (LWT), it flows through the snowpack at the rate of DYTIME (m h^{-1}), to form net water flow at the snowpack base. When air temperature goes below the base temperature, some of the liquid refreezes, reducing $liqw_{i,j}$ and adding to the snowpack through a freezing rate termed ice (mm h^{-1}). This is computed as in equation 4:

$$ice_i(t) = REFRZ.\min[0, (T_b - T_i(t))] \quad (4)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

Where, T_b is the threshold base temperature as given in equation 2 and $REFRZ$ (mm°C^{-1}h^{-1}) is considered as the freezing factor. When $we_{i,j}$ is less than a threshold (WETH), ice melt begins. The glacier melt is also computed as given in equation 2 but the snow albedo is replaced with the constant glacier albedo (ALBG). TOPMELT can simulate the full spatial distribution of the SWE, hence the output from the TOPMELT can be compared against the point ground observations and also with snow cover products like MODIS. The snow cover area (SCA) is calculated here by considering a pixel as snow-covered when the simulated SWE exceeds a threshold of 5 mm, based on [11].

2.4 Hydrological model: ICHYMOD

The ICHYMOD hydrological model, developed by [17], combines the TOPMELT snowpack model from [10] with a conceptual rainfall-runoff hydrological model at the basin scale. This model converts snowmelt and excess precipitation into runoff at the basin outlet and includes a snow routine, soil moisture routine, and flow routine. The soil moisture routine uses the Probability Distributed Model (PDM) by [18] to describe the spatial distribution of water storage capacity across the basin. The cubic law storage model is used to route base discharge from groundwater to the catchment outlet, while direct runoff from areas where storage capacity is exceeded is routed through a geomorphology-based distributed unit hydrograph consisting of two linear reservoirs in series. Depending on the glacial till imperviousness, ice melt runoff is either stored in the soil moisture or flows directly to the outlet through two linear reservoirs.

The Hargreaves method [19] is used to compute losses due to evapotranspiration, taking into account the status of soil moisture stored in the PDM. The model represents fast and slow response pathways using storage-based representations and calculates the total basin flow by summing the spatially lumped representations of fast and slow response at the outlet. Drainage to the slow flow path is a function of basin moisture storage, and the slow or base flow component of the total runoff is routed through an exponential store. Direct runoff from areas where storage capacity is exceeded is routed through a geomorphology-based distributed unit hydrograph, using a geomorphologic filter based on a threshold drainage area to distinguish hillslopes and channel networks. The routing time for each site in the basin is evaluated by assigning different typical velocity values to each pixel and classifying them as either hillslope or channel flow, with two velocities used to describe the flow routing.

2.5 Bias adjustment method

To simulate the regional snow dynamics realistically, it is necessary to correct the ERA5 precipitation and temperature. Various bias correction techniques can be used for this purpose. The simplest methods consist of adding the climatological difference between the ERA5 input and the in-situ meteorological station (the 'delta' method). This method is straightforward but implicitly assumes that the variability in ERA5 is unchanged. To overcome these limitations, a quantile-quantile mapping (QM) transformation, which is an empirical transformation [20], can be used. For a given variable, the cumulative density function (CDF) of ERA5 is first matched with the CDF of the references, generating a correction function depending on the quantile. Then, this correction function is used to unbias the ERA5 variable quantile by quantile. QM was applied separately for each month while considering precipitation. To avoid overfitting due to the small sample size of monthly values included in the calibration, the quantile adjustment was computed by considering deciles instead of centiles and applied by linearly interpolating the empirical distribution. A wet-day correction equalizing the fraction of days with precipitation between the observed and the modeled data was applied. The transfer functions were obtained for each ERA5 grid cell from a calibration period (1992-2005) and then applied to the ERA5 variables (precipitation). A validation period (2005-2019) was used

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

to examine the quality of the correction scheme. The implemented QM scheme was based on the R package qmap [21]. In case of temperature, Pettitt test [22] was used to detect changes in ERA5 temperature. Coincidentally, the division of ERA5 temperature into two periods corresponds to the calibration and validation periods selected before the Pettitt test was applied. Based on this evidence, the QM procedure was applied separately for the two periods. Two QM-corrected ERA5 inputs are considered: one where only precipitation is corrected (termed Precipitation-corrected ERA5, PC-ERA5) and another one where both precipitation and temperature are corrected (termed Precipitation-Temperature-Corrected ERA5, PTC-ERA5). Additionally, a blind test to evaluate the effectiveness of using error information in the calibration period, the temperature error from the calibration period is applied for the entire study, which is termed as PTC-cali-ERA5. The results of this study are limited to SWE error calculations after removing seasonality.

2.6 Reference precipitation at ERA5 spatial resolution

To investigate the scale-dependency of ERA5 errors, mean areal precipitation estimates were generated at the ERA5 spatial resolution using station precipitation data. This input data was referred to as Station Input at ERA5 Resolution (SIER). The Thiessen polygon method was used to redistribute the observed hourly station precipitation data over each ERA5 grid footprint. For temperature, the observed dataset was used without any modifications. The computation of errors regarding different ERA5 inputs, for both precipitation and SWE, were compared considering SIER as reference in order to disregard the error associated with the spatial scale of ERA5. However, to quantify the error of SIER itself, the SIER is compared with observed precipitation and its' input SWE simulation.

2.7 Comparison statistics

The comparison is carried out using the Kling-Gupta Efficiency (KGE) [23] as calculated using Equation 5:

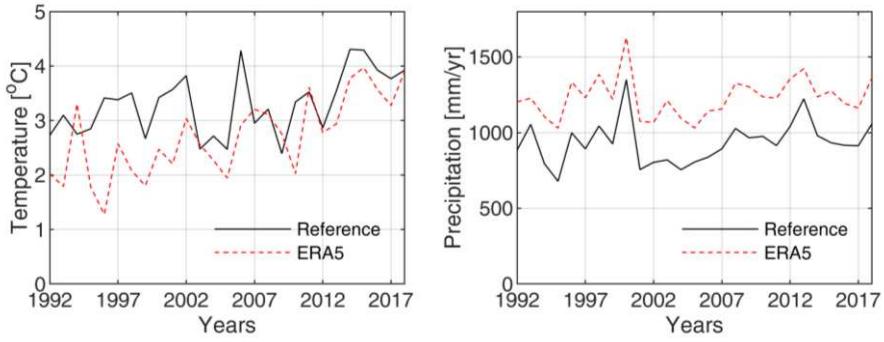
$$\mathrm{KGE} = 1 - \sqrt{(\gamma - 1)^2 + (\beta - 1)^2 + (\alpha - 1)^2} \quad (5)$$

Here, γ represents the correlation component, represented by Pearson's correlation coefficient; β is the bias ratio, represented by the ratio of estimated and reference means; and α is the variability component, represented by the ratio of the estimated and reference coefficients of variation. KGE ranges from negative infinity to one, where a value of one indicates a perfect match between the two series. Values of γ, β, or α smaller/larger than one indicate under/overestimation of the mean value or variability of observations. Note that log coordinates are used for the drainage area on the x-axis of the figures.

3. Results

Figure 2a, b presents the annual mean values of temperature and precipitation, respectively, for the Adige river basin closed at Bronzolo, reporting both the reference and ERA5 values. An almost constant overall bias of 1.36 affects the annual ERA5 precipitation totals, resulting from over-prediction of smaller precipitation and under-prediction of larger precipitation events. In contrast, ERA5 temperature data exhibit a non-stationary behavior, with two different biases observed for the first and second half of the data. Applying the Pettitt test to detect changes in ERA5 temperature, it showed a change point in 2005, with a bias of -0.91 °C for the 1991-2005 period and a bias of -0.29 °C for the period 2005-2019.

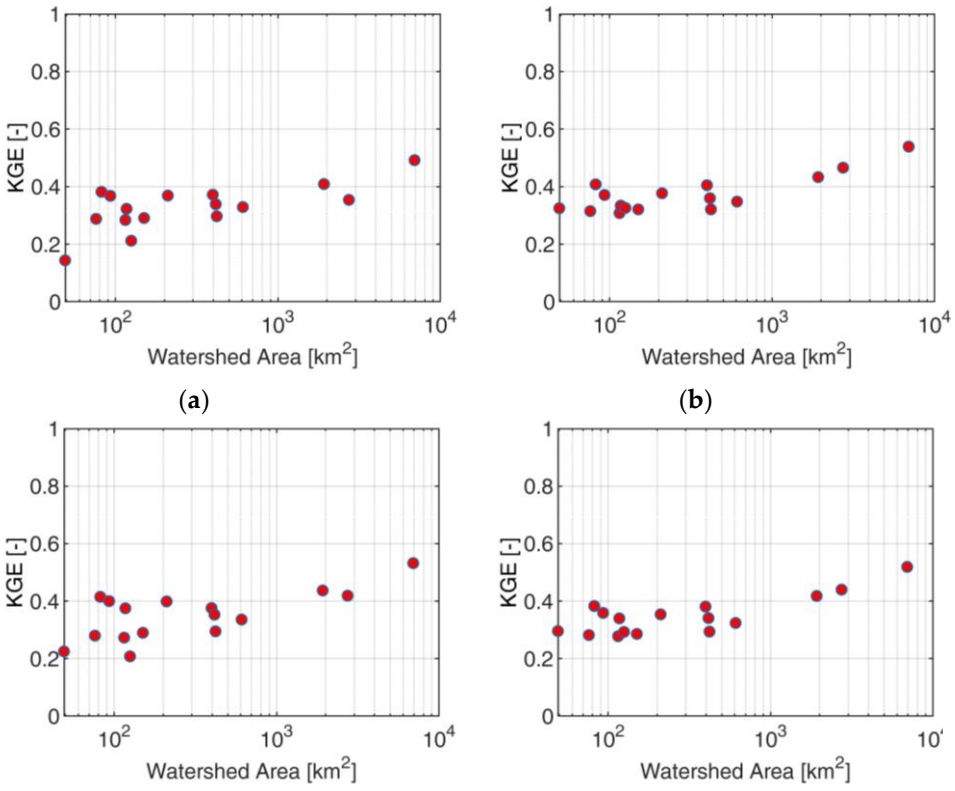
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1



(a) (b)

Figure 2. Adige at Bronzolo: mean annual (a) temperature and (b) precipitation

Figure 3a, b show a comparison of uncorrected and QM-corrected reference and ERA5 precipitation for the calibration period. Hourly mean areal precipitation of the 16 study basins is considered. Figure 3c, d show the same comparison for the validation period.T comparison is carried out using the Kling-Gupta Efficiency (KGE) applied to all data points except intervals where both data sources are zero. The KGE values of ERA5 precipitation for both periods indicate similar performances, which highlights the robustness of the QM procedure employed in this study.



(a) (b)

(c) (d)

Figure 3. KGE of Precipitation (a) ERA5-calibration (b) PC-ERA5-calibration (c) ERA5- validation (d) PC-ERA5-validation

The KGE shows a clear increasing trend with the basin size in all cases, with the smallest basin (Rio Plan at Plan) exhibiting a KGE less than 0.2 and the largest basin (Adige at Bronzolo) showing a KGE just over 0.5. PC-ERA5 demonstrates a slight improvement in KGE performance for both periods, with all basins exhibiting an increase in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

performance, particularly the smaller ones. The QM technique applied to ERA5 precipitation reduces the bias of the ERA5 precipitation while maintaining the correlation intact. However, the overestimation of the variability of the observations only results in a marginal improvement in the KGE of PC-ERA5.

Figure 4a, b report the KGE for SIER precipitation for the calibration and validation period, respectively as compared with the observation. As expected, the figure shows a clear scale dependence in the errors generated by aggregating the reference precipitation at the ERA5 resolution. For basins larger than 1000 km² KGE is close to one, whereas it decreases in a markedly for basin smaller than 1000 km² and even more remarkably for basin less than 100 km², with values around 0.6 for the smallest basin.

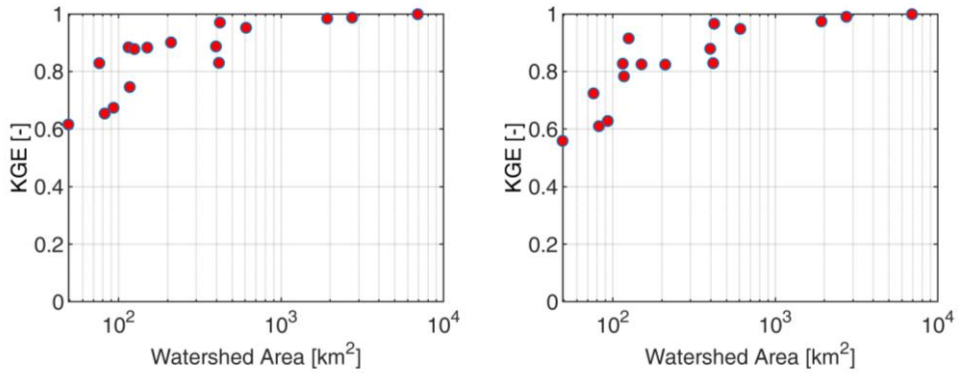


Figure 4. KGE of SIER precipitation with station (a) calibration (b) validation

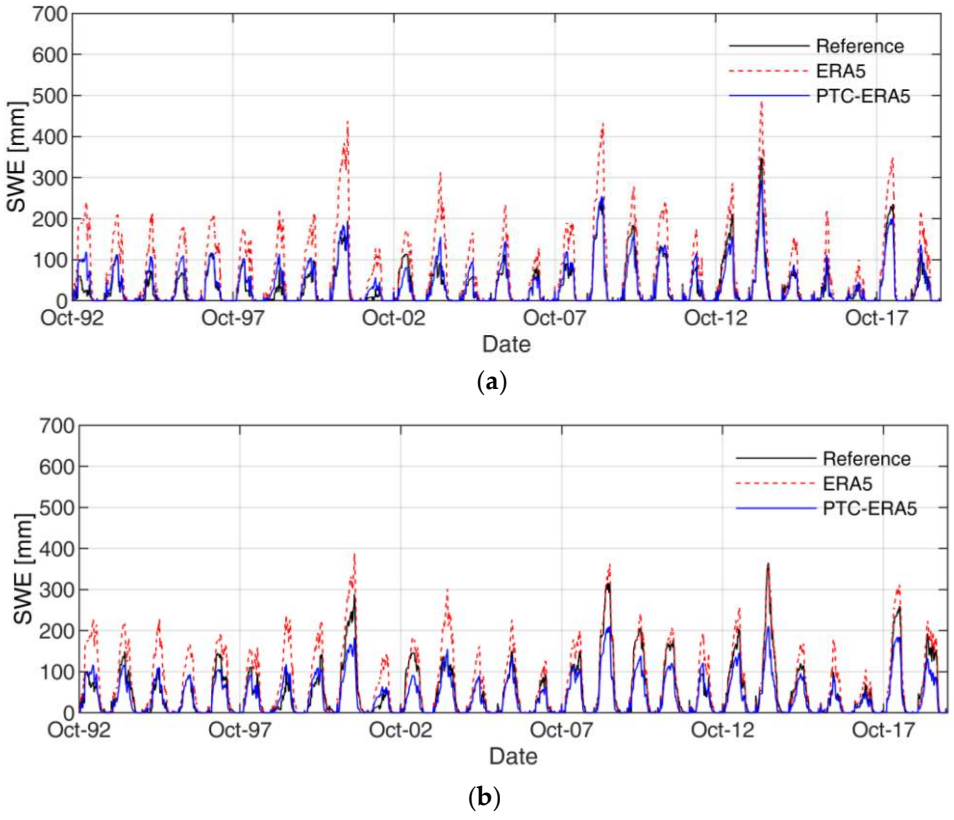


Figure 5. SWE time-series for different inputs for (a) Braies (b) Adige at Bronzolo

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1

Figure 5a, b show the SWE simulations for Braies and Adige at Bronzolo, respectively, using three different inputs: reference, ERA5, and PTC-ERA5. Braies is selected because it shows the worst performance (KGE=-0.351) for the ERA5 SWE simulation for the entire study period. The Adige at Bronzolo, even though it does not have the best performance (KGE=0.471), is selected as it represents all the study basins.

It is evident for both basins that ERA5 overpredicts the SWE, as expected due to its positive precipitation bias characterized by the overestimation of lower intensity precipitation (drizzle problem) and underestimation of higher intensity precipitation. The ERA5 drizzle precipitation produces smaller false snowfall events and when coupled with lower temperature from ERA5, the SWE is sustained in the catchment for a longer period. The QM correction for ERA5 precipitation helps to reduce the wet bias. For PTC-ERA5, the results are comparable to the reference for both basins. In some cases, the SWE peak is underestimated compared to the reference due to the underprediction of larger precipitation events and correction of the temperature errors in ERA5.

Figure 6 shows the mean bias ratio and KGE performance of SWE for different inputs during the validation period after removing seasonality. When the precipitation correction is applied, there is a slight improvement for both the mean bias ratio and KGE. However, the individual temperature correction for the calibration and validation period leads to a significant improvement in both the performances. It should be noted, however, that when the temperature correction of the calibration period is applied to both periods (blind test), the ERA5 performance decreases compared to when the temperature correction is applied to both periods separately.

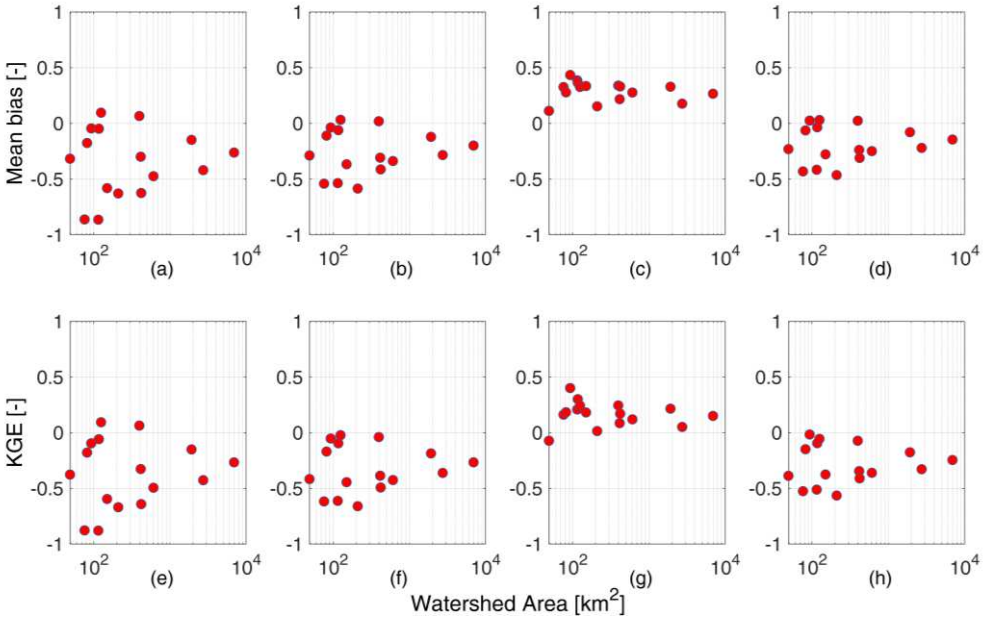
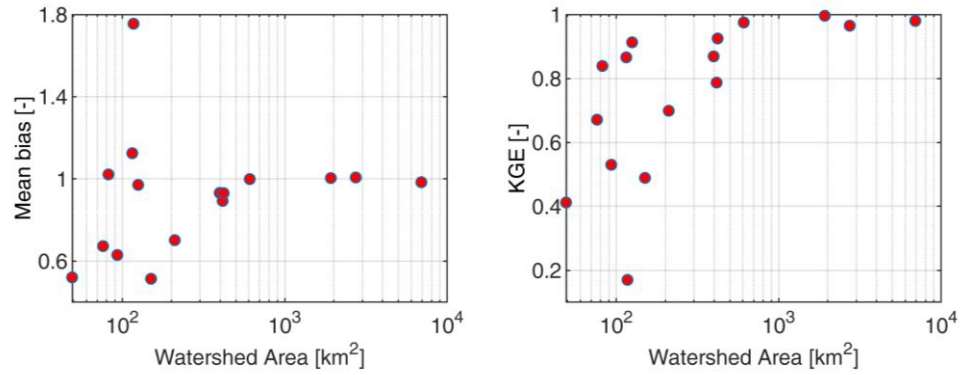


Figure 6. SWE after removing seasonality for validation period:Mean bias ratio for (a)ERA5 (b) PC-ERA5 (c) PTC-ERA5 (d) PTC-cali-ERA5 and KGE for (e) ERA5 (f) PC-ERA5 (g) PTC-ERA5 (h) PTC-cali-ERA5

Figure 7 displays the mean bias ratio and KGE of SWE simulation with SIER input for the validation period compared with the reference SWE considering observation. As seen in Figure 4 with the performance of SIER precipitation, the SIER SWE simulation for smaller basins performs poorly while larger basins remain unaffected.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 May 2023 doi:10.20944/preprints202305.0762.v1



(**a**) (**b**)

Figure 7. SIER input SWE with station (a) Mean bias ratio (b) KGE

Figure 8 reports the Fractional snow cover area (FSCA) simulation for the Adige at Bronzolo for the different inputs considered. FSCA is the percentage of the overall study area covered by snow and helps to understand the impact of various inputs on the spatial snow distribution. The FSCA for ERA5 input simulation is exaggerated, primarily due to the drizzle problem of ERA5 precipitation. As there is a significant number of false precipitation events in ERA5, it artificially increases snowfall in the basin. The lower temperature of ERA5 also delays the melting process, falsely increasing the number of snow cover days. For PTC-ERA5, the FSCA resembles the reference, and the number of days with 100% FSCA is comparable to the reference.

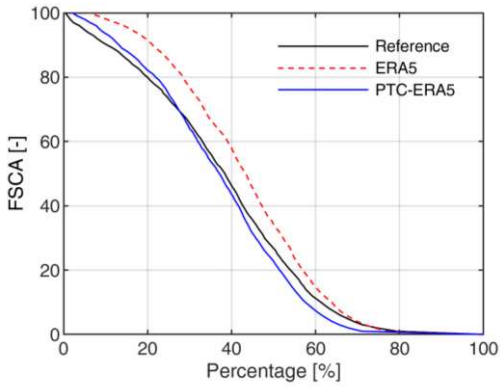


Figure 8. FSCA of Adige at Bronzolo for different inputs

4. Discussion

This study presents a comprehensive evaluation of ERA5 meteorological forcing for simulating SWE using the TOPMELT snow pack model in the Alpine catchment of Northern Italy from 1992 to 2019. The initial findings reveal a positive bias in ERA5 precipitation resulting from lower precipitation events and a negative ERA5 temperature bias compared to observations. The study employs SIER as a reference to comprehend the impact of ERA5 spatial scale on both precipitation and SWE simulation. SIER results show that ERA5's spatial scale affects only smaller basins, whereas the effect is negligible for larger basins. Both the precipitation and SWE performance of SIER decline for basin sizes below 1000 km², with a distinct effect observed for basins with an area less than 100 km².

The precipitation bias in ERA5 causes overestimation of snow, while lower temperature further delays the melt in the study area. The ERA5 temperature bias's impact is

 doi:10.20944/preprints202305.0762.v1

evident in SWE calibration and validation periods. As the precipitation bias remains relatively constant across different years, the validation period's relatively warmer temperature exhibits comparable SWE peaks, whereas the colder ERA5 temperature in calibration periods results in largely overestimated SWE peaks. The study considers using QM to correct ERA5 meteorological variables. Correcting precipitation and temperature largely reduces the bias in SWE simulation, but it is only after seasonality removal that the temperature correction's distinct role is observed. Correcting precipitation for some smaller basins marginally improves the KGE value, but larger basins do not benefit from this correction. However, temperature correction significantly improves the KGE performance for a range of basin sizes, with smaller basins showing comparatively greater performance enhancement. Although the temperature's role in SWE is evident regardless of basin size, it should be noted that temperature is crucial for SWE simulation in smaller basins. The blind test of temperature correction highlights the importance of applying different correction factors for different periods, as ERA5 displays distinct temperature behavior in the study period's first and second halves. As small temperature variations can significantly alter the snow process in high-altitude, smaller catchments, proper correction is recommended not only for precipitation but, most importantly, for ERA5's temperature dataset.

5. Conclusions

In conclusion, this research has identified the limitations and appropriate use of ERA5 meteorological variables for simulating snow processes in watersheds of varying sizes. It has also highlighted the challenge of addressing the ERA5 drizzle issue through monthly quantile mapping and emphasized the significance of temperature correction for accurate snow simulation. The study has quantified the static error caused by the spatial scale of ERA5 and its impact on precipitation and snow simulation for watersheds of different sizes. Despite these limitations, ERA5 reanalysis can serve as a valuable alternative dataset for simulating snow processes in areas with limited data availability, provided that appropriate corrections are made to the precipitation and temperature data before use in real-world hydrological applications.

Furthermore, the methodology developed in this study can be easily adapted to other similar watersheds with slight modifications to the TOPMELT model. This work also creates opportunities to explore the use of ERA5 dataset with other modelling approaches for simulating various hydrological processes. As a continuation of this study, the researchers plan to investigate the performance of the ERA5 dataset along with the ICHYMOD hydrological model in predicting flood events in the Adige River basin.

Author Contributions: Conceptualization, Susen Shrestha, Mattia Zaramella and Marco Borga; Formal analysis, Susen Shrestha; Investigation, Susen Shrestha; Methodology, Susen Shrestha, Mattia Zaramella and Marco Borga; Project administration, Felix Greifeneder; Software, Mattia Zaramella; Supervision, Mattia Callegari, Felix Greifeneder and Marco Borga; Validation, Susen Shrestha, Mattia Zaramella, Mattia Callegari, Felix Greifeneder and Marco Borga; Visualization, Susen Shrestha; Writing – original draft, Susen Shrestha and Marco Borga; Writing – review & editing, Susen Shrestha, Mattia Zaramella, Mattia Callegari and Marco Borga.

Funding: The publication of this research was funded by NEXOGENESIS (grant agreement number 101003881), which is a 4-year European collaborative project financed by the European Commission under the H2020 programme. While the whole work was carried out under the SECLI-FIRM project.

Data Availability Statement: ERA5 data used in this study are openly available in the Climate DataStore of the Copernicus Climate Change Service (https://cds.climate.copernicus.eu/#!/home, accessed on 5 May, 2023)

Acknowledgments: This research was the part of the European Union's Horizon 2020 research and innovation program under grant agreement No 776868 (SECLI-FIRM project).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. R. C. Bales, N. P. Molotch, T. H. Painter, M. D. Dettinger, R. Rice, and J. Dozier, "Mountain hydrology of the western United States," *Water Resour. Res.*, vol. 42, no. 8, 2006.
2. Y. Gao, H. Xie, N. Lu, T. Yao, and T. Liang, "Toward advanced daily cloud-free snow cover and snow water equivalent products from Terra–Aqua MODIS and Aqua AMSR-E measurements," *J. Hydrol.*, vol. 385, no. 1–4, pp. 23–35, 2010.
3. K. A. Dressler, G. H. Leavesley, R. C. Bales, and S. R. Fassnacht, "Dr," *Hydrol. Process. An Int. J.*, vol. 20, no. 4, pp. 673–688, 2006.
4. H. Hersbach *et al.*, "The ERA5 global reanalysis," *Q. J. R. Meteorol. Soc.*, vol. 146, no. 730, pp. 1999–2049, 2020.
5. M. W. Baba, A. Boudhar, S. Gascoin, L. Hanich, A. Marchane, and A. Chehbouni, "Assessment of MERRA-2 and ERA5 to Model the Snow Water Equivalent in the High Atlas (1981–2019)," *Water*, vol. 13, no. 7, p. 890, 2021.
6. E. Alonso-González, E. Gutmann, K. Aalstad, A. Fayad, M. Bouchet, and S. Gascoin, "Snowpack dynamics in the Lebanese mountains from quasi-dynamically downscaled ERA5 reanalysis updated by assimilating remotely sensed fractional snow-covered area," *Hydrol. Earth Syst. Sci.*, vol. 25, no. 8, pp. 4455–4471, 2021.
7. M. Tarek, F. P. Brissette, and R. Arsenault, "Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America," *Hydrol. Earth Syst. Sci.*, vol. 24, no. 5, pp. 2527–2544, 2020.
8. M. Raimonet *et al.*, "Evaluation of gridded meteorological datasets for hydrological modeling," *J. Hydrometeorol.*, vol. 18, no. 11, pp. 3027–3041, 2017.
9. S. Terzago *et al.*, "Sensitivity of snow models to the accuracy of meteorological forcings in mountain environments," *Hydrol. Earth Syst. Sci.*, vol. 24, no. 8, pp. 4061–4090, 2020.
10. M. Zaramella, M. Borga, D. Zoccatelli, and L. Carturan, "TOPMELT 1.0: a topography-based distribution function approach to snowmelt simulation for hydrological modelling at basin scale," *Geosci. Model Dev.*, vol. 12, no. 12, pp. 5251–5265, 2019.
11. N. Di Marco, M. Righetti, D. Avesani, M. Zaramella, C. Notarnicola, and M. Borga, "Comparison of MODIS and model-derived snow-covered areas: Impact of land use and solar illumination conditions," *Geosciences*, vol. 10, no. 4, p. 134, 2020.
12. N. Di Marco, D. Avesani, M. Righetti, M. Zaramella, B. Majone, and M. Borga, "Reducing hydrological modelling uncertainty by using MODIS snow cover data and a topography-based distribution function snowmelt model," *J. Hydrol.*, p. 126020, 2021.
13. G. Formetta, F. Marra, E. Dallan, M. Zaramella, and M. Borga, "Differential orographic impact on sub-hourly, hourly, and daily extreme precipitation," *Adv. Water Resour.*, vol. 159, p. 104085, 2022.
14. A. Galletti, D. Avesani, A. Bellin, and B. Majone, "Detailed simulation of storage hydropower systems in a large Alpine watershed.," in *Journal of Hydrology*, 2021, vol. 603.
15. L. Laiti *et al.*, "Testing the hydrological coherence of high-resolution gridded precipitation and temperature data sets," *Water Resour. Res.*, vol. 54, no. 3, pp. 1999–2016, 2018.
16. B. W. Brock, I. C. Willis, and M. J. Sharp, "Measurement and parameterization of albedo variations at Haut Glacier d'Arolla, Switzerland," *J. Glaciol.*, vol. 46, no. 155, pp. 675–688, 2000.
17. D. Norbiato, M. Borga, S. Degli Esposti, E. Gaume, and S. Anquetin, "Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins," *J. Hydrol.*, vol. 362, no. 3–4, pp. 274–290, 2008.
18. R. J. Moore, "The PDM rainfall-runoff model," *Hydrol. Earth Syst. Sci. Discuss.*, vol. 11, no. 1, pp. 483–499, 2007.

 doi:10.20944/preprints202305.0762.v1

19. G. H. Hargreaves and R. G. Allen, "History and evaluation of Hargreaves evapotranspiration equation," *J. Irrig. Drain. Eng.*, vol. 129, no. 1, pp. 53–63, 2003.

20. H. A. Panofsky and G. W. Brier, *Some applications of statistics to meteorology*. Earth and Mineral Sciences Continuing Education, College of Earth and …, 1968.

21. L. Gudmundsson, J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, "Downscaling RCM precipitation to the station scale using statistical transformations–a comparison of methods," *Hydrol. Earth Syst. Sci.*, vol. 16, no. 9, pp. 3383–3390, 2012.

22. A. N. Pettitt, "A non-parametric approach to the change-point problem," *J. R. Stat. Soc. Ser. C (Applied Stat.*, vol. 28, no. 2, pp. 126–135, 1979.

23. H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez, "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling," *J. Hydrol.*, vol. 377, no. 1–2, pp. 80–91, 2009.