UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Administrative unit: University of Padova

Department: Land, Environment, Agriculture and Forestry (LEAF)

PhD Program: Land, Environment, Resources and Health (LERH)

Batch: XXXIV

SEASONAL STREAMFLOW FORECASTS IN AN ALPINE BASIN BY INTEGRATION OF REANALYSIS DATA AND SEASONAL CLIMATE FORECASTS

Thesis financially supported by EURAC Research; Bolzano, Italy

PhD Program Coordinator: Prof. Marco Borga
Supervisor: Prof. Marco Borga

Ph.D. Candidate : Susen Shrestha

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento Territorio e Sistemi Agro-Forestali (TESAF)

CORSO DI DOTTORATO DI RICERCA: Land, Environment, Resources, Health (LERH)

Ciclo: XXXIV

Previsioni stagionali dei deflussi in un bacino alpino mediante integrazione di dati di rianalisi e previsioni climatiche stagionali

Thesis financially supported by EURAC Research; Bolzano, Italy

Coordinatore: Prof. Marco Borga
Supervisore: Prof. Marco Borga

Dottorando : Susen Shrestha

Table of content

List of Figures

List of Tables

List of acronyms

ECMWF: European Centre for Medium-Range Weather Forecasts

ERA5: ECMWF 5th Reanalysis

QM: Quantile Mapping

KGE: Kling-Gupta Efficiency

SWE: Snow water equivalent

ESP: Extended Streamflow prediction

NSW: National Weather service

IHC: Initial hydrological condition

BC: Bias correction

NSE: Nash Sutcliff efficiency

CRPS: Continuous ranked probability score

CRPSS: Continuous ranked probability skills score

SPRerr: Spread error

WRF: Weather and forecasting

ICAR: Intermediate Complexity Atmospheric Research Model

C3S : Copernicus Climate Change Service

DTM: Digital terrain model

CMF: Combined melting factor

RMF: Rain melt factor

NMF: Night melt factor

MODIS: Moderate Resolution Imaging Spectroradiometer

SCA: Snow cover area

PDM: Probability distributed model

Cdf: Cumulative density function

SIER: Station input at ERA5 resolution

FSCA: Fractional snow cover area

PC-ERA5: Precipitation corrected ERA5

PTC-ERA5: Precipitation temperature corrected ERA5

NWP: Numerical weather prediction

SF: Seasonal forecast

LT: Lead times

D: Days

W: Weeks

M: Months

IQR: Interquartile range

BF: Baseflow

TF: Total flow

SEAS5: System 5

Summary

Seasonal forecasts provide information on climate conditions several months ahead and therefore they could represent a valuable support for water resources decision making, warning systems as well as for the optimization of industry and energy sectors. However, seasonal forecasts can be affected by systematic biases and have horizontal resolutions which are typically coarser than the spatial scales of the practical hydrological applications. For these reasons, the reliability of forecasts needs to be carefully assessed before applying and interpreting them for specific applications. Reanalysis data such ERA5 represent a very interesting dataset for correcting seasonal forecasts. Building on these premises, this research aims to i) evaluate the use of ERA5 reanalysis data for snow and runoff prediction in small and medium-size Alpine basins, and ii) assess the use of System-5 ECMWF seasonal forecasts for runoff prediction in the same basins. The research is focused on the upper Adige river basin, where 16 basins ranging from 49to 6924 km^2 are considered, and on the use of the ICHYMOD hydrological model, which includes a snow accumulation and melt module.

A first study aims to assess to impact of ERA5 errors on snow water equivalent (SWE) by considering different catchment spatial scales. Reference SWE is obtained by using ICHYMOD fed with station precipitation and temperature data. It is found that ERA5 precipitation is characterized by an overall overestimation (around 36%), due to the well know drizzle problem which inflates lower precipitation amounts. ERA5 temperatures, on the other side, display a colder bias (around -0.65°C). These biases propagate to produce positive biases in SWE. The ERA5 precipitation and temperature biases are removed by using the Quantile Mapping (QM) technique. The QM technique is able to remove effectively the bias associated to the ERA5 precipitation, whereas improvements measured by the KGE statistic are only slight. On the other hand, the temperature bias removal is effective, and the SWE simulations after the ERA5 precipitation and temperature correction show comparable performances with the reference.

A second study aims to analyze the effectiveness of using ERA5 precipitation and temperature data for runoff and flood simulations in small to medium size alpine basins. It is found that the ERA5 precipitation bias is composed by an overestimation of small and very small precipitation events (the drizzle effect) and an underestimation of high intensity precipitation events. Whereas the precipitation bias translates in general poor runoff simulation, which are effectively corrected by the application of the QM technique, the analysis of the flood events makes a more complex, and interesting, case. Indeed, the 'drizzle' effect generates an overestimation of the soil moisture conditions, including those preceding flood events. This compensates in an effective way the negative bias in intense precipitation, in a way that makes the ERA5-simulated flood almost unbiased, event though with relatively poor simulation efficiency. It is not surprising that the application of the QM technique is not able to improve the simulation quality for flood events.

The third and last study focused on the examination of System-5 ECMWF seasonal forecasts for runoff prediction over the Upper Adige River basin. It is found that System-5 also displayed a

positive bias irrespective of the lead time. After the correction with quantile mapping considering the stations, the system-5 showed comparable precipitation values. Moreover, this correction assisted in a better runoff forecast as compared to the traditional Extended Streamflow Prediction (ESP) for lead time concerning up to 2 months. For a longer lead times, ESP and System 5 show comparable statistics. In the case of baseflow forecasting, there was a gain of one-week lead time irrespective of the forecasting method. Lastly, since the performance of the forecasting methodology depended on the initialization month, the role of the initial watershed wetness state was found to be crucial for both the forecasting scheme.

Riassunto

Le previsioni stagionali forniscono informazioni sulle condizioni climatiche a distanza di diversi mesi e pertanto possono rappresentare un valido supporto per il processo decisionale nell'ambito delle risorse idriche, i sistemi di allerta nonché per l'ottimizzazione dei settori industriali ed energetici. Tuttavia, le previsioni stagionali possono essere influenzate da distorsioni sistematiche e avere risoluzioni orizzontali che sono tipicamente più grossolane delle scale spaziali delle applicazioni idrologiche pratiche. Per questi motivi, l'affidabilità delle previsioni deve essere attentamente valutata prima di applicarle e interpretarle per applicazioni specifiche. I dati di rianalisi come ERA5 rappresentano un set di dati molto interessante per la correzione delle previsioni stagionali. Partendo da queste premesse, questa ricerca mira a i) valutare l'uso dei dati di rianalisi ERA5 per la previsione di neve e deflusso nei bacini alpini di piccole e medie dimensioni e ii) valutare l'uso delle previsioni stagionali ECMWF System-5 per la previsione del deflusso negli stessi bacini. La ricerca è focalizzata sull'alto bacino del fiume Adige, dove si considerano 16 bacini compresi tra 40 e 6900 km^2, e sull'utilizzo del modello idrologico ICHYMOD, che include un modulo di accumulo e scioglimento della neve.

Un primo studio mira a valutare l'impatto degli errori ERA5 sulla equivalente in acqua della neve (SWE) considerando diverse scale spaziali di bacino. Il valore SWE di riferimento è ottenuto utilizzando ICHYMOD alimentato con i dati di precipitazione e temperatura della stazione. Si è riscontrato che le precipitazioni di ERA5 sono caratterizzate da una sovrastima complessiva (circa il 40%), a causa del ben noto 'drizzle effect' che porta a sovrastime delle precipitazioni più deboli. Le temperature ERA5, d'altra parte, mostrano bias freddo (circa – 0.5 °C). Questi errori si propagano per produrre distorsioni positive nella predizione di SWE. Le distorsioni di precipitazione e temperatura ERA5 vengono rimosse utilizzando la tecnica Quantile Mapping (QM). La tecnica QM è in grado di rimuovere efficacemente il bias associato alla precipitazione ERA5, mentre i miglioramenti misurati dalla statistica KGE sono solo lievi. D'altra parte, la rimozione del bias di temperatura è efficace e le simulazioni SWE dopo la precipitazione ERA5 e la correzione della temperatura mostrano prestazioni comparabili con il riferimento.

Un secondo studio mira ad analizzare l'efficacia dell'utilizzo dei dati di precipitazione e temperatura ERA5 per simulazioni di deflusso e di piena in bacini alpini di piccole e medie dimensioni. Si è

riscontrato che il bias di precipitazione ERA5 è composto da una sovrastima di eventi di precipitazione piccoli e molto piccoli (‘drizzle effect’) e una sottostima di eventi di precipitazione ad alta intensità. Mentre la distorsione delle precipitazioni si traduce in una simulazione generale del deflusso scadente, che viene corretta efficacemente tramite l'applicazione della tecnica QM, l'analisi degli eventi alluvionali costituisce un caso più complesso e interessante. Il ”drizzle effect’, infatti, genera una sovrastima delle condizioni di umidità del suolo, comprese quelle precedenti gli eventi alluvionali. Ciò compensa in modo efficace la distorsione negativa nelle precipitazioni intense, in un modo che rende l'inondazione simulata da ERA5 quasi indistorta, anche se con un'efficienza di simulazione relativamente scarsa. Non sorprende che l'applicazione della tecnica QM non sia, in questo caso, in grado di migliorare la qualità della simulazione per gli eventi di piena.

Il terzo e ultimo studio è concentrato sull'esame delle previsioni stagionali dell'ECMWF System-5 per la previsione dei deflussi nell’alto bacino dell’Adige. Si è riscontrato che anche i dati da System-5 sono caratterizzati da una distorsione positiva indipendentemente dal lead time. Dopo la correzione con QM considerando i dati da stazione, System-5 ha mostrato valori di precipitazione comparabili. Inoltre, questa correzione ha contribuito ad una migliore previsione del deflusso rispetto alla tradizionale Extended Streamflow Prediction (ESP) per orizzonti temporali di previsione fino a 2 mesi. Per orizzonti temporali più lunghi, ESP e System 5 mostrano statistiche comparabili. Nel caso della previsione del flusso di base, i risultati mostrano un guadagno di una settimana di anticipo indipendentemente dal metodo di previsione. Infine, la forte dipendenza della qualità delle previsioni dal mese di inizializzazione indica il ruolo cruciale delle condizioni iniziali per lo schema di previsione.

1. Introduction

There has been a remarkable shift in the temperature variability (Gourdji et al., 2013; Rahmstorf & Coumou, 2011) and extremes and patterns of precipitation (C. Liu & Allan, 2013) due to recent global warming. The human-induced global warming is increasing at 0.2°C per decade accordingly to the intergovernmental Panel on Climate Change (IPCC, 2018). The change in mean and the trend of extreme temperatures are primarily attributed to greenhouse gases (Christidis et al., 2012) and are expected to continue in the coming decades. These trends as compared to climatological averages place a threat both to the society and the environment (Rummukainen, 2012). The European continent also remains vulnerable to the effect of changing climate in the coming decades (Kreibich et al., 2014), in which, the southern countries will be exposed to drought and wildfires due to the increasing frequency of heatwaves (Guerreiro et al., 2018; Jacob et al., 2014). On the other hand, extreme precipitations will be frequented especially in central (Zeder & Fischer, 2020) and northern (Casanueva et al., 2014) regions of Europe, which may increase flood hazard.

There are various branches of the economy that face repercussions due to weather and climate variability, with the energy industry recognized as one of the most vulnerable sectors (Wieczorek-Kosmala, 2020). Both short and long-term climatic variability has an impact on demand, supply, transport, distribution of energy (Dubus et al., 2018). Most notably, the effect of meteorological events is distinct for renewable energy sources such as wind, solar, and hydropower. Given the nature of climate changes, it is no longer sufficient to rely on the historical observation as traditionally done (Costa-Saura et al., 2022). Rather a robust climatic prediction is required to facilitate the users, managers and decision-makers (Bruno Soares et al., 2018).

Climate reanalyses combine past observations with models to generate consistent time series of multiple climate variables (Caroletti et al., 2019). Reanalyses are among the most-used datasets in the geophysical sciences. They provide a comprehensive description of the observed climate as it has evolved during recent decades, on 3D grids at sub-daily intervals (Yang et al., 2022). However, reanalyses data are still characterized by errors, which require ground-based observations as a reference for validation puroposes However, the downside of ground-based observations is its' sparse density; especially in complex orography, which hinders the assessment of reanalyses data (Kidd et al., 2017).

A reanalysis is a systematic approach that combines observation and climate model results, to provide a dynamically consistent gridded dataset spanning decades (Stocker et al., 2013). Reanalysis products are considered to assess the impact of changes in observation data, evaluate the progress in modeling and assimilation capabilities, and acquire state-of-the-art climatologies to quantity forecast error anomalies (Hersbach et al., 2020). Reanalysis also supports a wide range of applications ranging from inter-governmental assessments of climate change to understanding specific local weather (Bell et al., 2021) as well as applications in business such as energy and agriculture (Gelaro et al., 2017). Some of the examples of the latest reanalysis products are MERRA-

2 reanalysis (Gelaro et al., 2017) from the NASA GMAO, JRA-55 ((Kobayashi et al., 2015)) developed by the JMA and CFSR (version 2) produced by NCEP (Saha et al.,2014), ERA5 produced by ECMWF.
ECMWF has ample experience in production of reanalysis dataset and ERA5 is the fifth state-of-art iteration of their product. The latest product prior to ERA5 was ERA-Interim which was operational from 2006 until 2019. ERA5 as compared to its' predecessor and its' counterparts has a better horizontal and vertical resolution. It also benefits from the 10 years of model and data assimilation (DA) developments, it has enhanced number of output (like 100m wind) and has hourly high-resolution outputs. It is also timely available within 5 days of real time, while, due to the presence of the 10 ensemble component available at half the resolution and 3 hourly output, ERA5 allows for the uncertainty estimates as well (Bell et al., 2021). The data assimilation of the different variables involves satellite observations, in-situ dataset and ground based radar for rainfall since 2009. Overall ERA5 provides the land and oceanic climate variables with data availability spanning from 1979 to the present, having a spatial resolution of 0.25 degrees and temporal resolution of 1 hour at the global scale (Hersbach et al., 2020).
Each of the reanalysis products provide only the climatic information about the present and past but not of the future. Hence, forecast information is required to understand how the climate is going to evolve in the future based on observations as initial condition. Seasonal climate forecast lies between the short-term weather forecast and longer timescale (like climate change prediction and inter-annual prediction) (Kirtman et al., 2013). The seasonal forecast provides a probabilistic spread (ensemble prediction) of the average conditions (like temperature and rainfall) and its' evolution in the future, which can range from 1-month lead time upto a year (Rickards et al., 2014). There has been burgeoning interest in the application of seasonal forecast primarily due to development of enhanced numerical computing (Capa-Morocho et al., 2016). Some of the climate prediction centers around the world, which generate climate seasonal forecasts at global, regional and national scale, are International Research Institute for Climate and Society (IRI), ClimatePrediction Center (CPC), UK Met office, SEAS5 ECMWF.
Seasonal forecasts provide information on climate conditions several months ahead and therefore they could represent a valuable support for water resources decision making, warning systems as well as for the optimization of industry and energy sectors. However, seasonal forecast systems can be affected by systematic biases and have horizontal resolutions which are typically coarser than the spatial scales of the practical applications. For this reason, the reliability of forecasts needs to be carefully assessed before applying and interpreting them for specific applications. In addition, the use of post-processing approaches is recommended in order to improve the representativeness of the large-scale predictions of regional and local climate conditions. The development and evaluation downscaling and bias-correction procedures aiming at improving the skills of the forecasts and the quality of derived climate services is currently an open research field.
Extended streamflow prediction (ESP) (Day, 1985) method pioneered by the United States National Weather Service (NWS) in the 1970s has found extensive use in seasonal forecasting in the recent years as well due to its' simple and cost-effective approach (Harrigan, et al., 2018). ESP forecast ensembles are obtained by forcing a hydrological model - initialized with the initial hydrological condition (IHC) - with ensembles of historical observation (Day, 1985). Even though ESP forecast suffers from absence of information about the future climatic condition, nevertheless, it should be

considered along with the other state-of-art forecasting method as a benchmark to quantify the added skills coming from available forecasting products (Crochemore et al., 2017, Arnal et al., 2018). On the other hand, the seasonal forecasting products suffer from important bias when compared with the observations regardless of their data assimilation technique (Randall et al., 2007). These biases could originate from the imperfect conceptualization, discretization, and spatial averaging within grid cells (Soriano et al., 2019). As a result, the raw forecast has a certain amount of risk for industrial applications (e.g., energy, hydrology, agriculture), and post-processing is required to get finer resolution output for impact assessment and statistical features that are comparable to the observations (Torralba et al., 2017). In order to remove the systematic model errors, the available technique ranges from simpler approaches like bias correction (BC), working directly on interest variable, to more complex statistical downscaling, building on large –scale predictors (Maraun et al., 2010). To make the raw model output consistent with the local climatology, the BC method maps it to the analogous long historical observational reference. These strategies, which do not take into account the temporal correlation between model outputs and observations, ranging from basic adjustments in the mean and/or variance to more advanced quantile mapping (QM) alternatives, that can accommodate higher-order moments or even the entire distribution (Manzanas et al., 2019). Since the QM can be applied for the dataset that does not require a standard (e.g. Gaussian) distribution - regular criteria for the operational agencies - its' application has gained recent interest (Bedia et al., 2018).

There have been multiple runoff forecasting studies done in the context of Europe, for instance, the UK (Harrigan et al., 2018), France (Crochemore et al., 2016), Austria (Santos et al., 2021), Switzerland (Bogner et al., 2018). However, the latest iteration of the forecast products ECMWF system-5, except for a few studies (Sutanto & Van Lanen, 2021, Sánchez-García et al., 2022), has found limited implications in the context of European alpine basins. Most crucially, their application for baseflow forecasting has not been fully realized. Baseflow is termed as the contribution of groundwater that sustains the river discharge between precipitation events (Brutsaert, 2008). Since the baseflow remains substantial for the ecohydrological sustainability, farming, and hydropower production during the seasons with no rain, its' skillful forecast could help to mitigate (Ayers et al., 2021) the repercussion related to the low flow in the European alpines.

To address the gap in the research, the prime objective of this work is to exploit the meteorological information of ECMWF system-5 for the total and baseflow forecast for the Upper Adige river basin, situated in the northern Italian Alps. The application of ICHYMOD hydrological model for the ECMWF hindcast period of 1993-2016 with up to 6 months of the forecast horizon will be extensively reviewed in the context of baseflow. Prior to its' application, the bias associated with the ECWMF dataset is corrected with quantile mapping (QM) with the help of an observational dataset (STAT-ECMWF) and ERA5 dataset (ERA5-ECMWF). The performance of ERA5 for snow and runoff simulation will also be evaluated initially in order to understand how the ERA5 meteorological error propagates to the hydrological variables. To remove the error associated with the ICHYMOD hydrological model and to have seamless runoff dataset in case of missing runoff observation, the runoff simulation with observation as input will be used as a reference for error metrics. Further, the catchment initial states are based on the reference simulation for driving 25-ensemble members of the ECMWF variables for forecasting. The possible improvement in the baseflow forecasting for

the STAT-ECMWF datasets against the traditional ESP method will be quantified using metrics like bias, NSE, CRPSS, and Spread error with climatology as reference.

1.1 Research questions and objectives

Given the premise reported in the introduction, the main objective of this thesis is to forecast the seasonal runoff in the alpine watershed of the Northern Italy with the integration of the bias corrected ECMWF system-5 seasonal meteorological forecast dataset to the ICHYMOD hydrological model.

The specific objective are as follows:

- Scale dependence of errors in snow water equivalent simulations using ERA5 reanalysis over alpine basins
- Assessing ERA5 forcing data for flood modelling in alpine basins: impact of bias correction
- Sub-seasonal streamflow predictions in alpine catchments by combining numerical weather models and reanalysis data.

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

Shrestha, S., [1], Zaramella, M. [1], Callegari, M. [2], Greifeneder, F. [2], Borga, M. [1]

[1]Department of Land, Environment, Agriculture, and Forestry, University of Padova, Italy

[2]EURAC, Institute for Earth Observation, Bolzano, Italy

2.1 Abstract

This paper evaluated ERA5 reanalysis as a potential reference dataset for snow water equivalent (SWE) modelling in the Alpine basins with respect to increasing basin size. This is achieved by considering the ERA5 precipitation and temperatures as proxies for observations to simulate the semi-distributed TOPMELT model over basins ranging from 49 km^2 to 6924 km^2 for the Upper Adige river basin situated in the Eastern Italian Alps. This study focused, firstly, to compute the errors in the ERA5 meteorological variables and, secondly, to assess its' effect on SWE computation for the period of 1992 to 2019. We observed that ERA5 precipitation products generally overestimated the observation with over prediction of low-intensity rainfall ('drizzle problem') and underprediction of the high rainfall intensity, while the ERA5 temperature underestimated the observation. To overcome the data unavailability of SWE in the region and to remove the error associated with the

TOPMELT snow model, a model simulation of station precipitation and temperature is taken as a reference. The ERA5 drizzle led to producing fictitious low-intensity snowfall events, and when, these events were coupled with the colder ERA5 temperature, there was delayed snowmelt in the study basin. As a result, the SWE and consequently, the number of snow cover days factitiously increased over the study area. The precipitation bias linked with ERA5 was removed using the Quantile Mapping (QM) technique. For the QM, the period from 1992-to 2005 was considered as the calibration period and a monthly quantile mapping function was obtained using the station data. The same mapping function was applied for the validation period from 2005 to 2019. Even though there was an insignificant improvement in the KGE value, the QM technique was able to remove the bias associated with ERA5 precipitation. On the other hand, the use of the Pettitt test assisted in detecting the change in temperature bias for calibration and validation period. The temperature error was removed by applying unique additive constant for the two periods. The SWE model simulated after the ERA5 precipitation and temperature correction showed comparable performance with the reference and the number of fictitious snow cover days decreased considerably. The use of monthly moving mean was also considered to remove the seasonality associated with the SWE, which showed the important role of temperature, rather than precipitation, correction to simulate the SWE for the range of basin size. However, the temperature correction shows comparatively better performance for the SWE simulation concerning smaller basins as opposed to the larger basins.

2.2 Introduction

Mountain snowpack and its melt dominate the surface hydrology of many regions, with implications for water supply, hydropower, ecological processes, weather, and regional and global climate (Bales et al., 2006; Gao et al., 2010). 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 (Dressler, et al., 2006). 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 (Hersbach et al., 2020) as an input to snow pack models, has been shown to provide reasonable simulations of SWE for mountain basins (Baba et al., 2021, Alonso-González et al., 2021).

(Baba et al., 2021) compared the performance of MErRA-2 and ERA-5 reanalysis meteorological forcing with the distributed SnowModel in the high Atlas region. Both the reanalysis product 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, (Alonso-González et al., 2021) 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 WRF to develop a 1km regional-scale snow reanalysis. The results showed a good temporal and spatial correlation of the snow variables with the MODIS fractional 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 (Tarek et al., 2020). 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 (Alonso-González et al., 2021; Raimonet et al., 2017; Terzago et al., 2020).

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 km^2-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 (Zaramella et al., 2019) 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 (Zaramella et al., 2019). 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 (Di Marco et al., 2021, 2020).

2.3 Materials and methods

2.3.1 Study area and data

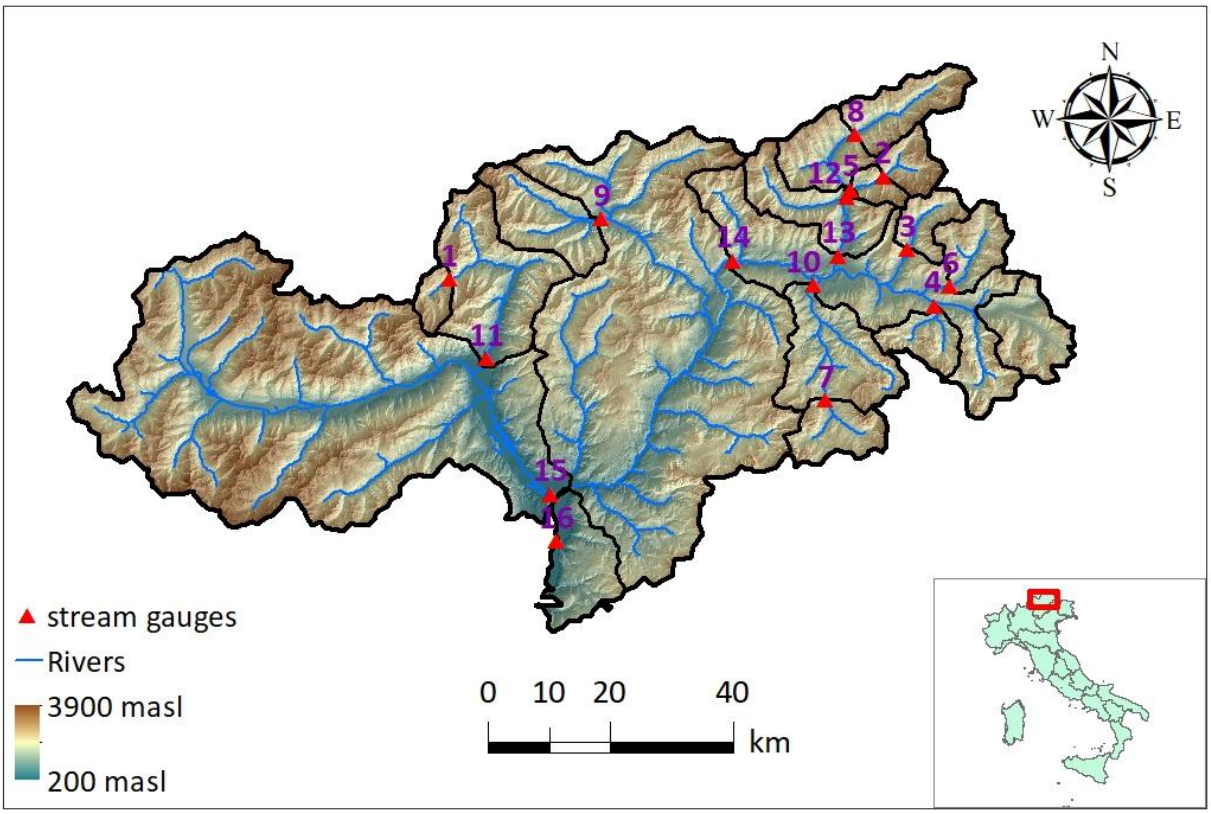

Figure 2.1 The upper Adige River basin closed at Bronzolo

The study basins are located in the upper Adige river basin closed at Bronzolo, in the Eastern Italian Alps (Figure 2.1). This is an alpine catchment with a drainage area of approximately 6924 km^2. 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 (Formetta et al., 2022) with the precipitation ranging from 500mm in the northwest region to 1600mm in the southern region (Galletti et al., 2019). During the winter season, the precipitation is stored as snow and the streamflow is minimum. The streamflow shows two maxima: first due to snowmelt in the early summer and the second due to intense storms in autumn (Laiti et al., 2018). The main agricultural areas in the northern part of the catchment are located in the Venosta valleys, and the cultivation comprise 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 km^2) and 124 temperature gauges (1 per 55 km^2) covering the study region. The hourly temperature, precipitation, and runoff data from 1991 till 2019 is made available from the Hydrographic Office of Bozen, Bolzano. To assess the performance of the model at varying basin size areas, sixteen watersheds with drainage area ranging from 49 km^2 to 6924 km^2 were selected (Figure 2.1, Table 2.1). These watersheds were selected for the analysis because the

relative minor impact of operations from artificial reservoirs allowed in these cases the use of a hydrological model and the ensuing comparison of simulated versus observed discharges.

SN	Name	Elevation range(m)	Mean elevation(m)	Area (Km^2)
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

Table 2.1 Drainage area and elevation of the study basins

2.3.2 ERA-5 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 (Hersbach et al., 2020). 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 variables with data availability spanning from 1979 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 on ERA5 is reported 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 cells that lie within and around 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. 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.

2.3.3 Ground data

The study basin has a dense network of meteorological stations. There are 88 rain gauges (1 per 72 km^2) and 124 temperature gauges (1 per 55 km^2) within the study basin. The hourly temperature,

precipitation, and runoff data from 1991 till 2019 is collected from the Hydrographic Office of Bozen, Bolzano. If there is a missing hourly dataset, the model only considers only the stations for both precipitation and temperature lapse rate calculation. All the analysis focus the period of 1992-2019, however in case of runoff data, there was missing continuous dataset for some basins hence the calibration and validation of the runoff result was limited to the year available in the table 3.2.

S.N.	Basins	Year
1	Rio Plan	1994-2019
2	Rio Riva a Seghe	2001-2008
3	Braies	2013-2019
4	Rio Riva a Caminata	2003-2019
5	Gadera a Pedraces	2011-2011
6	Gadera Mantana	1992-2018
7	Passirio	2013-2019

Table 2.2: List of basins with limited data

2.3.4 The snowpack model: TOPMELT

The snowpack model used in this work is TOPMELT (Zaramella et al., 2019). 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 (Di Marco et al., 2021; Zaramella et al., 2019). The TOPMELT is integrated with the ICHYMOD semi-distributed basin-scale conceptual rainfall-runoff hydrological model. The ICHYMOD model consists of a snow routine, a soil moisture routine, and a flow routine. For the computation of losses due to evapotranspiration, the Hargreaves method (Hargreaves & Samani, 1982) is used. The detail of the TOPMELT and ICHYMOD model is outlined in annex A and annex B respectively. The parameters of TOPMELT has been calibrated and validated using the observed runoff which has been described in annex C.

In this study the model has not been calibrated or validated with the ground based snow stations but relies on the work done by (Di Marco et al., 2020). In the study the author used the TOPMELT model to simulate the snow depth in the western part of the upper Adige river basin and made the comparison with the twenty-eight ground-based snow depth stations. The analysis showed a median NSE value equal to 0.91 across the stations. This signifies that TOPMELT is a robust tool to simulate snowpack distribution and its related complex spatial variability over the UARB study basin. The analysis carried out by (Di Marco et al., 2020) should be taken as a reference for the validation of TOPMELT performance as in this paper, the TOPMELT is directly applied for the SWE simulation without validating with any observed snow variables.

2.3.5 Bias adjustment method

To simulate realistically the regional snow dynamics, ERA5 precipitation and temperature should be corrected. Different bias correction techniques may be used. The simplest methods consist of adding the climatological difference between ERA5 input and the reference data (the ‘delta’ method). This method is straightforward, but implicitly assume that the variability in ERA5 is unchanged. A quantile–quantile mapping (QM, hereinafter) transformation (the empirical

transformation of (Panofsky & Brier, 1968) may be used to overcome these limitations. 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.

More specifically, ERA5 and reference distributions are matched by establishing a quantile-dependent correction function that translates simulated quantiles into their reference counterparts. This function is then used to translate the modeled time series into bias-adjusted values with a distribution representative of the reference data. QM was applied separately for each month. The transfer functions were obtained for each ERA5 grid cell from a calibration period (1992-2005) and then applied to the ERA5 variables (precipitation and temperature). In order 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 modelled data was applied. A validation period (2005-2019) was used to examine the quality of the correction scheme. The simulation for the year 1991-1992 is only considered as a TOPMELT model setup period. The implemented QM scheme was based on the R package qmap (Gudmundsson et al., 2012). The use of the monthly quantile was considered to provide an opportunity for the replicability of this work in data scare area as well.

2.3.6 Reference precipitation at ERA5 spatial resolution

In order to examine the scale dependence of ERA5 errors, the station precipitation data were used to generate estimates of mean areal precipitation at the ERA5 spatial resolution. This input data was termed Station Input at ERA5 Resolution (SIER). To achieve this, the Thiessen polygon method was used to redistribute the observed hourly station precipitation data over each ERA5 grid footprint.

2.3.7 Removing seasonality for SWE statistical analysis

Seasonality provides a strong signal to SWE simulations which may obscure the impact of ERA5 scale and errors on SWE simulations. SWE seasonality was identified by calculating a time-averaged daily cycle of SWE simulations based on a moving window of 30 days across the whole period of analysis. Then, the seasonality was removed by subtracting the SWE seasonality to each SWE data.

2.4 Results

Figure 2.2 a,b shows the annual mean values of temperature and precipitation, respectively, reporting both the reference and the ERA5 values. Whereas an almost constant overall relative bias of 1.40 is affecting annual ERA5 precipitation totals, generated by over-prediction of small totals and under-prediction of large totals, a non-stationary behavior is affecting ERA5 temperature data, with the temperature showing two different biases for the first and for the second half of the data. The application of the Pettitt test (Pettitt, 1979) to detect changes in ERA5 temperature bias led to rejecting the null hypothesis of stationarity, showing change point on 2005, with a bias of -0.9098 °C for the 1991-2005 period and a bias of -0.2841 °C for the period 2005-2019. Out of mere

coincidence, the division of ERA5 temperature into two periods coincides with the calibration and validation period, which was selected prior to the application of the Pettitt test. Based on this evidence, the QM procedure was applied separately for the two periods. In the following, two QM-corrected ERA5 input are considered: one where only the 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).

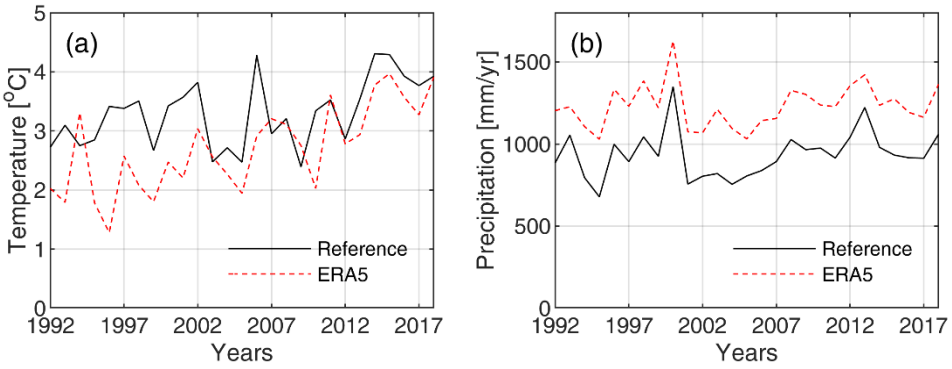

Figure 2.2 Adige at Bronzolo: mean annual (a) temperature and (b) precipitation

Figure 2.3 a,b shows a comparison between reference and ERA5 precipitation, both uncorrected and QM-corrected, respectively, for the calibration period. The comparison is carried out by considering the hourly mean areal precipitation of the 16 different study basin by using the Kling-Gupta Efficiency (KGE) (Gupta et al., 2009) applied to all data points but the intervals where both the data sources are zero. The Kling-Gupta Efficiency (KGE) is as follows:

$$KGE = 1 - \sqrt{(\gamma - 1)^2 + (\beta - 1)^2 + (\alpha - 1)^2} \qquad \text{Equation 2.1}$$

where γ is 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. If two series exactly match, the KGE is one. γ, β or α value smaller/larger than one indicates that the mean value or variability of observations is underestimated/overestimated. Please note the use of log coordinates for the drainage area on the x axis of the figures. Figure 2.3 c,d reports the same comparison for the validation period.

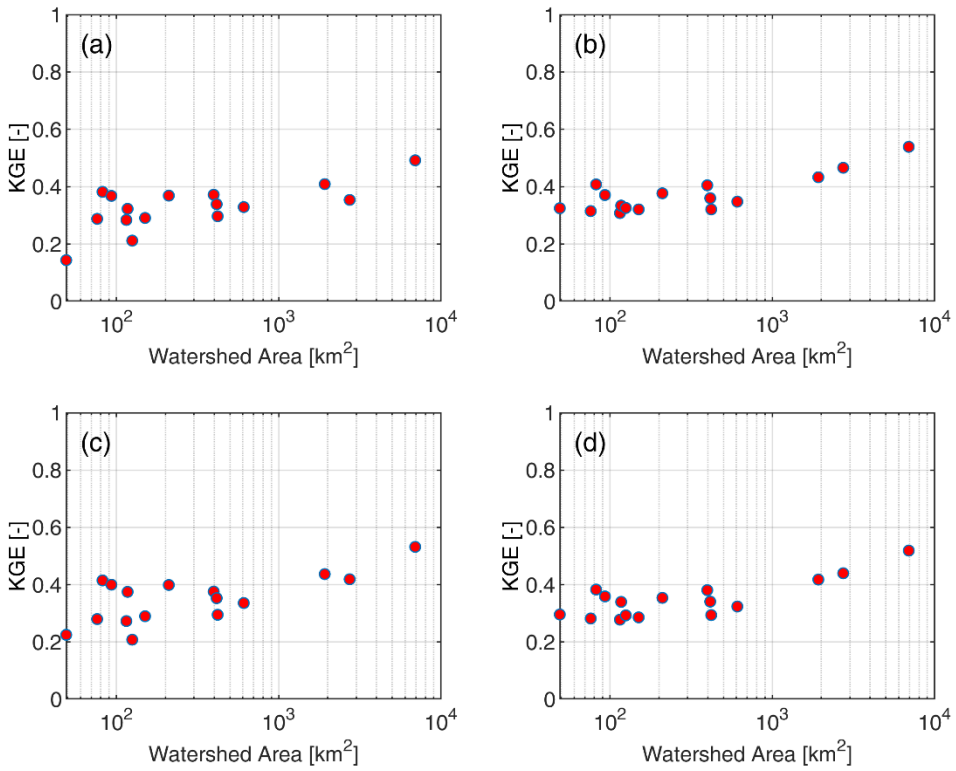

Figure 2.3: KGE of Precipitation (a) ERA5-calibration (b) PC-ERA5-calibration (c) ERA5- validation (d) PC-ERA5-validation

The KGE of ERA5 precipitation for both periods shows similar performances, which points to the robustness of the QM procedure applied here. There is a clear trend of KGE increasing with the basin size in all the cases, with KGE around or less than 0.2 for the smallest basin (Plan at Plan) and around or over 0.5 for the largest basin (Adige at Bronzolo). The PC-ERA5 shows a slight improvement in the KGE performance for both periods. All basins show an increase in performance, with the smallest ones showing a larger gain, as expected. The QM technique applied to ERA5 precipitation helps to decrease the bias of the ERA5 precipitation while keeping the correlation intact. Nevertheless, the overestimation of the variability of the observations causes the KGE of PC-ERA5 to improve only slightly.

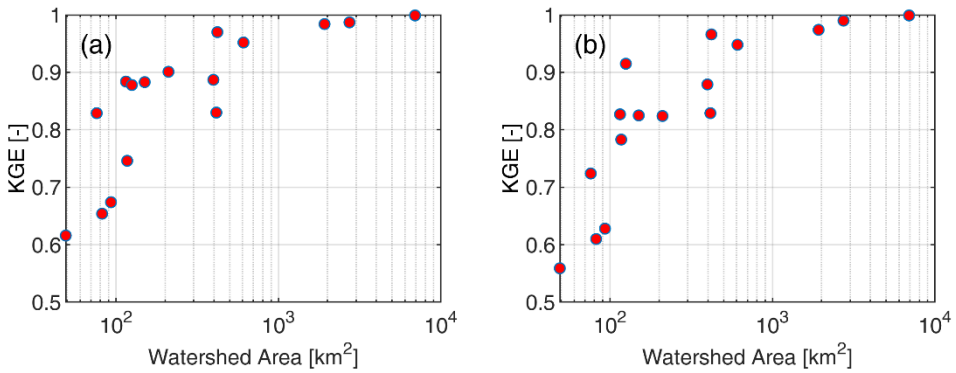

Figure 2.4 KGE of SIER precipitation with station (a) calibration (b) validation

Figure 2.4 a, b report the KGE for SIER precipitation for the calibration and validation period, respectively. 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 600 km^2 KGE is close to one, whereas it decreases in a marked way for basin smaller than 600 km^2 and even more remarkably for basin less than 100 km^2, with values around or less than 0.6 for the smallest basin.

Since it was evident that ERA5 precipitation was marred with bias, it was important to understand where this bias was arising from. The Quantile-Quantile (QQ) plot of ERA5 hourly precipitation as shown in Figure 2.5 was able to highlights this inherent ERA5 drizzle problem. Here only the study basin of Braies and Adige a Bronzolo is plotted but the error is consistent in all the other basins as well. As shown in QQ plot, the lower intensity rainfall less than 1mm is always overpredicted by ERA5 while the larger precipitation events is under-represented by ERA5. This figure sufficiently proves that the ERA5 precipitation bias was originating from lower precipitation events even though the event precipitation was under-represented.

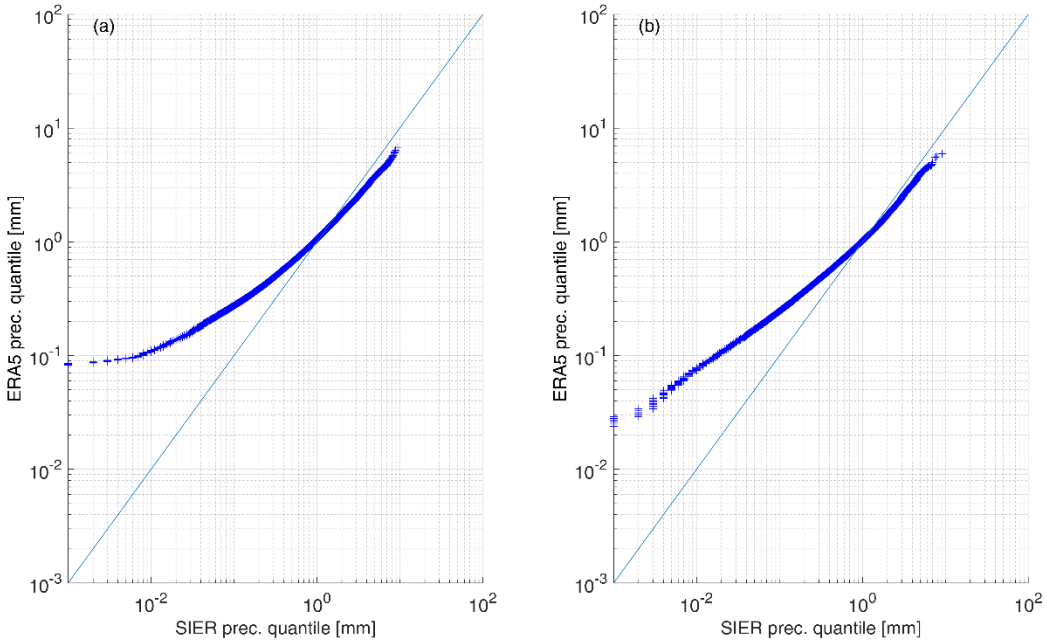

Figure 2.5 Quantile-Quantile plot for ERA5 hourly precipitation for (a) Braies (b) Adige a Bronzolo

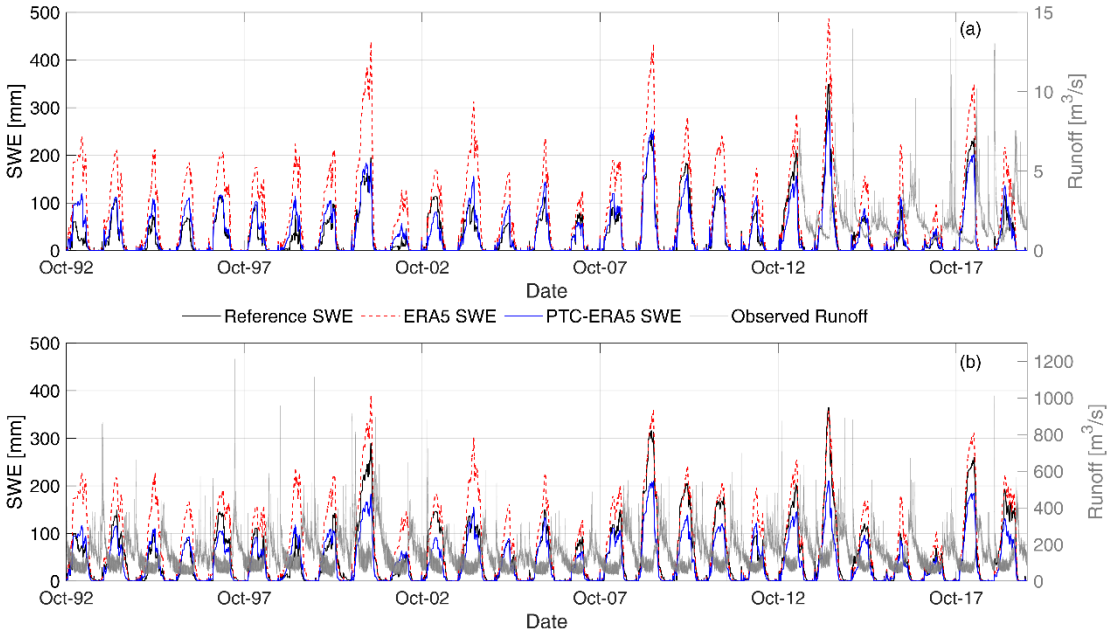

Figure 2.6 SWE and observed runoff time-series for different inputs for (a) Braies (b) Adige at Bronzolo

Figure 2.6 a,b shows observed runoff and the SWE simulations for Braies and Adige at Bronzolo, respectively, using three different input: reference, ERA5, and PTC-ERA5. The figure helps to visualize the timing of water availability in the study basin as either water stored in form of snow water equivalent or as the river discharge output. It is clear that during the period of high SWE, there is minimum runoff while runoff increases with decrease in SWE. Even though the effect of precipitation is not available here, the influence of SWE as a short-term storage to replenish the alpine river system is distinctively visible here. Braies is selected here because it shows the worst performance (KGE=-0.351) for the ERA5 SWE simulation for the whole study period. The Adige at Bronzolo even though does not have the best performance (KGE= 0.471), is selected as it represents the whole study basins. For both the basins, it is evident that the ERA5 is overpredicting the SWE. This is expected as the ERA5 has a positive precipitation bias, which is 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 it is 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 the PTC-ERA5, the results are comparable with the reference for both the basins. The SWE peak in some cases is underestimated compared to the reference due to the underprediction of the larger precipitation events and correction of the temperature errors in ERA5.

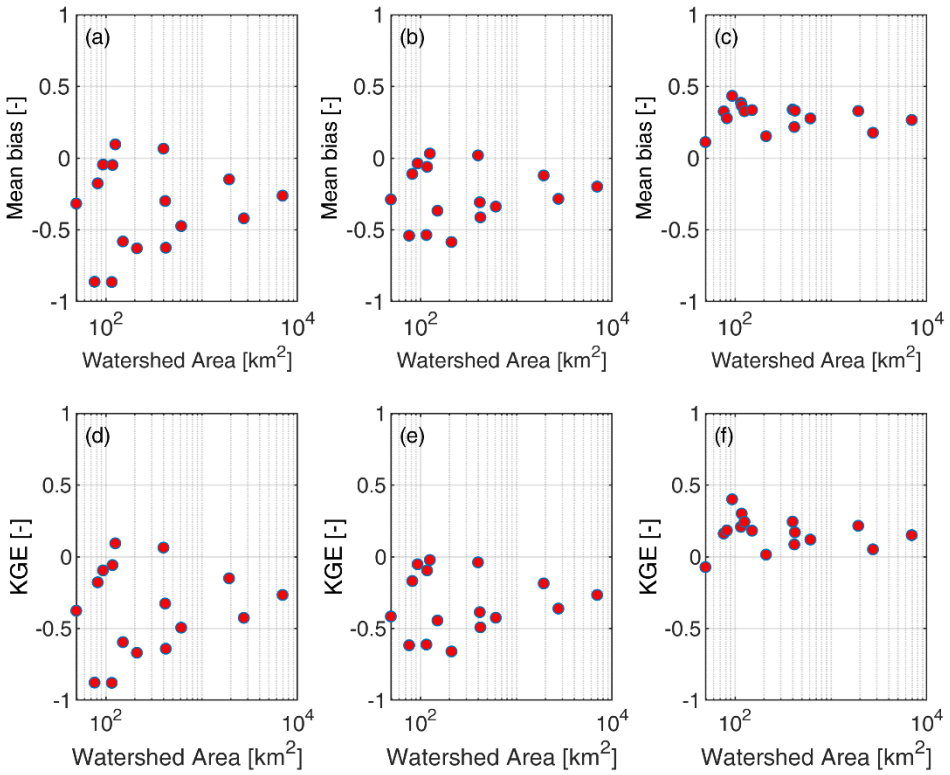

Figure 2.7 SWE after removing seasonality for the validation period: Mean bias ratio for (a) ERA5 (b) PC-ERA5 (c) PTC-ERA5. KGE for (d) ERA5 (e) PC-ERA5 (f) PTC-ERA5

Figure 2.7 reports the mean bias ratio and KGE performance of SWE for different inputs and for the validation period after removing seasonality. There is only a slight improvement for both the mean bias ratio and KGE when only the precipitation correction is applied. On the other hand, the application of the temperature correction, leads to a significant improvement in both the performances.

Figure 2.8 represents the mean bias ratio and KGE SWE simulation with SIER-input for the validation period. The results agree almost perfectly with the SIER precipitation performances reported in Figure 2.4. The SWE simulation for the smaller basins performs poorly while the larger basins remain unaffected.

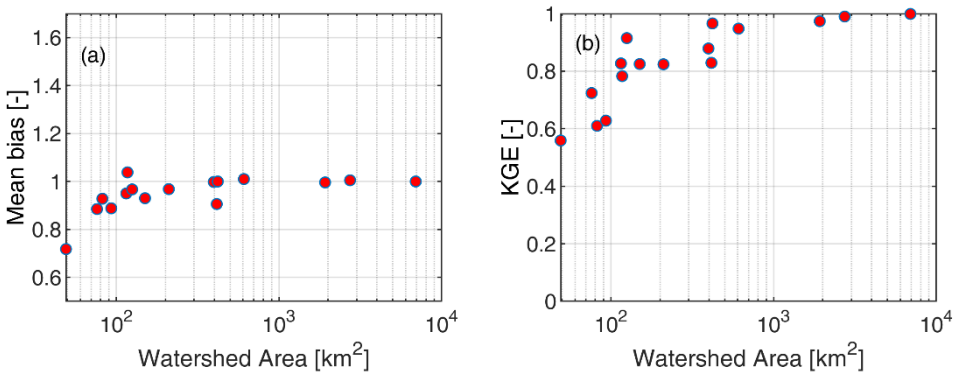

Figure 2.8 SIER input SWE with station (a) Mean bias ratio (b) KGE

Figure 2.9 reports the 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 it helps to understand the impact of the various input on the spatial snow distribution. The FSCA for ERA5 input simulation is exaggerated, which is primarily due to the drizzle problem of ERA5 precipitation. As there is a significant number of false precipitation in ERA5, it helps to artificially increase the snowfall in the basin. The lower temperature of ERA5 also helps to delay the melting process and as a result the number of snow cover days falsely increases. For the PTC-ERA5, the FSCA resembles the reference, and the number of days with 100% FSCA is also comparable with the reference.

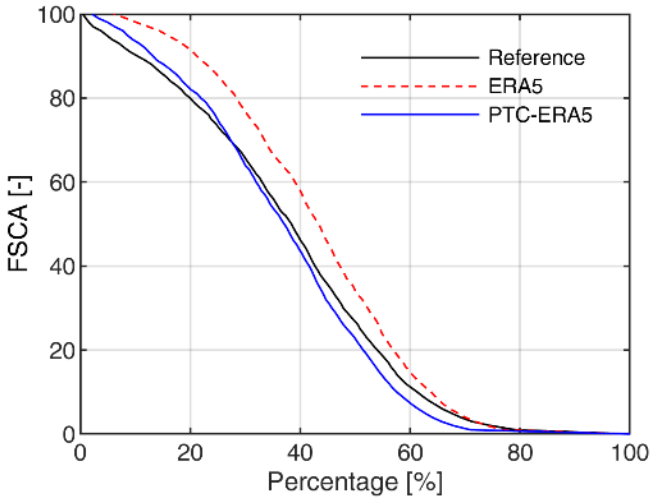

Figure 2.9 FSCA of Adige at Bronzolo for different inputs

2.5 Discussion

This study provides an evaluation of ERA5 meteorological forcing to simulate SWE using TOPMELT snow pack model in the Alpine catchment of Northern Italy from 1992 to 2019. The initial study shows a positive bias in ERA5 precipitation originating from the lower precipitation events and a negative ERA5 temperature bias as compared to the observation. The SIER is considered a reference to understand the role of the ERA5 spatial scale on both precipitation and SWE simulation. The SIER results (precipitation and SWE) show the spatial scale of ERA5 affecting only the smaller basins while for the larger basins the effect of spatial scale is negligible. The precipitation bias in ERA5 helps to

overestimate the snow and the lower temperature further delays the melt in the study area. The effect of ERA5 temperature bias is seen in the SWE concerning calibration and validation period, since the precipitation bias remains relatively similar for different years, the comparative warmer period in validation shows comparable SWE peaks while for calibration periods the peaks are largely overestimated due to colder ERA5 temperature. The use of QM is considered to correct the ERA5 precipitation while an additive constant, unique to the calibration and validation period, is used to correct the ERA5 temperature error. The correction of precipitation and temperature largely reduces the bias in SWE simulation but it is only after removing the seasonality, that the distinct role of the temperature correction is seen. The correction of precipitation for a few smaller basins shows little improvement in KGE value but the larger basins do not benefit from this correction. However in the case of temperature correction, there is an improvement in the KGE performance for the range of basin size, notably, the increase in performance for the smaller basins is comparatively higher than the larger basins. Even though the role temperature for the SWE is evident irrespective of the basin size, it must be duly noted that it is much more crucial for SWE simulation concerning smaller basins. As a small variation in temperature could largely alter the snow process in the high altitude smaller catchment, a proper correction is recommended not only for the precipitation but most importantly, also for the temperature of the ERA5 dataset.

2.6 Conclusion

To conclude, this research work identified the shortcoming and application of ERA-5 meteorological variables for snow simulation. It also shed light on the difficulty in removing the ERA5 drizzle problem by using monthly quantile mapping and also highlighted the importance of temperature correction for the simulation of snow processes. Nonetheless, the ERA5 reanalysis has the potential to be used as a substitute dataset in places where there is sparse data availability. However, it is of utmost importance to correct the precipitation and temperature data prior to its application to solve real-world hydrological problems.

Finally, this work could be easily reproduced to other similar watersheds with few adjustments in the TOPMELT model. This work also opens door to ample opportunities to consider the ERA5 dataset along with the modelling approach to simulate other hydrological processes. As the continuation of this work, we look forward to applying the ERA5 dataset along with ICHYMOD hydrological model to access its performance during the flood events in the Adige river basin.

2.7 References

Alonso-González, E., Gutmann, E., Aalstad, K., Fayad, A., Bouchet, M., & Gascoin, S. (2021). Snowpack dynamics in the Lebanese mountains from quasi-dynamically downscaled ERA5 reanalysis updated by assimilating remotely sensed fractional snow-covered area. Hydrology and Earth System Sciences, 25(8), 4455–4471.

Baba, M. W., Boudhar, A., Gascoin, S., Hanich, L., Marchane, A., & Chehbouni, A. (2021). Assessment of MERRA-2 and ERA5 to Model the Snow Water Equivalent in the High Atlas (1981–2019). Water, 13(7), 890.

Bales, R. C., Molotch, N. P., Painter, T. H., Dettinger, M. D., Rice, R., & Dozier, J. (2006). Mountain hydrology of the western United States. Water Resources Research, 42(8).

Brock, B. W., Willis, I. C., & Sharp, M. J. (2000). Measurement and parameterization of albedo variations at Haut Glacier d'Arolla, Switzerland. Journal of Glaciology, 46(155), 675–688.

Da Ros, D., & Borga, M. (1997). Use of digital elevation model data for the derivation of the geomorphological instantaneous unit hydrograph. Hydrological Processes, 11(1), 13–33.

Di Marco, N., Avesani, D., Righetti, M., Zaramella, M., Majone, B., & Borga, M. (2021). Reducing hydrological modelling uncertainty by using MODIS snow cover data and a topography-based distribution function snowmelt model. Journal of Hydrology, 126020.

Di Marco, N., Righetti, M., Avesani, D., Zaramella, M., Notarnicola, C., & Borga, M. (2020). Comparison of MODIS and model-derived snow-covered areas: Impact of land use and solar illumination conditions. Geosciences, 10(4), 134.

Dressler, K. A., Leavesley, G. H., Bales, R. C., & Fassnacht, S. R. (2006). Dr. Hydrological Processes: An International Journal, 20(4), 673–688.

Formetta, G., Marra, F., Dallan, E., Zaramella, M., & Borga, M. (2022). Differential orographic impact on sub-hourly, hourly, and daily extreme precipitation. Advances in Water Resources, 159, 104085.

Galletti, A., Avesani, D., Bellin, A., & Majone, B. (2019). Detailed simulation of storage hydropower systems in a large Alpine watershed. Geophysical Research Abstracts, 21.

Gao, Y., Xie, H., Lu, N., Yao, T., & Liang, T. (2010). Toward advanced daily cloud-free snow cover and snow water equivalent products from Terra–Aqua MODIS and Aqua AMSR-E measurements. Journal of Hydrology, 385(1–4), 23–35.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Downscaling RCM precipitation to the station scale using statistical transformations–a comparison of methods. Hydrology and Earth System Sciences, 16(9), 3383–3390.

Hargreaves, G. H., & Allen, R. G. (2003). History and evaluation of Hargreaves evapotranspiration equation. Journal of Irrigation and Drainage Engineering, 129(1), 53–63.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., … Schepers, D. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society,

146(730), 1999–2049.

Laiti, L., Mallucci, S., Piccolroaz, S., Bellin, A., Zardi, D., Fiori, A., … Majone, B. (2018). Testing the hydrological coherence of high-resolution gridded precipitation and temperature data sets. Water Resources Research, 54(3), 1999–2016.

Moore, R. J. (2007). The PDM rainfall-runoff model. Hydrology and Earth System Sciences Discussions, 11(1), 483–499.

Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E., & Anquetin, S. (2008). Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins. Journal of Hydrology, 362(3–4), 274–290.

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

Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. Journal of the Royal Statistical Society: Series C (Applied Statistics), 28(2), 126–135.

Raimonet, M., Oudin, L., Thieu, V., Silvestre, M., Vautard, R., Rabouille, C., & Le Moigne, P. (2017). Evaluation of gridded meteorological datasets for hydrological modeling. Journal of Hydrometeorology, 18(11), 3027–3041.

Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrology and Earth System Sciences, 24(5), 2527–2544.

Terzago, S., Andreoli, V., Arduini, G., Balsamo, G., Campo, L., Cassardo, C., … von Hardenberg, J. (2020). Sensitivity of snow models to the accuracy of meteorological forcings in mountain environments. Hydrology and Earth System Sciences, 24(8), 4061–4090.

Zaramella, M., Borga, M., Zoccatelli, D., & Carturan, L. (2019). TOPMELT 1.0: a topography-based distribution function approach to snowmelt simulation for hydrological modelling at basin scale. Geoscientific Model Development, 12(12), 5251–5265.

3. Assessing ERA5 forcing data for flood modelling in alpine basins: impact of bias correction

Shrestha, S.[1], Zaramella, M. [1], Crespi, A. [2], Callegari, M. [2], Greifeneder, F.[2], Borga, M.[1]

[1]Department of Land, Environment, Agriculture, and Forestry, University of Padova, Italy

[2]EURAC, Institute for Earth Observation, Bolzano, Italy

3.1 Abstract

The European Center for Medium-Range Weather Forecasts (ECMWF) has recently released its most advanced reanalysis product, the ERA5 dataset. It was designed and generated with methods giving it multiple advantages over the previous release, the ERA-Interim reanalysis product. Notably, it has a finer spatial resolution, is archived at the hourly time step, uses a more advanced assimilation system and includes more sources of data. This paper aims to evaluate the ERA5 reanalysis as a potential reference dataset for hydrological modelling by considering the ERA5 precipitation and temperatures as proxies for observations using a semi-distributed hydrological model. The study further investigates the usefulness of the ERA5 products for the flood modelling based on the increasing size ranging from 40km^2 to 6900 km^2 over the Upper Adige river basin in the Eastern Italian Alps. This study shows that ERA5-based precipitation product is affected by a significant bias which translates to biased runoff at all spatial scales considered in the study. We observed that ERA5 precipitation product generally overestimate low-intensity rainfall and underestimate high rainfall intensity in the region. We analysed how this affects simulation of annual max floods over the study area. The results show that flood simulations are in general surprisingly good for the basin size larger than 600 km^2, as they result from the combination of two cascading errors: i) overestimation of the soil moisture conditions at the start of the event and ii) the underestimation of the event forcing rainfall. However, the better flood results concerning larger basins are immediately lost when a unitary basin size is considered for the analysis. Even though the analysis of the bias corrected ERA5 variables show improved hourly runoff performance, the correction leads to further decrease in the flood simulation capabilities of ERA5 irrespective of the basin size. Differences between ERA5 and observation datasets are mostly linked to precipitation, as temperature only marginally influences the hydrological simulation outcomes.

3.2 Introduction

River flooding is a natural process but poses a significant socioeconomic hazard, causing human distress and damage to properties and infrastructure. In Europe, floods caused approximately EUR 147 billion in economic damage between 1980 and 2019 (EEA, 2021). Moreover, the economic

losses associated with flood events have been on the increase in the past decades (since 1970), partly due to changing weather patterns (IPCC, 2014) but mainly driven by socioeconomic developments such as population growth, increasing wealth and ongoing urbanization in flood-prone areas (Barredo, 2009, Bouwer, 2011, Koks et al., 2014). Flood prediction is therefore needed for a range of societal purposes. Both continuous and flood-event hydrological models are often used for flood estimation and for real-time flood forecasting. The specific issue in this case is whether the models and the input data are able to faithfully reproduce the processes associated with changing runoff generation mechanisms.

Over the last decades, the emergence of near-global and high-resolution gridded products has introduced new possibilities for hydrologic modelling and for flood simulation in data-scarce regions, in particular with the availability of the ERA-5 global weather reanalysis dataset (Baez-Villanueva et al., 2021; Faghih et al., 2022, Probst & Mauser, 2022; Tarek, et al., 2021, 2020). Despite these products still being affected by systematic, random, and detection errors (Sevruk et al., 2009; Zambrano-Bigiarini et al., 2017, Baez-Villanueva et al., 2018), which are more pronounced over mountainous regions(Beck et al., 2019), these studies have shown that ERA5 reanalysis has as a good potential to represent a reference dataset for hydrological modelling. Tarek et al., (2020) used data from 3138 North American catchments by considering the ERA5 precipitation and temperatures as proxies for observations in the hydrological modelling process, using two lumped hydrological models, and showed that ERA5-based hydrological modelling performance is equivalent to using observations over most of the basins. (Sun et al., 2021) examined results from application of ERA5 for hydrological modelling of 11 basins in the complex High Mountain Asia region, reporting rather large biases in the case the reanalysis data were used without bias correction. Probst and Mauser, (2022) exploited ERA5 for driving an uncalibrated setup of the hydrological model PROMET for the period 1980–2016 over the Danube river basin. They showed that bias correction was essential for ERA5 to provide good results in this hydrologically complex region. These studies have shown that ERA5 precipitation products generally overestimate low-intensity rainfall and underestimate high rainfall intensity. These precipitation biases are transmitted to the runoff simulations by generating errors which depend on the runoff model used, on the basin hydro-climatology and on the specific segment of the runoff hydrograph considered (Baez-Villanueva et al., 2021). Bias correction procedures are essential, but they are not a panacea. Given the complex structure of ERA5 bias, correction procedures typically weigh more certain types of bias than others, and therefore have a limited correction power.

Overall, a literature review shows that more efforts were focused so far on analyzing the performances of ERA5 for continuous runoff modelling, rather than for flood modelling more specifically. It is expected that the examination of ERA5 performance for flood modelling could be particularly revealing, because of the typical ERA5 precipitation bias which over-estimates the low-intensity precipitation while under-predicts the high-intensity precipitation.. Indeed, in the case of rainfall-induced floods, the positive bias of low intensity precipitation could affect the prediction of the antecedent soil moisture conditions, where the negative bias could affect more the flood volume and peak. The overall errors in flood modelling are therefore expected to be determined by the balance between the two biases.

Given this background, this work aims to quantify how the errors in hourly ERA5 precipitation and temperature input data propagates to the runoff and flood performances in a set of alpine watersheds located in the Eastern Italian Alps. Analysis of flood modelling is carried out by considering in a separate way antecedent soil moisture conditions (as surrogated by the runoff prediction at the start of the event) and flood peak. More specifically, the following objectives are considered: i) how ERA5 input errors affect the modelling of antecedent soil moisture conditions and flood peak individually, and their combined final impact; ii) the effects of bias error correction procedures. We expect that results from this study may help to better understand the impact of ERA5 input errors on flood modelling and prediction and on devise better error model correction.

3.3 Materials and methods

3.3.1 Study area and data

The study basins are located in the upper Adige river basin closed at Bronzolo, in the Eastern Italian Alps (Figure 3.1). This is an alpine catchment with a drainage area of approximately 6924 km^2. 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

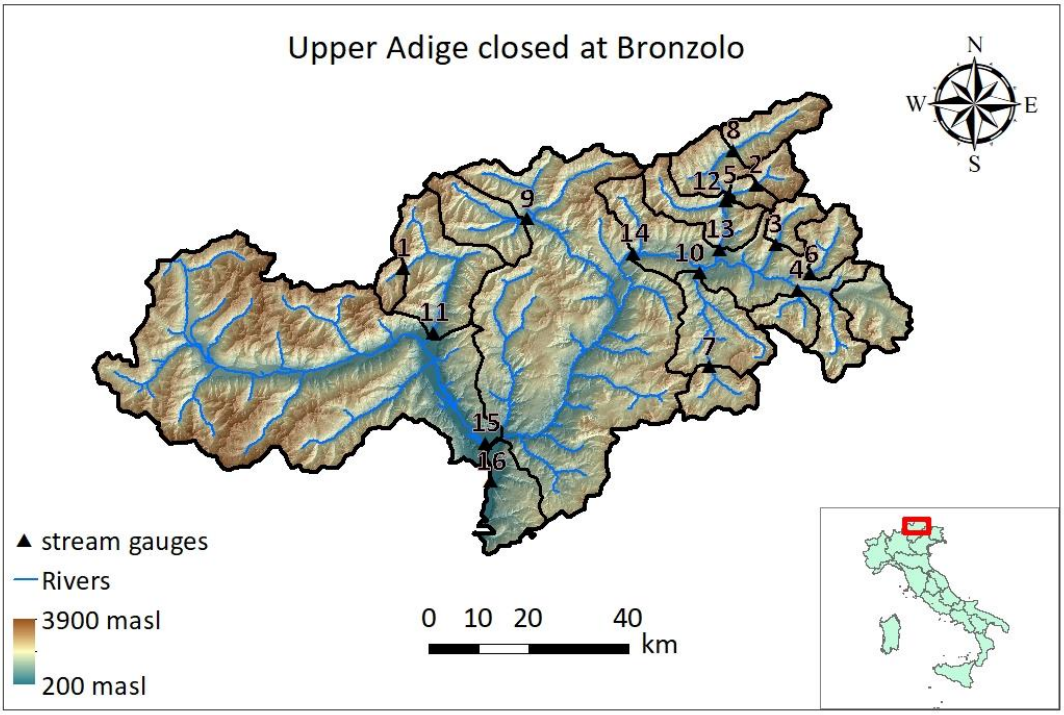

Figure 3.1: The upper Adige River basin closed at Bronzolo

elevation gradients govern the spatial precipitation distribution (Formetta et al., 2022) with the precipitation ranging from 500mm in the northwest region to 1600mm in the southern region

(Galletti et al., 2019). During the winter season, the precipitation is stored as snow and the streamflow is minimum. The streamflow shows two maxima: first due to snowmelt in the early summer and the second due to intense storms in autumn (Laiti et al., 2018). The same sixteen basins as the SWE works has been considered for the runoff and flood analysis in this work as well.

3.3.2 ERA-5 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 (Hersbach et al., 2020). 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 variables with data availability spanning from 1979 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 on ERA5 cis reported 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 cells that lie within and around 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. ERA5 air temperature is scaled to the mean elevation of each study basin based on the climatological monthly lapse rate valid for the region.

3.3.3 Ground data

The study basin has a dense network of meteorological stations. There are 88 rain gauges (1 per 72 km^2) and 124 temperature gauges (1 per 55 km^2) within the study basin. The hourly temperature, precipitation, and runoff data from 1991 till 2019 is collected from the Hydrographic Office of Bozen, Bolzano. If there is a missing hourly dataset, the model only considers only the stations for both precipitation and temperature lapse rate calculation. All the analysis focus the period of 1992-2019, however in case of runoff data, there was missing continuous dataset for some basins hence the calibration and validation of the runoff result was limited to the years as described in the previous section.

3.3.4 Hydrological model

The conceptual, semi-distributed hydrological model ICHYMOD (Norbiato et al., 2008; 2009,Zaramella et al., 2019)was used for runoff simulation at hourly time resolution. For a comprehensive explanation of the model, the reader is directed to the annex A. Briefly, ICHYMOD is a continuous, conceptual, semi-distributed rainfall-runoff model that operates on an hourly or daily time step. It includes modules that describe snow buildup and melting, glacier melting, soil moisture, groundwater, and flow generation. Snow and glacier melt are considered using distribution function approach with combined radiation degree day concept (Cazorzi & Dalla

Fontana, 1996). The potential evapotranspiration via the Hargreaves formula (Hargreaves & Samani, 1982) while the simulation of surface and subsurface flows is carried out by means of the Probability Distribution Model (PDM) from Moore (2007). The detail of the model is provided in Annex B with further detail regarding model parameters, calibration and validation in Annex C.
The ICHYMOD hydrological model has been successful applied to the Adige river basin by different authors. For instance, (Puspitarin et al., 2020) used to model for the Adige basin at Ponte Adige to explore the impact of glacier shrinkage and climate change on the streamflow pattern and its' effect on power generation. Likewise, (Stergiadi et al., 2020) considered the ICHYMOD model to assess the impact of geology on the seasonal hydrological predictability in the similar alpine setting. Other authors ((Avesani et al., 2022; François et al., 2017; Mei et al., 2016) have successfully applied the ICHYMOD for the different purpose in the alpine basins. Hence these previous works provide a robust background to consider the ICHYMOD model for this study as well.

3.3.5 Bias adjustment method

Corollary to previous work on SWE, the ERA5 precipitation and temperature were corrected for the runoff simulation as well. Different bias correction techniques may be used. The simplest methods consist of adding the climatological difference between ERA5 input and the reference data (the 'delta' method). This method is straightforward, but implicitly assume that the variability in ERA5 is unchanged. A quantile–quantile mapping (QM, hereinafter) transformation (the empirical transformation of (Panofsky & Brier, 1968) may be used to overcome these limitations. 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.

More specifically, ERA5 and reference distributions are matched by establishing a quantile-dependent correction function that translates simulated quantiles into their reference counterparts. This function is then used to translate the modeled time series into bias-adjusted values with a distribution representative of the reference data. QM was applied separately for each month. The transfer functions were obtained for each ERA5 grid cell from a calibration period (1992-2005) and then applied to the ERA5 variables (precipitation and temperature). In order 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 modelled data was applied. A validation period (2005-2019) was used to examine the quality of the correction scheme. The simulation for the year 1991-1992 is only considered as a TOPMELT model setup period. The implemented QM scheme was based on the R package qmap (Gudmundsson et al., 2012).

3.3.6 Reference precipitation at ERA5 spatial resolution

In order to examine the scale dependence of ERA5 errors, the station precipitation data were used to generate estimates of mean areal precipitation at the ERA5 spatial resolution. This input data was

termed Station Input at ERA5 Resolution (SIER). To achieve this, the Thiessen polygon method was used to redistribute the observed hourly station precipitation data over each ERA5 grid footprint.

3.4 Results

Figure 3.2 a, b shows the annual mean values of temperature and precipitation, respectively, reporting both the reference and the ERA5 values. Whereas an almost constant overall bias of 1.36 is affecting annual ERA5 precipitation totals, generated by over-prediction of small totals and under-prediction of large totals, a non-stationary behavior is affecting ERA5 temperature data, with the temperature showing two different biases for the first and for the second half of the data. The application of the Pettitt test (Pettitt, 1979) to detect changes in ERA5 temperature bias led to rejecting the null hypothesis of stationarity, showing change point on 2005, with a bias of -0.9098 °C for the 1991-2005 period and a bias of -0.2841 °C for the period 2005-2019. Out of mere coincidence, the division of ERA5 temperature into two periods coincides with the calibration and validation period, which was selected prior to the application of the Pettitt test. Based on this evidence, the QM procedure was applied separately for the two periods. In the following, two QM-corrected ERA5 input are considered: one where only the 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).

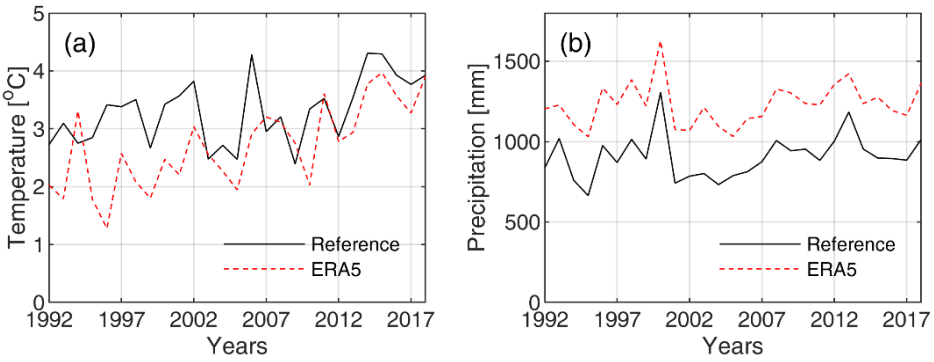

Figure 3.2: Adige at Bronzolo yearly (a) Temperature (b) Precipitation

Figure 3.3 a, b shows a comparison between reference and ERA5 precipitation, both uncorrected and QM-corrected, respectively, for the calibration period. The comparison is carried out by considering the hourly mean areal precipitation of the 16 different study basin by using the Kling-Gupta Efficiency (KGE) (Gupta et al, 2009) applied to all data points but the intervals where both the data sources are zero. The Kling-Gupta Efficiency (KGE) is as follows:

$$KGE = 1 - \sqrt{(\gamma - 1)^2 + (\beta - 1)^2 + (\alpha - 1)^2} \qquad \text{Equation 3.1}$$

where γ is 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. If two series exactly match, the KGE is one. A β or α value smaller/larger

than one indicates that the mean value or variability of observations is underestimated/overestimated. Please note the use of log coordinates for the drainage area on the x axis of the figures. Figure 3.3 c, d reports the same comparison for the validation period.

The KGE of ERA5 precipitation for both periods shows similar performances, which points to the robustness of the QM procedure applied here. There is a clear trend of KGE increasing with the basin size in all the cases, with KGE around or less than 0.2 for the smallest basin (Plan at Plan) and around or over 0.5 for the largest basin (Adige at Bronzolo). The PC-ERA5 shows a slight improvement in the KGE performance for both periods. All basins show an increase in performance, with the smallest ones showing a larger gain, as expected. The QM technique applied to ERA5 precipitation helps to decrease the bias of the ERA5 precipitation while keeping the correlation intact. Nevertheless, the overestimation of the variability of the observations causes the KGE of PC-ERA5 to improve only slightly.

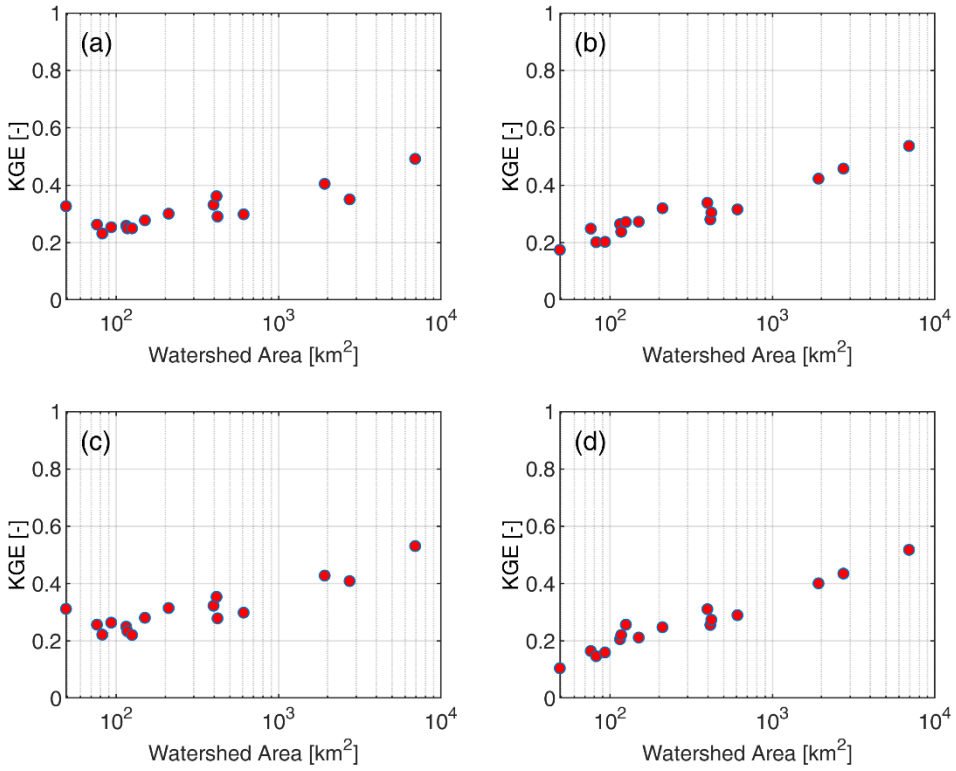

Figure 3.3: KGE of Precipitation (a) ERA5-calibration (b)PC-ERA5-calibration (c) ERA5- validation (d) PC-ERA5-validation

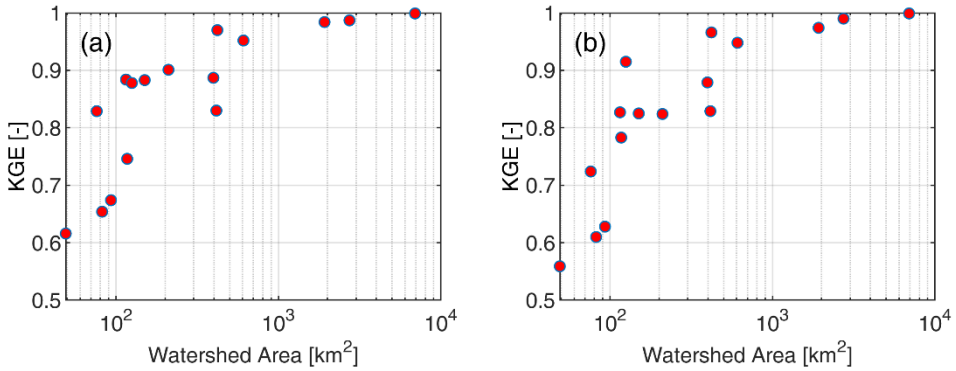

Figure 3.4: KGE of SIER precipitation with station (a) Calibration (b) validation

Figure 3.4 a, b report the KGE for SIER precipitation for the calibration and validation period, respectively. 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 600 km^2 KGE is close to one, whereas it decreases in a marked way for basin smaller than 600 km^2 and even more remarkably for basin less than 100 km^2, with values around or less than 0.6 for the smallest basin.

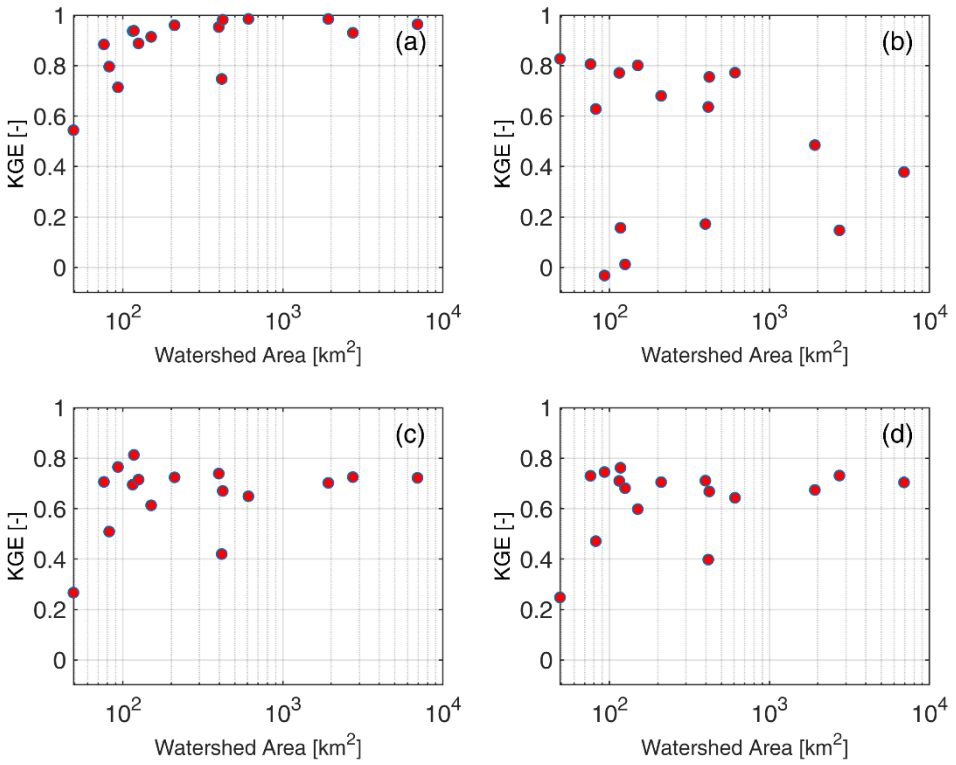

Figure 3.5: KGE of hourly runoff of (a) SIER, (b) ERA5, (c) PC-ERA5 (d) PTC-ERA5

Figure 3.5 shows the hourly runoff KGE performance of SIER and different ERA5 inputs with the station input for the period of 1992-2019. The SIER runoff performance is depended on the basin size; the KGE value is increasing from less than 0.6 for the smallest basin to KGE value greater than 0.8 concerning the basins larger than 100 km^2. As similar to the performance of SIER precipitation performance, the SIER spatial scale influences the runoff performance of smaller basins (<100 km^2)

as compared to the larger ones. While in case of ERA5 input runoff, the KGE is decreasing with the increase in the basin size, with KGE value equaling 0.8 for smaller and medium basins but for larger one, the value dips to less than zero in the worst case.

This result is counterintuitive to the precipitation performance of ERA5 but this should be attributed to the effect of ERA5 precipitation on runoff. The drizzle in ERA5 is overpredicting the overall precipitation and considerably increasing the soil moisture content. This ERA5 initial catchment state mimics the hydrological condition of the basin lying in the higher elevation, which in reality receive higher rainfall and wets the ground due to snow accumulation. Hence the runoff shows good KGE performance for smaller basins due to the exaggerated ERA5 precipitation. On the other hand, since the larger basins consist of the combination of multiple smaller basins, the error due to the soil moisture content also gets accumulated with increase in basin size. This leads to low KGE performance of the larger basins. After the precipitation correction of ERA-5(PC-ERA5), there is an improvement in KGE for all the basins with value ranging between 0.4 and 0.8 for all the basins (except for RioPlan KGE ~0.2), whereas, the gain in skill for the larger basin is comparatively better than or smaller basins. The PC-ERA5 reduces the runoff bias and variability ratio while keeping the correlation intact and thus increasing the KGE value. Since the temperature (PTC-ERA5) correction has visually no change in the performance of PC-ERA5, it could be deducted that temperature correction has no role in runoff simulations.

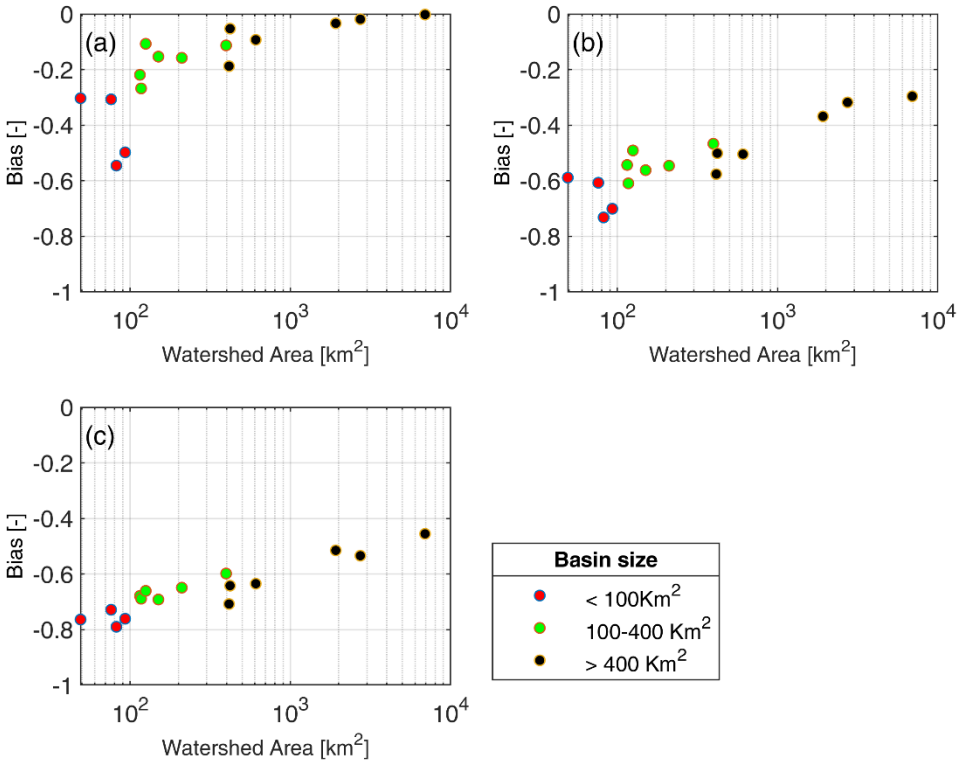

Figure 3.6: Relative bias of maximum precipitation for (a) SIER, (b) ERA5, (c) PC-ERA5

The Figure 3.6 reports the maximum precipitation bias within the 4-day flood study. Similar to previous analysis, the all the inputs considered is compared with the station precipitation. Corollary to previous results, in this case also the spatial scale of ERA5 only affects the event precipitation of

smaller basins, producing negative bias of almost 0.6 while for the larger basins the affect is none existing. It is also interesting to note that the ERA5 always under-predicting the reference station with bias close to -0.7 for the smallest group of basins while the bias decreases with increase in basin size with value of around -0.3 for basins larger than 600 km^2. After the precipitation correction, the bias of PC-ERA5 is further decreased as compared to ERA5.

The first day runoff relative bias is plotted in Figure 3.7. The SIER input as compare to the station input runoff showed similar bias for the varying basin size, however for basin <100 km^2 displays larger negative bias. In case of ERA5 input, the first day runoff is always overpredicted by ERA5, which is interesting and also an expected finding. As ERA5 is characterized by the drizzle problem, the false ERA5 precipitation wets the ground and resulting into higher initial moisture condition than the reference station with resulting average bias of 0.42. Since, PC-ERA5 diminishes the overall ERA5 precipitation, which also decreases the initial moisture condition and has reduced first day runoff bias of -0.09. While the temperature correction doesnot play any role in alteration of initial soil moisture condition.

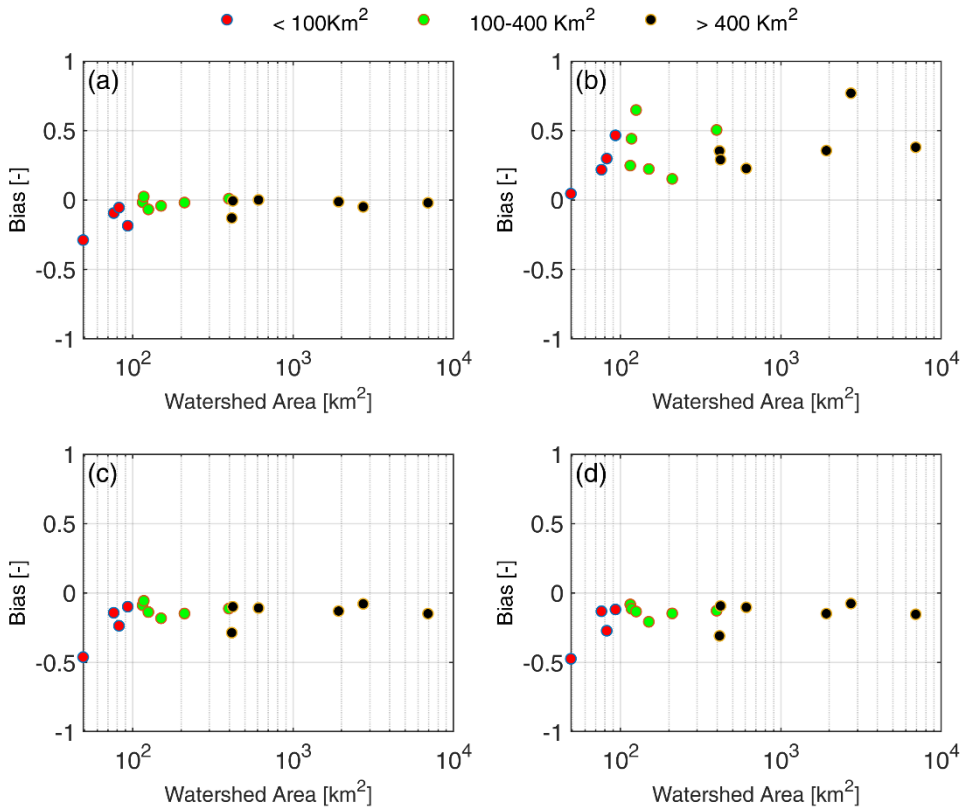

Figure 3.7: First day runoff bias for (a) SIER, (b) ERA5, (c) PC-ERA5 (d) PTC-ERA5

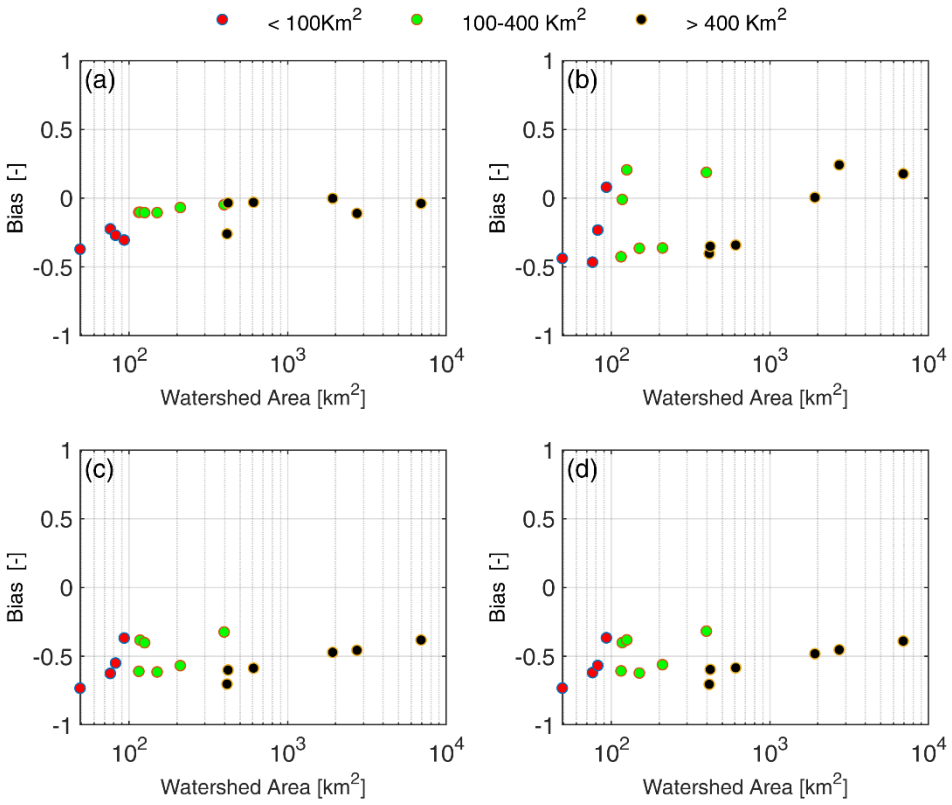

Figure 3.8: Flood peak bias for (a) SIER, (b) ERA5, (c) PC-ERA5 (d) PTC-ERA5

The Figure 3.8 shows the maximum runoff bias within the 4 day window for different input considered. The SIER results with the reference shows the smallest basin suffering the most with value close to -0.4, while the bias decreases with the increase in basin size, it never goes past 0. In case of ERA5 input, the bias is comparable with the reference station-input simulation and in some cases overpredicts the reference, which is always true for basins >600km^2. This nature of maximum runoff performance of ERA5 can be attributed to two cascading effects; first due to the ERA5 drizzle, which increases the soil moisture content prior to the flood event and it is also explained in the first day runoff comparison. Second, due to the underprediction of the event precipitation. Even though the maximum precipitation is under predicted by ERA5, the over prediction of the initial soil moisture compensate for the event precipitation which helps for the comparable flood performance. The PC-ERA5 underpredicts the flood event due to the underprediction of both the initial soil moisture and the event precipitation. The PTC-ERA5 performance doesnot change much as compare to PC-ERA5, showing limited impact of the temperature on the flood event.

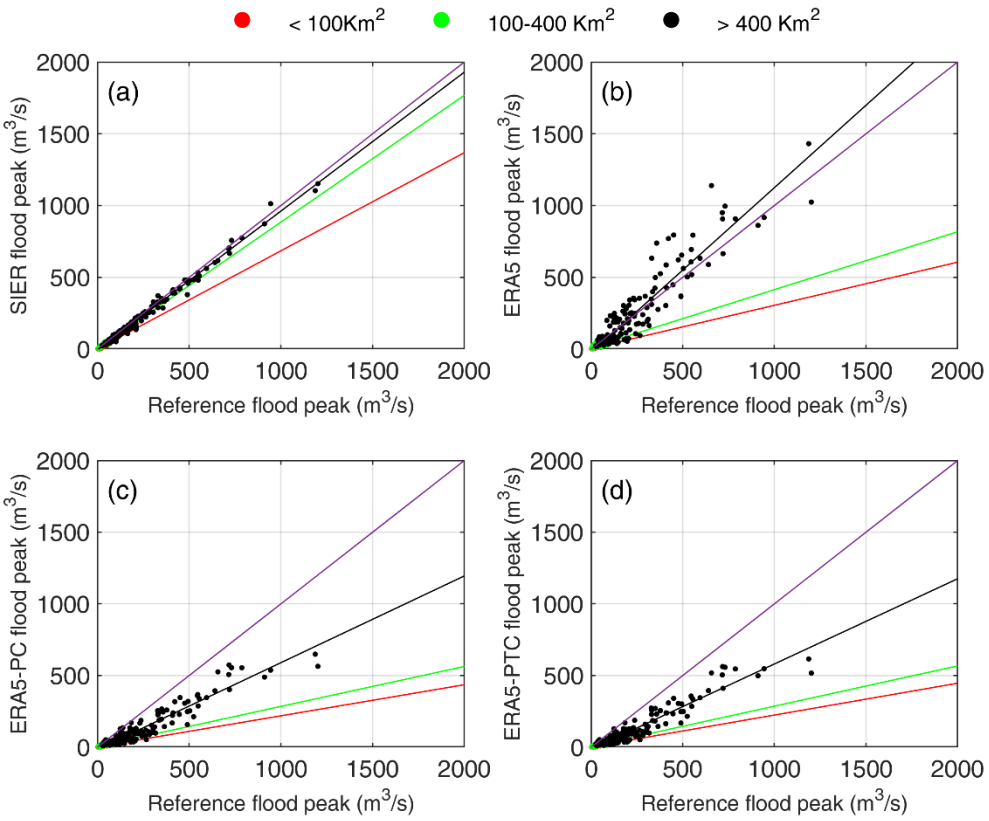

Figure 3.9 maximum runoff scatter plot for (a) SIER, (b) ERA5, (c) PC-ERA5 (d) PTC-ERA5

Figure 3.9 shows the scatter plot considering the flood peak simulation. The SIER performs similarly with the reference simulation, however as seen in the previous analysis the smaller basins under predict the flood with slope of 0.68, while there is gradual improvement in the value for medium and large basins with slope of 0.88 and 0.97 respectively. In case of ERA-5 simulation, the slope of regression line for the larger basin overpredicts the reference with value of 1.14, while the small and medium basins have similar slope value of 0.30 and 0.40 respectively, displaying limited flood capabilities of ERA5 for these basins. These slope values further deplete when the precipitation correction (PTC-ERA5) is applied with value of 0.22 and 0.28 for small and medium basins whereas for large basin, it remains 0.61. The temperature correction renders no again in information for the flood peaks as the slope value for the different basin size does not change in comparison to PC-ERA5.

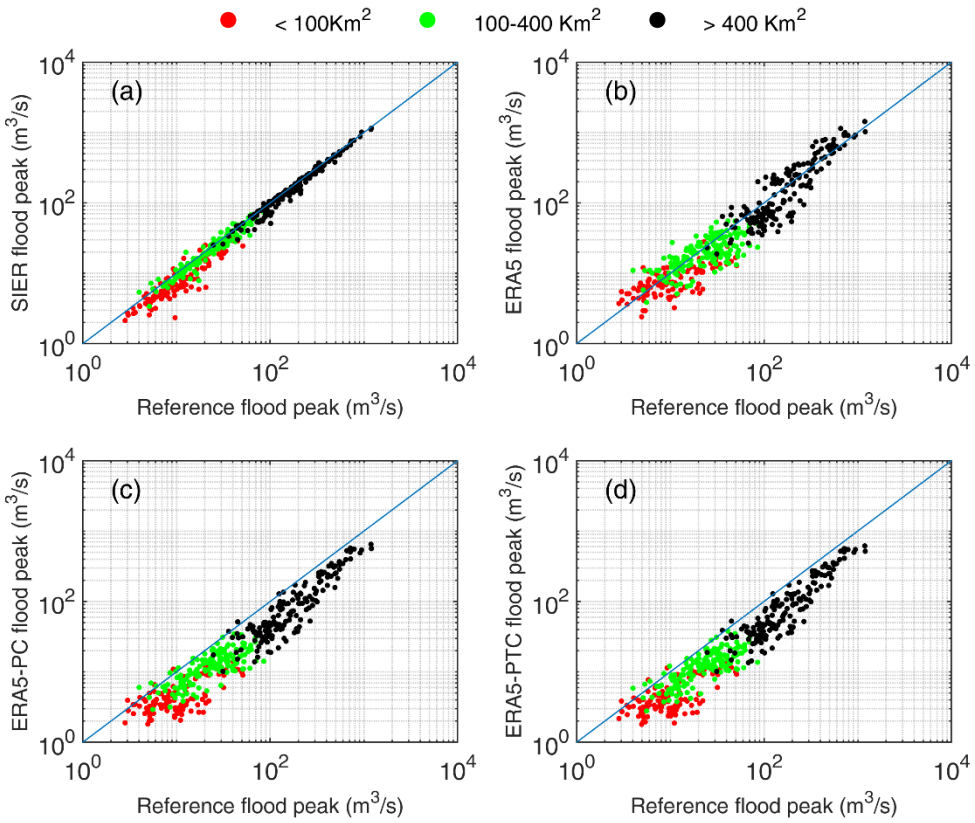

Figure 3.10 Log scale maximum runoff scatter plot for (a) SIER, (b) ERA5, (c) PC-ERA5 (d) PTC-ERA5

Since it was not possible to plot the line intercept and visualize the flood peaks in the same normal plot, Figure 3.10 uses the same information as in Figure 3.9 but plots in log scale for better representation. This figure comprehensively shows what was already explained in the paragraph above. The ability of SIER to predict the flood events for larger basin is better however, this skill degrades with decrease in basin size. It also shows the capability of ERA5 to predict flood especially for the larger basins but the performance decreases with decrease in basin size. After the precipitation correction of ERA5, the flood detection property of ERA5 is further lost while the temperature correction does not have a significant effect on the flood simulation.

The Figure 3.11 shows the unitary max runoff scatter plots for the varying inputs. For SIER, the large and medium basins show comparable performance with slope of 0.84 and 0.81 while for the smaller basin size, the SIER input largely underpredicts the reference with slope of 0.65. As oppose to the maximum runoff results as reported in Figure 3.9, it is interesting to see the inability of ERA5 to properly capture the unitary maximum flood. Where the large and medium basins both show comparable slope value of 0.26, while in case of the small basin, the slope is found to be 0.31.It could be argued that the overeprediction of ERA5 flood for larger basin was solely based on the basin size and if it was taken out of the equation, ERA5 would lose its' ability to predict the flood event. PC-ERA5 and PTC-ERA5 shows similar performance and does not hold much information to report any useful narrative regarding the unitary maximum runoff. Both the correction report the similar slope value of 0.20, 0.20, 0.16 for small, medium and large basins.

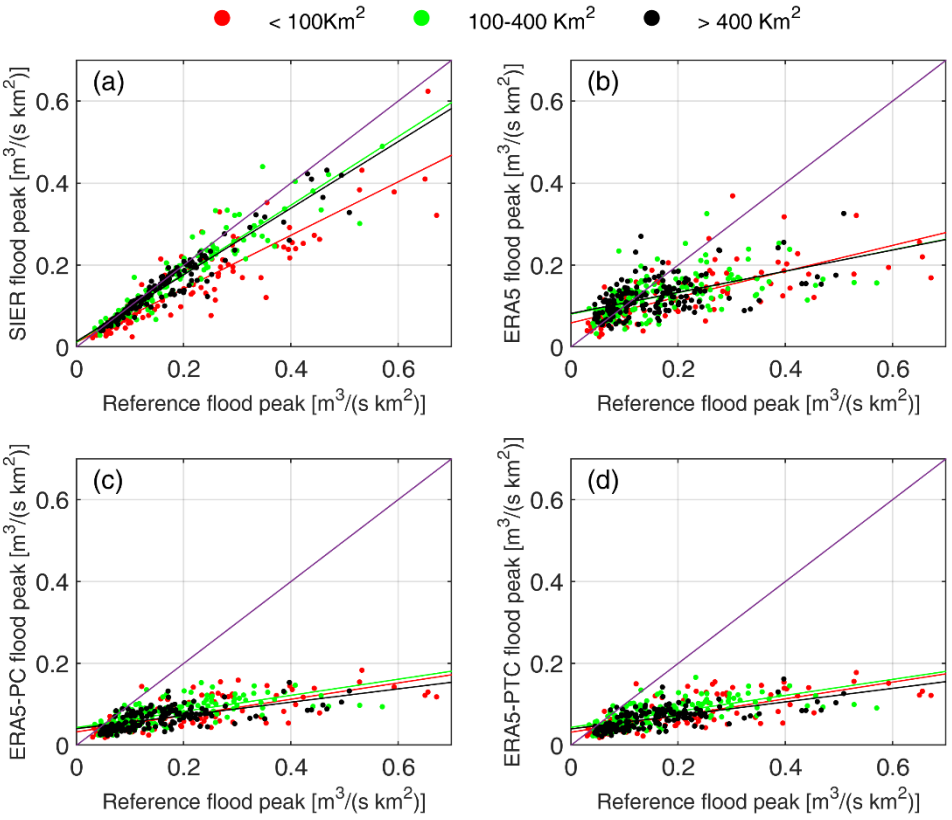

Figure 3.11: Unitary maximum runoff scatter plot for (a) SIER, (b) ERA5, (c) PC-ERA5 (d) PTC-ERA5

In Figure 3.12, the flood error is computed for different return periods. The SIER performs comparable with the station input for different flood return period, in case of shorter return periods, the error is close to zero. ERA5 flood performance for varying return periods has mixed results, where ERA5 overpredicts/ underpredicts the reference station irrespective of the return period. However on closer look there is some cases in shorter return period (<30 years), where ERA5 outperforms the reference flood. PC-ERA5 and PTC-ERA5 as similar with all the previous results, shows comparable performance with each other and mostly, both underpredict the flood for the given return periods.

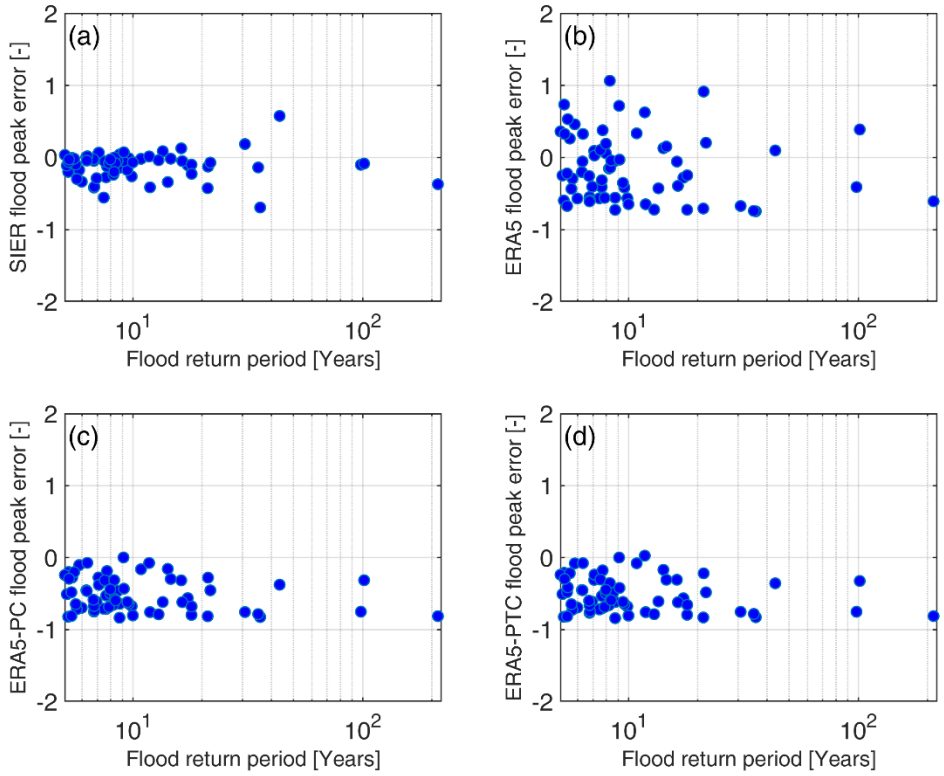

Figure 3.12: Different return period flood error for (a) SIER, (b) ERA5, (c) PC-ERA5 (d) PTC-ERA5

In Figure 3.12, the flood error is computed for different return periods. The SIER performs comparable with the station input for different flood return period, in case of shorter return periods, the error is close to zero. ERA5 flood performance for varying return periods has mixed results, where ERA5 overpredicts/ underpredicts the reference station irrespective of the return period. However on closer look there is some cases in shorter return period (<30 years), where ERA5 outperforms the reference flood. PC-ERA5 and PTC-ERA5 as similar with all the previous results, shows comparable performance with each other and mostly, both underpredict the flood for the given return periods.

3.5 Discussion

This research work outlines the usage of ERA5 dataset for the runoff modelling in the alpine catchment of northern Italy using ICHYMOD hydrological model. Furthermore, the suitability of the flood simulation using ERA5 dataset in its' original state and after bias correction with respect to basin size has been extensively reviewed. The spatial scale of ERA5 seem to affect only the basin with smaller area for results concerning precipitation and runoff. The ERA5 helps to provide a comparable runoff results for smaller basins, however for the larger basins, the runoff performance decreases invariably. This behavior should be linked with the 'drizzle' problem in ERA5 since, the smaller precipitation events creates comparable initial moisture content for the high altitude snow dominate catchment, which leads to better results. However for the larger basins, these errors gets accumulated and hence affects the runoff results. Interestingly, this error helps to show comparable results of flood for basin with larger size due to overestimation of the initial moisture state and

under prediction of the event precipitation. Sadly, the ERA5 flood detection capability for larger basin is lost when the unitary basin size is considered, which highlights the effectiveness of the ERA5 for flood simulation was only relating to its' basin size. The QM correction of precipitation helps considerably to improve the runoff performance across different basin size but the gain in performance for larger basin is comparatively higher as compared to the smaller basins. However the temperature correction doesn't play any role in improving the runoff performance. Further, the bias correction of precipitation decrease the ERA5 ability to detect floods; even though the performance was substantial for a niche larger basins.

3.6 Conclusion

To conclude, this work provides a procedure to integrate the ERA5 dataset to model the hourly runoff in the alpine basins. On one hand, the ERA5 drizzle error proves useful for runoff simulation concerning smaller basins, while on the other hand, the same error provides detrimental results for the larger basins. Surprisingly, this error when coupled with under prediction of event precipitation, provides a comparable flood values for the basins with larger area. However, the flood capabilities is lost when the flood value is reduced to unit basin size. The correction of precipitation provides comparable results of hourly runoff simulation but the temperature correction doesn't affect the runoff simulation at all. Irrespective of the basin size, the bias correction further reduces any flood detection property the raw ERA5 initially displayed.

3.7 References

Baez-Villanueva, Oscar M, Zambrano-Bigiarini, M., Mendoza, P. A., McNamara, I., Beck, H. E., Thurner, J., … Thinh, N. X. (2021). On the selection of precipitation products for the regionalisation of hydrological model parameters. Hydrology and Earth System Sciences, 25(11), 5805–5837.

Baez-Villanueva, Oscar Manuel, Zambrano-Bigiarini, M., Ribbe, L., Nauditt, A., Giraldo-Osorio, J. D., & Thinh, N. X. (2018). Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. Atmospheric Research, 213, 34–50.

Barredo, J. I. (2009). Normalised flood losses in Europe: 1970–2006. Natural Hazards and Earth System Sciences, 9(1), 97–104.

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., … Adler, R. F. (2019). MSWEP V2 global 3-hourly 0.1 precipitation: methodology and quantitative assessment. Bulletin of the American Meteorological Society, 100(3), 473–500.

Bouwer, L. M. (2011). Have disaster losses increased due to anthropogenic climate change? Bulletin of the American Meteorological Society, 92(1), 39–46.

Da Ros, D., & Borga, M. (1997). Use of digital elevation model data for the derivation of the geomorphological instantaneous unit hydrograph. Hydrological Processes, 11(1), 13–33.

Di Marco, N., Avesani, D., Righetti, M., Zaramella, M., Majone, B., & Borga, M. (2021). Reducing hydrological modelling uncertainty by using MODIS snow cover data and a topography-based distribution function snowmelt model. Journal of Hydrology, 126020.

EEA. (2021). Economic losses from climate-related extremes in Europe. Retrieved from https://www.eea.europa.eu/ims/economic-losses-from-climate-related

Faghih, M., Brissette, F., & Sabeti, P. (2022). Impact of correcting sub-daily climate model biases for hydrological studies. Hydrology and Earth System Sciences, 26(6), 1545–1563.

Formetta, G., Marra, F., Dallan, E., Zaramella, M., & Borga, M. (2022). Differential orographic impact on sub-hourly, hourly, and daily extreme precipitation. Advances in Water Resources, 159, 104085.

Galletti, A., Avesani, D., Bellin, A., & Majone, B. (2019). Detailed simulation of storage hydropower systems in a large Alpine watershed. Geophysical Research Abstracts, 21.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Downscaling RCM precipitation to the station scale using statistical transformations–a comparison of methods. Hydrology and Earth System Sciences, 16(9), 3383–3390.

Hargreaves, G. H., & Allen, R. G. (2003). History and evaluation of Hargreaves evapotranspiration equation. Journal of Irrigation and Drainage Engineering, 129(1), 53–63.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., … Schepers, D. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146(730), 1999–2049.

Koks, E. E., de Moel, H., Aerts, J. C. J. H., & Bouwer, L. M. (2014). Effect of spatial adaptation measures on flood risk: study of coastal floods in Belgium. Regional Environmental Change, 14(1), 413–425.

Laiti, L., Mallucci, S., Piccolroaz, S., Bellin, A., Zardi, D., Fiori, A., … Majone, B. (2018). Testing the hydrological coherence of high-resolution gridded precipitation and temperature data sets. Water Resources Research, 54(3), 1999–2016.

Moore, R. J. (2007). The PDM rainfall-runoff model. Hydrology and Earth System Sciences Discussions, 11(1), 483–499.

Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E., & Anquetin, S. (2008). Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins. Journal of Hydrology, 362(3–4), 274–290.

Pachauri, R. K., Allen, M. R., Barros, V. R., Broome, J., Cramer, W., Christ, R., … Dasgupta, P. (2014). Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change. Ipcc.

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

Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. Journal of the Royal Statistical Society: Series C (Applied Statistics), 28(2), 126–135.

Probst, E., & Mauser, W. (2022). Evaluation of ERA5 and WFDE5 forcing data for hydrological modelling and the impact of bias correction with regional climatologies: A case study in the Danube River Basin. Journal of Hydrology: Regional Studies, 40, 101023.

Sevruk, B., Ondrás, M., & Chvíla, B. (2009). The WMO precipitation measurement intercomparisons. Atmospheric Research, 92(3), 376–380.

Sun, H., Su, F., Yao, T., He, Z., Tang, G., Huang, J., … Chen, D. (2021). General overestimation of ERA5 precipitation in flow simulations for High Mountain Asia basins. Environmental Research Communications, 3(12), 121003.

Tarek, M., Brissette, F., & Arsenault, R. (2021). Uncertainty of gridded precipitation and temperature reference datasets in climate change impact studies. Hydrology and Earth System Sciences, 25(6), 3331–3350.

Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. Hydrology and Earth System Sciences, 24(5), 2527–2544.

Xie, A., Ren, J., Qin, X., & Kang, S. (2007). Reliability of NCEP/NCAR reanalysis data in the Himalayas/Tibetan Plateau. Journal of Geographical Sciences, 17(4), 421–430.

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., & Ribbe, L. (2017). Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. Hydrology and Earth System Sciences, 21(2), 1295–1320.

Zaramella, M., Borga, M., Zoccatelli, D., & Carturan, L. (2019). TOPMELT 1.0: a topography-based distribution function approach to snowmelt simulation for hydrological modelling at basin scale. Geoscientific Model Development, 12(12), 5251–5265.

4. Article 3: Subseasonal streamflow predictions in alpine catchments by combining numerical weather models and reanalysis data.

Shrestha, S.[1], Zaramella, M [1], Crespi, A. [2], Bogner, K.[3], Zappa, M.[3], Callegari, M.[2], Greifeneder, F.[2], Borga, M.[1]

[1]Department of Land, Environment, Agriculture, and Forestry, University of Padova, Italy

[2]EURAC, Institute for Earth Observation, Bolzano, Italy

[3]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, Switzerland

4.1 Abstract

The is an increasing interest in the application of hydrological forecast for the management of water resources (e.g. hydropower and agricultural) at a seasonal scale. The goal of this paper is to make a preliminary assessment of the skill of post-processed Numerical Weather Predictions (NWPs) in contrast to traditional ESP. Their performance are assessed with the forecasts based on climatology with daily resolutions. The total runoff and particularly the base flow runoff will be extensively analyzed for the northern Italian Alps. The results of spatially aggregated predictions of the variables are compared to daily simulations and long-term daily averages of simulations driven by the meteorological observations (i.e. climatology). The performance of the forecast skill will be quantified based on the statistics like bias, Nash-Sutcliff efficiency (NSE), Continous Rank Probability Score (CRPSS), and Spread error ratio (SPRerr). The results of post-processed NWP show better total runoff skill in the first few weeks as compared to the ESP, which loses its' skill after the first week. However, for the longer timescale of months, even though both the forecast does not have credible skill values, the ESP shows a comparatively better performance. The baseflow also shows a similar pattern of performance with skills scores that are comparatively better than total flow However skillful forecasting of the sub-surface runoff process beyond two weeks remains challenging. This could be attributed to the difficulty in modeling thin and highly variable soil variables characteristics of the alpine head basins and also to the limited ability of the NWP to predict the chaotic nature for a longer timescale.

4.2 Introduction

The recent advancement in computing capacity has allowed Numerical Weather prediction (NWP) models to incorporate the inherently chaotic nature of climatic forecasts (Bogner et al., 2018). For seasonal timescale, to integrate the error associated with the initial conditions, probabilistic prediction is needed, which is possible through ensemble forecasts (Van Schaeybroeck & Vannitsem, 2018). These forecasts when combined with the hydrological model serve as a robust

tool for streamflow and groundwater forecasting studies (J. Liu et al., 2022), which are useful for water resource management, energy production, and early warning system (Monhart et al., 2019).

Extended streamflow prediction (ESP) (Day, 1985) method pioneered by the United States National Weather Service (NWS) in the 1970s has found extensive use in seasonal forecasting in the recent years as well due to its' simplistic and cost-effective approach (Harrigan, et al., 2018). ESP forecast ensembles are obtained by forcing a hydrological model - initialized with the initial hydrological condition (IHC) - with ensembles of historical observation (Day, 1985). Even though ESP forecast suffers from inadequate information about the future climatic condition, nevertheless, it should be considered along with the other state-of-art forecasting method as a "best" (Yuan et al., 2015) forecast methodology has not been established yet. ESP also serves as the benchmark to quantify the added skills coming from available forecasting products (Crochemore et al., 2017, Arnal et al., 2018).

On the other hand, the forecasting products suffer from important bias when compared with the observations regardless of their data assimilation technique (Randall et al., 2007). These biases could originate from the imperfect conceptualization, discretization, and spatial averaging within grid cells (Soriano et al., 2019). As a result, the raw forecast has a certain amount of risk for industrial applications (e.g., energy, hydrology, agriculture), and post-processing is required to get finer resolution output for impact assessment and statistical features that are comparable to the observations (Torralba et al., 2017). In order to remove the systematic model errors, the available technique ranges from simpler approaches like bias correction (BC), working directly on interest variable, to more complex statistical downscaling, building on large –scale predictors (Maraun et al., 2010). To make the raw model output consistent with the local climatology, the BC method maps it to the analogous long historical observational reference. These strategies, which do not take into account the temporal correlation between model outputs and observations, ranging from basic adjustments in the mean and/or variance to more advanced quantile mapping (QM) alternatives, that can accommodate higher-order moments or even the entire distribution (Manzanas et al., 2019). Since the QM can be applied for the dataset that does not require a standard (e.g. Gaussian) distribution - regular criteria for the operational agencies - its' application has gained recent interest (Bedia et al., 2018).

There have been multiple runoff forecasting studies done in the context of Europe, for instance, the UK (Harrigan et al., 2018), France (Crochemore et al., 2016), Austria (Santos et al., 2021), Switzerland (Bogner et al., 2018). However, the latest iteration of the forecast products ECMWF system-5, except for a few studies (Sutanto & Van Lanen, 2021, Sánchez-García et al., 2022), has found limited implications in the context of European alpine basins. Most crucially, their application for baseflow forecasting has not been fully realized. Baseflow is termed as the contribution of groundwater that sustains the river discharge between precipitation events (Brutsaert, 2008). Since the baseflow remains substantial for the ecohydrological sustainability, farming, and hydropower production during the seasons with no rain, its' skillful forecast could help to mitigate (Ayers et al., 2021) the repercussion related to the low flow in the European alpines.

To address the gap in the research, the prime objective of this work is to exploit the meteorological information of ECMWF system-5 for the total and baseflow forecast for the Upper Adige river basin, situated in the northern Italian Alps. The application of ICHYMOD hydrological model for the ECMWF hindcast period of 1993-2016 with up to 6 months of the forecast horizon will be extensively reviewed in the context of baseflow. Prior to its' application, the bias associated with the ECWMF dataset is corrected with quantile mapping (QM) with the help of an observational dataset (STAT-ECMWF). To remove the error associated with the ICHYMOD hydrological model and to have seamless runoff dataset in case of missing runoff observation, the runoff simulation with observation as input will be used as a reference for error metrics. Further, the catchment initial states are based on the reference simulation for driving 25-ensemble members of the ECMWF variables for forecasting. The possible improvement in the baseflow forecasting for the STAT-ECMWF datasets against the traditional ESP method will be quantified using metrics like bias, NSE, CRPSS, and Spread error with climatology as reference.

4.3 Study area and data

The study basins are located in the upper Adige river basin closed at Bronzolo, in the Eastern Italian Alps (Figure 4.1). This is an alpine catchment with a drainage area of approximately 6924 km^2. 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 (Formetta et al., 2022) with the precipitation ranging from 500mm in the northwest region to 1600mm in the southern region (Galletti et al., 2019). During the winter season, the precipitation is stored as snow and the streamflow is minimum. The streamflow shows two maxima: first due to snowmelt in the early summer and the second due to intense storms in autumn (Laiti et al., 2018).

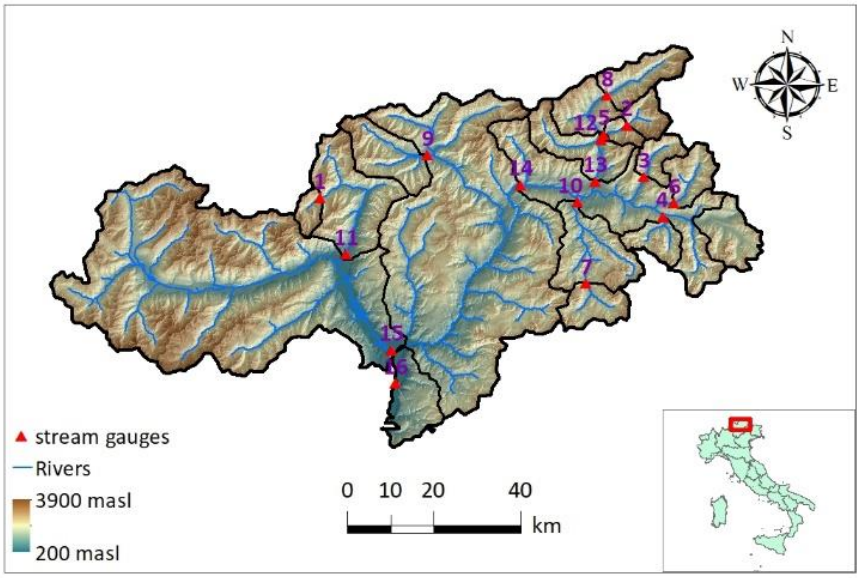

Figure 4.1: The upper Adige River basin closed at Bronzolo

4.3.1 ECMWF forecast

Each of the reanalysis products provide only the climatic information about the present and past but not of the future. Hence, forecast information is required to understand how the climate is going to evolve in the future based on reanalysis as initial condition. Seasonal climate forecast lies between the short-term weather forecast and longer timescale (like climate change prediction and inter-annual prediction) (Kirtman et al., 2013). The seasonal forecast provides a probabilistic spread (ensemble prediction) of the average conditions (like temperature and rainfall) and its' evolution in the future, which can range from 1-month lead time upto a year (Rickards et al., 2014).
Even though short term weather and seasonal forecasts try to quantify the future climatic states, they are part of a different paradigm which requires treating them in a different way. Seasonal predictions begin with an observable condition of the climate system which evolves over the course of a few months. Errors that exist at the beginning of the prediction (due to inaccurate measurement of the initial circumstances and assumptions assumed in the model formulation) remain or, more frequently, expand through the model integration, reaching magnitudes equal to those of the predictable signals. Some of these errors are random, and the impact they have on the result is assessed by using ensembles. Rather than talking specifically about the weather in a particular location in a particular day, seasonal forecasts will provide us with information on the likelihood that the upcoming season will be wetter, drier, warmer or colder than "usual" for that time of year. This type of long-term forecasting is possible because some parts of the Earth's system behave predictably and with slower rates of evolution than the atmosphere (such as the ocean and cryosphere), which allows their effect on the atmosphere to provide an audible signal ("ECMWF," n.d.).

The ECMWF provides the forecast information of meteorological variables named SEAS5 (System-5), which is the latest iteration following the previous System 4. SEAS5 compared to its' predecessor has improved atmosphere and ocean models with higher resolution with upgraded atmosphere and ocean models. The ERA5 meteorological variables provide the boundary conditions for the SEAS5 (Wang et al., 2019). The SEAS5 is initialized on the first day of each month and the forecast horizon leads up to 7 months. The re-forecast period is from 1981-to 2016 which has an ensemble size of 25 while, the real-time forecasts started from 2017 till 2019 with 51 members ensemble size. In Figure 4.2, an example of the 25 member daily ensemble runoff forecast with the median value is plotted with the station input model simulated daily runoff as reference runoff. The figure is plotted for Adige at Bronzolo for the initialization day of October 1, 1993 with leadtime of 180 days.

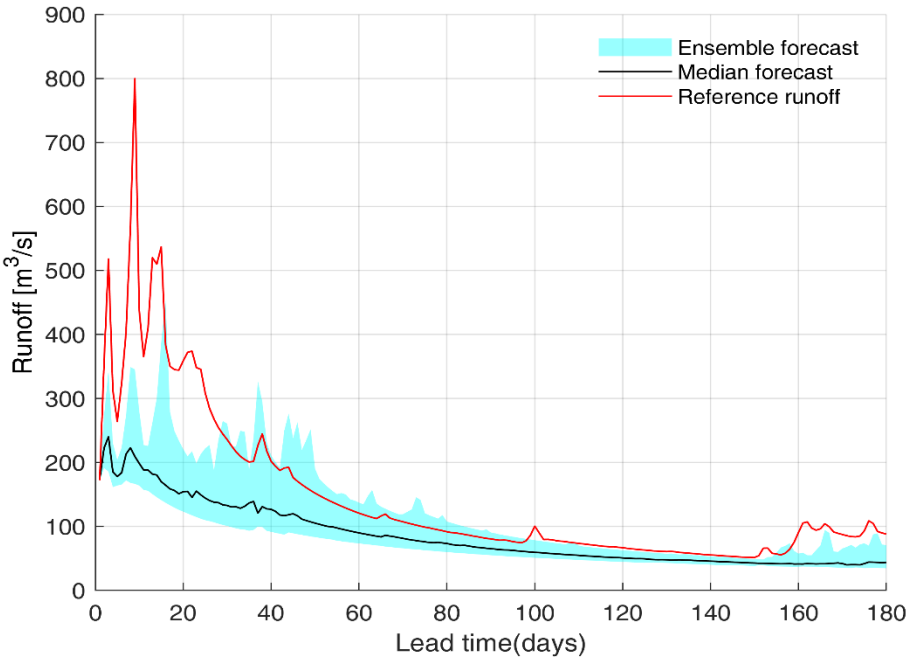

Figure 4.2: ECMWF forecast for October 1, 1993 initialisation with Leadtime upto 180 days

For our study, the daily temperature and precipitation forecast issued by ECMWF system5 will be used which has a global spatial resolution of 1° × 1°. In which, the forecast issued for day 1 (24 hours) until day 186 (4464 hours) starting from the first day of each month is used. The data were downloaded at the following link:

https://cds.climate.copernicus.eu/cdsapp#!/dataset/seasonal-original-single-levels?tab=form.

4.3.2 Ground data

4.3.3 Ground data

The study basin has a dense network of meteorological stations. There are 88 rain gauges (1 per 72 km^2) and 124 temperature gauges (1 per 55 km^2) within the study basin. The hourly temperature, precipitation, and runoff data from 1991 till 2019 is collected from the Hydrographic Office of Bozen, Bolzano. If there is a missing hourly dataset, the model only considers only the stations for both precipitation and temperature lapse rate calculation. All the analysis focus the period of 1992-2019, however in case of runoff data, there was missing continuous dataset for some basins hence the calibration and validation of the runoff result was limited to the years as reported in previous article.

4.4 Methodology

4.4.1 Hydrological model

The conceptual, semi-distributed hydrological model ICHYMOD (Norbiato et al., 2008; 2009, Zaramella et al., 2019)was used for runoff simulation at hourly time resolution. For a comprehensive explanation of the model, the reader is directed to the annex A. Briefly, ICHYMOD is a continuous,

conceptual, semi-distributed rainfall-runoff model that operates on an hourly or daily time step. It includes modules that describe snow buildup and melting, glacier melting, soil moisture, groundwater, and flow generation. Snow and glacier melt are considered using distribution function approach with combined radiation degree day concept (Cazorzi & Dalla Fontana, 1996). The potential evapotranspiration via the Hargreaves formula (Hargreaves & Samani, 1982) while the simulation of surface and subsurface flows is carried out by means of the Probability Distribution Model (PDM) from Moore (2007). The detail of the model is provided in Annex B with further detail regarding model parameters, calibration and validation in Annex C. The performance of the ICHYMOD model has been well replicated by other work that has been done in the alpine basin (Avesani et al., 2022; François et al., 2017; Mei et al., 2016; Puspitarini et al., 2020; Stergiadi et al., 2020). Hence these work provide a strong background to apply the ICHYMOD model for the purpose of this work.

4.4.2 Daily model simulation and reference datasets

The ICHYMOD hydrological model performs at hourly timestep, however, to integrate the daily ECMWF forecast dataset, the model was modified to the 'daily' resolution. To achieve this, the hourly observed precipitation was summed to obtain the daily value and again, the precipitation was equally divided into 24 values representing the hours of the day. In the case of temperature, the daily average temperature was considered for the 24 hours of the day. Based on the updated variables, the model was run at hourly timestep and the values were averaged to obtain the daily

Sn	Name	KGE	
		Hourly	Daily
1	Rio Plan	0.832	0.831
2	Rio Riva a Seghe	0.824	0.795
3	Rio Anterselva Bagni	0.807	0.825
4	Braies	0.686	0.705
5	Rio Riva a Caminata	0.783	0.772
6	Rio Casies Colle	0.866	0.859
7	Gadera a Pedraces	0.825	0.804
8	Aurino a Cadipietra	0.879	0.878
9	Ridanna a Vipiteno	0.899	0.870
10	Gadera Mantana	0.810	0.802
11	Passirio	0.842	0.814
12	Aurino a Caminata	0.892	0.881
13	Aurino a S.Giorgio	0.900	0.899
14	Rienza Vandoies	0.875	0.885
15	Ponte Adige	0.849	0.840
16	Adige a Bronzolo	0.873	0.870

Table 4.1: KGE performance of hourly and daily model

runoff values. Even though in essence the daily model is run at an hourly timescale, this slight modification is required to assimilate the ECMWF dataset. Since there was a minimum change in the KGE performance between the 'daily model' and the 'hourly model', as shown in table 4.1 the daily model was deemed suitable for all the analyses considered hereafter. The comparison was done taking timeperiod of 1992 till 2019. While for the purpose of runoff forecasting , the daily model was simulated from October 1993 to October 2016, to include only the hindcast period of the ECMWF.

To remove the error associated with the hydrological model and also to have a seamless runoff value in case of a missing gauge dataset, a surrogate dataset of daily runoff simulation was required. The reference is considered as the station data input daily runoff simulation, as mentioned in the previous paragraph. The model parameter errors can be reduced with proper calibration, which allows simulated results to be close to the observations (Moradkhani & Sorooshian, 2009).

ICHYMOD remains the core of the forecast system, which will be run in reference mode taking meteorological observations as input for predicting past and initial conditions and in forecast mode taking the meteorological forecast as driving forces for predicting future hydrological states. All the error metrics will consider this reference as a benchmark simulation.

4.4.3 Bias-Correction of ECMWF (STAT-ECMWF)

The temperature and precipitation variables from ECMWF were bias-corrected by quantile mapping considering the station (STAT-ECMWF) and ERA5 renalysis (ERA-ECMWF) dataset before its' application for runoff forecasting.

Figure 4.3 (a and b) show the bias present in the ECMWF precipitation (RAW-ECMWF) forecast for lead time (1 and 6 months). To put into perspective the concept of leadtime (LT) horizon, for LT of 1 month the January plot represents all the values initialized on January 1st, while for LT of 6 months, it represents the value initialized on August 1st. This figure visualizes the monthly precipitation boxplot based on the daily values for the respective inputs.

The hindcast seasonal forecasts (SF) of temperature, and precipitation from ECMWF System5 over the period 1993 – 2016 were collected and the monthly aggregates were computed. The original SF covers the study basin with a spatial resolution of 1°x1°, composed of 25 members and 6 months LT. Since there are two reference dataset used to bias correct the SF dataset, two quantile mapping computation taking ERA and station dataset as calibration of adjustments is considered. In case of the STAT-ECMWF, it is upscaled to 0.25 degrees spatial resolution. To achieve this, the Thiessen polygon method was used to redistribute the observed hourly station precipitation data to each 0.25 degrees cell center.The Snow Correction Factor (SCF) is set equal to 1.4, in agreement with those reported in previous works for the Alps (Carturan et al.,2012). SCF is used to consider the rain gauges' inability to catch snow during the snowfall period. There are 23 stations grid covering the study area as shown in Figure 4.4. In the case of the temperature, it remains the same as the daily stations' lapse rate. While in case of reference as ERA, the precipitation from the same sets of ERA grids as STAT is used in its' original form to correct the SF. Since the previous analysis shows ERA5 also has overall bias as compared to the stations, i.e. around 1.36 for the basin at Bronzolo, the bias

correction from ERA is further divided by this constant. The resulting precipitation is termed ERA-ECMWF.

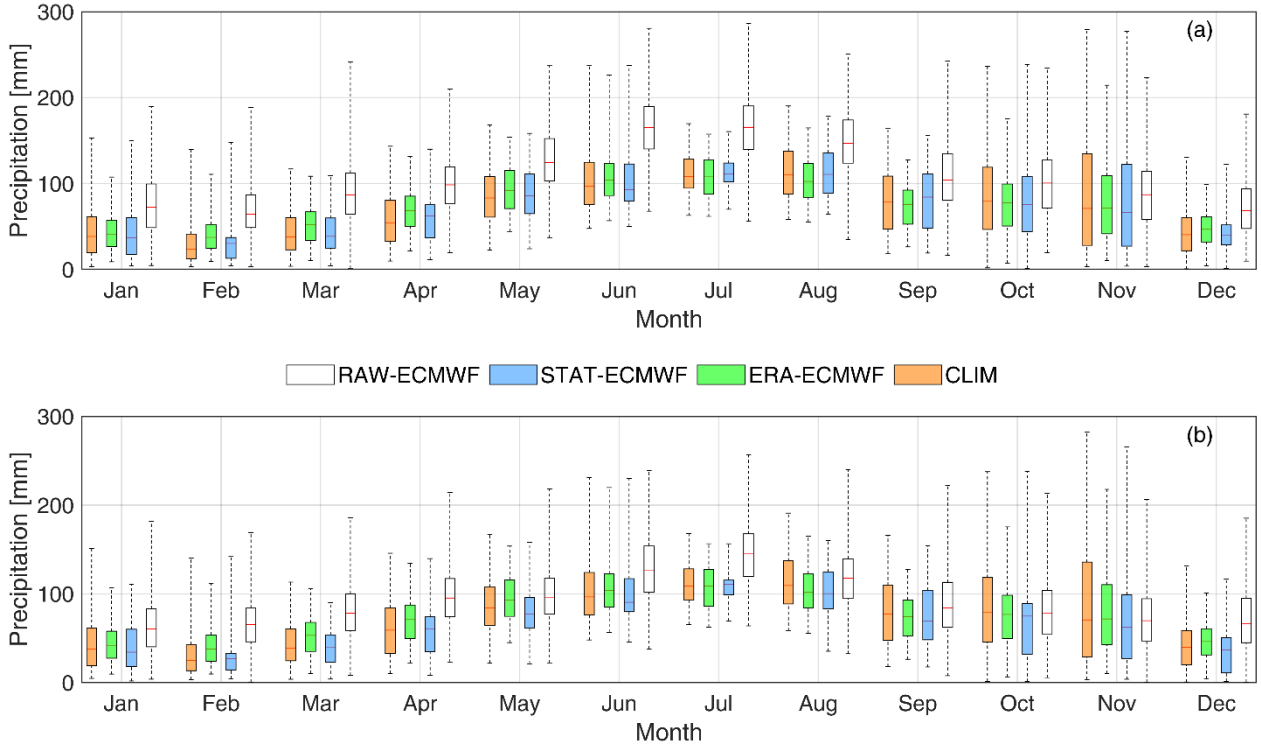

Figure 4.3: Precipitation for RAW-ECMWF, STAT-ECMWF, and CLIM for lead time (a) one month (b) six months

Five spatial ECMWF grids as shown in Figure 4.4 cover the study area. Firstly, all forecast dataset for each lead time and member were spatially disaggregated from their original 1°x1° grid onto the same grid of ERA5 data at 0.25°x0.25° spatial resolution in order to match with ERA5 grid. The spatial disaggregation was performed by applying a bilinear interpolation of the original SF fields. Bilinear interpolation estimates the value at each point of the target grid by means of a weighted average of the values at the four nearest points in the original 1°x1° grid. The weights of the four nearest cells are determined by their distance from the target point.

The bias correction of resampled SF fields was based on the quantile-mapping (QM) procedure. This method aims to define a quantile-specific transfer function determined by the mismatch between modeled and reference empirical cumulative distribution functions over a common calibration period. The transfer function is determined at a set of physical values corresponding to a span of discrete quantiles and then interpolated to correct all physical values.

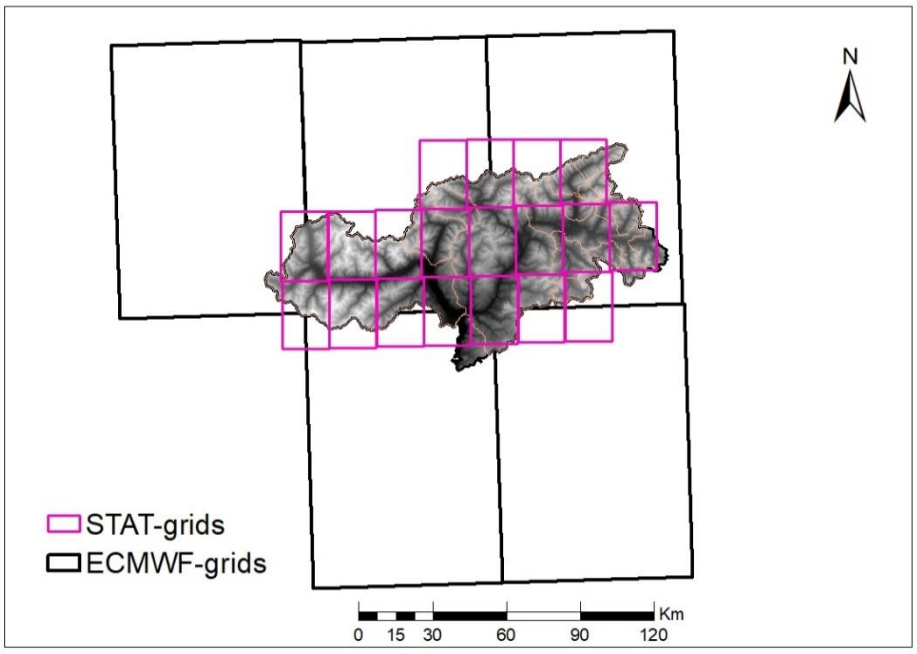

Figure 4.4: The ECMWF and Station grid cell covering the study basin

Firstly, all monthly SF for each lead time and member were spatially disaggregated from their original 1°x1° grid onto the same grid of ERA5 data at 0.25°x0.25° spatial resolution. The spatial disaggregation was performed by applying a bilinear interpolation of the original SF fields. Bilinear interpolation estimates the value at each point of the target grid by means of a weighted average of the values at the four nearest points in the original 1°x1° grid. The weights of the four nearest cells are determined by their distance from the target point.

Secondly, the bias correction of resampled SF fields was based on the quantile-mapping (QM) procedure. This method aims to define a quantile-specific transfer function determined by the mismatch between modeled and reference empirical cumulative distribution functions over a common calibration period. The transfer function is determined at a set of physical values corresponding to a span of discrete quantiles and then interpolated to correct all physical values. For each LT and month, the transfer function was calibrated for each point of the 0.25°x0.25° grid by comparing all monthly station values from 1993– to 2016 (23 values) with the SF monthly values of all members over the same interval (23 x 25 values). The correction was applied to the monthly values of all members, for the LT and month under consideration, over the whole period 1993–2016. Both ERA (ERA-ECMWF) and station (SIER-ECMWF) were considered to bias corrected the forecast dataset.

The bias correction made for the Raw-ECMWF precipitation by stations and ERA is also shown in Figure 4.3 concerning both the LT of one month and six months. SEAS5 re-forecasts, before January 1, 2017, have the atmosphere initialized by ERA-interim (Johnson et al., 2019). Even though ERA5 has a better data assimilation technique, in both time and space, compared with ERA-interim, the two reanalyses are expected to perform similarly (Calì et al., 2022). The behavior of ERA re-analysis is also portrayed in the ECMWF forecast dataset as well, in which there is a positive precipitation

bias regardless of the months and the forecast horizon. The application of quantile-mapping (QM) considering the station data, substantially reduces the bias present in the raw forecast.

4.4.4 Seasonal forecast STAT-ECMWF

To forecast the runoff based on the STAT-ECMWF, for a given forecast day, the initial state of the model is provided by the reference simulation. For each forecast day, to simulate the daily model, the daily STAT-ECMWF precipitation is equally divided into 24 hourly values, the STAT-ECMWF daily temperature is provided as hourly values, and thus obtained hourly runoff is average to obtain the daily runoff. Starting from the forecast day, the ICHYMOD model is forced with the 25 ensembles of meteorological variables with a lead time of 1 day to 6 months, resulting in 25 ensemble members of runoff values for each forecast horizon.

4.4.5 Seasonal forecast ERA-ECMWF

The ERA-ECMWF runs on the same model setup as the STAT-ECMWF, however in case of ERA-ECMWF, the model is simulated with the bias corrected meteorological variables of SF with ERA. For each LT, there remains 25 ensemble members of runoff forecast as similar with STAT-ECMWF.

4.4.6 ESP (Extended Streamflow prediction)

Similar to the ECMWF input, the reference simulation provides the initial catchment state. The model is then forced with the past/future observed dataset until the 6-month forecast horizon. However, it should be noted that the data concerning the forecast horizon is excluded from the analysis. For instance, to forecast the 5th day (5 January 1994), starting from January 1, 1994, all the 5th January data from 1991 till 2016 is considered however the forecasted day (5 January 1994) data itself is excluded. In this study the ESP considers data set from 1991 till 2016, hence excluding the forecasted day variables, there will be 24 ensemble runoff forecasts.

4.4.7 Error Metrics

To quantify the goodness of the agreement between the forecasted runoff ensembles and the observed states, there are few error metrics considered in this study. We have considered both the deterministic and probabilistic measures of skill to assess the forecast performance as done by (Monhart et al., 2019).

The raw forecast precipitation product due to its' coarse spatial representation distorts the vital local features like heterogeneous topography, which leads to notable bias (Anghileri et al., 2019). The forecast also suffers from a lead-time-dependent bias termed drift (Manzanas, 2020). Even though the bias in the ECMWF meteorological variable is corrected prior to its' application for runoff forecasting, there will be always a certain amount of error associated with it. Since the precipitation bias translates to the runoff values as well, the bias of the runoff forecast needs to be evaluated.

The runoff bias for a given LT as shown in equation 4.1 is computed by averaging all the ensemble members of the forecast and compared with the reference (model simulated runoff).

$$\mathrm{Bias(LT)} = \frac{1}{\mathrm{Y}}\sum_{\mathrm{y=1}}^{\mathrm{Y}} \frac{\mathrm{Q}_{\mathrm{mean-m,y}}^{\mathrm{FOR}}(\mathrm{LT}) - \mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}(\mathrm{LT})}{\mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}(\mathrm{LT})} \qquad \text{equation 4.1}$$

Where,

LT: Lead time (days)

$\mathrm{Q}_{\mathrm{mean-m,y}}^{\mathrm{FOR}}$: Mean runoff over the m ensemble members for month y. m=24 for ESP, m=25 for ECMWF

$\mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}$: Reference simulation

Y: Number of years (1993-2016)

Here, the bias value of one suggest no error in the forecast whereas, value largerthan one suggests overestimation of the reference and vice-versa. Nash–Sutcliff efficiency (NSE) is generally applied to assess the performance of the hydrological model with the observations (Nash & Sutcliffe, 1970). Whereas, in this study, NSE is used to quantify the forecast runoff error based on its' ensemble mean. NSE is defined as the normalized value, which calculates the relative strength of the residual variance ("noise") against the variance of the reference ("information") as given in Equation 4.2. The value of the NSE ranges from -∞ until 1. For the perfect forecast, there is no "noise" in the forecast and the NSE equals 1. Conversely, if the forecast produces the residual variance equal to the reference variance, it results in an NSE of 0, which indicates the mean forecast has the same predictive skill as the mean of climatology. An efficiency of less than zero (NSE < 0) occurs when the climatology is a better predictor than the forecast.

$$\mathrm{NSE(LT)} = 1 - \frac{\sum_{\mathrm{y=1}}^{\mathrm{Y}}\left(\mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}(\mathrm{LT}) - \mathrm{Q}_{\mathrm{mean-m,y}}^{\mathrm{FOR}}(\mathrm{LT})\right)^2}{\sum_{\mathrm{y=1}}^{\mathrm{Y}}\left(\mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}(\mathrm{LT}) - \mathrm{Q}_{\mathrm{mean}}^{\mathrm{REF}}(\mathrm{LT})\right)^2} \qquad \text{equation 4.2}$$

Where,

LT: Lead time (days)

$\mathrm{Q}_{\mathrm{mean-m,y}}^{\mathrm{FOR}}$: Mean runoff over the m ensemble members for month y. m=24 for ESP, m=25 for ECMWF

$\mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}$: Reference simulation

$\mathrm{Q}_{\mathrm{mean}}^{\mathrm{REF}}$: Mean of reference simulation

Y: number of years (1993-2016)

Moriasi et al., 2015 provided a range of NSE values as given in table 4.2, for hydrological analysis considering daily, monthly and, annual timescale.

NSE Criteria	Unsatisfactory	Satisfactory	Good	Very good
	NSE=< 0.50	0.50 < NSE =< 0.70	0.70 < NSE =< 0.80	NSE > 0.80

Table 4.2:Range of NSE values

Even though the bias and NSE can describe the average behavior of the ensemble values, these deterministic error metric fails to address the substantial skill coming from the spread of the forecast. Hence better metrics of comparison are required to quantify the added value from the spread; especially at the seasonal scale (Kumar et al., 2014). To benefit from the probabilistic nature of the forecast, two other statistics are used i.e. continuous Ranked Probability Skill Score (CRPSS) and spread to error ratio (SprERR).

The continuous Ranked Probability Score (CRPS) is a probabilistic verification score that considers the full ensemble i.e. forecast error and forecast spread, to understand how much the forecasts agree with the reference (Hersbach, 2000). CRPS calculates the area between the cumulative distribution function (CDF) of the forecasted ensembles and the reference. The calculation of CRPS is given in equation 4.3 and visually represented in Figure 4.5.

$$CRPS(F_t - y_t) = \int_{-\infty}^{+\infty} (F_t(y) - H(y - F_t))^2 dy \qquad \text{Equation 4.3}$$

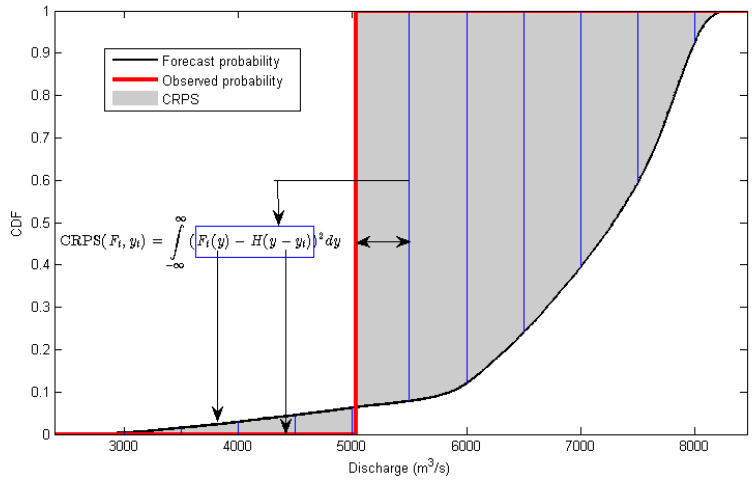

Figure 4.5: CRPS calculation representation

The CRPS value ranges from zero until ∞, where the null value corresponds to a perfect forecast while any increase in the value represents the forecast degradation.

Whereas the CRPSS (Continuous Ranked Probability skill score), as given in equation 4.4, quantifies the usefulness of a given forecast with respect to the benchmark (climatology). In this study, only the calculation of CRPSS is done. CRPSS is a dimensionless value, which ranges from −∞ to 1 (perfect forecast). The CRPSS equal to 0 indicates the forecast to be as good as the benchmark and the negative value indicates the forecast to be worse than the benchmark (Anghileri et al., 2019).

$$\mathrm{CRPSS(LT)} = 1 - \frac{\sum_{y=1}^{Y} \mathrm{CRPS}_y^{\mathrm{FOR}}(\mathrm{LT})}{\sum_{y=1}^{Y} \mathrm{CRPS}_y^{\mathrm{BEN}}(\mathrm{LT})} \qquad \text{equation 4.4}$$

Where for a given leadtime, $\mathrm{CRPS}_y^{\mathrm{FOR}}(\mathrm{LT})$ represents the CRPS of a given forecast scheme.

$\mathrm{CRPS}_{\mathrm{y}}^{\mathrm{BEN}}(\mathrm{LT})$ represents the benchmark CRPS value based on climatology.

Harrigan et al., 2018 provided the range of CRPSS values as given in table 4.3 to evaluate any skills coming from the forecast. Additionally, the CRPSS value near zero i.e. within ± 0.05 is termed as neutrally skillful (Bennett et al., 2017).

CRPSS	<0	0	0 to 0.25	0.25:0.5	0.5 to 0.75	0.75 to1
Skill	Negative	No	Low	Moderate	High	Very high

Table 4.3: CRPSS based skill

SprErr represents the reliability of forecast (Hopson, 2014), in which the ratio between the forecast ensemble i.e. forecast spread, and the mean squared error of the ensemble forecast i.e. forecast error is calculated as given in equation 4.5. SprErr value of 1 corresponds to a reliable forecast, in which the forecast spread and error are equal. For cases when the SprErr value is less than 1, it signifies overconfidence (larger errors compared to spread) and when the SprErr value is greater than 1 indicates overdispersion (larger spread compared to error) (Monhart et al., 2019).

$$\mathrm{SprER}(\mathrm{LT}) = \frac{1}{\mathrm{Y}}\sum_{\mathrm{y}=1}^{\mathrm{Y}} \frac{\sum_{\mathrm{m}=1}^{\mathrm{M}}\left(\mathrm{Q}_{\mathrm{y,m}}^{\mathrm{FOR}}(\mathrm{LT}) - \mathrm{Q}_{\mathrm{mean-m,y}}^{\mathrm{FOR}}(\mathrm{LT})\right)^2}{\sum_{\mathrm{m}=1}^{\mathrm{M}}\left(\mathrm{Q}_{\mathrm{y,m}}^{\mathrm{FOR}}(\mathrm{LT}) - \mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}(\mathrm{LT})\right)^2} \qquad \text{equation 4.5}$$

LT: Lead time (days)

$\mathrm{Q}_{\mathrm{y,m}}^{\mathrm{FOR}}$: Forecast for given month and ensemble.

$\mathrm{Q}_{\mathrm{mean-m,y}}^{\mathrm{FOR}}$: Mean runoff over the m ensemble members for month y. m=24 for ESP, m=25 for ECMWF

$\mathrm{Q}_{\mathrm{y}}^{\mathrm{REF}}$: Reference simulation

$\mathrm{Q}_{\mathrm{mean}}^{\mathrm{REF}}$: Mean of reference simulation

Y: number of years (1993-2016)

4.5 Results

The results include the hydrologically relevant variables of total runoff and baseflow concerning STAT-ECMWF, ERA-ECMWF and ESP forecast compared to the climatology. It should be noted that the results of the sixteen basins are consolidated based on the specific lead days as represented in table 4.4. In the case of days (D), the result specifies the given target day, while for weeks (W) and months (M), the results are averaged for a given range of days, for instance, M2 represents the second target month in which the error metrics are averaged from day 29 until 56th day.

Title	D2	D4	D6	W2	W3	W4	M2	M4	M6
Day	2nd	4th	6th	8 to 14	15 to 21	22 to 28	29 to 56	85 to 112	141 to 168

Table 4.4: Benchmark leadtime

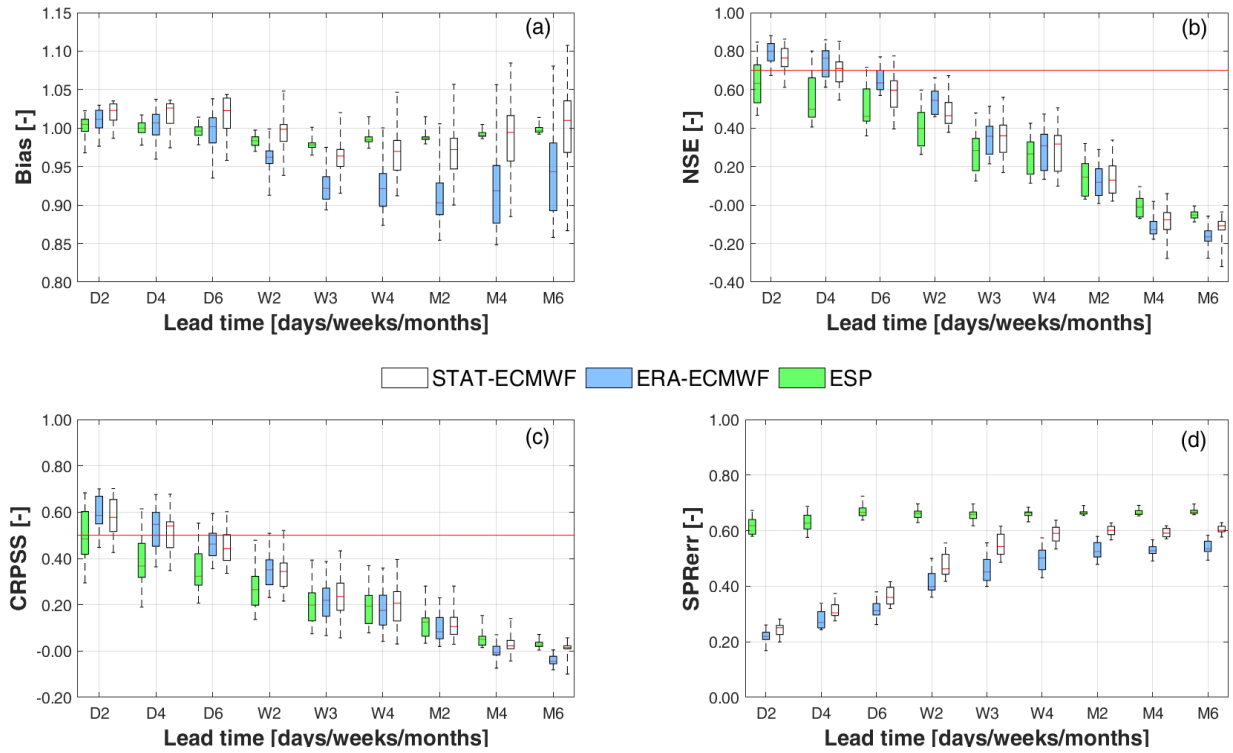

Figure 4.6:Total flow performance of (a) Bias (b) NSE (c) CRPSS (d) SPRerr

Figure 4.6 shows the performance of different error metrics concerning total flow with the benchmark LT. The bias (Figure 4.6a) for the ESP remains close to the one suggesting it to be closer to climatology. In case of ERA-ECMWF and STAT-ECMWF, the bias shows similar performance until 1st week (D6) with value remaining slightly greater than 1, while from W2 onwards, the nature of bias starts differentiating. From W2 till M6, STAT-ECMWF remains relatively closer to one as compare to ERA-ECMWF. However, the nature of bias shows similar behavior with a gradual decrease in value below one until W4, and after, there is an increase in bias until M6. This increase is comparatively sharper for STAT-ECMWF and for M6 the bias is greater than 1, while for ERA-ECMWF, the bias is always less than zero from W2 onwards. Interesting, ESP also follows the similar routine as both ECMWF correction but the spread for ESP is considerably small irrespective of the LT and the value is always much closer to 1.

Figure 4.6b shows the apparent decay of NSE with an LT for the traditional ESP and for both the ECMWF corrections. The NSE value is within the accepted range (~0.5) until the second week for the ECMWF (STAT and ERA) while for ESP, the NSE falls below 0.5 after the first week itself. The ESP underperforms ECMWF until the first month but from LT 2 months onwards, the NSE performs slightly better as compared to ECMWF (STAT and ERA). It should be noted that NSE value shows no credible skills for all the forecast considered from LT M2. Even though the difference in STAT and ERA corrections is negligible, ERA-ECMWF shows a better performance until W2 and for further time horizon, STAT-ECMWF is better.

The CRPSS value follows the same routine as the NSE value. The skills remain moderate (~0.5) until the first week for ECMWF (STAT and ERA) but for ESP the value goes below 0.5 within the first few days (D4). Corollary to the NSE values, the ECMWF (STAT and ERA) shows better performance

against the ESP method until one month (W4), while ESP has comparatively better performance after the 2nd month (M2). While comparing STAT and ERA, until W2, both behave similarly with ERA performing slightly better whereas from W2, the STAT leads the performance until M6. It is key to note, the median value of ERA-ECMWF goes below zero from M4 onwards but for STAT-ECMWF, it is always positive. Likewise, the CRPSS value of ESP the value never goes below 0.

The SPRerr shows distinctly different characteristics for the ESP and the ECMWF variables. For ESP, the value remains within 0.6 and 0.7 for all the LT, and also the interquartile range (IQR) diminishes with LT. On the other hand, for both the ECMWF, the initial days are characterized by smaller values with an increase in value until the first month (W4) and after that, the values remain constant (~0.6) for STAT-ECMWF and (~0.5) for ERA-ECMWF. While the SPRerr value of ESP remains higher than the ECMWF in all cases, STAT-ECMWF has comparatively higher SPRerr value than ERA-ECMWF. For both the forecasts (ESP and ECMWF) as the SPRerr value is less than 1, it indicates overconfidence. As for ESP, since the forecast is based on the past meteorological data, it can be expected that the ensemble forecast always remains within the certain range. This results in the similar forecast spread leading to similar SPRerr values. While in case of ECMWF, in the initial days the low value of SPRerr can be associated to the small forecast error since the forecast spread is small. However with increasing LT, the uncertainty keeps increasing including range of forecast. The correction made by station and ERA helps to keep the SPRerr close to ESP after 1 month LT of forecast.

In general, the added benefit of using ECMWF (STAT and ERA) as compared to ESP is seen in the first month but after that, the ESP seems to be the better method for the total flow forecast. The forecast skill in terms of statistics shows useful information for both ECMWF until the second week while for ESP, there is an immediate loss of skill after the first week (D6). While comparing ERA-ECMWF with STAT-ECMWF, even though their performance looks similar, ERA correction perform marginally better in terms of NSE and CRPSS until W2, whereas after W3, STAT-ECMWF perform better to some extent.

Figure 4.7 shows the performance considering only the baseflow. As similar to total flow, the baseflow bias (Figure 4.7a) for ESP remains close to one regardless of the LT with a small IQR. For the ECMWF (STAT and ERA), the bias is close to one until W2 and the value disperses from one until the end of forecast horizon. The IQR for both ECMWF corrections is increasing with the LT and is much higher as compared to ESP. For STAT-ECMWF, the bias is > 1 (W2), and after, it decreases further until (M4), and for M6, the value slightly improves with value > 1. For ERA-ECMWF, the nature of graph is similar to STAT-ECMWF until D6 and the value dips below one for M2 and even though there is a slight improvement for M4 and M6, the value never exceeds 1.

The NSE (Figure 4.7b) for the 2nd day looks comparable for all the forecasts with a median NSE ranging above 0.8. As seen with the result of total flow, there is a sharp decrease in NSE for ESP which can be considered skillful (~0.5) until the second week while for the ECMWF, the NSE remains valuable (~0.5) until the 3rd week. After M2, the ESP outweighs ECMWF performance until the end end of six month forecast. As similar to total flow, ERA-ECMWF performs slightly better than STAT-ECMWF until W4 and after, STAT-ECMWF leads the NSE value.

For CRPSS (Figure 4.7c) as well the performance is gradually decreasing until the end of six months for all the forecast. The ESP is deemed acceptable (~0.5) until W1 whereas for the ECMWF, the permissible CRPSS skill is seen until the 2nd week (W2). After performing similarly to ECMWF for W4, ESP outperforms both ECMWF correction for LT of 2 and 6 months. While comparing ERA and STAT correction, STAT-ECMWF slightly leads the ERA-ECMWF until W2, but after, STAT-ECMWF has relatively higher CRPSS value. The CRPSS value never becomes negative for ESP and for STAT-ECMWF, the median value is always positive however for ERA-ECMWF, the median value is negative after M4.

The SPRerr (Figure 4.7d) for the ESP is higher as compared to the ECMWF for all the LTs, where the value for ESP gradually increases from ~0.5 (D2) to ~0.65 (D6) and remains similar until M6. While for STAT-ECMWF, the SPRerr gradually increases from <0.2(D2) and reaches a stable value (~0.6) after W4. For ERA-ECMWF, the value is similar to STAT-ECMWF until W2 and after, the value diverges with a similar value (~0.5) from W4 onwards.

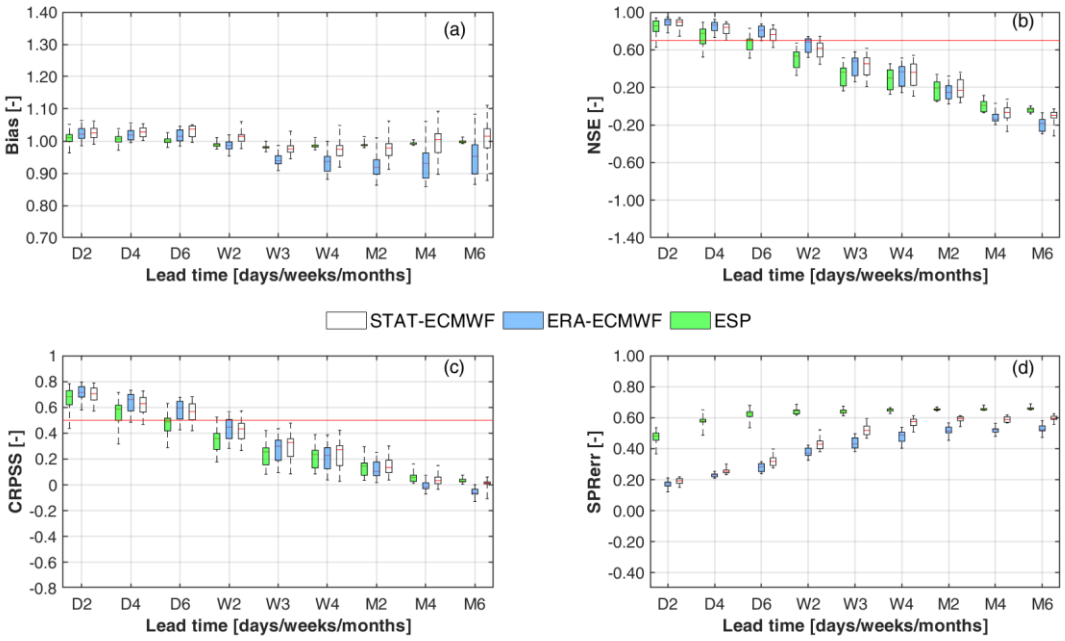

Figure 4.7:Baseflow performance of (a) Bias (b) NSE (c) CRPSS (d) SPRerr

The baseflow was computed subtracting the surface runoff with the total runoff. Briefly, the baseflow performances are better as compared to the total flow concerning all the inputs. Especially, there is an extension of the one-week time window as compared to total flow, in which the forecast could be deemed skillful. The ESP and ECMWF (STAT and ERA) remain skillful in terms of NSE until the end of 2nd and 3rd week respectively. When the CRPSS value is taken as reference, the ESP and ECMWF (STAT and ERA) remain skillful until the end of 1st and 2ndweek respectively. While comparing ERA-ECMWF with STAT-ECMWF, the baseflow results look similar to the result of totalflow. Even though the performance of ERA-ECMWF and STAT-ECMWF looks similar, ERA

correction perform slightly better in terms of NSE and CRPSS, until W2, whereas after W3, STAT-ECMWF shows slightly better results amongst the two corrections.

Figure 4.8 shows the NSE, CRPSS, SPRerr, bias for the total and baseflow comparison for the STAT-ECMWF values. Even though, both the bias (Figure 4.8a) shows similar value, for baseflow the bias is slightly higher. The bias always remains within the limit of ±5%, which remains positive until the W2 for both the forecast. The bias gradually increases from a negative value in W3 to a positive value for M4 (BF) and M6 (TF). The range of bias remains similar for both forecast in which the spread in bias gradually increases from W4 onwards.

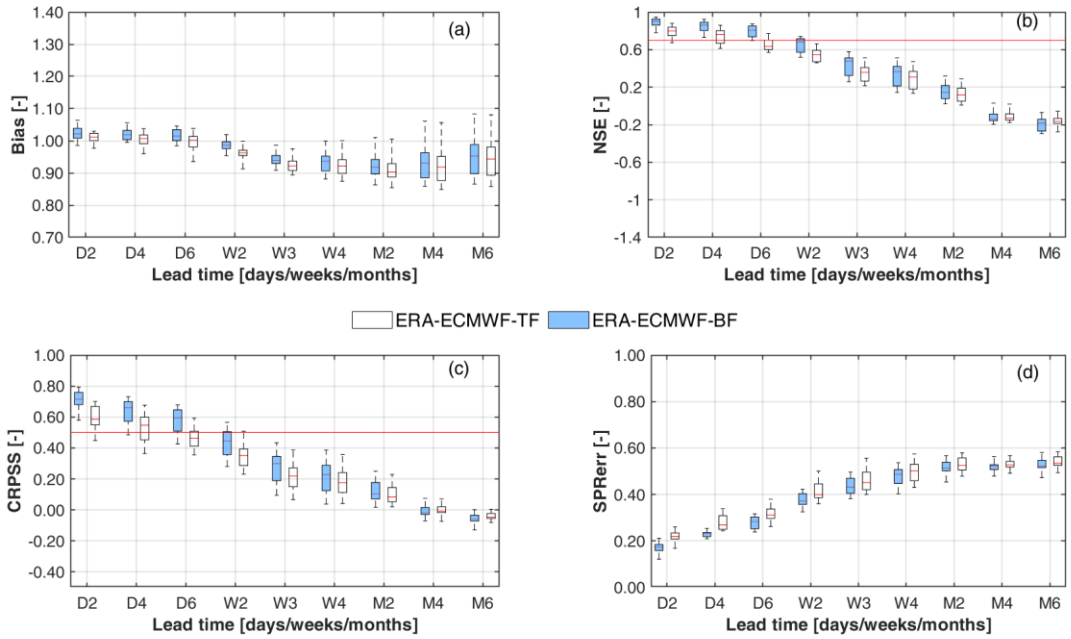

Figure 4.8:Total and baseflow performance of STAT-ECMWF of (a) Bias (b) NSE (c) CRPSS (d) SPRerr

For both the flows, there is a gradual decrease of NSE (Figure 4.8b) from D2 until M6 in which the baseflow is performing better than total flow until M2. After M4, both the forecast perform similarly with values below 0. As reported previously, the NSE value remains acceptable (~0.5) until W2 and W3 for total flow and baseflow respectively. The CRPSS (Figure 4.8c) also follows a similar pattern as NSE where the baseflow performs slightly better considering all the LT. The acceptable CRPSS value (~0.5) is W2 for baseflow while for total flow it is D6. For both the flows, the IQR of CRPSS is largest for the W3 which constantly diminishes until M6 and becomes negative for M4 and M6.The SPRerr (Figure 4.8d) of the total flow is slightly higher than baseflow for all the LT. For both the variables, the SPRerr is gradually increasing until W4 and it remains constant (~0.6) until M6.

To summarize, until the second month the baseflow runoff performs (NSE, CRPSS) better than the total flow while for longer LT (M4, M6) their performance remains similar. The nature of the graph (IQR) also remains similar for all the statistics considered.

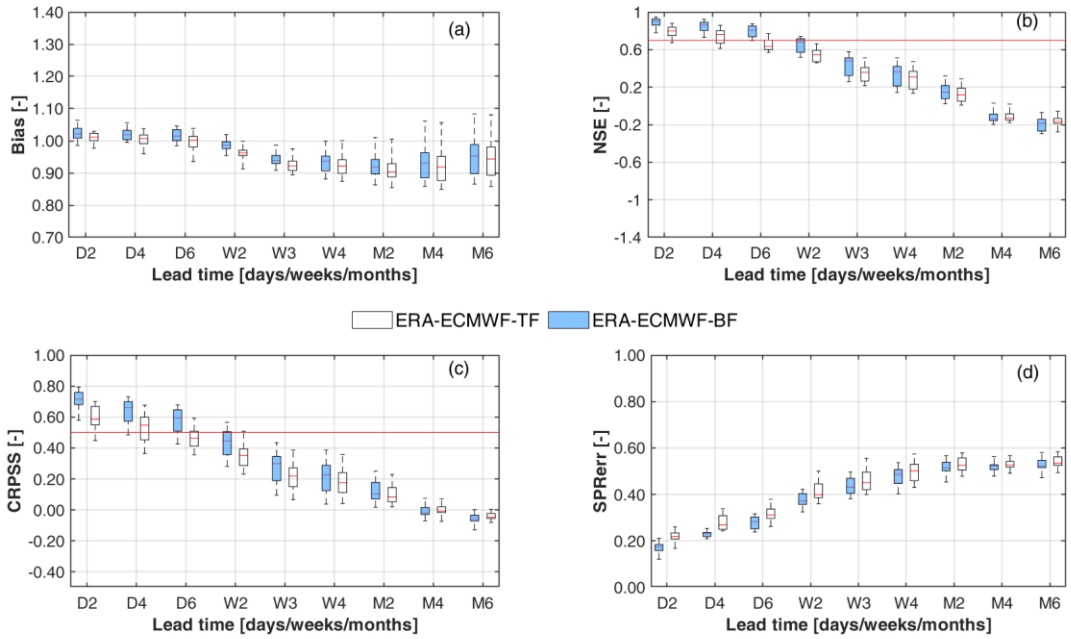

Figure 4.9:Total and baseflow performance of ERA-ECMWF of (a) Bias (b) NSE (c) CRPSS (d) SPRerr

Figure 4.9 shows the NSE, CRPSS, SPRerr, bias for the total and baseflow comparison for the ERA-ECMWF values. Even though, both the bias (Figure 4.9a) shows similar value, for baseflow the bias is only slightly higher. Both the bias has similar value close to one until D6 and the value decreases close to 0.9 (M2) and there is a slight improvement for M4 and M6. With the increase in LT, the range of the bias also increases for both the flows.

For both the flows, there is a gradual decrease of NSE (Figure 4.9b) from D2 until M6 in which the baseflow is performing better than total flow until M2. From M4, both the forecast perform similarly with negative values. As reported previously, the median NSE remains acceptable (~0.5) until W2 and W3 for total flow and baseflow respectively. The CRPSS (Figure 4.9c) also follows a similar pattern as NSE where the baseflow performs better than total flow until M2 and after, TF is only slightly better. The median acceptable CRPSS value (~0.5) is W2 for baseflow and D6 for total flow. The IQR of CRPSS is largest for the W3 which constantly diminishes until M6. The CRPSS value for both the flows becomes negative for M4 and M6. The SPRerr (Figure 4.9d) of the total flow is slightly higher than baseflow for all the LT. For both the variables, the SPRerr is gradually increasing until W4 and it remains constant (~0.5) until M6.

To summarize, until the second month the baseflow runoff performs (NSE, CRPSS) better than the total flow while for longer LT (M4, M6) their performance remains similar. The nature of the graph (IQR) also remains similar for all the statistics considered.

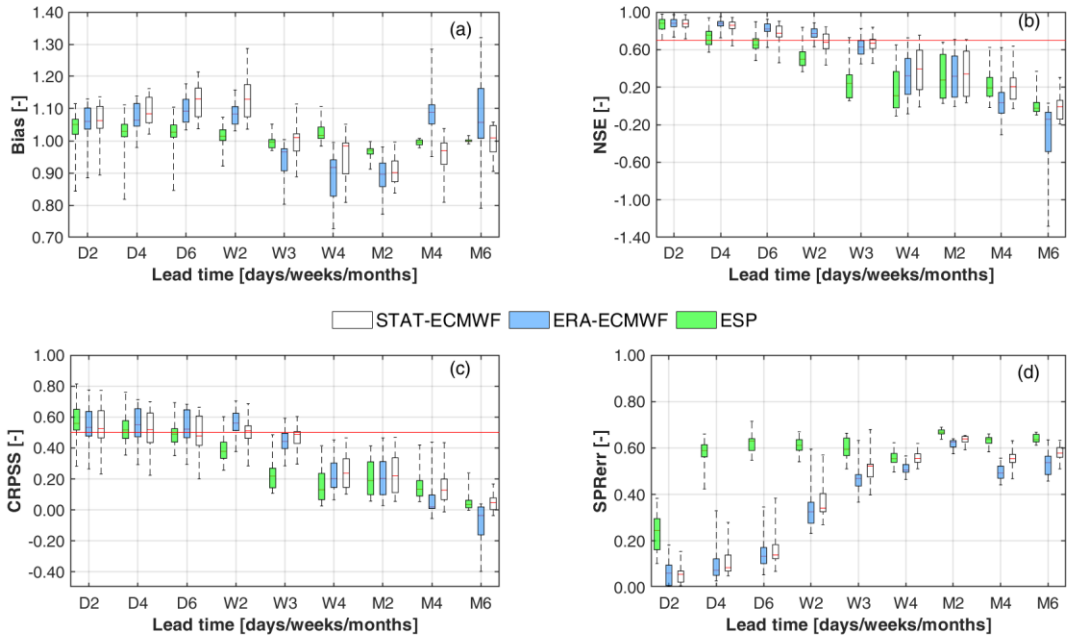

Figure 4.10: Baseflow performance of STAT-ECMWF, ERA-ECMWF, ESP for March initialization (a) Bias (b) NSE (c) CRPSS (d) SPRerr

Since the forecast is highly dependent on initial hydrological state of a given watershed, two experiment were considered with March and August as the initialization month. Normally for the study area, February/March signifies the termination of the snow accumulation and with increasing temperature from March onwards induces runoff owing to snow melt. The temperature starts dipping after August and rain is expected which leads to snow storage. Hence given the two varying nature of the initialization month, the two experimental setup were considered for their effect in baseflow simulation.

Figure 4.10 shows the baseflow comparison considering the initialization month of March. For D2 all the forecast show similar bias (Figure 4.10a), however for all the other LTs, the nature of bias is contrasting. In case of ESP, the bias remains always close to one for all the LT, however, the IQR is large for the initial days which diminishes with the LT. The nature of the graph for the ECMWF (ERA and STAT) is similar for all LTs, with bias > 1 until W2 which decreases to ~0.9 for M2 and after that, the value starts increasing close to 1. However, the STAT-ECMWF has higher bias than ERA-ECMWF from D4 onwards upto W4 and after performing similarly for M2, STAT-ECMWF displays lower bias as compared to ERA-ECMWF for M4 and M6. Another feature that is distinctive is the higher spread of ERA-ECMWF bias as compared to STAT-ECMWF. With the exception for W2, the ERA-ECMWF always has higher range in bias as compared to STAT-ECMWF and especially for M4 and M6, the bias variability of ERA-ECMWF is exceptionally high.

The NSE (Figure 4.10b) value shows a comparable value for the D2 for all the forecast considered. The NSE value shows ESP having useful information (NSE~0.5) until W2 and there is a gradual decrease until the W3. After W3, the NSE hovers within 0.1 and 0.3 until the end of M4 and the value becomes negative for M6. In comparison, the ECMWF (STAT and ERA) shows robust NSE value

(>0.6) until the W3, while for W4 and M2, the NSE remains ~0.4. The NSE shows gradual degradation with value ~0.2 (M4) and ~0 (M6). In case of ERA-ECMWF, for M6 the range of NSE is quite large in comparison to STAT-ECMWF. For LT of M4 and M6, ESP and STAT-ECMWF show comparable performance with ERA-ECMWF slightly underperforming as compared to the other two. ERA-ECMWF also displays a large IQR for NSE for M6 which is quite distinctive as compared to the other two forecasts.

The CRPSS (Figure 4.10c) have similar values for the D2 concerning all the forecast, in the case of ESP the value decreases until the W3 and it remains similar until M4 (~0.2) while CRPSS is closer to zero for the 6th month. For ECMWF (ERA and STAT), the CRPSS value is acceptable (~0.5) until the 3rd week while this value is D6 for ESP. For W4 and M2, the value remains similar (0.2) concerning all forecasts. For M4 and M6, the ESP performs slightly better than the other two forecasts. As seen in bias and NSE, the CRPSS of M6 of ERA-ECMWF shows the value reaching -0.40 which is extremely low and differentiating characteristics as compared to the other two forecasts.

The SPRerr (Figure 4.10d) for ESP is always higher than the STAT-ECMWF for all the LT. ESP SPRerr value jumps from ~0.25 (D2) to ~0.6 (D4) and this value is sustained until M6. This nature is different for the ECMWF (STAT and ERA), where the value is ~0.1 (D2), which gradually increases to ~0.5 (W3) and it remains similar till M6, except for M2 (~0.6).

In a contrast to the March initialization, the August initialization (Figure 4.11) shows a comparable performance for ESP and ECMWF (STAT and ERA). The bias (Figure 4.11a) is close to one for ESP with relatively small IQR as seen in previous analyses. For ECMWF (STAT and ERA), the value remains similar until M2, while for M4 and M6, the values are widely differing. For STAT-ECMWF, the bias is close to one for M4 and ~1.15 for M6, whereas, for ERA-ECMWF the value is close to 0.85 for both M4 and M6. Comparatively, for all the LT, the ERA-ECMWF has higher spread of the bias nonetheless, for M4 and M6, the STAT-ECMWF also shows higher bias spread.

In case of NSE (Figure 4.11b), all the forecast shows decay with time however the ECMWF (STAT and ERA) performs better than ESP until W4. From M2, the value turns negative for all the forecast, with ESP showing comparatively better performance. For ECMWF (STAT and ERA), the acceptable NSE (~0.5) is there until W2 but for ESP it is D6. The ERA and STAT show comparable performance until D6, but the STAT-ECMWF perform slightly better from W2 till M4 while for M6, ERA-ECMWF performs better than STAT-ECMWF. The spread of the NSE of STAT-ECMWF for M6 is quite large.

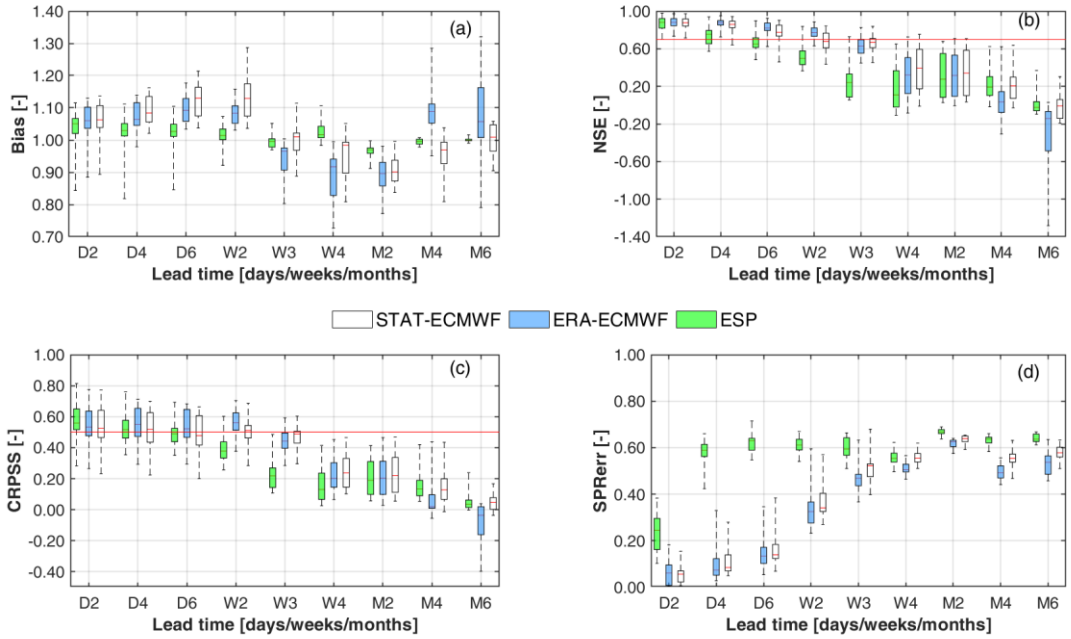

Figure 4.11: Baseflow performance of STAT-ECMWF, ERA-ECMWF, ESP for August initialization (a) Bias (b) NSE (c) CRPSS (d) SPRerr

A similar result is followed in CRPSS (Figure 4.11c) as well in which, until W4 the ECMWF(STAT and ERA) is performing better than ESP however after M2, ESP outperforms STAT-ECMWF. The value concerning ERA-ECMWF becomes negative for M2, while the STAT-ECMWF has negative value only for M6. For ESP, the value never goes below zero. The acceptable CRPSS (~0.5) is D6 for ECMWF (STAT and ERA) while it is D2 for ESP.

The spread error (Figure 4.11d) of ESP always remains within 0.6 to 0.7 and it is always higher than the ECMWF (STAT and ERA) for all the LT. For both ECMWF, the value is gradually increasing from D2 (~0.4) until W2 (~0.5) but differs hereafter till M6. For STAT-ECMWF, the value increases to ~0.6 and remains similar until M6, while for ERA-ECMWF, the SPRerr value of ~0.5 remains similar from W2 till M6. The value of ESP and STAT-ECMWF remain similar after M2.

4.6 Discussion

This study quantifies the skills in runoff forecast from NWP (ECMWF system-5 dataset) and ESP methodology against the climatology for the Northern Italian Alps concerning the period of 1993 until 2016. The bias correction of the ECMWF dataset with the station and ERA is considered prior to its' application. The quantile mapping technique applied to the raw dataset provides acceptable results with reference until the six-month LT horizon. Here the results of total flow, but most importantly, the baseflow are thoroughly analyzed with the help of ICHYMOD hydrological model.

The total runoff result shows an immediate loss in performance for ESP while the NWP provides reliable value until the end of the second week. This performance is validated by the deterministic NSE value and with the CRPSS, which takes the probabilistic approach for forecast verification.

Considering the median value of 0.5 as a reference for acceptable skill in NSE and CRPSS. The NSE value illustrates the skill of ESP until the end of 1st week, while the CRPSS value hints at the loss of skill immediately after the 2nd day. Similarly, the NWP provides reliable results until the 2nd week based on NSE and until the end of the 1st week for CRPSS. While in comparison, for both the methodology, the baseflow results show an extension of the one-week time horizon, in which the error metrics shows comparable results, for instance: The W3 of baseflow results look similar to the W1 of the total flow of river. Consequently, the acceptable skill (NSE, CRPSS) for baseflow extends one week further as compared to total flow. The gain in the skill of one week window for sub-surface flow as compared with the total flow was also shown by (Bogner et al., 2018), which also studied the runoff forecasting of Alpine basins situated in Switzerland. For both the flows, the decay of the performance continues with ESP showing better values from LT of 2nd month onwards. This also highlights the inefficiency of the NWP to model the chaotic climate for longer lead times. The median bias shows similar results to climatology for both the input with NWP displaying increasing IQR with the lead times. However, the spread error shows completely different nature, for ESP remains within certain range but for NWP, there is a gradual increase in the value with a consistent value after the month. This nature could be attributed to the ensemble spread of NWP forecast which remains pretty small at the beginning but with LT, the spread becomes larger when it tries to incorporate varying climate scenarios. Whereas for ESP, the ensemble spread since is based on past values, there is not much room to include the chaotic climate nature, which limits itself to a certain range of values. It is clear that there are some problems of forecasting the sub-surface process at longer leadtime (beyond two week), which could be attributed to the alpine head basin characteristics of thin and highly variable soil variables.

In general these temporal analysis highlight the dependency of the information gain on the reaction time of the investigated model variable. The slower the hydrological variable is reacting (i.e. the more persistent the water stays in the system), the longer the skill of the NWP forecast will be superior to climatology. This similarity between memory and predictability has also been found by (Orth & Seneviratne, 2013). It should be however noted that the type of catchment type and the forecast initialization month dictates the forecast skill. This study is limited by the skill dependency of the forecast based on the geology, none the less, the result based on the initialization month does reveal some of the forecast capabilities. Ideally, the ESP should perform better when initialized for the low flow season (December-March) and melting season (April-June), where the initial catchment states dictate the seasonal flow. Whereas, when the precipitation controls the hydrological regime like late summer (August), and autumn (September-November), the future meteorological information controls the forecast skill.

Here results of two representative initialization months (March and August) are considered to argue the claim made in the previous paragraph. When the CRPSS values are considered, the March initialization does show a better performance for ESP until D6, and for longer LT (M2-M6) the

performance is similar to STAT-ECMWF. Generally, IC predictability in March is low compared to the rest of the spring months as a result of the depleted subsurface water levels encountered during this month as well as because of the negative temperatures that prohibit snow from melting. While in case of August initialization, it is evident that the skill of forecast (NSE and CRPSS) is better for STAT-ECMWF than for ESP until one month and after, both forecast perform similarly.

4.7 Limitation and Conclusion

The above results are in agreement with the analysis done by (Stergiadi et al., 2020), which consider two geological varying catchments property that is also included in this study i.e. Passirio and Gadera. These study basins characterized by dependency on the future meteorological variables (Passirio) and initial catchment states (Gadera) show contrasting results based on the initialization month and forecast procedure considered. Which highlights the importance of catchment property to understand the effectiveness of the forecast scheme considered. This study fails to acknowledge the geological information of the 16 basins and most importantly, since the results of the 16 basins are consolidated, it is impossible to analyze each basin.

Nonetheless, this study successfully carried out the assimilation of the 25 members NWP (ECMWF system-5) forecast information for the runoff prediction over the northern Italian Alpine basins. The QM bias correction of the forecast dataset for the runoff forecast proves critical as compared to traditional ESP forecast methodology specially for low flow initialization months. The NWP provides added skill as compared to ESP for lead time up to 2 months and after both forecast shows similar performance. There is also an added skill arising for baseflow as compared to the total flow, however, the thin and highly variable soil variables alpine characterstics of the study area makes it difficult to forecast sub-surface runoff beyond two weeks with acceptable skill scores. Nonetheless, it should be stressed that any slight improvement in the forecast skill could manifest in significant economical gains, especially in the case of flood forecasting or decisions concerning management of hydro-power.

In general both verification measures (deterministic and probabilistic) outline the profit of NWPs for water management purposes especially in the field of agriculture, where sub-surface water flows and storage are highly relevant. These processes show improved forecast skills in comparison to climatology for the whole forecast period because of the persistence and memory effects of the water in the soil. But also the management of hydro-power plants could gain from extended-range stream-flow forecasts allowing a greater flexibility in decisions regarding the regulation of reservoir in- and out-flows and residual flows.

5. General conclusion

This thesis aims to understand the seasonal runoff forecasting for the Alpine basin using traditional ESP method and the state-of-the-art C3S ECMWF forecast information. Since ECMWF uses the ERA5 as the boundary condition, this study exploits the ERA-5 information for primarily snow simulation and secondarily for flood modelling. The semi-distributed TOPMELT snow model is considered for

the snow simulation while ICHYMOD hydrological model, which integrates the TOPMELT snow routine, is considered for the hydrological analysis. This study considers the range of basin size from 10s of Kms to 1000s to understand how the performance changes with change in spatial scale.

The first objective of snow simulation showed ERA-5 precipitation having positive bias originating from the low-intensity rainfall (termed drizzle) while the temperature showed a colder bias. These errors resulted in overprediction of SWE primarily due to the drizzle and secondarily due to the ERA5 lower temperature which helped to sustain the snow in the ground. The spatial scale of ERA5 affected the smaller basins while the larger basins remain unaffected. Eventhough the application of QM is applied to the meterological variables, it was impossible to get rid of the drizzle in ERA5. The QM of ERA5 precipitation does help to reduce the SWE bias but the important role of temperature for SWE simulation was seen after removing seasonality using monthly moving mean. The role of temperature correction for SWE for the smaller basins was highly evident as compare to the larger basins.

The second objective follows the same framework as the first objective with runoff and flood as the desired result. Primarily same as with the SWE simulation, the spatial scale of ERA5 only affected the runoff performance of the smaller basins. The drizzle error in ERA5 actually helped to produce comparable runoff result for the smaller basins due to increase in soil moisture as it makes comparable initial states as the high altitude snow watershed. In case of larger basins, since the drizzle errors get accumulated with increase in basin size, the ERA5 induced runoff didn't perform comparatively well for the larger basins. It is however interesting to note that the flood results concerning ERA5 showed comparable results for larger basins due to over estimation of initial soil moisture states and underprediction of the event precipitation. These encouraging results were sadly no more fruitful when considering unit watershed area, hence revealing the effectiness of ERA5 only pertaining to the basin size. The QM of ERA5 precipitation help improve the hourly runoff performance but the flood capabilities further decreased and there was no role of temperature in runoff and flood simulation.

The third objective looks into the skills gained in using the bias corrected ECMWF and ESP for the total and baseflow against the climatology for the same study area as for the prior two objectives. The QM bias correction is used considered both stations and ERA5 dataset. Comparitively, the ECMWF shows better performance and gain in skill of 1 week for baseflow and few days for total flow. However for longer time scale of more than 2 months, the ECMWF underperforms the ESP hinting towards the chaotic climatic nature for longer LT. During March initialization when the initial catchment states plays a major, there is comparable performance of ESP with the ECMWF but for August initialization, when the skill comes from the information of forecast, the ECMWF performs better than ESP until the second month.

Overall, this thesis was successful in integrating the large ERA5 and ECMWF dataset for the hydrological process simulation in the Adige basin. The dense meteorological network makes the upper Adige basin suitable to quantify the error in ERA5 and ECMWF dataset and its' affect on snow

and runoff simulation. The study highlights the need of correction of these dataset before its' application which is performed by rather simpler monthly quantile mapping. This method was considered since the study aspires to be the benchmark in this field of work and replicate the results in other data-scarce mountainous catchments as well.

6. Software and output data

Primarily MATLAB computer programming was used to run the TOPMELT and ICHYMOD model. All the data analysis and figure representation was also done in the same platform. The use of ArcMAP was considered to prepare the map of the study area. In some cases, the excel spreadsheet was used to construct tables and visualize dataset. Python language was considered in the initial phase of data download of ERA5 and ECMWF. All of the data and the script relating to this work are saved in a personal computer of the author and can be provided immediately upon request.

Annex A

The snowpack model: TOPMELT

The snowpack model used in this work is TOPMELT (Zaramella et al., 2019). 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 (Di Marco et al., 2021; Zaramella et al., 2019).The 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 to and energy class, the model cells within $n_b \times n_c$ is provided with corresponding glacier area fraction.

Hence 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. Zaramella et al. (2019) 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 Thiessen polygon method to estimate the mean precipitation over the basin while the air temperature data was used to calculate the unique hourly vertical lapse rate for the whole basin. The precipitation data are corrected with the snow correction factor (SCF) to account the inefficiency of the gauge during snow periods. 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 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 vertical precipitation gradient, which considers increased precipitation over elevation. This is obtained by considering precipitation gradient G (km^{-1}) as given in equation 4.1.

$$\mathrm{p_i} = \mathrm{pbasin}.\left(1 + \mathrm{G}.\frac{\mathrm{h_i} - \mathrm{h_{ref}}}{1000}\right) \qquad \textit{Equation 0.1}$$

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 vertical temperature lapse rate help to obtain the mean air temperature over the i^{th} elevation. With the help of the threshold temperature T_c, the estimation of precipitation phase(solid or liquid) is performed.

Please refer to (Zaramella et al., 2019) for details about the methodology of precipitation estimation, which accounts also for the variability with elevation. 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 $\mathrm{F_{i,j}\,(t)[mm\ h^{-1}]}$ at time t is calculated as follows:

$F_{i,j} = CMF.RI_{i,j}\,(t).\left(1 - albs_s(t)\right).max\{0; T_i(t) - T_b\}$ Equation 0.2

where CMF is the combined melting factor$[\mathrm{mm.\,m^2.\,°C^{-1}.\,MJ^{-1}}]$, RI(t) is the clear sky solar radiation $[\mathrm{MJ.\,m^{-2}.\,h^{-1}}]$ at time t, $\mathrm{alb_s(t)}$ is the snow albedo [-] at time t (Brock et al., 2000), $\mathrm{T_i(t)}$ is the air temperature at the i-th elevation band at time t while $\mathrm{T_b}$ =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 from (Brock et al., 2000) as

$albi(t) = ALBS - \beta_2[log10 \sum_k T_i\,(t_k)]$ Equation 0.3

Where, ALBS(-) is albedo of fresh snow, β_2(-) is a dimensionless parameter and $\sum_k T_i\,(t_k)$ is the summation of the positive hourly temperatures which 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 and $liqw_{i,j}$ is reduced and gets added to the snowpack through a freezing rate, termed ice (mm $^{h-1}$). This is computed as in equation 2.3.

$ice_i(t) = REFRZ.min[0, (T_b - T_i(t))]$ Equation 0.4

Where, T_b is the threshold base temperature as given in equation 2.1 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 2.1 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.

Annex B

ICHYMOD hydrological model

ICHYMOD hydrological model (Norbiato, 2008) integrates the TOPMELT snowpack model (Zaramella et al., 2019) with the semi-distributed basin-scale conceptual rainfall-runoff hydrological model. The hydrological model transforms melt from snow and excess precipitation into runoff at the outlet of the basin. The ICHYMOD model consists of a snow routine, a soil moisture routine, and a flow routine.

The soil moisture routine exploits the Probability Distributed Model (PDM) (Moore, 2007) to describe the spatial variation of water storage capacity across a basin. While with the help of the cubic law storage model, the base discharge is routed from groundwater to the catchment outlet. While the direct runoff from the proportion of the basin, where storage capacity has been exceeded is routed by a geomorphology-based distributed unit hydrograph (Da Ros & Borga, 1997), consisting of a cascade of two linear reservoirs in series. Ice melt runoff is routed to the outlet in two ways, depending on glacial till imperviousness. Some part of meltwater from ice influences the soil moisture storage, while the rest flows directly to the outlet as a cascade of two linear reservoirs.

For the computation of losses due to evapotranspiration, the Hargreaves method (Hargreaves & Allen, 2003) is used for potential evapotranspiration while the status of soil moisture stored in the PDM is considered as well. Storage-based representations of the fast and slow response pathways yield a spatially lumped representation of fast and slow response at the basin outlet which, when summed, gives the total basin flow.

Drainage to the slow flow path is represented by a function of basin moisture storage. The slow or base flow component of the total runoff is assumed to be routed through an exponential store. Direct runoff from the proportion of the basin where storage capacity has been exceeded is routed by means of a geomorphology-based distributed unit hydrograph. With this procedure, a geomorphologic filter based on a threshold drainage area is used to distinguish hillslopes and channel network starting from the space-filling representation of the drainage system directly obtainable from DEMs (Da Ros and Borga, 1997a). The routing time of each site in the basin is evaluated assigning different typical velocity values in each pixel pertaining to the basin and classified as hillslope or channel. The two velocities used to describe the flow routing.

Annex C

Model Calibration and validation

In order to have the best model parameter sets for both the TOPMELT and ICHYMOD hydrological model, the simulated runoff was compared with the observed runoff stations at each of the sixteens river outlet. The study was divided into two periods of calibration (October 2003 - September 2019) and validation (October 1992 – September 2003). For the calibration period, 5000 Monte Carlo simulation of the station input hourly runoff was compared with the observed runoff, to obtain the best parameter sets corresponding to highest NSE value with the lowest bias. Further, these same parameter sets were considered for the validation period as well. It should be noted that due to lack of continuous runoff data for some basins there was a need to consider different sets of calibration and validation period. The information of the different calibration and validation period along with the corresponding runoff NSE and bias are also listed in the table a. The NSE performance ranges from 0.617 till 0.86 for all the basins and periods, in most cases the NSE exceeds over 0.7. In the same setup, the relative bias ranges from -0.164 till 0.075, with bias mostly within ±0.1 and largely around ±0.05. These sound performance of the model make it suitable to perform all the runoff analysis here after.

The list of basin parameters which have different value for the basins are given in Table B. While the general parameters that are same for all the basins are listed in Table C. All the parameters remain same for the simulation concerning stations, ERA5 and ECMWF dataset except for PCF which remains 1 for ERA5 and ECMWF dataset.

		Calibration Period			Validation Period		
Sn	Name	Period	NSE	Bias	Period	NSE	Bias
1	Rio Plan	2003 - 2019	0.759	0.009	1994 - 2003	0.752	-0.106
2	Rio Riva a Seghe	2003 - 2008	0.736	-0.044	2001 - 2003	0.771	-0.164
3	Rio Anterselva Bagni	2003 - 2019	0.729	0.038	1992 - 2003	0.719	-0.061
4	Braies	2013 - 2016	0.666	-0.016	2016 - 2019	0.617	-0.084
5	Rio Riva a Caminata	2003 - 2011	0.786	0.075	2011 - 2019	0.743	0.029
6	Rio Casies Colle	2003 - 2019	0.694	0.003	1992 - 2003	0.82	-0.047
7	Gadera a Pedraces	2003 - 2011	0.699	0.025	2011 - 2018	0.702	0.009
8	Aurino a Cadipietra	2003 - 2019	0.776	0	1992 - 2003	0.79	0.004
9	Ridanna a Vipiteno	2003 - 2019	0.805	0.015	1992 - 2003	0.82	-0.029
10	Gadera Mantana	2003 - 2018	0.679	-0.077	1992 - 2003	0.682	-0.098
11	Passirio	2013 - 2016	0.832	-0.05	2016 - 2019	0.812	0.042
12	Aurino a Caminata	2003 - 2019	0.786	-0.003	1992 - 2003	0.86	-0.039
13	Aurino a S.Giorgio	2003 - 2019	0.804	0.007	1992 - 2003	0.814	0.056
14	Rienza Vandoies	2003 - 2019	0.738	0.032	1992 - 2003	0.767	-0.006
15	Ponte Adige	2003 - 2019	0.694	-0.046	1992 - 2003	0.73	-0.065
16	Adige a Bronzolo	2003 - 2019	0.818	-0.011	1992 - 2003	0.808	-0.054

Table A: NSE and bias performance for calibration and validation period.

SN	Parameters	Description	Range
1	RMF	Rain Melt Factor	0.3-0.58
2	CMF	Combined Melt Factor	0.0041-0.0184
3	PCF	Precipitation correction factor	0.92-1.14
4	ERF	Evaporation reduction factor	1
5	NMF	Night Melt factor	0.16-0.16
6	$PRECGRAD$	Precipitation Gradient	0-0.49
7	ALBSNOW	Snow albedo	0.73-0.94
8	ALBGLAC	Glacier albedo	0.3-0.3
9	$BETA2$	Dimensionless parameter for albedo computation	0.092
10	CMIN	Minimum surface storage capacity	0-0.5
11	CMAX	Maximum surface storage	150-900
12	B	Pareto exponent	0.01-0.5
13	KG	Groundwater recharge constant	2219-58000
14	$STMIN$	Minimum storage	10-191
15	$K1$	Mean residence time of the first surface runoff reservoir	3.4:4.5
16	$K2$	Mean residence time of the second surface runoff reservoir	3.4:4.5
17	KS	Groundwater discharge constant	0.0001:0.087
18	BG	Groundwater recharge exponent	1:1.8
19	BE	Exponent of the evaporation function	1:2
20	$Q0$	Background flow constant	0:4
21	M	Backgound flow exponent	0:100
22	$KGLAC$	Mean residence time of the 2 glacial surface flow reservoirs	3.4:4.5
23	CEL	Muskingum-Cunge kinetic wave celerity	3.3
24	$DISP$	Muskingum-Cunge hydrodynamic dispersion	20000:60000

Table B: List of basin parameters.

SN	Parameters	Description	Value
1	SCF	Snow correction factor	1.4
2	Tsnow	Snow rain threshold	1.5°C
3	Tmelt	Base temperature of snow	0°C
4	Liquid water	Water holding capacity of water equivalent	0.1
5	Delaytime	Speed of water propagation through snow pack	$3mh^{-1}$
6	Refreezing	Liquid water refreezing factor	0.03 $mm\text{-}dC^{-1}h^{-1}$
7	ECF	Evaporation correction factor	0.7

Table C: List of general parameters

6.1 References

Anghileri, D., Monhart, S., Zhou, C., Bogner, K., Castelletti, A., Burlando, P., & Zappa, M. (2019). The value of subseasonal hydrometeorological forecasts to hydropower operations: How much does preprocessing matter? Water Resources Research, 55(12), 10159–10178.

Arnal, L., Cloke, H. L., Stephens, E., Wetterhall, F., Prudhomme, C., Neumann, J., … Pappenberger, F. (2018). Skilful seasonal forecasts of streamflow over Europe? Hydrology and Earth System Sciences, 22(4), 2057–2072.

Ayers, J. R., Villarini, G., Schilling, K., & Jones, C. (2021). Projected changes in monthly baseflow across the US Midwest. International Journal of Climatology, 41(12), 5536–5549.

Bedia, J., Golding, N., Casanueva, A., Iturbide, M., Buontempo, C., & Gutiérrez, J. M. (2018). Seasonal predictions of Fire Weather Index: Paving the way for their operational applicability in Mediterranean Europe. Climate Services, 9, 101–110.

Bennett, J. C., Wang, Q. J., Robertson, D. E., Schepen, A., Li, M., & Michael, K. (2017). Assessment of an ensemble seasonal streamflow forecasting system for Australia. Hydrology and Earth System Sciences, 21(12), 6007–6030. https://doi.org/10.5194/hess-21-6007-2017

Bogner, K., Liechti, K., Bernhard, L., Monhart, S., & Zappa, M. (2018). Skill of hydrological extended range forecasts for water resources management in Switzerland. Water Resources Management, 32(3), 969–984.

Brutsaert, W. (2008). Long-term groundwater storage trends estimated from streamflow records: Climatic perspective. Water Resources Research, 44(2).

Crochemore, L., Ramos, M.-H., Pappenberger, F., & Perrin, C. (2016). Seasonal streamflow forecasting by conditioning climatology with precipitation indices. Hydrology and Earth System Sciences, 21(3), 1573–1591.

Crochemore, L., Ramos, M.-H., Pappenberger, F., & Perrin, C. (2017). Seasonal streamflow forecasting by conditioning climatology with precipitation indices.

Da Ros, D., & Borga, M. (1997). Use of digital elevation model data for the derivation of the geomorphological instantaneous unit hydrograph. Hydrological Processes, 11(1), 13–33.

Day, G. N. (1985). Extended streamflow forecasting using NWSRFS. Journal of Water Resources Planning and Management, 111(2), 157–170. https://doi.org/10.1061/(ASCE)0733-9496(1985)111:2(157)

Formetta, G., Marra, F., Dallan, E., Zaramella, M., & Borga, M. (2022). Differential orographic impact on sub-hourly, hourly, and daily extreme precipitation. Advances in Water Resources, 159, 104085.

Galletti, A., Avesani, D., Bellin, A., & Majone, B. (2019). Detailed simulation of storage hydropower systems in a large Alpine watershed. Geophysical Research Abstracts, 21.

Hargreaves, G. H., & Allen, R. G. (2003). History and evaluation of Hargreaves evapotranspiration equation. Journal of Irrigation and Drainage Engineering, 129(1), 53–63.

Harrigan, S., Prudhomme, C., Parry, S., Smith, K., & Tanguy, M. (2018). Benchmarking ensemble streamflow prediction skill in the UK. Hydrology and Earth System Sciences, 22(3), 2023–2039.

Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather and Forecasting, 15(5), 559–570.

Hopson, T. M. (2014). Assessing the ensemble spread–error relationship. Monthly Weather Review, 142(3), 1125–1142.

Kumar, S., Dirmeyer, P. A., & Kinter III, J. L. (2014). Usefulness of ensemble forecasts from NCEP Climate Forecast System in sub-seasonal to intra-annual forecasting. Geophysical Research Letters, 41(10), 3586–3593.

Laiti, L., Mallucci, S., Piccolroaz, S., Bellin, A., Zardi, D., Fiori, A., … Majone, B. (2018). Testing the hydrological coherence of high-resolution gridded precipitation and temperature data sets. Water Resources Research, 54(3), 1999–2016.

Liu, J., Yuan, X., Zeng, J., Jiao, Y., Li, Y., Zhong, L., & Yao, L. (2022). Ensemble streamflow forecasting over a cascade reservoir catchment with integrated hydrometeorological modeling and machine learning. Hydrology and Earth System Sciences, 26(2), 265–278.

Manzanas, R. (2020). Assessment of model drifts in seasonal forecasting: Sensitivity to ensemble size and implications for bias correction. Journal of Advances in Modeling Earth Systems, 12(3), e2019MS001751.

Manzanas, R., Gutiérrez, J. M., Bhend, J., Hemri, S., Doblas-Reyes, F. J., Torralba, V., … Brookshaw, A. (2019). Bias adjustment and ensemble recalibration methods for seasonal forecasting: a comprehensive intercomparison using the C3S dataset. Climate Dynamics, 53(3), 1287–1305.

Monhart, S., Zappa, M., Spirig, C., Schär, C., & Bogner, K. (2019). Subseasonal hydrometeorological ensemble predictions in small-and medium-sized mountainous catchments: benefits of the NWP approach. Hydrology and Earth System Sciences, 23(1), 493–513.

Moore, R. J. (2007). The PDM rainfall-runoff model. Hydrology and Earth System Sciences Discussions, 11(1), 483–499.

Moradkhani, H., & Sorooshian, S. (2009). General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis. In Hydrological modelling and the water cycle (pp. 1–24). Springer.

Moriasi, D. N., Gitau, M. W., Pai, N., & Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. Transactions of the ASABE, 58(6), 1763–1785.

Nash, J. E., & Sutcliffe, J. V. (1970). River forcasting using conceptual models, 1. A discussion of principles. J. Hydrol, 10, 280–290.

Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E., & Anquetin, S. (2008). Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins. Journal of Hydrology, 362(3–4), 274–290.

Orth, R., & Seneviratne, S. I. (2013). Predictability of soil moisture and streamflow on subseasonal timescales: A case study. Journal of Geophysical Research: Atmospheres, 118(19), 10–963.

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., … Srinivasan, J. (2007). Climate models and their evaluation. In Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC (FAR) (pp. 589–662). Cambridge University Press.

Sánchez-García, E., Abia, I., Domínguez, M., Voces, J., Sánchez, J. C., Navascués, B., … Pastor, F. (2022). Upgrade of a climate service tailored to water reservoirs management. Climate Services, 25, 100281.

Santos, I. M., Herrnegger, M., & Holzmann, H. (2021). Seasonal discharge forecasting for the Upper Danube. Journal of Hydrology: Regional Studies, 37, 100905.

Soriano, E., Mediero, L., & Garijo, C. (2019). Selection of bias correction methods to assess the impact of climate change on flood frequency curves. Water, 11(11), 2266.

Stergiadi, M., Di Marco, N., Avesani, D., Righetti, M., & Borga, M. (2020). Impact of geology on seasonal hydrological predictability in alpine regions by a sensitivity analysis framework. Water, 12(8), 2255.

Sutanto, S. J., & Van Lanen, H. A. J. (2021). Streamflow drought: implication of drought definitions and its application for drought forecasting. Hydrology and Earth System Sciences, 25(7), 3991–4023.

Torralba, V., Doblas-Reyes, F. J., MacLeod, D., Christel, I., & Davis, M. (2017). Seasonal climate prediction: a new source of information for the management of wind energy resources. Journal of Applied Meteorology and Climatology, 56(5), 1231–1247.

Van Schaeybroeck, B., & Vannitsem, S. (2018). Postprocessing of long-range forecasts. In Statistical postprocessing of ensemble forecasts (pp. 267–290). Elsevier.

Wang, Q. J., Shao, Y., Song, Y., Schepen, A., Robertson, D. E., Ryu, D., & Pappenberger, F. (2019). An evaluation of ECMWF SEAS5 seasonal climate forecasts for Australia using a new forecast calibration algorithm. Environmental Modelling & Software, 122, 104550.

Yuan, X., Wood, E. F., & Ma, Z. (2015). A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development. Wiley Interdisciplinary Reviews: Water, 2(5), 523–536.

Zaramella, M., Borga, M., Zoccatelli, D., & Carturan, L. (2019). TOPMELT 1.0: a topography-based distribution function approach to snowmelt simulation for hydrological modelling at basin scale. Geoscientific Model Development, 12(12), 5251–5265.