Politecnico di Milano

School of Civil, Environmental and Land Management Engineering

Master of Science in Civil Engineering

Performance evaluation of the MOLOCH meteorological model for discharge forecasting over the Seveso-Olona-Lambro river basins

---

Nicolás Andrés Chaves González

Tutor:
Prof. Giovanni Ravazzani

Advisor:
Prof. Alessandro Ceppi

2020-2021

To anyone who thinks is worthy of this dedicatory. But surely to my family and very close friends. Also to my country, because I know we are living hard times and we need professionals capable of building solid bases which allow us to thrive as a fair society, because I want to help with that task, and I feel optimistic about it.

A todo aquel que se crea merecedor de esta dedicatoria. Pero fijo fijo a mi familia y amigos muy cercanos. También a mi país, porque sé que estamos viviendo tiempos difíciles y que necesitamos profesionales capaces de construir bases sólidas que nos permitan prosperar como una sociedad equitativa, porque quiero ayudar en esa tarea y me siento optimista al respecto.

Abstract

Floods are disastrous natural severe weather events that cause large damage worldwide each year, inducing loss of human lives, destruction of infrastructure and economical losses. Consequently, forecasting this type of events through hydrological modelling is of great importance from a civil protection point of view since it allows institutions to reduce the generated hydrological risk by means of early warning systems. Nevertheless, ungauged basins where there is lack of direct measurements of meteorological information to force the models is one encountered problematic affecting the forecasts. In the present study is evaluated the possibility of using meteorological predictions coming from MOLOCH model to force the FEST-WB hydrological model to perform discharge forecasting. This, under the hypothesis that the committed error of the prediction is negligible when using forecasts up to 24 h. The study is done in a well-known area such as the Seveso-Olona-Lambro river basins located in northern Italy. Thus, the main hydro-meteorological variables are analysed by carrying out a comparison between spatialized observed meteorological data coming from ARPA and meteonetwork weather stations and meteorological predictions. Moreover, a sensitivity analysis following the one-factor-at-time methodology is accomplished with the aim of defining which forcing mostly affects flowrate forecasts. It is shown that discreet correspondence of information with an underestimation trend – particularly for large values - is verified for discharge at hourly and daily scale, and that underestimation of precipitation - especially in summer - and overestimation of solar radiation are the main reasons of this. Hence, to improve the predictability when coupling the two models, some actions should be evaluated to enhance the correspondence of the meteorological forcings.

Riassunto

Le inondazioni sono disastri naturali che causano una grande quantità di danni intorno tutto il mondo ogni anno, inducendo perdite di vite umane, distruzione dell'infrastruttura e perdite economiche. Di conseguenza, la previsione di questi tipi di eventi attraverso i modelli idrologici è di estrema importanza da un punto di vista della protezione civile poiché permette alle istituzioni di ridurre il rischio idrologico creato con l'utilizzo dei sistemi di allertamento. Tuttavia, i cosiddetti ungauged basins, cioè, i bacini idrografici ove esiste una mancanza di misure dirette d'informazione meteorologica per l'inizializzazione dei modelli, è una problematica che influisce le previsioni. Nel presente studio viene valutata la possibilità di usare le previsioni meteorologiche ottenute tramite il modello MOLOCH per inizializzare il modello idrologico FEST-WB e così realizzare le previsioni di portata. Questo, ipotizzando che l'errore commesso nella previsione di portata risulti essere trascurabile quando vengono utilizzate previsioni meteorologiche fino alle prime 24 ore. Lo studio viene sviluppato in un'area ben conosciuta come i bacini dei fiumi Seveso-Olona-Lambro situati nel nord d'Italia. In tal modo, le principali variabili idro-meteorologiche vengono analizzate svolgendo un confronto fra l'informazione meteorologica osservata e spazializzata, ottenuta tramite le stazioni meteo di ARPA e meteonetwork, e quella prevista. Inoltre, viene eseguita un'analisi di sensitività con l'utilizzo della metodologia "un fattore alla volta" (one-factor-at-time) con lo scopo di definire quale delle forzanti influisce di più la previsione di portata. Viene evidenziato il fatto che per le portate a scale oraria e giornaliera esiste una corrispondenza dell'informazione discreta con una tendenza alla sottostima, in particolare per i valori elevati, e che la sottostima della precipitazione (soprattutto in estate) e la sovrastima della radiazione solare sono le ragioni principali di questo comportamento. Pertanto, per migliorare la prevedibilità quando vengono accoppiati i modelli sopra citati, azioni aggiuntive devono essere valutate per innalzare la corrispondenza delle forzanti meteorologiche.

Resumen

Las inundaciones son desastres naturales que causan una gran cantidad de daño alrededor del mundo cada año, induciendo pérdidas de vidas humanas, destrucción de infraestructura y pérdidas económicas. En consecuencia, la predicción de este tipo de eventos a través de modelos hidrológicos es de suma importancia desde el punto de vista de la protección civil ya que permite a las instituciones reducir el riesgo hidrológico creado, por medio de sistemas de alerta temprana. Sin embargo, cuencas hidrográficas sin instrumentación en las que existe una falta de mediciones directas de información meteorológica para inicializar los modelos, es una problemática encontrada que afecta a las predicciones. En el presente estudio se evalúa la posibilidad de usar predicciones meteorológicas obtenidas con el modelo MOLOCH para inicializar el modelo hidrológico FEST-WB y así realizar predicciones de caudal. Lo anterior se realiza hipotetizando que el error cometido en la predicción de caudal resulta ser insignificante cuando se utilizan predicciones meteorológicas hasta las primeras 24 h. El estudio se realiza en un área ampliamente estudiada como la cuenca de los ríos Seveso-Olona-Lambro localizados en el norte de Italia. De este modo, las principales variables hidro-meteorológicas son analizadas llevando a cabo una comparación entre información meteorológica observada y espacializada, obtenida a partir de las estaciones climáticas de ARPA y meteonetwork, y las predicciones meteorológicas. Adicionalmente, se efectúa un análisis de sensibilidad con la metodología "un factor a la vez" (one-factor-at-time) con el objetivo de definir cuál de las forzantes afecta mayoritariamente las predicciones de caudal. Se evidencia que para los caudales a escalas horaria y diaria existe una correspondencia de información discreta con una tendencia a la subestimación, particularmente para valores elevados, y que, la subestimación de la precipitación (especialmente en verano) y la sobreestimación de la radiación solar son las razones principales de este comportamiento. Por consiguiente, para mejorar la predictibilidad cuando se acoplan los dos modelos mencionados, acciones adicionales deben ser evaluadas para realzar la correspondencia de las forzantes meteorológicas.

Contents

List of Figures

Chapter 1

Introduction

1.1 State of the art

Floods are the most common and disastrous natural severe weather events, they induce large damage in population, infrastructure, and generate economic losses. In 2020 it has been estimated that 1.47 billion people are exposed to intense flood risk, being one third in poverty conditions [1]. Moreover, from 1970 to 2012, the 79 % of the weather, climate and water-related disasters worldwide were storms and floods, causing around 1 million deaths and approximately US$ 2 trillion of economic losses [2]. Hence, forecasting floods is of great interest since – under the civil protection framework – it allows us to prevent the risk in a passive way, for example, with early alert systems. Consequently, hydrological models forced with observed meteorological variables have been a tool to properly monitor discharges in rivers where high risk of flooding is present.

Nevertheless, problems affecting discharge prediction arise due to different factors, for instance, lack of observed hydro-meteorological information because of poor network of weather and hydrological stations, or short concentration times due to the characteristics of the basin – especially mountainous ones. Therefore, coupling meteorological and hydrological models or using additional ways – besides direct measurements – to monitor the meteorological variables that influence the discharge forecast, are topics in which interest has been put on. For instance, in [3] were coupled a numerical weather prediction model with a hydrological model to perform discharge forecasts. It was shown that the system was able to reproduce hydrological processes and the flood peaks during the calibration and validation period. It was also noticed that errors in precipitation were noticeable for large thresholds of discharge, whilst for small precipitations events the system had better performance. Similar procedures were performed by [4] and [5] where different flood events were analysed when initializing a hydrological model with deterministic and ensemble weather forecasts. Furthermore, information coming from multiple sources has also been used. For example, in [6] a dataset given by multiple numerical weather predictions (NWPs) coming from a global weather prediction system were used for flood warning in a catchment. It was shown that uncertainties propagate in the forecast chain and that NWPs did not properly represent spatial variability of precipitation. Additionally, a more recent example is given by [7] where flood forecasting was made by coupling a distributed hydrological model with a global ensemble precipitation data – which is a project that collects forecast products from major forecast centres in the world.

With this, it is evident that further problems concerning the meteorological forcings of the hydrological models exist, for instance, the presence of phenomena difficult to forecast such as convective precipitations which affect the discharge prediction, the existence of flood events at the sub-daily scale [8], the uncertainty propagation through the forecasting chain [9] [10], or the lack of information in the studied areas. Consequently, their study has been of great importance to improve the performance on discharge forecasts. Thus, in studies such as [11] it is said that the lack of rainfall and discharge information are factors influencing the flood forecasting. It was evaluated the influence of rainfall errors – in terms of volume and duration - on the performance of a model calibrated with limited discharge data, and it was shown that calibrating the model with a limited number of discharge events is useful to perform flood forecast with uncertain rainfall data. Moreover, volume rainfall errors affect the most the performance of the model, and large volume and duration errors present at the same time create a compensation leading to a good flood prediction.

In respect of the last factor influencing the discharge forecasts, it is, the lack of hydro-meteorological information, studies such as [12] assessed the issue of defining the amount of discharge measurements needed to properly parametrize a model providing discharges. On the other hand, usefulness of other type of precipitation data such as remote measurements were evaluated in [13] from a water resources perspective in which it was assessed the suitability of common precipitation products over two scarce, complex mountainous terrains. Similar procedures were followed in [14] where different precipitation satellite products were used to force a physically-based distributed hydrological model, and it was found that different products have diverse qualities in terms of statistical indexes. Furthermore, other types of methodologies such as regionalization are used to perform discharge forecasts in ungauged basins. For instance, in [15] different regionalization (of parameters involved in the hydrological modelling) methodologies are tested in a catchment situated in Norway. Finally, more recent studies such as [16] evaluated the performance of different reanalysis datasets, satellite product, and the Weather Research Forecasting (WRF) to represent heavy rainfall over a region in Himalaya. This was done by comparing the datasets with rain gauge data. Moreover, this information was used to initialize a hydrological model to assess their ability to reproduce floods in the area. It was found that the reanalyses perform unsatisfactorily, that the variability in rainfall was evidenced in the modelled floods and, the output obtained from WRF model or TMPA (Multisatellite Precipitation Analysis) datasets could be used in the future to perform flood forecasting in early detection and warning.

Finally, making a further step and considering the above-described problematics, especially the one concerning the influence of meteorological uncertainties into the hydrological simulations, some procedures have been tested to perform a pre-processing or bias correction to the input data of the hydrological models. For example, in the already mentioned study [9] location correction and bias adjustment of rainfall data was done by using a "best match" approach and applying a multiplicative factor, respectively. In [17] general circulation model (GCM) daily rainfall simulations were transformed by applying a bias-correction method which consists of two steps: frequency correction, and intensity correction. Moreover, a multiplicative shift bias correction of the mean monthly values of rainfall was performed for comparison with the last methods. Additionally, to account for uncertainty in rainfall predictions due to model errors, in [18] three different multi-model post-processing methods of deterministic precipitation in order to estimate the forecast rainfall probabilities were evaluated. In [19] the quality of precipitation and streamflow forecasts were assessed when applying bias-correction methods in the precipitation data at seasonal scale. Furthermore, concerning meteorological variables, in [20] - under the framework of grass growth management - it was performed a verification of how accurate the European Centre for Medium-Range Weather Forecasts (ECMWF) deterministic forecasts were. Different variables were considered: air and soil temperature and rainfall. Systematic biases were observed and corrected-bias methodologies (seven methods) were applied in order to evaluate the improvement of the forecasts by re-assessing the accuracy. Finally, in [21] a bias-correction method for precipitation considering orographic characteristics was proposed. Additionally, to evaluate the robustness of the method, a cross-validation approach was followed.

1.2 Aims of the study

It has been shown that when performing discharge forecasts some problems affecting the results arise due to different factors, for instance, lack of direct meteorological information, difficulty in predicting some types of phenomena, uncertainty propagation through the forecasting chain, the characteristics of the catchment, the temporal scales at which the floods may happen, and so on. In the present study is of particular interest the problem related to the lack of hydro-meteorological information in river basins when there exists the necessity of performing discharge forecasting. Consequently, and considering the possibility of couple meteorological and hydrological models to predict flowrates, the following question is asked: are direct measurements of meteorological variables necessary to perform flood forecasts? Or is it possible to use only meteorological forecasts as if they were observed data? And to this concern it is made the hypothesis that the committed error in the discharge forecast is negligible when using meteorological forecasts up to 24 hours so applicable flowrate prediction can be done in data scarce or ungauged river basins.

Thus, following the state of the art above described, it becomes clear that some actions can be carried out with the purpose of trying to overcome this difficulty and improve the discharge forecasts. Particularly, one possible method could consider the following: a treatment of the available observed data with the scope of homogenize it, the analysis of the forecasted information, the identification of the atmospheric variables that mostly affect the flowrate forecast, and finally, the assessment of possible bias correction methodologies applied in the already identified variables. In the present study it is tried to understand a first part of the problem consisting in the first three steps. Hence, three objectives are set: first, homogenize and analyse observed hydro-meteorological information over the Seveso-Olona-Lambro (SOL) river basins located in northern Italy; second, analyse the main hydro-meteorological variables, in the area of study, between observed weather data coming from ARPA Lombardia (Lombardy Regional Environmental Protection Agency) and meteonetwork (meteorological network) weather stations, and MOLOCH meteorological model forecasts; and third, evaluate how much can we trust in meteorological MOLOCH forecasts when used as input of FEST-WB hydrological model for discharge forecasting in ungauged basins.

The present study is structured in the following way. First, the area of study description where the main characteristics of the SOL river basins are given. Second, materials and methods, where the meteorological and hydrological models, the observed and forecasted meteorological variables and its treatment, the coupling strategy of both models, and the statistical analysis to be done are described. Third, the results and discussion, in which a climatological characterization derived from the observed data treatment, the comparison between observed and forecasted hydro-meteorological variables, and a sensitivity analysis are presented. And finally, the conclusions of the study.

Chapter 2

Area of study

The area of interest in the present study contains the basins of Olona, Seveso and Lambro rivers, located in the north of Milano city, in Lombardy region of Italy.

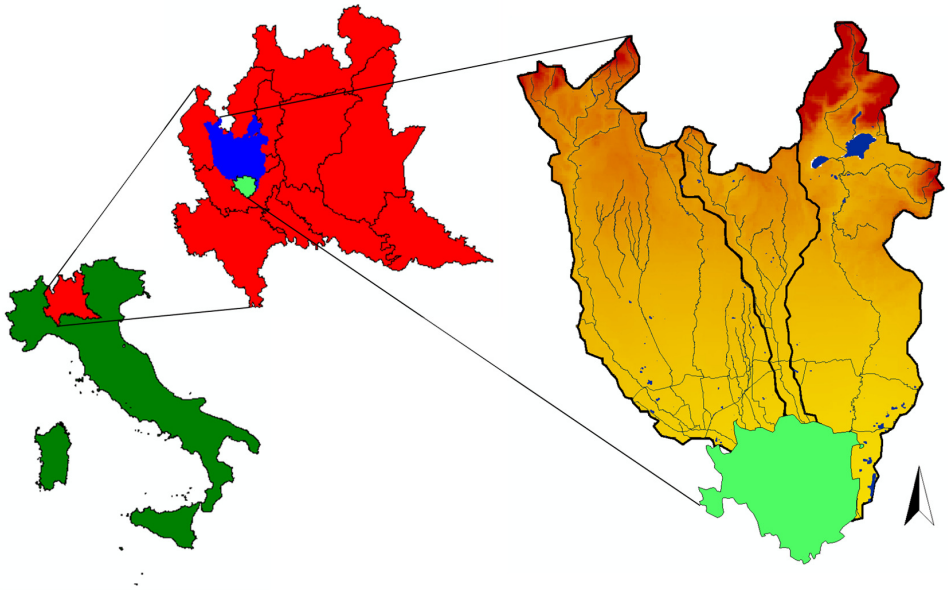


Figure 2.1: General location of the area of study

Within this ensemble of basins all water courses have a consolidated configuration with a discharge capacity progressively inferior from up to downstream, and all of them converge in the urban zone of Milano. The principal water courses growing from the pre-Alps belt from east to west are: Vettabbia, Redefossi, Lambro Meridionale, Olona, Bozzente, Lura, Guisa, Nirone, Pudiga, Garboger, Seveso, and Lambro. Additionally, in this region are also present the North-West overflow channel (Canale scolmatore di Nort-Ovest – CSNO) and the Olona deviation, both playing a role in the protection system of Milano [22] [23] [24].

Concerning the different characteristics of the studied area, it is important to notice that the Olona-Seveso-Lambro basins are localized in Milano province and, consequently, in the Pianura Padana (Po Valley). For instance, it is a rich-water territory sited in the threshold between the high and the low Pianura Padana, characterized by important natural springs. From a lithological viewpoint, the north zone is constituted by coarse materials - such as gravels – covered by an acidic clay soil not suitable for crops, while in the central and south areas finer materials are present – such as sands and silts. Additionally, one relevant characteristic is that this is one of the most urbanized zones of Italy, especially along Olona and Lambro rivers. Therefore, the agriculture is of interest only in the irrigation zones in the western side of the region [24].

With respect to the climate of the zone, in a further section of this document a characterization is made based on available observed data. Nevertheless, it is possible to say that the climate is the one of Pianura Padana and, consequently, of moderate continental type with warm summers and cold foggy winters. However, continental climate is mitigated by the Alps which protect the area of study from the arrival of cold air masses, while it feels the influence of the Mediterranean basin [23]. In the following paragraphs each basin conforming the studied area is described.

2.1 Olona basin

Olona basin is in the western part of the studied region, it is contained in Italy and Switzerland with a surface of 911 km^2 of which 99 % is within Italian borders. Additionally, 11 % of the basin is mountainous [25]. In respect of Olona river, it has different sources, from which the main one is located at approximately 1000 m a.s.l. in Rasa di Verese, and others in the hillsides of the pre-Alps in the northern zone of Varese – a municipality in north-west of Lombardy region – such as Pinzella and Legnone mountains. Moreover, another source is found in Valganna – a town city of Varese -, which gives rise to the eastern reach of Olona River. Then, it flows 60 km in south direction until reaching Milano city, and, after trespassing the urban zone, it goes out with the name Lambro Meridionale which, in SantAngelo Lodigiano municipiality flows into Lambro river [26] [27].

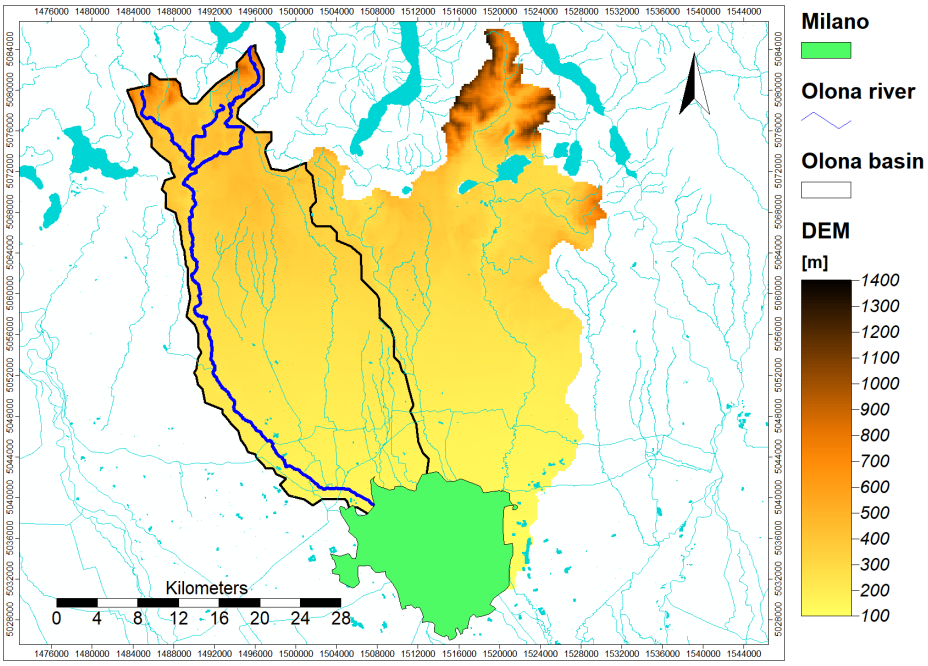


Figure 2.2: Olona basin

It is possible to subdivide the catchment into two main zones: the mountainous and the plains. The former contains the region from the source of the river until Ponte Gurone. Here the water body has a "Y" shape in which the western reach is characterized by a strong urbanization given by Varese and Induno Olona municipalities. On the other hand, although some small urban areas are present, eastern reach is mainly constituted by agricultural terrain and forests. Both reaches are merged in a point located at Molini Trotti locality [25].

Concerning the prairie zone, it starts at Ponte Gurone and ends when it reaches Milano. Here, the basin becomes narrow and long in north-south direction and, along the river there are urban areas alternating with non-occupied zones by settlements, such as forests and agricultural terrain. The Olona valley ends when the river crosses the Milano-Varese highway, after which the basin becomes flat and the river enters into a highly urbanized zone, crossing Castellanza and Legnano municipalities. Then, downstream there is once again alternation between agricultural regions and urban settlements until reaching the limits of Rho municipality, in which there exists a deviation structure called "Olona 1",

trough which floods are conducted into the North-West overflow channel (Canale Scolmatore di Nort Ovest – CSNO) [25] [26] [27]. Furthermore, in north of Milano it is located the Olona Deviation, first thought with the scope of deviate Olona floods being connected directly with CSNO. It conducts part of the discharge collected in the CSNO into the Lambro Meridionale, downstream Milano city, having a maximum capacity of $54\,m^3/s$. With this, the maximum capacity of Olona river is $58\,m^3/s$ in the north of Milano [26].

2.2 Seveso basin

Seveso torrent grows in Monte Pallanza at 490 m a.s.l. in the south of Como province. It trespasses different urban settlements located in Brianza - a region at the foot of the Alps between Milano and Lake Como – for over $50\,km$ in south direction until Milano Niguarda, where it enters into Naviglio della Martesana - the subsurface channels of Milano – localized in Milano downtown. Then, from the inner channel network of the city it arrives at Lambro Settentrionale [28] [29] [30].

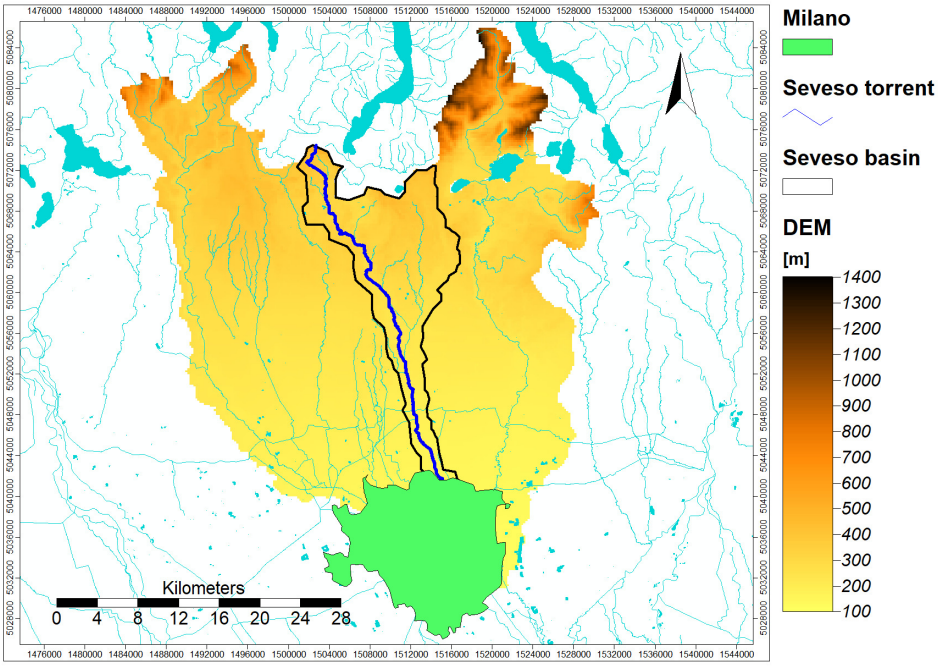


Figure 2.3: Seveso basin

The hydrographic basin can be subdivided into three main areas [30] [29] [28]:

- North Milano city, in which the septentrional zone has a leaf shape with a network constituted by a tributary system with consecutive order, in Como province. Then, in Milano province zone, the basin is narrow and long without main tributaries. Furthermore, within this area four sections can be identified: the natural Seveso, in the northern part starting from the source of Seveso torrent until Lentate municipality where there is little urbanization and high slopes are present; the natural Certesa, in the east of the last section, concerning Certesa torrent - which is the main tributary of Seveso – starting from its source until Terrò torrent confluence and with similar characteristics as natural Seveso; the urban Certesa, which starts from the end of the last section until the confluence with Seveso, here the slopes are lower and there are large urban areas; finally the urban Seveso, starting from Lentate until the beginning of the subsurface network of Milano, this zone is characterized by a flat surface and high urbanization.
- The trespassing of Milano city.
- The area downstream Milano where the torrent flows in south-west direction until Lambro confluence.

In respect of the entire Seveso basin with closure section at via Ornato, where the subsurface channel network of Milano starts, the surface is 277 km^2 from which 100 km^2 belong to urban areas [29]. Additionally, within the basin there are different morphologic characteristics, generally speaking, one part has its extension in mountainous areas with a maximum and minimum altitudes of 600 m a.s.l. and 200 m a.s.l., respectively, and an area of 155 km^2 which also considers the surface of Certesa torrent (62 km^2). Another area develops in a prairie zone which substantially is found within Milano province with altitudes between 135 m a.s.l. and 200 m a.s.l., and 175 km^2 of extension [30] [28].

Concerning Certosa torrent basin, some problems of geomorphological instabilities are present and flood problems are common. The ground has low permeability in the surface while is highly permeable in lower stratums. Additionally, in the natural regions above-described, a large part of the hydrographic network remains dry when little precipitation is verified. Nevertheless, Seveso torrent has a continuous flux from which a fraction is guaranteed by discharges coming from Carimate and Fino Mornasco purifiers [30]. Finally, a relevant characteristic of the dynamics present in this torrent is that, in flood conditions, Seveso is partially deviated into the North-West overflow channel (Canale Scolmatore di Nort Ovest – CSNO) which is 34 km long and has a capacity of 30 m^3/s. It discharges water coming from Seveso in Paderno Dugnano municipality, into the Olona Deviation and, in very extreme cases, into Ticino river [29] [30].

2.3 Lambro basin

Lambro river grows with the name Lambrone at Triangolo Lariano in the pre-Alps localized in the southern area of Como lake, in Magreglio municipality, with an altitude of 1300 m a.s.l.. It follows a path of about 130 km in which, after crossing Milano city, it flows into Po river at Senna Lodigiana municipality [31] [23]. The basin has an entire surface of 1980 km^2 (corresponding to 3 % of river Po's surface, the largest river in Italy) from which only 5 % is mountainous. On the other hand, with closure section at Redefossi deviation confluence localized in the south of Milano city, the basin has 553 km^2 from which 199 km^2 belong to urban areas, while 354 km^2 to extra-urban areas. Additionally, the catchment is characterized by a complex jointed hydrographic network. The different water courses, situated north Milano city, flow in north-south direction, and are interconnected by an artificial channel network made with irrigation or flood protection purposes [32] [31].

Concerning the path followed by the river, a first reach is developed in north-south direction until Canzo municipality, then in east-west direction until Ponte Lambro, and once again in north-south direction until its discharge into Pusiano lake. Thereafter, it continues in south direction until Villasanta, before crossing Parco Monza, and flowing in the lowest part of a valley [31]. Going downstream, three different sections can be identified: between Parco Monza and Sesto San Giovanni until the beginning of Milano, where the course is characterized by meanders and is surrounded by urban and production settlements that have subtracted area for river expansion. The second section downstream in the countryside east Milano, Lambro river flows in a reach with a straight tendency, and contained with artificial levees. Finally, in the section between the Redefossi deviation confluence and Po river, where the Lambro Meridionale flows into Lambro Settentrionale [31] [27].

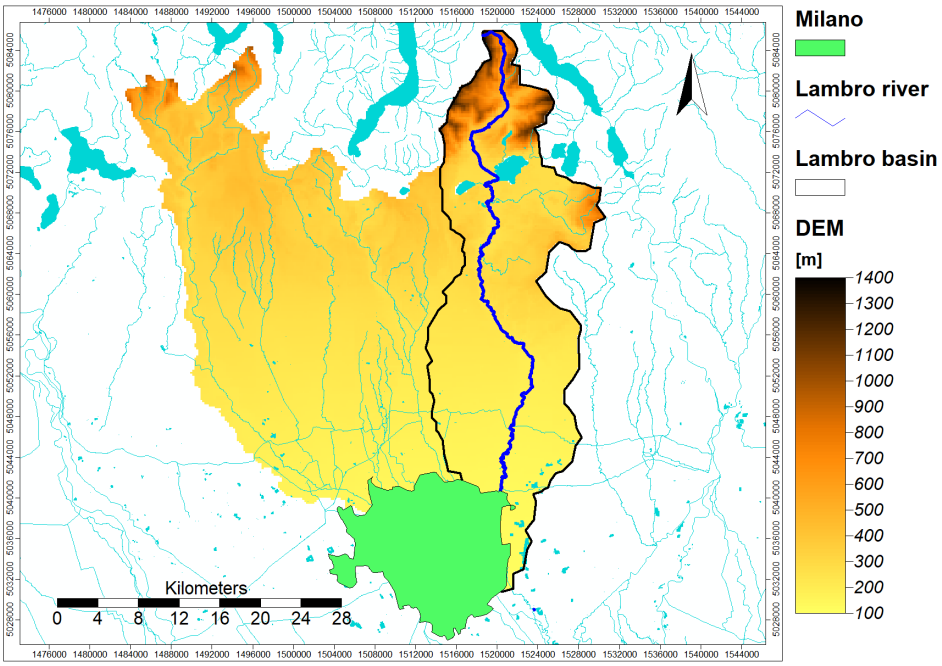


Figure 2.4: Lambro basin

The basin can be subdivided in 4 zones [31]:

- Lake, upstream Pusiano lake, where many small water courses with torrential regime are present, with high slopes. At the end of this zone the river has a smaller slope until its confluence in Pusiano lake. Furthermore, in this area it can also be included the basin correspondent to Alserio lake with similar behaviour as Pusiano and sited in the west side of the catchment.
- Natural Lambro, between Pusiano lake and Villasanta municipality, characterized by high and moderate slopes and reduced urbanization.
- Urban Lambro, between Monza municipality and south-east border of Milano city, until the confluence with Redefossi deviation. The slopes are mild and there is a strong urbanization.
- Irrigation Lambro, between Redefossi deviation confluence and Po river, it is constituted by a plain where a large irrigation network is present. Within this zone there is the confluence between Lambro and Lambro Meridionale rivers.

Finally, it is important to notice that the flux of the Lake zone section is influenced by the reservoir dynamics – filling and emptying – of Pusiano and Alserio lakes, since they have a surface of about $8\,km^2$. Additionally, in the natural Lambro section, the discharge practically depends on the geomorphologic characteristics of the basin, while urban Lambro zone is influenced principally by the discharge capacity of the sewage system of the municipalities [31] [27].

Chapter 3

Materials and methods

Once selected the area of interest different activities were done to achieve the aims of the present study. In the following sections are described the materials used to perform the correspondent analysis, such as the meteorological and hydrological models, the available observed and forecasted data with their correspondent given treatment and processing. Then, are also described the coupling strategy between the models, the performed hydrological simulations, and finally, the statistical analysis.

3.1 Meteorological model: MOLOCH

In the present study it was used the information obtained by a meteorological model capable of performing forecasts of the meteorological forcings in the studied area and in the studied period. Particularly, MOLOCH model is chosen since it has high spatial resolution to better describe phenomena occurring in the area, and because it is a model that has been adopted in Italy by different agencies to perform real time forecasting. Furthermore, there is availability of the forecast information in the studied period. In the following paragraphs the model is described as shown in [33] and [34].

The model was developed at CNR-ISAC (National Research Council of Italy, Institute of Atmospheric Sciences and Climate), it is implemented over Italy with a daily operational chain that also comprises the hydrostatic model BOLAM, which initial conditions are derived from the analyses (00 UTC) and forecasts of the Global Forecast System (GFS, NOAA/NCEP, USA) global model. Therefore, MOLOCH model is normally nested (1-way) into the BOLAM runs performed at coarser resolution and initialized with 3-h BOLAM forecast to avoid downscaling based on pure interpolation from the global model.

Concerning general characteristics of the model and its operability, it was written in Fortran 90, it integrates non-hydrostatic, fully compressible equations for the atmosphere within a grid size of $1.25\,km$, 60 atmospheric levels and 13 soil levels (the spatial resolution of the model has had two main improvements, changing from $2.2\,km$ to $1.55\,km$ in March 2014, and from $1.55\,km$ to $1.25\,km$ in October 2016). Additionally, it operates as short-range (12-48 h) weather forecasting model, starting the forecasts at 03:00 UTC of each day, and obtaining output fields with hourly frequency.

Referring to the model dynamics, MOLOCH is a non-hydrostatic, fully compressible, convection resolving model which integrates a set of atmospheric equations of the prognostic variables pressure, absolute temperature, specific humidity, horizontal and vertical components of wind velocity, turbulent kinetic energy, and five water species: cloud water, cloud ice, rain, snow, graupel/hail. The variables are spatially represented through latitude-longitude, optionally rotated, Arawaka C-grid, meaning that velocity components are stored in the grid faces, while other quantities can be stored in the corners. Moreover, for the integration – in spatial terms - it is employed a hybrid terrain-following vertical coordinates, depending on air density, relaxing smoothly to horizontal surfaces at a higher elevation from the ground. On the other hand, time integration is done with an implicit scheme for the vertical propagation of sound waves, and explicit forward-backward, time-split scheme for the horizontal propagation of gravity and sound waves. Furthermore, three-dimensional advection is computed using the Eulerian Weighted Average Flux scheme, and horizontal second order diffusion and a small divergence damping are included to prevent energy accumulation on the shorter space scales.

Finally, concerning the model physics, the scheme has 4 components. First, the atmospheric radiation, which is computed with a combined application of the Ritter and Geleyn, and the ECMWF schemes. Second, the sub-grid turbulence parametrization, which uses a scheme based on a E-l, order 1.5 closure theory, where turbulent kinetic energy (including advection) is evaluated. Third, the water cycle microphysics, based on a parametrization in which the spectral properties of hydrometeors are simulated assuming a generalized gamma function distribution. And fourth, the soil model with vegetation, including a soil model that uses 4-6 layers, whose depths increase moving downward. This model computes surface energy, momentum, water and snow balances, heat and water vertical transfer, and vegetation effects at the surface and the soil. It considers orography, the observed geographical distribution of different soil types and soil physical parameters. The soil model also includes treatment of water freezing and melting processes within the ground. Additional comments concerning forecast information obtained with MOLOCH model and the given use in the present study are done in further sections.

3.2 Hydrological model: FEST-WB

A hydrological model is an approximated mathematical representation of some physical processes involved in the hydrological cycle. They are useful to estimate the discharge flowing into a water course starting from meteorological variables.

According to different characteristics that these models may have, some classification is made. For instance, depending on the type of equations used to estimate discharge, models are divided into two categories: input-output and physically based. The former relates directly input and output variables without describing the physical processes present in the hydrological cycle. Conversely, the latter kind of models do through differential equations that assess mass balance and momentum, considering spatial and temporal variations [35]. Another division of models is made in terms of how they do represent the basin: lumped and distributed. The first one considers the basin as a unique element, which characteristics are mean values for the whole area; consequently, these models only return the discharge at the closure section of the watershed. Instead, distributed models represent the basin as a set of cells in which the catchment is discretized, therefore, heterogeneity of properties characterizing the area is considered. A further classification takes into account the modelled events: flood event and continuous models. Flood event models estimate the discharge in a limited time when precipitation events that create floods are verified. On the other hand, continuous models work even when there are no intense precipitations and, consequently, in addition to model flood events, they do model hydrological processes that allow to reconstruct the hydrograph even for low discharges.

A physically based, distributed hydrological model developed in Italy at Politecnico di Milano named FEST-WB, which acronym states for *Flash-flood Event-based Spatially-distributed rainfall-runoff Transformation – Water Balance model*, was used in the present study. It has as input the punctual hourly measurements of the following meteorological forcings: solar radiation, air temperature, air relative humidity, wind speed and precipitation. It also needs maps containing elevation, soil use and vegetation information [36]. With this, the model evaluates the main processes of the hydrological cycle: evapotranspiration, infiltration, surface runoff, flow routing, subsurface flow, and snow dynamics [37]. Particularly, for the present study a discretization of 200 m x 200 m cells was set. Model structure is shown in Fig. 3.1:

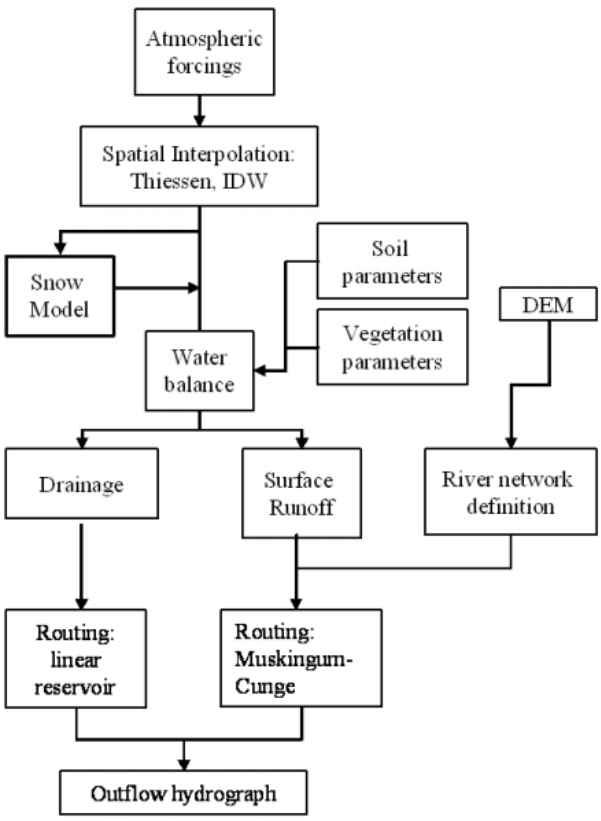


Figure 3.1: Diagram of FEST-WB processes [35]

Model's structure is divided into 5 main components described below [35] [38]:

- Flow paths and channel network definition: the first one is derived from the digital elevation model by means of a least-cost path algorithm. In reference to hillslope and channel network, the constant minimum support area concept is exploited; it is selected a minimum drainage area required to define a channel.
- Spatialization of site measured meteorological forcings: meaning the spatial interpolation of meteorological forcing measurements. In the present case it is used the inverse distance weighting method – IDW. A key aspect when interpolating air temperature is its reduction with altitude; a constant lapse rate of 0.0065 $°C/m$ is used to adjust the values. Therefore, to interpolate air temperature first it is necessary to report measured values to an altitude of reference by adjusting them with the lapse rate, then spatial interpolation through IDW method is performed and, finally, each interpolated value is reported on the ground by adjusting them again [36]. Additionally, thermal inversion phenomena are neglected. Last procedure can be seen in Fig. 3.2:

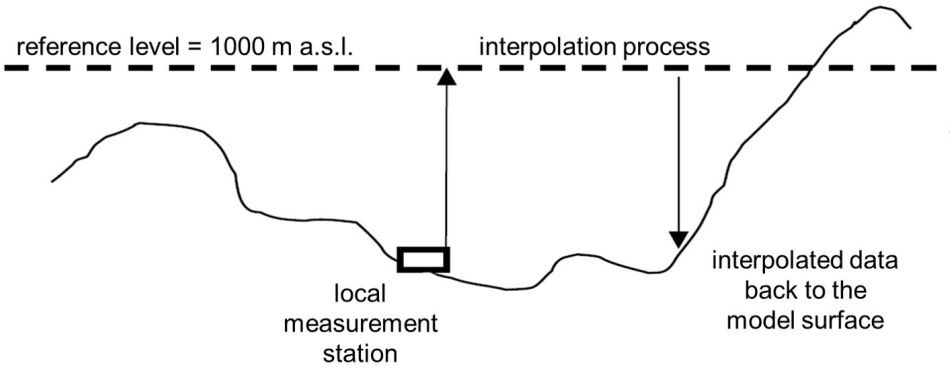


Figure 3.2: Interpolation process of temperature in FEST-WB [35]

With respect to solar radiation, longwave net radiation is estimated as function of air temperature, and shortwave net radiation is evaluated taking into account the effect of topography.

- Snowpack dynamics: considers the snow melt and the snow accumulation [37]. Precipitation P is divided into liquid, P_l, and solid, P_s, as function of air temperature, T_a:

$$\begin{aligned} P_l &= \alpha_P P \\ P_s &= (1-\alpha_P) P \end{aligned} \tag{3.1}$$

Where:

$$\begin{cases} \alpha_P = 0 & ; \quad T_a \leq T_{inf} \\ \alpha_P = 1 & ; \quad T_a \geq T_{sup} \\ \alpha_P = \frac{T_a - T_{inf}}{T_{sup} - T_{inf}} & ; \quad T_{inf} \leq T_a \leq T_{sup} \end{cases} \tag{3.2}$$

T_{inf} and T_{sup} are threshold temperatures in which snow, precipitation, or a combination of both can be verified. In the present study they are such that both temperatures are $0\,^{\circ}C$.

The degree day concept is used for the snow melt simulation. The melt rate M_s, in m/s, is evaluated as:

$$M_s = \begin{cases} C_m(T_a - T_b) & ; \quad T_a > T_b \\ 0 & ; \quad T_a \leq T_b \end{cases} \tag{3.3}$$

Where T_b is a threshold temperature set as $0\,^{\circ}C$ in the current study, C_m is an empirical coefficient function of meteorological and geographic location, usually ranging from 4.3x10^{-8} and 6.9x10^{-8} $m^{\circ}C^{-1}s^{-1}$ [35].

With respect to the movement of water over the surface when considering snow dynamics, it is assumed that the terrain under the snow is frozen and, consequently, melted water is not allowed to infiltrate into the soil. Moreover, liquid precipitation and melted water is supposed to flow following a reservoir routing scheme over the snowpack with a celerity of 1.67x10^{-3} m/s. This water only becomes part of the infiltration process when it reaches a location not covered by snow.

- Runoff computation: is performed in each cell of the domain. It is used the modified SCS-CN method extended for continuous simulations in which the potential maximum soil retention, S, is updated in each cell at the beginning of a precipitation event as:

$$S = S_1(1 - \varepsilon) \tag{3.4}$$

Where S_1 is the maximum retention for dry conditions (AMC 1), ε is the degree of saturation which depends on soil moisture:

$$\varepsilon = \frac{\theta - \theta_{res}}{\theta_{sat} - \theta_{res}} \tag{3.5}$$

Where θ_{res} is linked to the residual soil humidity, θ_{sat} is the saturation soil moisture, and θ is the soil moisture which dynamic is described by the water balance equation in a cell not covered by snow [38]:

$$\frac{\partial \theta}{\partial t} = \frac{1}{Z}(P_l - R - D - ET) \tag{3.6}$$

Where Z is the soil depth, P_l the liquid precipitation, R is surface runoff flux, D is drainage flux, and ET is the evapotranspiration rate. It is assumed that the soil moisture does not vary in time in locations covered by snow.

In respect of the effective evapotranspiration, it is estimated as a fraction of the potential rate of evapotranspiration, tuned by a function depending on soil moisture content. This potential rate is evaluated by means of the Priestley-Taylor radiation-based equation, which is a simplified method of Penman-Monteith [39].

- Overland and subsurface flow routing: overland is made through the diffusion wave scheme based on the Muskingum-Cunge method in its non-linear form with the time variable celerity. On the other hand, the subsurface in the cells not covered by snow is obtained by means of the linear reservoir routing scheme [38].

3.3 Observed weather data

Meteorological information contains the variables that are involved in the physical processes represented in the hydrological model. Therefore, they are the input data of the model. In the present study it has been used a database for the period 2003-2020 containing information at hourly resolution of the following meteorological variables:

- Precipitation - P $[mm]$: representing the amount of water falling at a particular place and in a period. One millimetre represents one litre of water falling over an area of one square metre. It should consider the melted snow and the rainfall but, unfortunately, not all the rain gauges are heated and, consequently, analyses were excluded when there was presence of snow cover.
- Temperature - T $[°C]$: proportional to the average kinetic energy of the atoms and molecules of which the air is made up of. It is relevant for the estimation of energy exchange between atmosphere and earth's surface, for the snow melting, and for the estimation of evapotranspiration.
- Relative humidity - RH [%]: expressing how close the air is of being saturated with water vapour. It has a strict relation with distribution and occurrence of precipitation, and it influences the estimation of evapotranspiration.
- Solar radiation - SR $[W/m^2]$: which represents the transfer of solar energy through electromagnetic waves. In particular, it is used the incoming solar radiation downward – also called shortwave radiation – which represents the incoming ultraviolet, visible, and a limited fraction of the infrared energy. It is the input of energy in the earth-atmosphere energy balance and it is relevant in estimating evapotranspiration and snow melting.
- Wind speed - W $[m/s]$: is the air in motion with respect to the earth's surface, it is the relation between distance covered by air and the required time to cover that distance. Notice that, in the present case, this variable do not influence the simulations since the potential evapotranspiration is estimated through the Priestley-Taylor equation.

Data for Como, Lecco, Milano, Monza e Brianza, and Varese provinces, were collected from the *Regional Agency for Environmental Protection (ARPA - Agenzia Regionale per la Protezione Ambientale)* database from Lombardy region. Moreover, starting from year 2013, information from *meteonetwork* (MNW) which is an open source database were data are collected by citizen scientists was included (for more information about MNW the reader can refer to: https://www.meteonetwork.it/rete/). In the following table the number of stations per variable are shown:

Table 3.1: Number of *ARPA* and *MNW* stations per variable.

Data	P	T	RH	SR	W
ARPA	89	76	61	23	48
MNW	142	144	130	130	130
Total	231	220	191	153	178

It is important to highlight that series of data of a given station do not necessarily cover the whole studied period (18 years), but they do cover some periods of time. Additionally, the information in the database is reported in UTC + 0, while information from ARPA in UTC + 1 and MNW in UTC + 1 and UTC + 2 in winter and summer, respectively. Thus, in the last case the information was homogenized and then, all the download data was shifted by one hour in such a way that everything was reported in UTC + 0. In Fig. 3.3 to Fig. 3.7 are shown the location of the stations in the studied region:

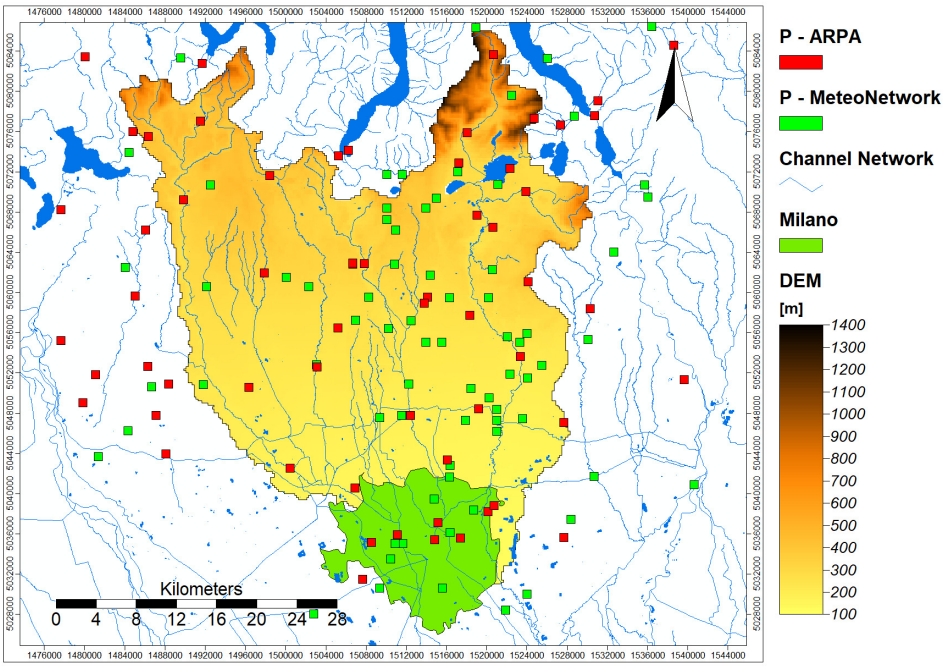


Figure 3.3: Pluviometric stations

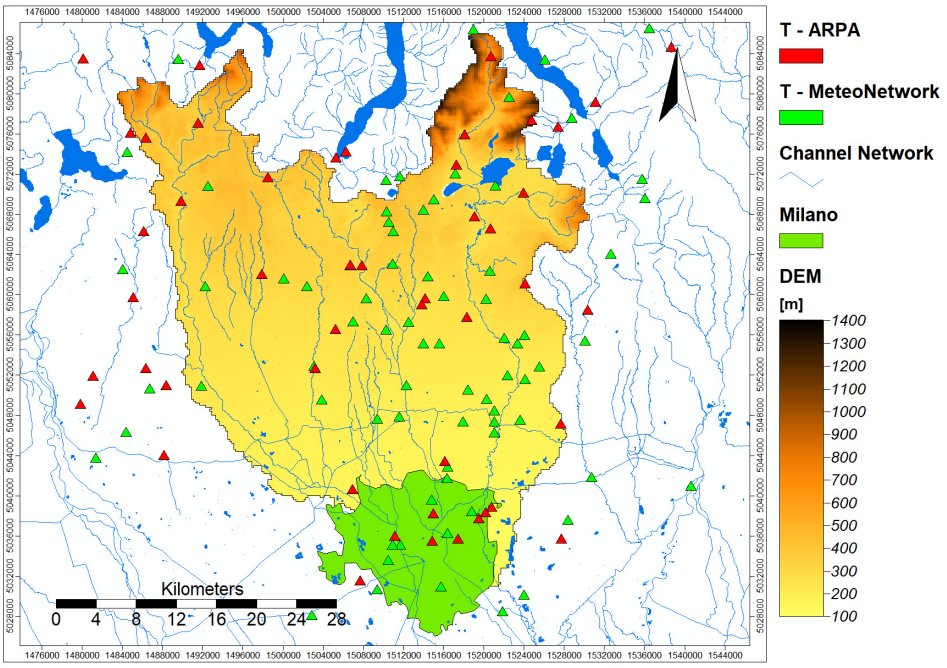


Figure 3.4: Thermometer stations

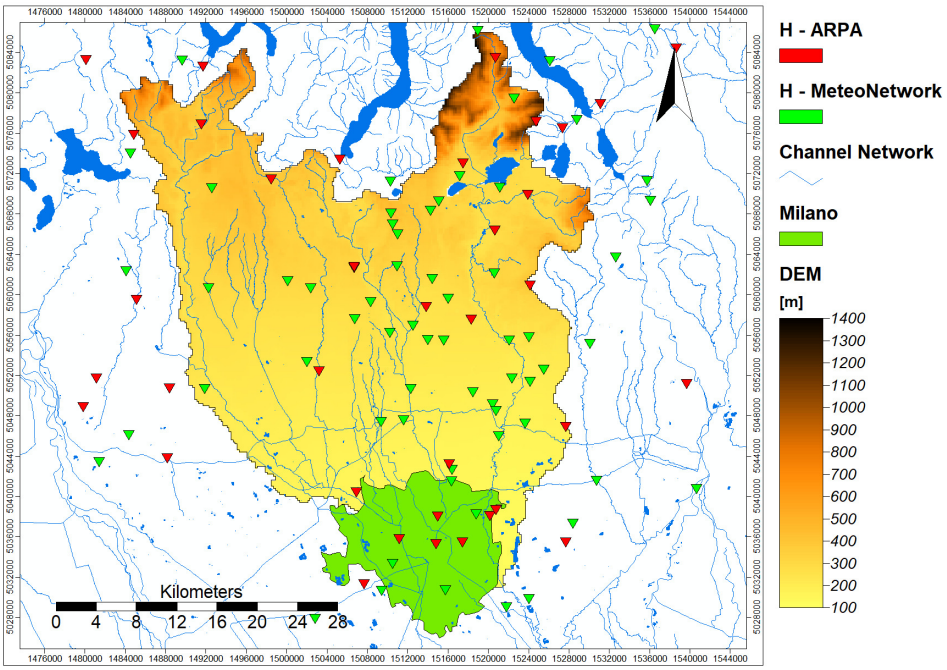


Figure 3.5: Hygrometer stations

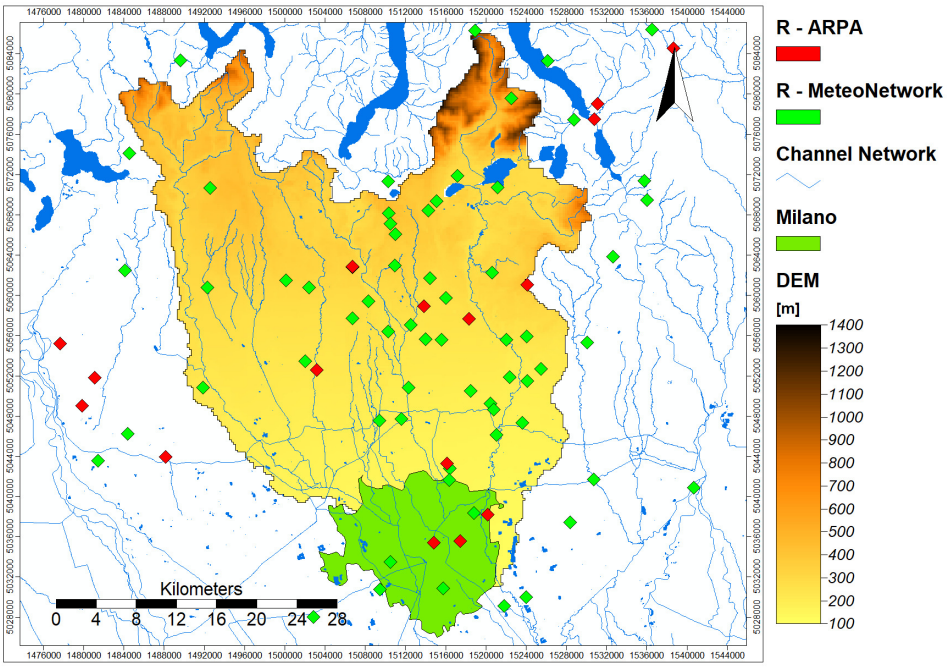


Figure 3.6: Solar radiation stations

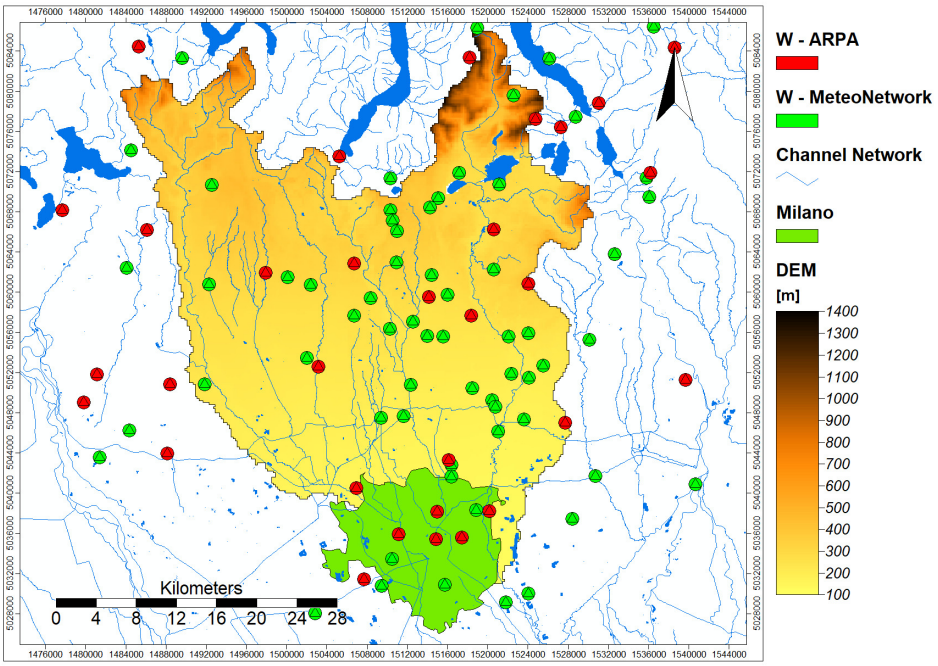


Figure 3.7: Wind speed stations

3.4 Quality control

When measuring meteorological variables in-situ by means of the corresponding instrumentation, some problems may arise leading to wrong data acquisition. Consequently, it is possible to perform a quality control procedure with the scope of identify and substitute data that may have not physical meaning, leading to better simulation results [36]. Then, by substituting manually this information, more accurate results from the hydrological modelling are expected. In the present study values fulfilling the following conditions were highlighted:

- Precipitation: negative values, hourly difference greater than $50 mm$, or values greater than $30 mm$.
- Temperature: values lesser than $-15\,°C$, values greater than $40\,°C$, or hourly difference greater than $8\,°C$.
- Relative humidity: values lesser than $10\,\%$.
- Solar radiation: negative values, or values greater than $1200\,W/m^2$.
- Wind speed: values greater than $15\,m/s$.

When identifying one of the above-described situations and when considering that there was not logical explanation for that value, it was substituted with non-available data (-999.9). This, except for relative humidity in which, if values lesser than $10\,\%$ followed a linear trend, then they were substituted with $10\,\%$ or not replaced if a logical explanation was found. Observed situations are described below:

- Increase of temperature greater than $8°C$ and drop of relative humidity to values lesser than 10% at the hourly scale, in similar date for several stations. In this case, values were not substituted since this is probably due to the presence of *föhn* winds - coming from the Alps - which are katabatic dry and warm winds that can increase temperature and reduce the relative humidity.
- Negative values of precipitation immediately after the same quantity but positive, for a given station. In this case negative values were substituted with $0\,mm$. This since some problems with the instrumentation may lead to reporting negative values after reading precipitation in one hour. It was decided to not substitute with non-available data (-999.9) in order to not to lose information related to the intermittence of precipitation.

- Wind speed larger than $300\,m/s$ were found in some stations, these were the only values that were substituted with -999.9.
- Values barely exceeding the limits described. In this case the information was not substituted since they did not affect the daily trend followed by each variable at hourly scale.

It is important to highlight that the quality control is not made by isolating each station, on the contrary, analysis was performed by considering the context given by nearby stations of the selected variable and the other variables.

3.5 Forecasted weather data

Forecasted MOLOCH weather data from year 2013 to 2020 is used in this study. Some procedures were followed in order to create a unique file before using the information as input of the hydrological model:

- Downloading from MOLOCH archive catalogue. It is obtained a .tgz file for each day of the year containing .grib2 files for each hour of forecast made in Italy region. The first 24 h of forecast have been taken, it is, from 04:00 UTC to 03:00 UTC of the next day.
- Post-processing of each file to reduce their size and make it readable by FEST-WB. This is done through Climate Data Operator – CDO program which allow us to merge files, convert them into netcdf files, reduce the area of interest, calculate wind speed from u and v components, set time references, remove x and y variables since latitude and longitude are already present, and remove all variables not used in the hydrological simulation (see Appendix A - Post-processing of MOLOCH data).
- Processing of missing data: in the downloaded files it might happen that there is no information of some days. Therefore, it was decided to use data from the available day immediately before the missing date. These days were highlighted and omitted when performing the corresponding analyses.

Additional to these procedures, it must be highlighted that – as stated in Section 3.1 – MOLOCH model has had two spatial resolution changes within the period of interest. Consequently, a re-grid operation from 2013 to October 2016 (when the last change of resolution happened) was performed with the aim of representing the information with $1.25\,km$ of spatial resolution. Then, after merging the data of each day, a unique netcdf file was obtained, making the forecasted information readable by FEST-WB.

3.6 Coupling strategy and hydrological simulations

The forecasting system applied in the present study consists in initializing the hydrological FEST-WB model with forecasted forcing variables obtained from the MOLOCH meteorological model. For instance, forecasted meteorological variables (solar radiation, temperature, relative humidity, precipitation, and wind speed) at hourly scale from 04:00 UTC of one day to 03:00 UTC of the next day are used as input in FEST-WB model in order to forecast hydrological quantities such as potential evapotranspiration, soil moisture, and discharge. Concerning the hydrological simulations, the following were done:

- Simulation initialized with observed weather data in the period 2003 to 2020. This allows us to create the reference benchmark with which perform further analysis. Moreover, since the hydrological model executes the spatialization of the observed information from the different weather stations, it is possible to characterize our studied region from a meteorological and climatological point of view.
- Simulation initialized with forecasted weather data for each day of the overlapping period between observed and forecasted information, it is, from January 2013 to December 2020. This is done in order to perform a comparison in terms of the hydro-meteorological variables between simulations initialized with observed and forecasted weather data.

- With the aim of identify which of the four meteorological forcings (temperature, solar radiation, relative humidity, and precipitation) influences the most the results given by the hydrological model, it is performed a sensitivity analysis. For instance, the one-factor-at-time (OAT) methodology [40] [41] was used, it is, performing hydrological simulations forced with all 3 observed input and 1 forecasted data input. Consequently, 4 simulations were done from 2013 to 2020, and afterwards, a statistically-based analysis is executed.

3.7 Statistical analysis

Some statistical indexes are assessed to evaluate how much the forecasted hydro-meteorological variables differ from the observed/simulated ones. In the following paragraphs are explained the statistical tools used to appropriately compare simulations forced with observed and forecasted data, and to perform the climatological characterization of the area of study.

In first place, statistical indexes evaluating the average error between forecasted and observed data, such as the mean error (ME), the mean absolute error (MAE), and the root mean square error (RMSE) are used. Expressions to estimate them are as follows:

$$ME = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i) \tag{3.7}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|F_i - O_i| \tag{3.8}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2} \tag{3.9}$$

In which N is the number of elements in the sample, and F_i and O_i are the *i-th* forecasted and observed values, respectively.

Concerning the characteristics of these indexes [42], all of them have a best score of 0. The ME represents the average forecast error, and in consequence, it indicates the direction of the estimation, it is, whether there is an average under or overestimation. Additionally, it ranges from $-\infty$ to ∞, it does not represent the magnitude of the error, and it is possible to obtain the best score when there exist compensating errors, meaning that it is an index that do not measure the correspondence between forecast and observations. The MAE and RMSE do represent the average magnitude forecast error, they range from 0 to ∞, and they do not indicate the direction of the estimation. Moreover, the RMSE is greatly influenced by large errors; therefore, it is a proper index when big differences between observed and forecasted data are unfavourable.

In second place, linear regressions and assessment of the coefficient of determination R^2 are used when comparing the information. The regression is executed following the least squares methodology, in which the sum of the square of the residuals is minimized. On the other hand, the coefficient of determination expresses how close the data is to a straight line when represented in a scatter plot, it is, it expresses the correspondence between observed and forecasted data. It has a range between 0 – which represents a useless regression – and 1 – which represents a perfect regression. The coefficient of determination is evaluated as [43]:

$$R^2 = \frac{SSR}{SST} \tag{3.10}$$

Where both the total sum of squares SST and the regression sum of squares SSR represent the sum of squared deviations of observed or regression predicted values around their mean:

$$SST = \sum_{i=1}^{N}(y_i - \overline{y})^2 \tag{3.11}$$

$$SSR = \sum_{i=1}^{N} [\widehat{y}(x_i) - \overline{y}]^2 \tag{3.12}$$

In which y_i and $\widehat{y}(x_i)$ are the *i-th* observed and regression predicted values, respectively, and $\overline{y}$ is the mean of the observation values.

In third place, one way to verify the forecasts is through plots that allow us to identify the correspondence with observed values. For example, scatter plots - in which the observed/simulated data is plotted against the forecasts - are a good way to have a visual hint of accurate forecasts if data points are aligned to the diagonal. Moreover, scatter plots between observed/simulated and its difference with the forecasts – the errors - are useful to identify some possible relationships such as over or underestimation. Additionally, bar plots expressing the percentage in which the ME or the MAE are within a range – bins of the errors – are useful to identify how big are the usual errors committed by the forecasts. Finally, bar plots expressing the number of times an observed/simulated and forecasted variables overpassed a threshold are applicable when trying to identify over or underestimation tendencies.

One further comment concerning the present study is that to properly manage the obtained information from the hydrological simulations, especially for the comparison and sensitivity analyses, three considerations were taken: first, since MOLOCH forecasts are given at UTC + 0 and ARPA and MNW information at UTC + 1, a shift in time of forecasts is necessary to set all information at UTC + 0, second, the missing days - previously identified from the forecast dataset – were removed from the series to be analysed, and third, it was only considered the profound flowrate when the snow coverage was greater than 0.1 %. Additionally, graphs were made with Matlab software.

Chapter 4

Results and discussion

In this chapter is presented the climatological characterization of the area of study in which a temporal analysis of the regionalized hydro-meteorological variables is performed. Then, it is discussed the comparison between the observed and forecasted data from FEST-WB simulations initialized with the observed dataset and MOLOCH forecasts in the period from 2013 to 2020, this with the aim of evaluate the performance of the meteorological model. Finally, the sensitivity analysis, performed with the one-factor-at-time (OAT) methodology, is shown with the scope of defining which one of the forecasted meteorological variables influences the most the simulated hydrological variables – especially the flowrate – when compared with the observed results.

4.1 Climatological characterization

With the observed meteorological information from the stations described in section 3.3 and considering that one of the processes performed by FEST-WB model is the spatialization of the information, it is possible to perform a climatological characterization of the studied area. Therefore, analyses were performed using spatialized information for the period 2003 to 2020, particularly, for the *Seveso-Olona-Lambro* (SOL) river basins in the case of the climatological characterization, and for the *Seveso* River with closure at *Bovisio-Masciago* in the case of the discharge analysis. In Fig. 4.1 is represented the studied region.

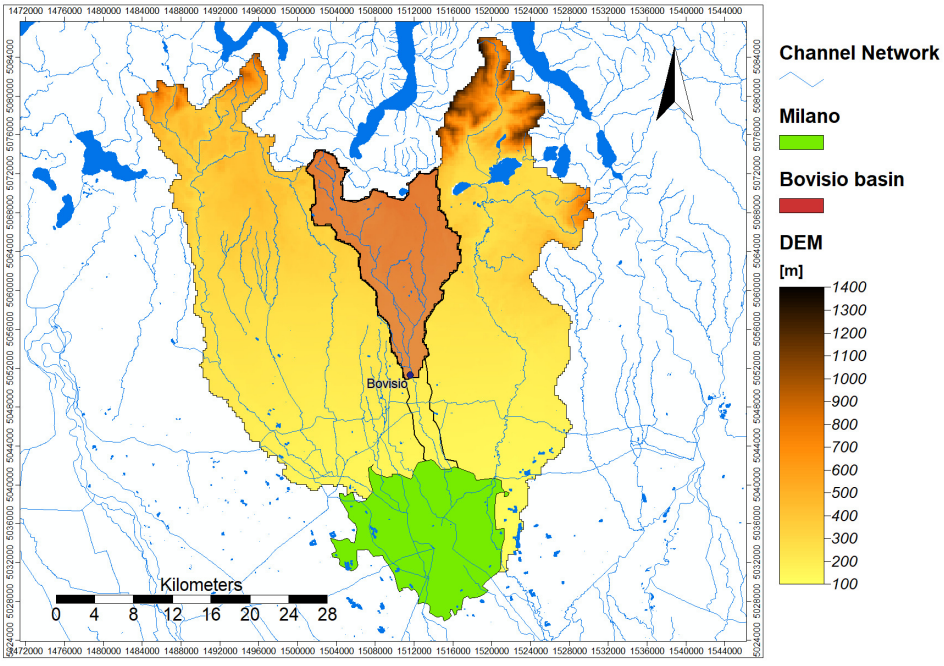


Figure 4.1: Studied region

In the following are described some results of temporal analysis for the meteorological forcing variables (precipitation, temperature, relative humidity, solar radiation, and wind speed) additional to the soil moisture and potential evapotranspiration. Then, some relations between the behaviour of the variables are highlighted to complete the characterization. Finally, temporal analysis is also performed for the simulated discharge at *Bovisio* station, which has been chosen since upstream that section there are no hydraulic works (in particular, as above-described, the overflow channel CSNO) and, consequently, it is possible to observe an undisturbed behaviour of the basin (although large urbanization is present in the area).

4.1.1 Temperature

This variable is the measure of the average kinetic energy of particles and molecules composing the air. Therefore, it is a parameter of the thermal state of the matter. It presents variability in terms of altitude and seasonality. The latter being a proper characteristic of middle latitude regions. The annual mean temperature distribution is shown in Fig. 4.2. It is observed a range of $2.0\,°C$ with temperatures varying from $12.0\,°C$ to $14.0\,°C$. The mean temperature in the studied period corresponds to $13.2\,°C$. Moreover, it is observed an increasing trend of temperature with a rate of $0.08\,°C/year$. Related to this, it is possible to observe that from year 2014, annual mean temperature has always been above the mean in the studied period.

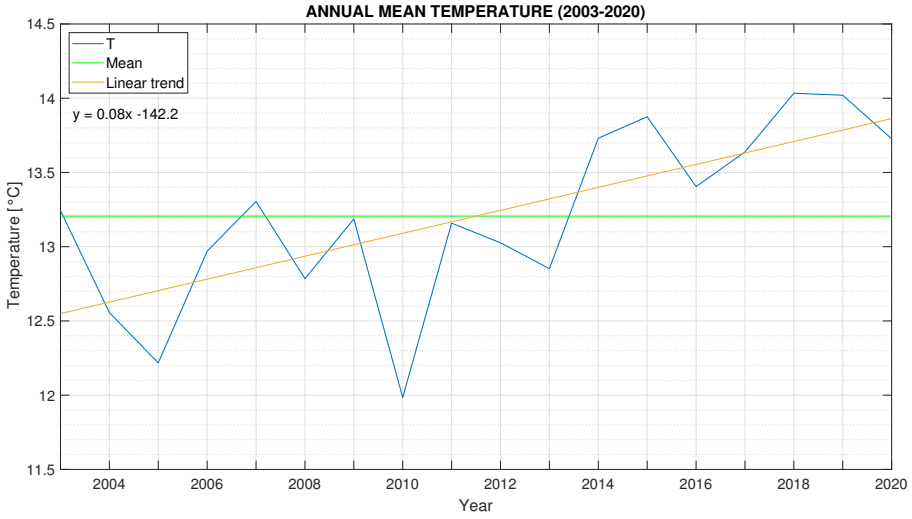


Figure 4.2: Annual mean temperature

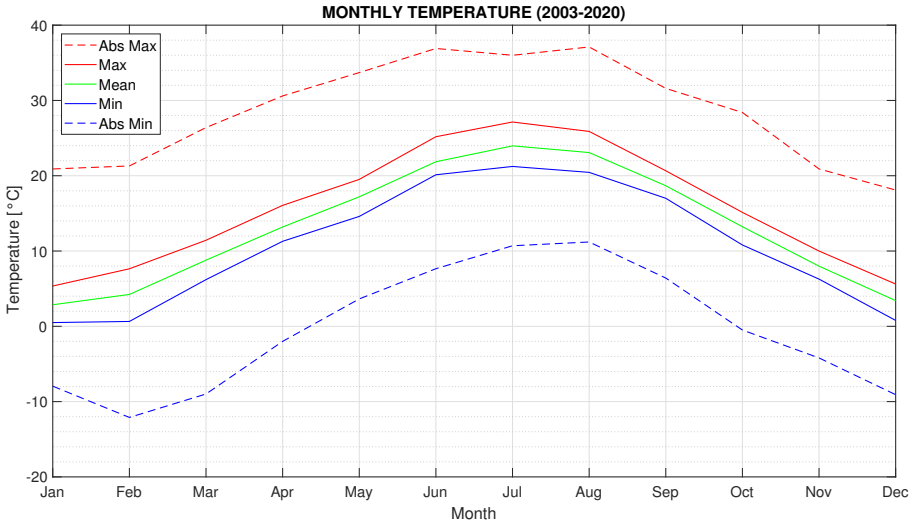


Figure 4.3: Monthly mean temperature

With respect to monthly variations, in Fig. 4.3 we can observe the absolute maximum, maximum mean, mean, minimum mean, and absolute minimum monthly temperatures. From this, it is possible to see the seasonality of this physical quantity, with a warm period from June to August and a cold period from December to February. Maximum absolute temperature is verified in August with 37.1 $°C$. Conversely, minimum absolute temperature is $-12.1\,°C$ in February. The annual range of temperature, it is, the difference of average temperature in the warmest and coldest months, is 21.1 $°C$; being the coldest month January with a mean temperature of 2.9 $°C$ and the warmest month July with 24.0 $°C$.

Concerning the absolute maximum and minimum temperatures per day, in Fig. 4.4 and Fig. 4.5 are shown the correspondent values per month in the studied period. With this, it is again possible to observe the seasonality of the temperature. We observe that from mid of November to mid of March the minimum temperatures are consistently below 0.0 $°C$. Then it starts increasing until it is consistently above 10.0 $°C$ from June to September. Similar comments can be made for the maximum temperature, being lesser than 20.0 $°C$ in the period from November to February, and greater than 30.0 $°C$ from the end of May to mid of September. One additional consideration is that the highest temperature is verified on August 11^{th} (2003) with 37.1 $°C$ and the lowest temperature on February 6^{th} (2012) with $-12.1\,°C$. Finally, it must be remembered that the values presented in this paragraph are absolute and at the basin scale; it is, they cover the entire area of study since spatialization has been performed.

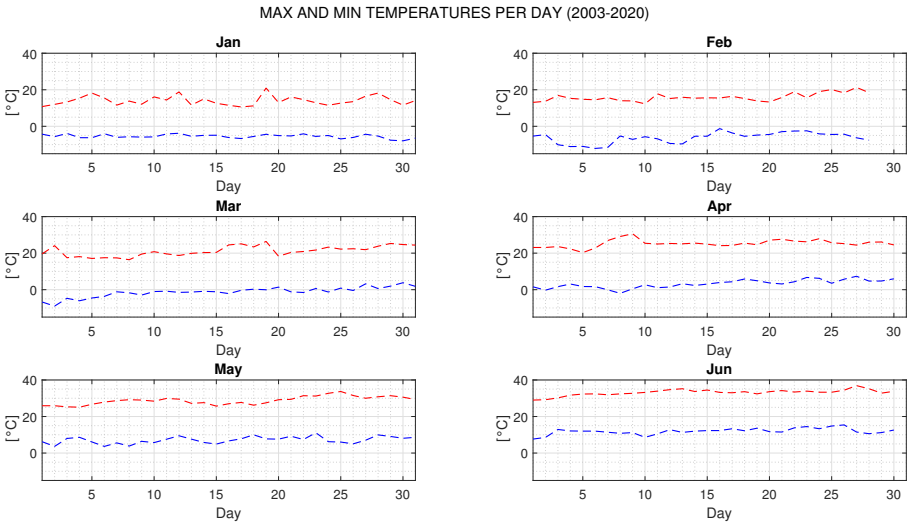


Figure 4.4: Maximum and minimum temperatures per day (Jan-Jun)

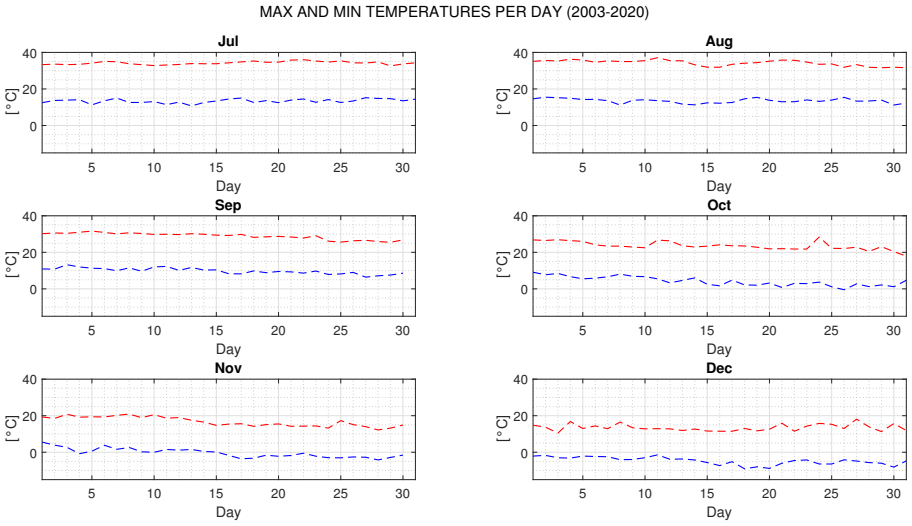


Figure 4.5: Maximum and minimum temperatures per day (Jul-Dec)

4.1.2 Solar radiation

Incoming solar radiation represents the transfer of solar energy by means of electromagnetic waves. It is related to evapotranspiration, snow melting, and it represents the input of energy in the earth-atmosphere energy balance. Particularly, it is used the incoming solar radiation downward, also called shortwave radiation. Annual maximum radiation is shown in Fig. 4.6 for the studied period. It varies between 911 W/m^2 to 1000 W/m^2, with a mean value of 951 W/m^2. With respect to its trend, it is maintained practically constant along the years since the rate of variation is 0.12 W/m^2. However, form year 2011 it is possible to observe a decreasing trend.

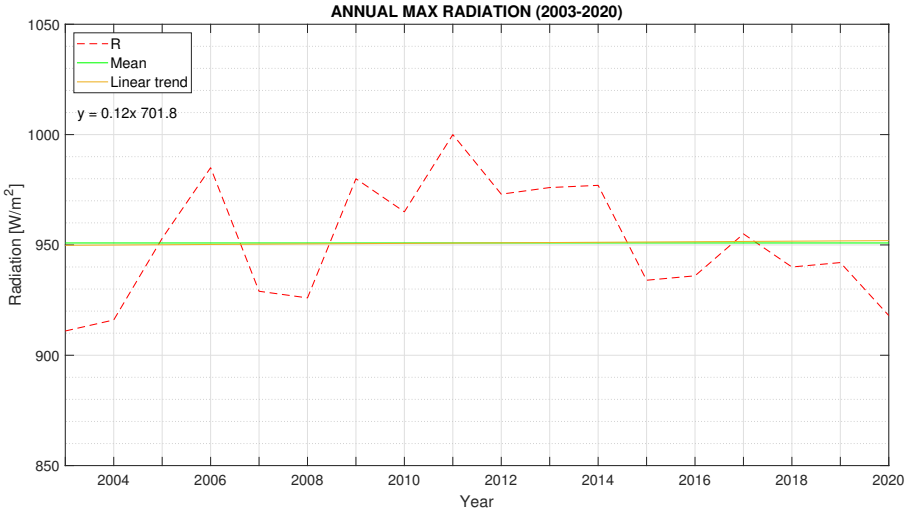


Figure 4.6: Annual maximum radiation

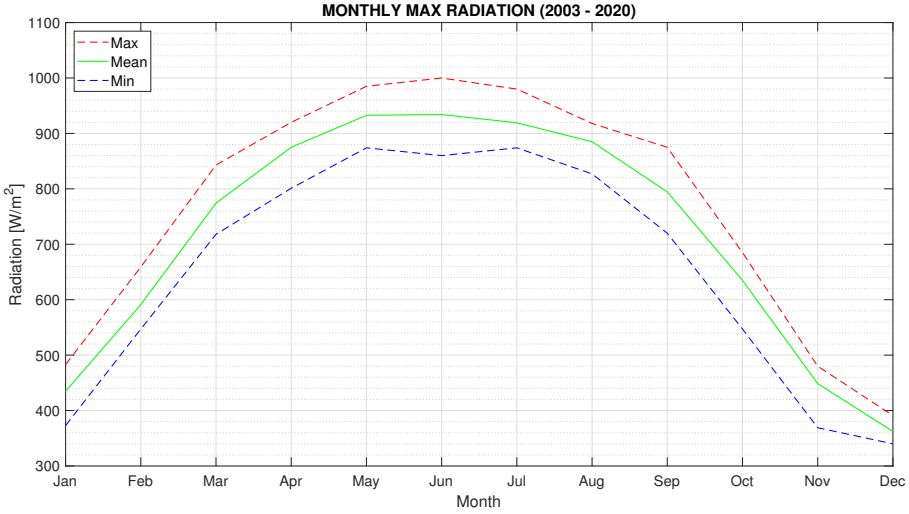


Figure 4.7: Monthly maximum radiation

With respect to the maximum, mean, and minimum monthly solar radiation in Fig. 4.7, again it is evident the seasonal nature of this quantity. Greater values are verified in the warm season with maximum mean in June (934 W/m^2). Conversely, lower values in the colder season with minimum in December (362 W/m^2). The annual range, for instance, 572 W/m^2. This trend is due to the tilting of the earth on its axis, making that – in the north hemisphere – mid latitudes receive more incoming solar radiation in June, while less radiation in December. This behaviour is also evident in Fig. 4.8 to Fig. 4.10 in which the maximum radiation per day in the studied period is drawn. Particularly, on Fig. 4.10 we can see the variation day by day, but also the overall trend in the entire year, again, following the above-described behaviour.

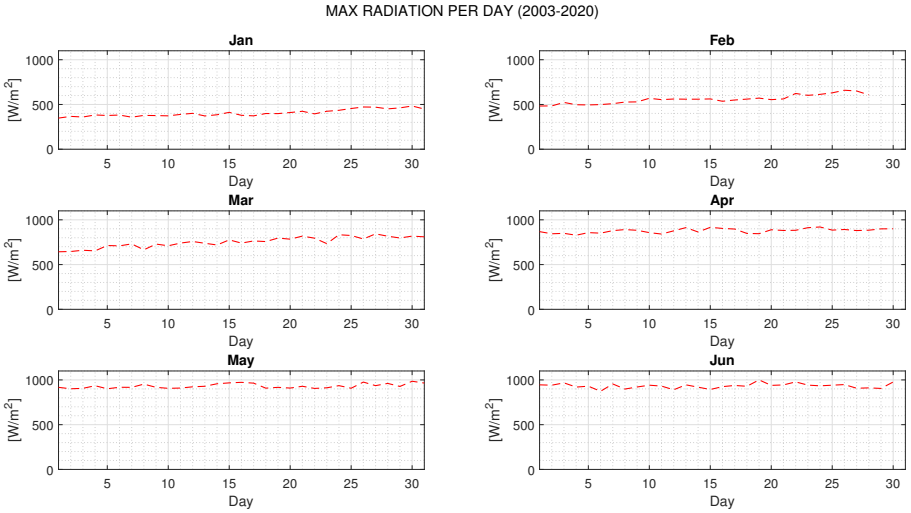


Figure 4.8: Maximum radiation per day (Jan-Jun)

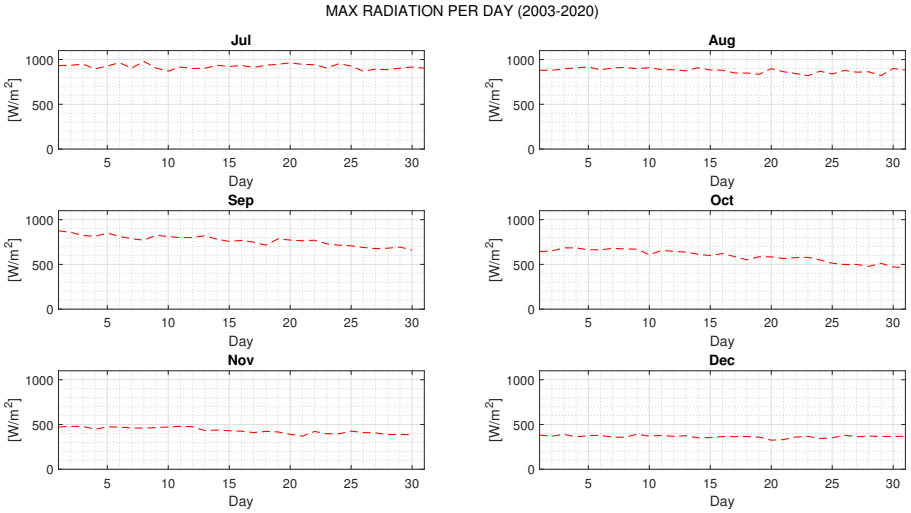


Figure 4.9: Maximum radiation per day (Jul-Dec)

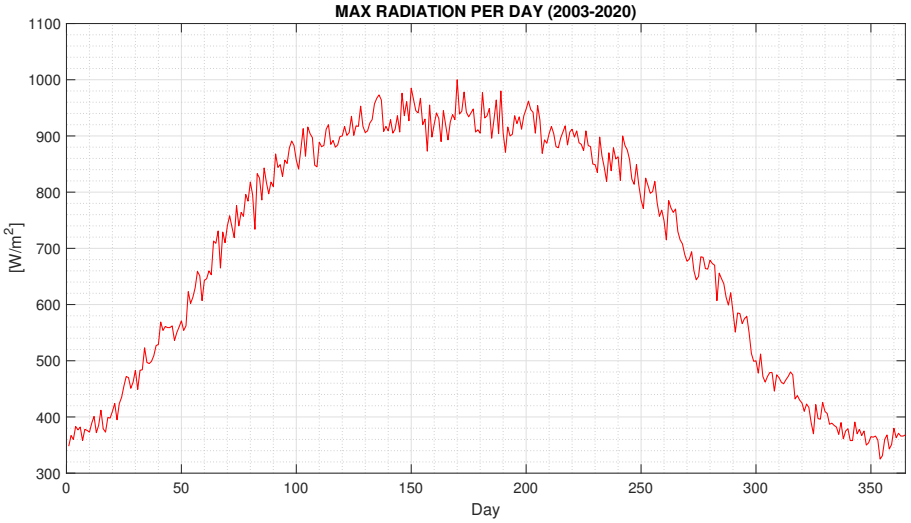


Figure 4.10: Maximum radiation per day

4.1.3 Relative humidity

Relative humidity is a measure of how close the air is of being saturated with water vapour. Consequently, it has a strict relation with distribution and occurrence of precipitation, and it is involved in evapotranspiration process. The annual mean humidity in the studied period is shown in Fig. 4.11. We observe a range of variation of approximately 67 % to 77 %, with a mean of 71 %. Values oscillate around the mean and, therefore, the increasing trend has a rate of 0.17 $\%/year$. These values allow us to say that the studied region tends to be more humid than dry, proper characteristic of northern Italy inland zones.

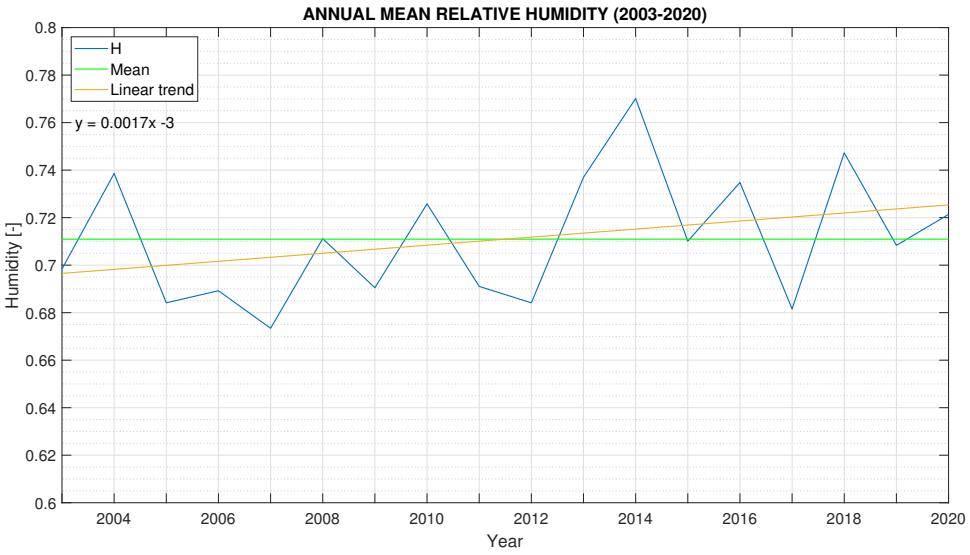


Figure 4.11: Annual mean relative humidity

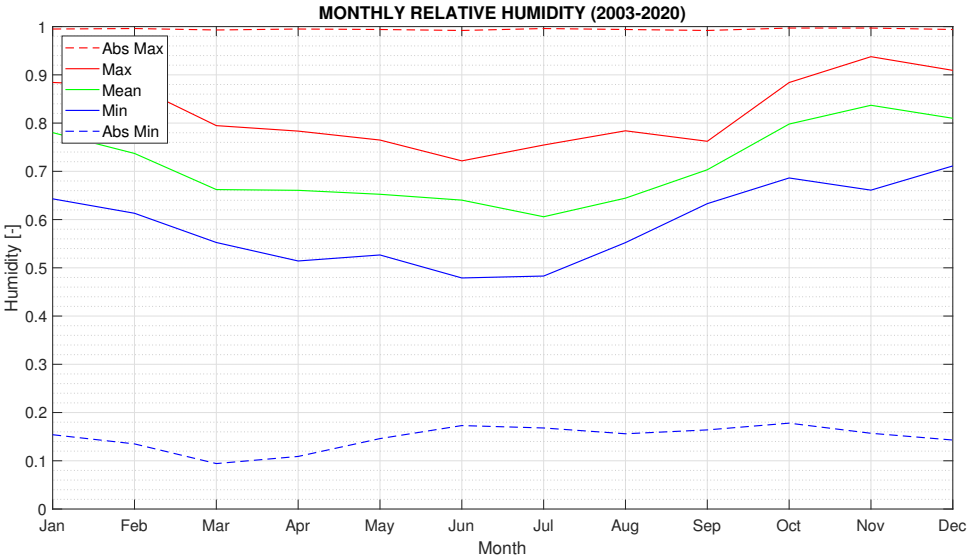


Figure 4.12: Monthly relative humidity

In respect of the monthly variations, in Fig. 4.12 are shown the absolute maximum, maximum mean, mean, minimum mean, and absolute minimum monthly relative humidity. We can observe that for the coldest period of the year humidity is larger than in the warmest period. The range of the mean value is 24 % with a minimum of 60 % in July and a maximum of 84 % in November. From this, it is possible to see that mean humidity varies in a range of high values and, therefore, in the studied region are expected humid summers and foggy winters. Absolute minimum humidity values vary between 9 % to 18 % and are verified – as stated in previous sections – due to the presence of dry winds coming from the Alps.

Concerning maximum, mean, and minimum humidity at daily scale per month, we can see in Fig. 4.13 and Fig. 4.14, the trend in the studied period. In this case it is also possible to observe the seasonal behaviour of this variable. Greater values in the colder season and lower values in the warmer period. It is important to highlight that this trend is verified since cold air masses need less water vapour to reach saturation, while the opposite happens for warm temperatures. Referring to maximum humidity, it is seen that values are all near to 100 %, meaning saturation of the air.

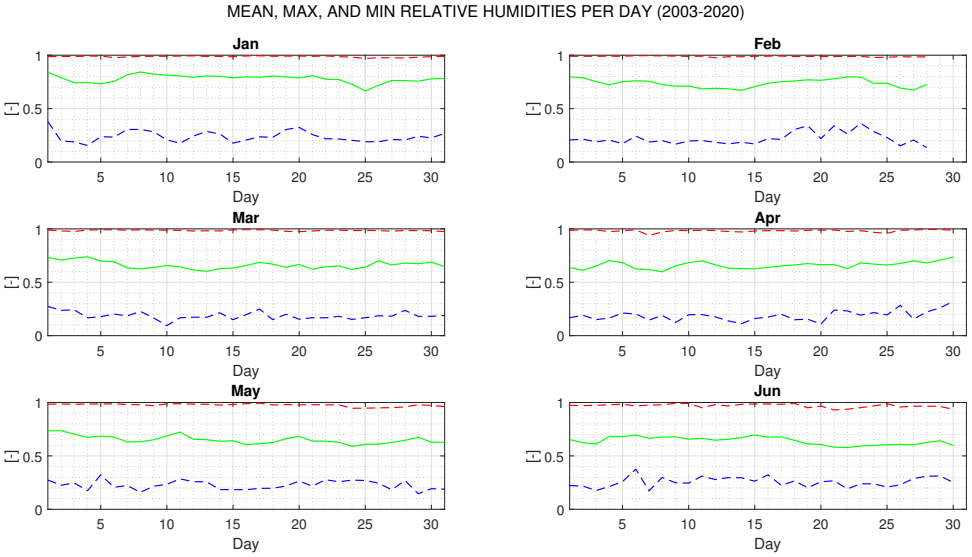


Figure 4.13: Mean, maximum, and minimum relative humidities per day (Jan-Jun)

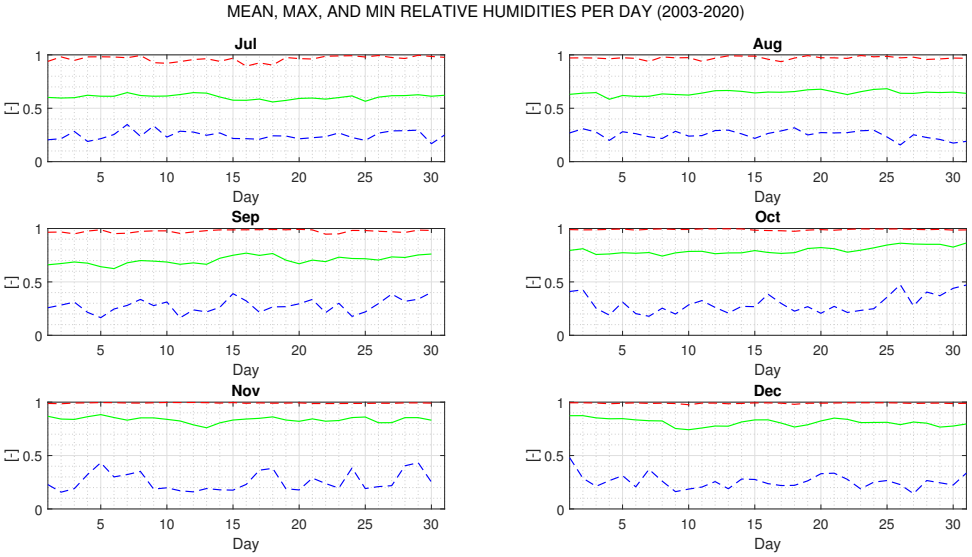


Figure 4.14: Mean, maximum, and minimum relative humidities per day (Jul-Dec)

4.1.4 Precipitation

Represents the amount of water that falls and arrives at the earth's surface in a particular place and in a period. It is a meteorological variable that presents huge spatial and temporal variation. In Fig. 4.15 is shown the annual precipitation in the studied period. The mean is 1200 mm and it is observed an increasing trend along the years with a rate of 14.2 $mm/year$. It is remarkable that the mean is a value which represents the average between a greater amount of precipitation in the mountainous zone of the area of study and the smaller amount of precipitation in the plain zone. Concerning maximum annual precipitation at daily and hourly scales in the period 2003-2020, in Fig. 4.16 we observe a mean of 60.4 mm and 13.7 mm, respectively. Their trends are practically constant along the years. Additionally, maximum daily precipitation ranges from 44.1 mm to 95.6 mm and, at hourly scale from 9.3 mm to 19.7 mm.

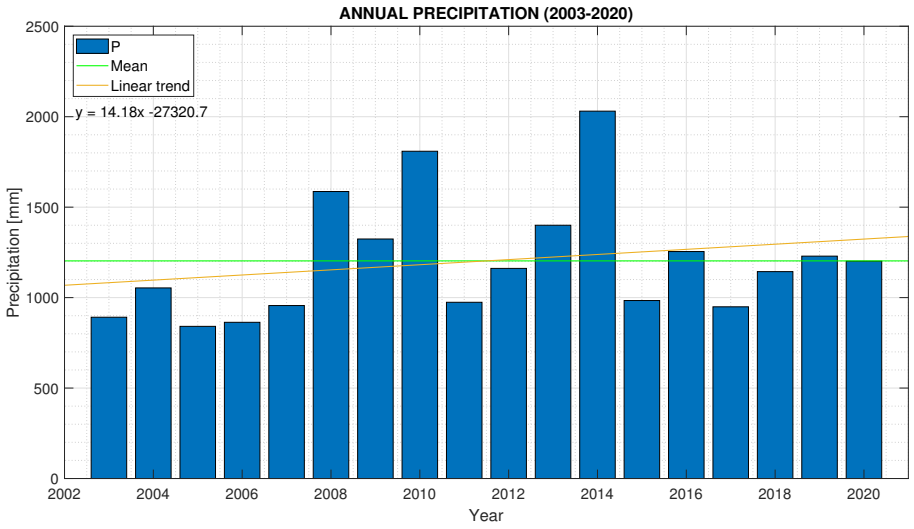


Figure 4.15: Annual precipitation

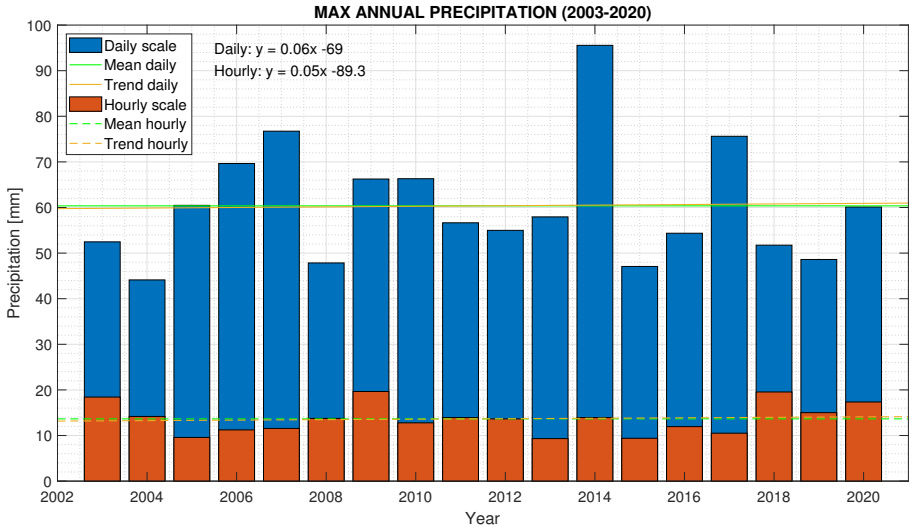


Figure 4.16: Maximum annual precipitation

In respect of monthly distribution in Fig. 4.17 are represented the maximum, mean, and minimum precipitations. A bi-modal behaviour is evident with two precipitation periods, one with larger precipitation heights at Autumn, and the other one at Spring. The annual range of precipitation is 101.4 mm, being January the month with the lowest mean precipitation (57.4mm) and November the one with the highest (158.8 mm).

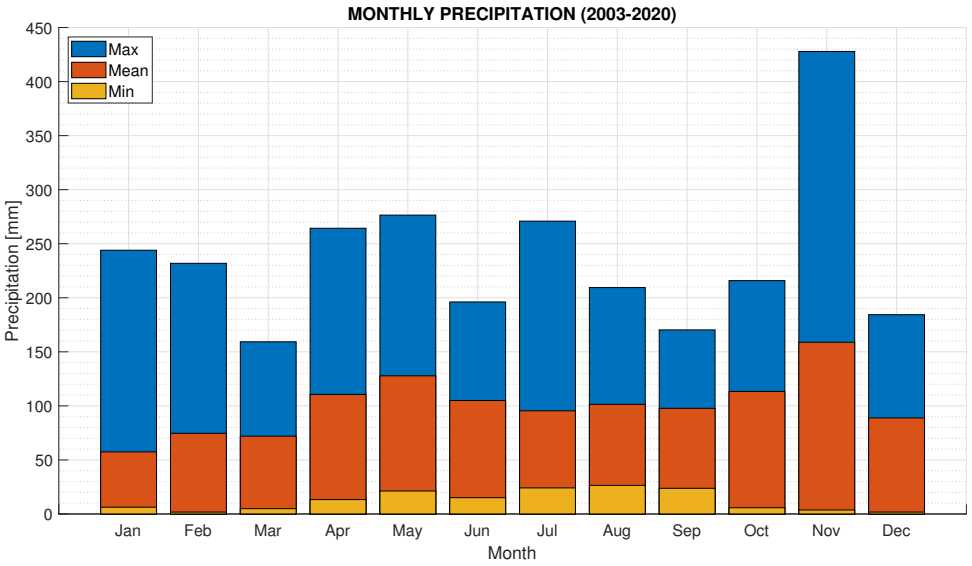


Figure 4.17: Monthly precipitation

A relevant comment is that, referring to minimum precipitation, values in the warm season of the year – from May to September – are greater than the ones in the colder season. This, meaning that, although Summer do not present the highest mean monthly precipitation - it does not belong to one of the precipitation periods -, minimum heights are larger during this season. This can be seen from Fig. 4.17, in which the trends of mean and minimum monthly precipitations are out of phase. In reference to maximum heights, it is evident that the maximum is verified in November with $427.9\,mm$. Additionally, it is relevant that maximum value of July, which do not belong to one of the precipitation periods, is similar to the ones verified in April and May. This relies in the fact that in Summer season stormy events are common at the end of the days.

Bi-modal behaviour is also evident in Fig. 4.18 representing the mean number of wet days per month in the studied period. A wet day is defined as a day in which precipitation height is larger or equal than $1\,mm$. Again, larger values occur in the rainy seasons, with a maximum of 12 days in May followed by 11 days in June and November. Moreover, the mean number of days with rain is at least 6 days for any month, being this value the one for January.

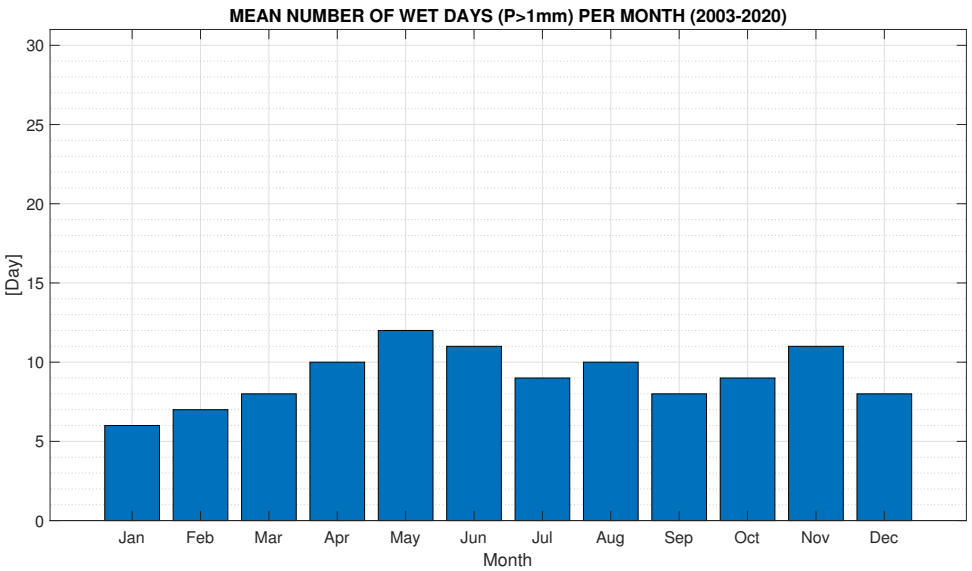


Figure 4.18: Mean number of wet days per month

Additional to the previous information, the daily occurrences of precipitation (considering values greater than 1 mm) and the maximum hourly precipitation per month at the daily scale, from 2003 to 2020, were estimated (Fig. 4.19 to Fig. 4.22). On first place, concerning the occurrences, we can observe that months in which the 50 % threshold is overpassed are April to July and, November and December, which is in concordance to the bi-model regime already identified in the previous paragraphs. No precipitation occurrence was verified on January 24$^{\text{th}}$ and September 20$^{\text{th}}$. Additionally, we can observe that in Winter the daily occurrences are lower, indicating that one characteristic of this season is that it is dry in terms of precipitation. On second place, concerning the maximum hourly precipitation, it is evident that higher values are verified in Summer. This is, once again, related to the stormy events that occur in this period of the year. The maximum value happened on July 7$^{\text{th}}$ 2009 with a precipitation of 19.7 mm. On the other hand, in Winter season values rarely overpass 6 mm, which confirms that one characteristic of this period of the year is that it is dry from the precipitation point of view.

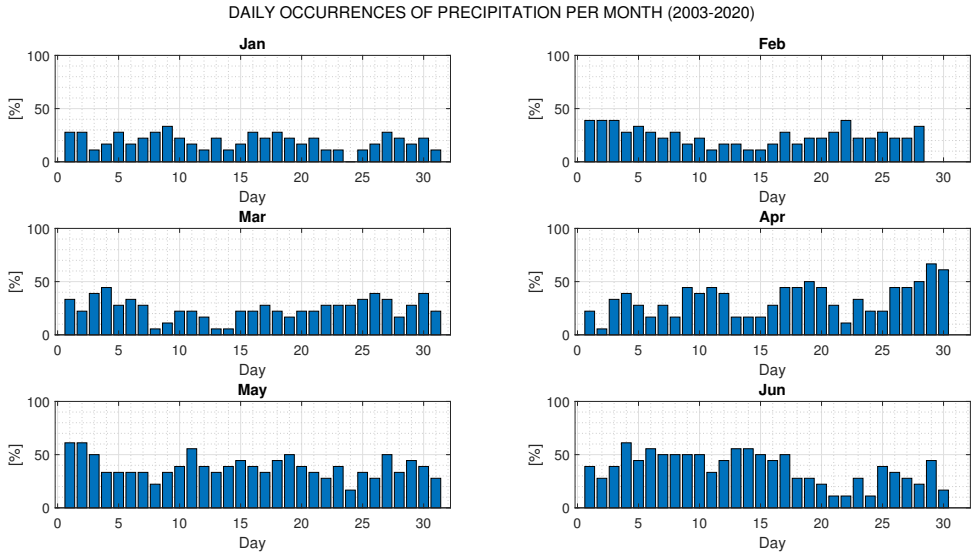


Figure 4.19: Daily occurrences of precipitation per month (Jan-Jun)

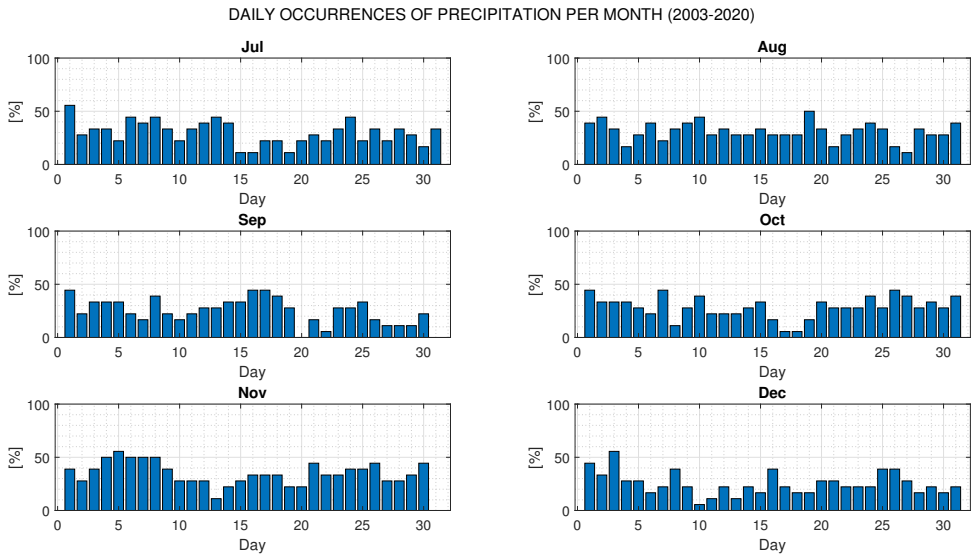


Figure 4.20: Daily occurrences of precipitation per month (Jul-Dec)

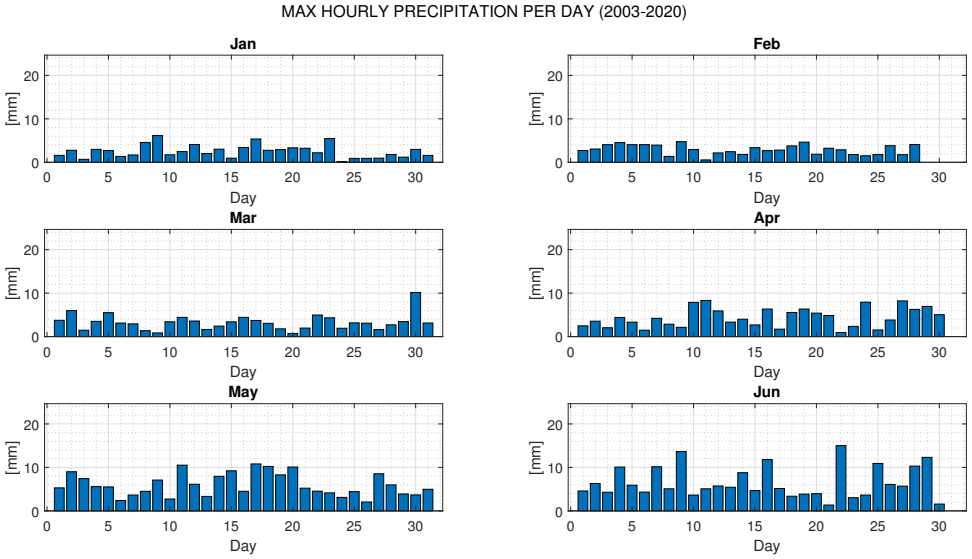


Figure 4.21: Daily occurrences of precipitation per month (Jan-Jun)

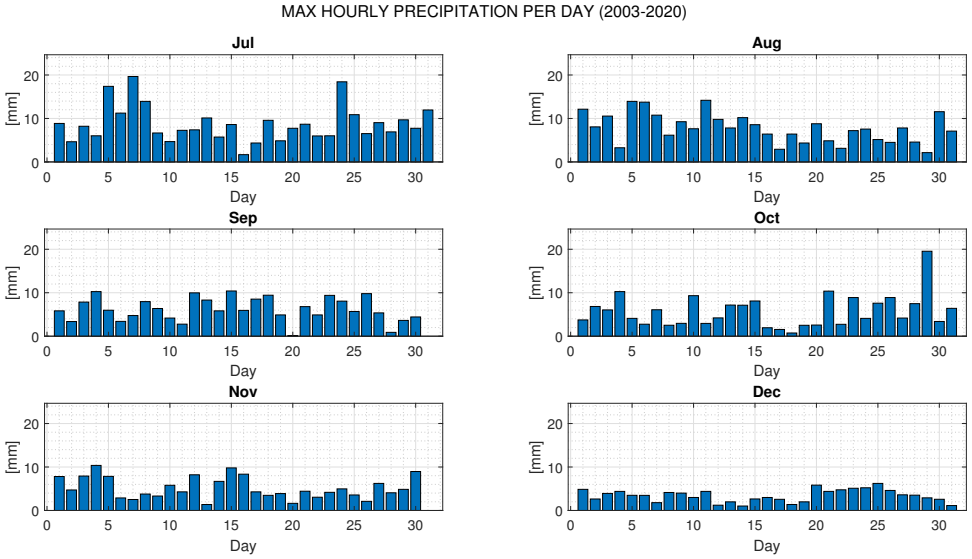


Figure 4.22: Daily occurrences of precipitation per month (Jul-Dec)

4.1.5 Wind speed

Represents the horizontal moving of the air, being the relation between covered distance and time needed to cover that distance. It is a variable that strongly influences the rate of evapotranspiration. In Fig. 4.23 and Fig. 4.24 are represented the hourly maximum, mean, and minimum wind speeds per day in the studied period. Values range from 0.1 m/s to 9.7 m/s, while mean values vary from 0.9 m/s to 2.0 m/s increasing from February and then decreasing from September. Meaning that wind speed has a moderate seasonality along the year.

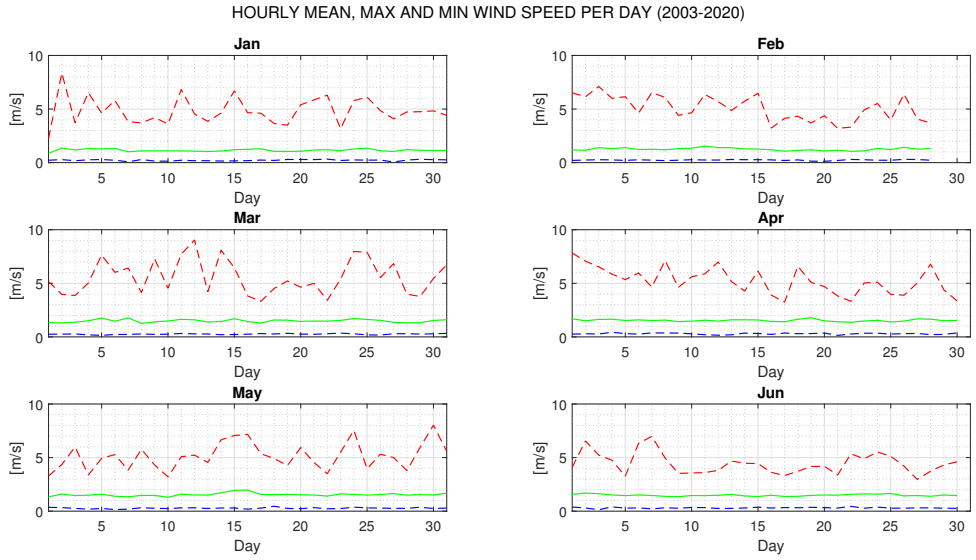


Figure 4.23: Hourly mean, maximum, and minimum wind speed per day (Jan-Jul)

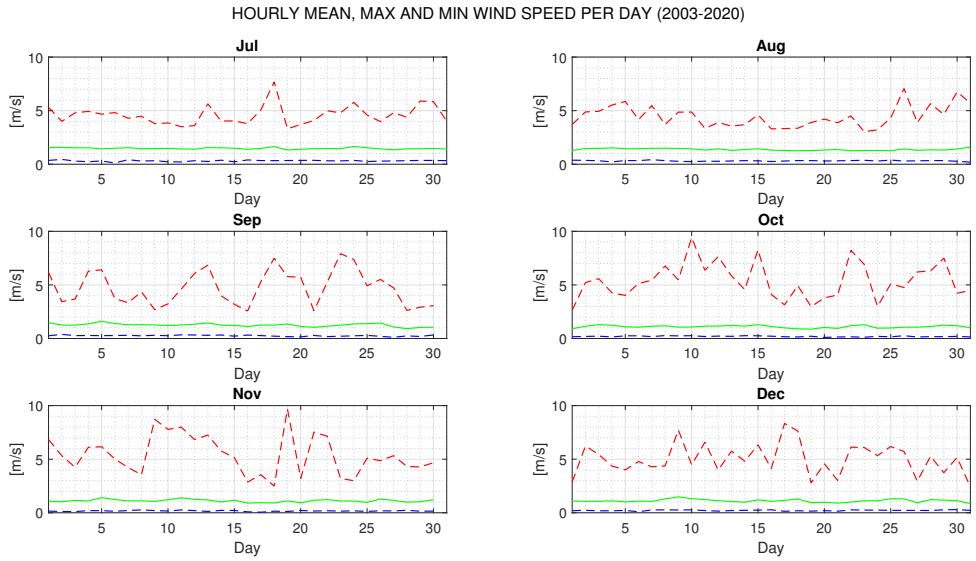


Figure 4.24: Hourly mean, maximum, and minimum wind speed per day (Jul-Dec)

4.1.6 Potential evapotranspiration (PET)

Evapotranspiration considers two phenomena that occurs at the surface: evaporation and transpiration. The former concerning the transformation of water in liquid state into gas state at surfaces such as soil, rivers, lakes, and wetted vegetation. The transpiration also considers the transformation from liquid to gas, but of the water contained in vegetation. Potential evapotranspiration represents the evapotranspiration that would happen if there were no restrictions on the available water. It is a quantity affected by temperature, solar radiation, and wind speed. Annual PET is presented on Fig. 4.25, values range from 510 mm to 622 mm with a mean of 569 mm. It is also observed an increasing trend along the years with a rate of 2.7 $mm/year$.

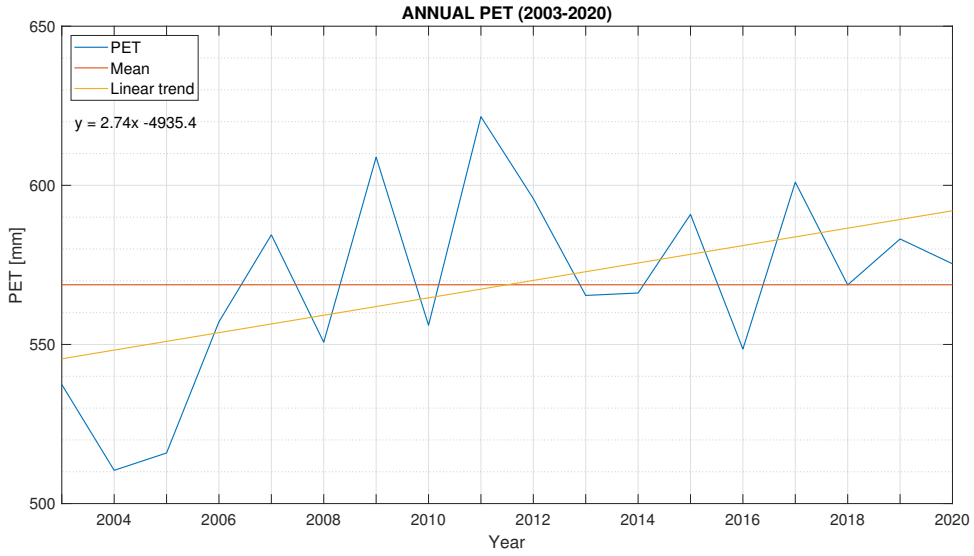


Figure 4.25: Annual PET

In Fig. 4.26 is shown the maximum, mean, and minimum PET in the studied period. Greater values are present – in concordance with solar radiation and temperature – in the warm season of the year; this because higher temperatures and greater radiations provoke a higher potentiality of evaporating water from the surface. On the other hand, smaller values of PET are present in the cold season. Moreover, the range of this quantity is 100 mm with maximum mean value in July (103 mm) and minimum mean value in December (3 mm).

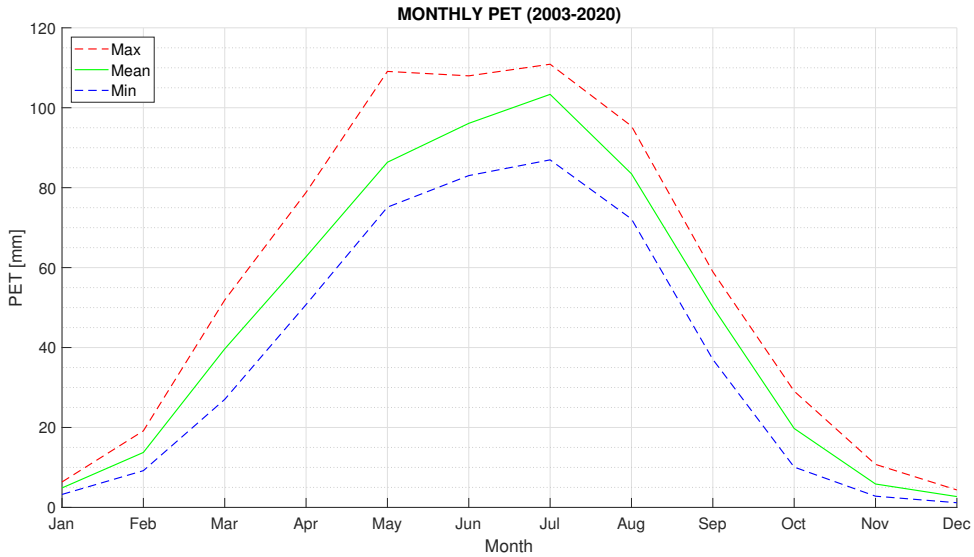


Figure 4.26: Monthly PET

Concerning the daily maximum of PET, on Fig. 4.27 and Fig. 4.28 it is possible to observe the trend. Once again, in this type of plots it is evident the seasonal nature of the quantity. Maximum is verified on 19^{th} June with a value of $5mm$, while the lowest maximum PET is seen on 5^{th} December with $0.2mm$.

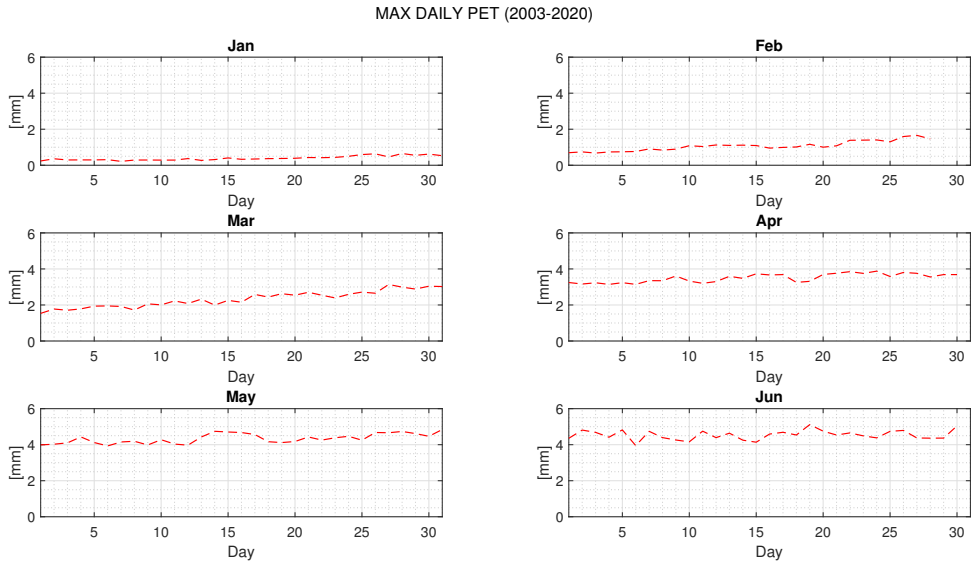


Figure 4.27: Maximum daily PET (Jan-Jun)

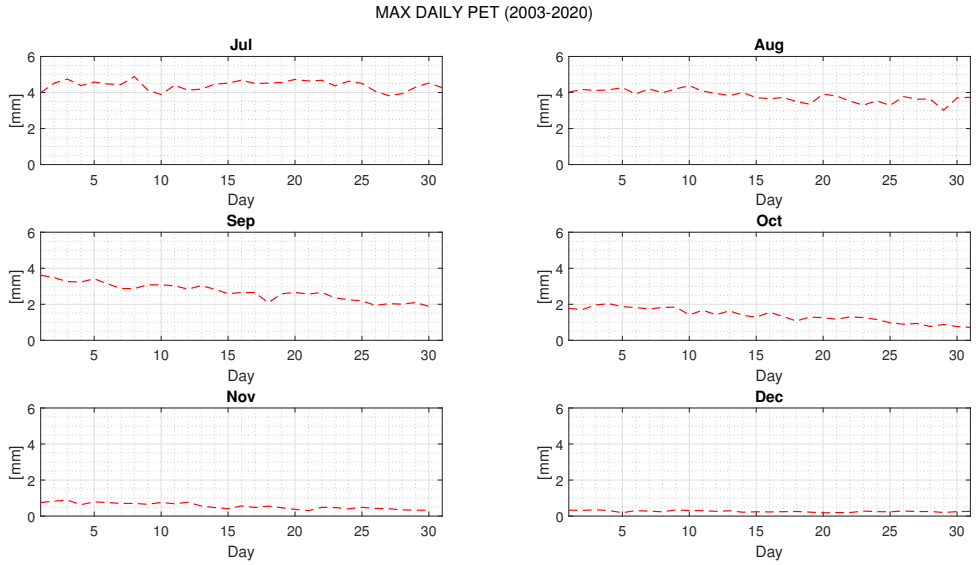


Figure 4.28: Maximum daily PET (Jul-Dec)

4.1.7 Soil moisture

Is the quantity of water contained in the soil expressed – from volumetric point of view - as ratio between the water volume in the soil and the total volume which considers liquid, solid and gas phases. Consequently, it is a dimensionless quantity that varies from 0 – meaning dry soil – to the porosity at saturation; being the porosity the fraction of void space into the soil. This is a relevant parameter in the infiltration process, and it is related to evapotranspiration and precipitation. Moreover, it must be mentioned that CN conditions are unchanged in this study since no significant differences have been found during the last 20 years, whilst great increase of urbanization was present from 1954 to 2000. The annual mean trend in the studied period in Fig. 4.29 shows a variation between 14.0 % to 16.7 % with a mean value of 15.3 %. The linear trend shown expresses an increasing tendency which however is very small, this since it has a rate of $0.04\%/year$. Consequently, it can be said that annual mean soil moisture is basically constant along the years.

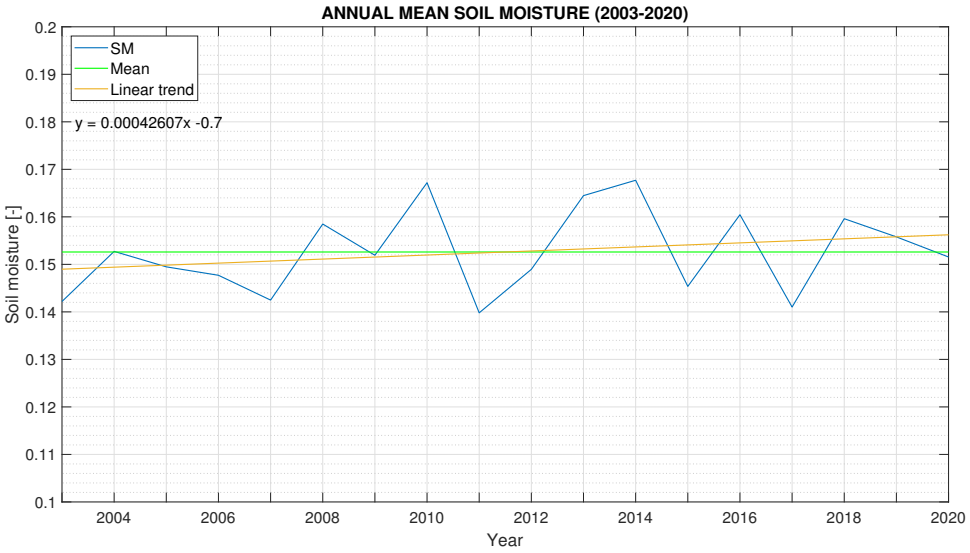


Figure 4.29: Annual mean soil moisture

In respect of the absolute maximum, maximum mean, mean, minimum mean, and absolute minimum monthly soil moisture in the studied period shown in Fig. 4.30, it is possible to observe that lower and high values are present in the warm and cold seasons, respectively. The annual range is about 7.2 %, with the lowest and highest mean being 11.9 % and 19.1 %, verified in July and November, respectively. Moreover, absolute minimum and maximum are within the band of 6.0 % to 28.6 %.

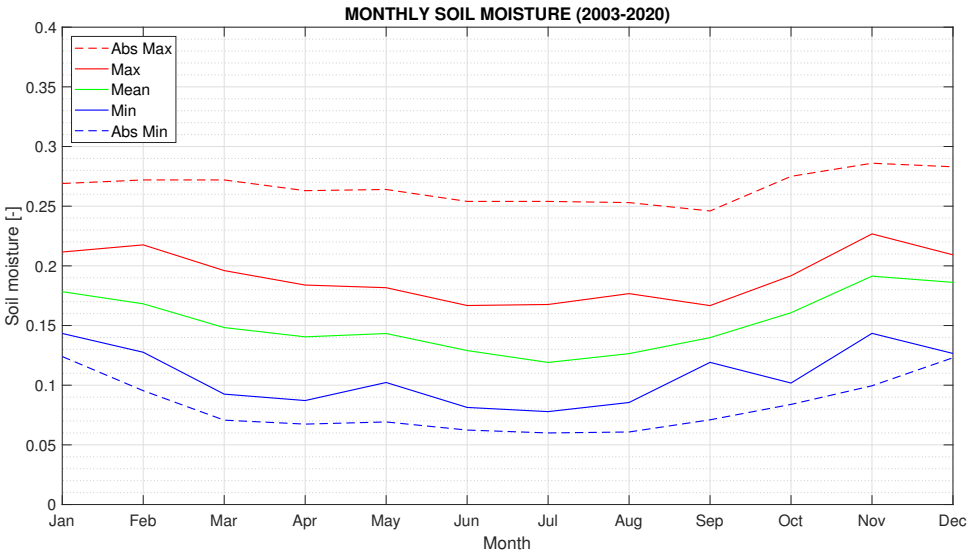


Figure 4.30: Monthly soil moisture

Maximum and minimum soil moisture per day in the studied period is shown in Fig. 4.31 and Fig. 4.32. It is also possible to observe the behaviour linked to the seasonality present in the variables that influence the water content of the soil. Maximum value occurs the 16$^{\text{th}}$ of November (28.6 %) and minimum the 23$^{\text{rd}}$ of July (6.0 %).

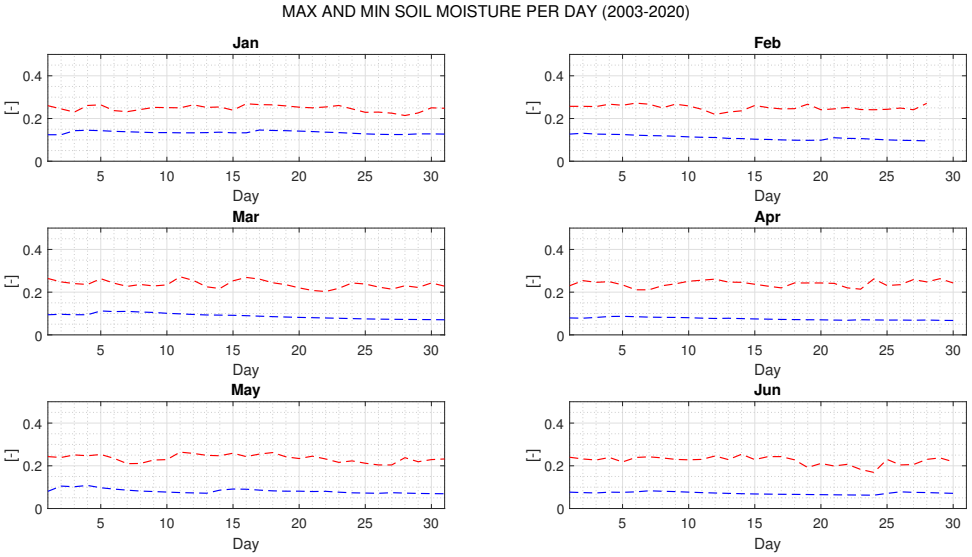


Figure 4.31: Maximum and minimum soil moisture per day (Jan-Jun)

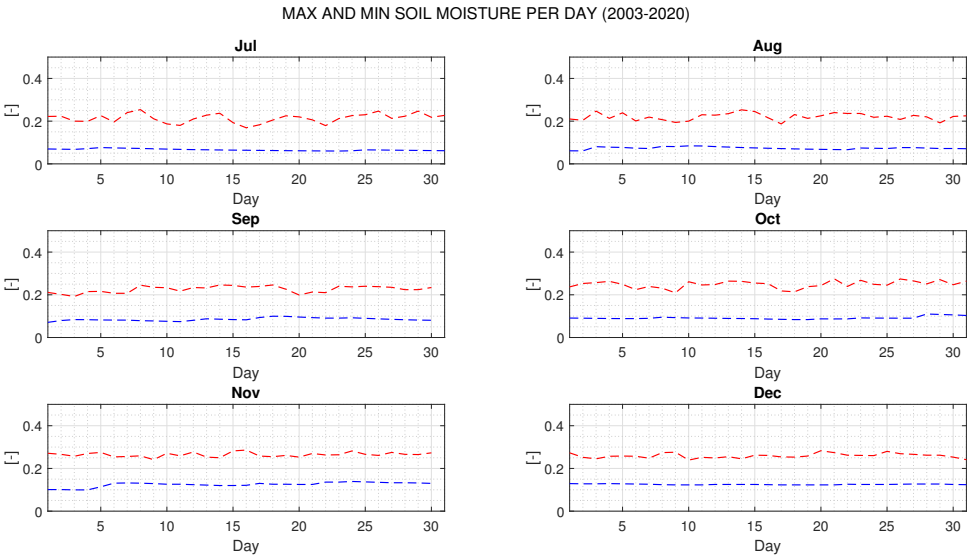


Figure 4.32: Maximum and minimum soil moisture per day (Jul-Dec)

4.1.8 Characterization overview

Once each variable has been described through the temporal analysis, it is plausible to give an overall description of the studied region from the meteorological and climatological viewpoint. *SOL* basin is located in Northern Italy and, consequently, is characterized by seasonality of the meteorological variables. Incoming solar radiation, as well as temperature, present larger values from June to August, and lower values from December to February; corresponding to Summer and Winter. This makes relative humidity to have the opposite behaviour, since greater temperatures lead the air to be warmer, and therefore, it is less easy to be saturated. Nevertheless, relative humidity has large values along the year, leading to humid summers and foggy winters. In respect of precipitation, a bi-modal behaviour is present in the region, with two precipitation periods, one with larger precipitation heights at Autumn, and the other one at Spring. Maximum hourly precipitations are verified in Summer, principally due to stormy events occurring at the end of the days. Finally, concerning wind speed, a moderate seasonality is observed along the year, with values increasing from February and decreasing from September. Moreover, the presence of dry winds coming from the Alps (*föhn* winds) make the relative humidity to drop to very low values, while increasing the temperature.

Potential evapotranspiration follows the same trend as temperature and incoming solar radiation. Explanation is that larger radiation leads to larger temperatures of the air, which at the same time makes the relative humidity to be lower. Consequently, rate of potential evapotranspiration is expected to be larger since air can hold more water vapour before getting saturated. Opposite case during the colder period, in which temperature is closer to the dewpoint, relative humidity is larger, and air cannot store too much water vapour, leading to condensation. It is also important to notice that the moderate seasonality of wind speed do not strongly affect this variable, whilst the factors above-described do. Concerning soil moisture, it is mainly affected by precipitation and potential evapotranspiration. The observed trend is that it is higher during the cold season, in which low precipitations and low potential evapotranspiration are verified. On the other hand, during warm season, relative low precipitation – not belonging to rainy periods of the bi-modal behaviour - and high potential evapotranspiration are present. Thus, it is clear the direct relation of evapotranspiration and soil moisture during the no-rainy periods of the year (Summer, and Winter).

4.1.9 Discharge at Bovisio

Boviso station has been chosen to make the discharge analysis since no hydraulic works are present upstream, allowing us to see the natural behaviour of the basin. Simulated hydrographs with FEST-WB hydrological model were obtained for each year of the studied period (Fig. 4.33). They have also been plotted thresholds of 25 m^3/s, 35 m^3/s (which is the first warning threshold) and 45 m^3/s, since they allow us to define how many times they have been overpassed and, consequently, understand how much possible the observation of an inundation event is.

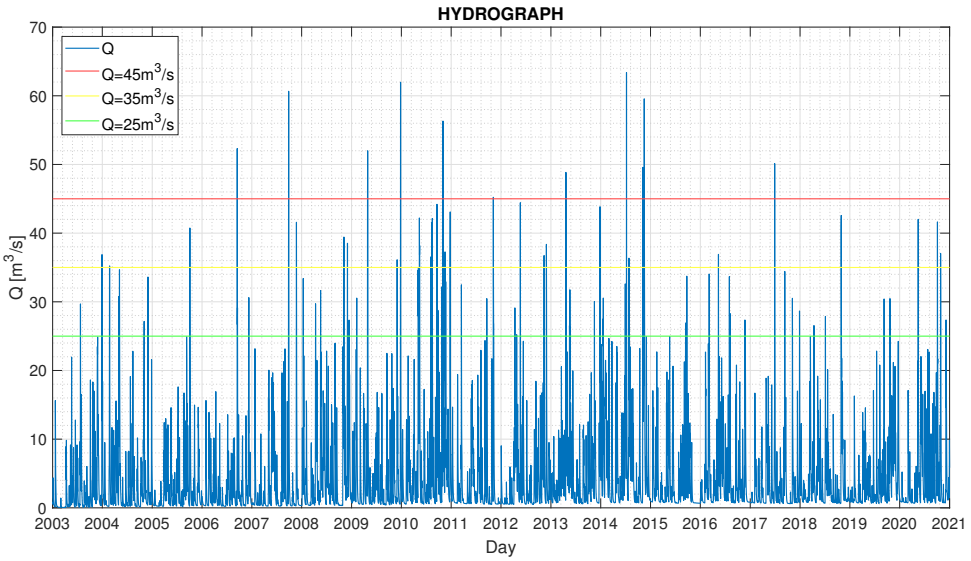


Figure 4.33: Hydrograph (2003-2020)

In Fig. 4.34 is shown the maximum flowrate at daily scale per year. The mean value is 45.6 m^3/s, and it is observed a decreasing trend with a rate of $-0.26\ m^3/s/year$. Maximum flowrate was 63.4 m^3/s in the year 2014, which – as described before – also correspond to the year with the maximum annual precipitation, and maximum annual precipitation at daily scale. On the other hand, the lowest maximum flowrate is 30.5 m^3/s verified in 2019 which, however, do not correspond to the year with the minimum annual precipitation.

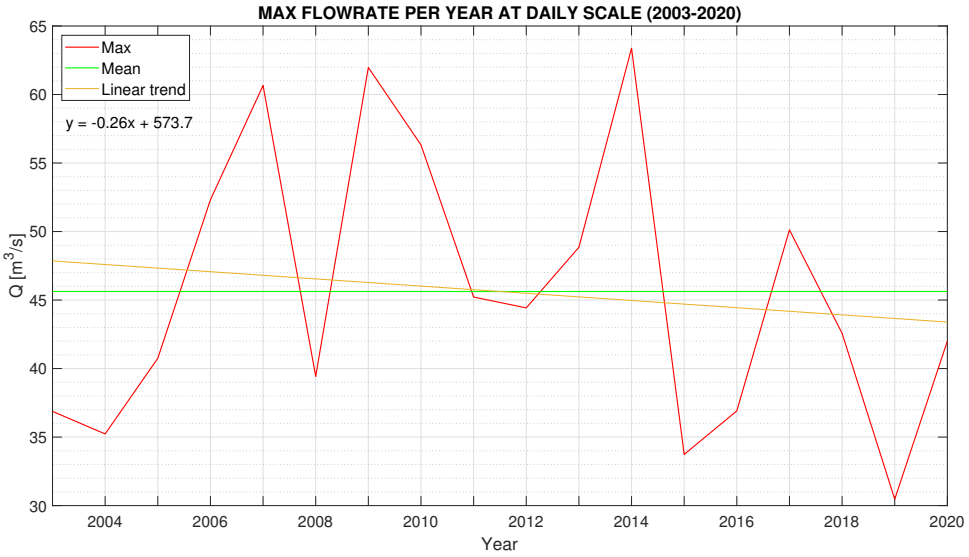


Figure 4.34: Maximum flowrate per year at daily scale

In respect of threshold overpassing, in Fig. 4.35 to Fig. 4.37 are represented the number of days per year in which flowrate has been larger than the threshold. As expected, the number of days decreases when increasing the flowrate to be surpassed. Mean values are 6.6 days, 2.6 days, and 0.7 days, for 25, 35 and 45 m^3/s, respectively. All three graphs show a decreasing trend along the years with a decreasing rate when the threshold is larger. Additionally, the 25 m^3/s flowrate was overpassed at least 2 days in each year belonging to the studied period, the maximum was observed in 2010 with 21 days. For 35 m^3/s, 16 out of 18 years have at least 1 day in which the threshold was surpassed, the maximum was also verified in 2010 with 11 days. Finally, for the highest flowrate – 45 m^3/s -, only in eight years the limit was exceeded, and the maximum number of days is 3 for 2009 and 2014.

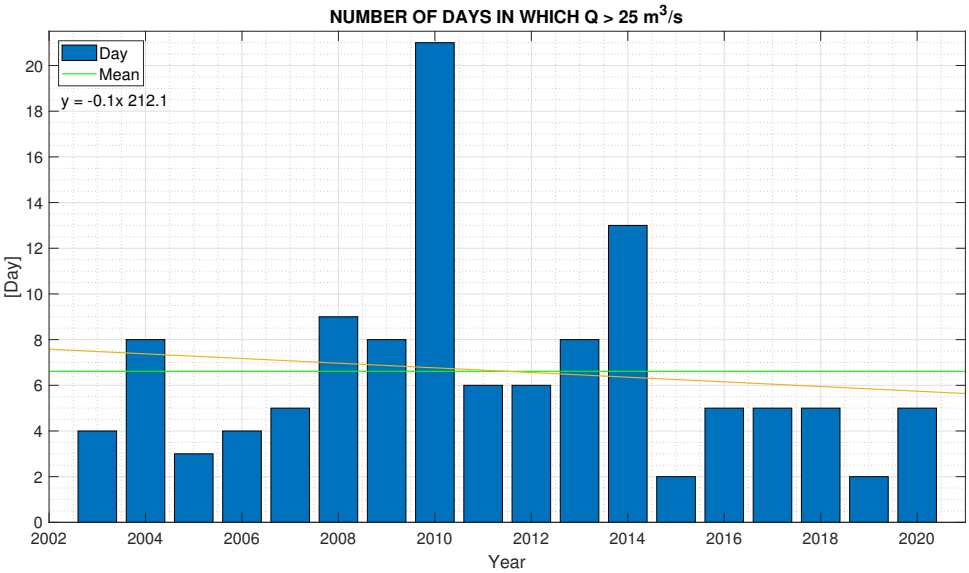


Figure 4.35: Number of days with $Q > 25 m^3/s$

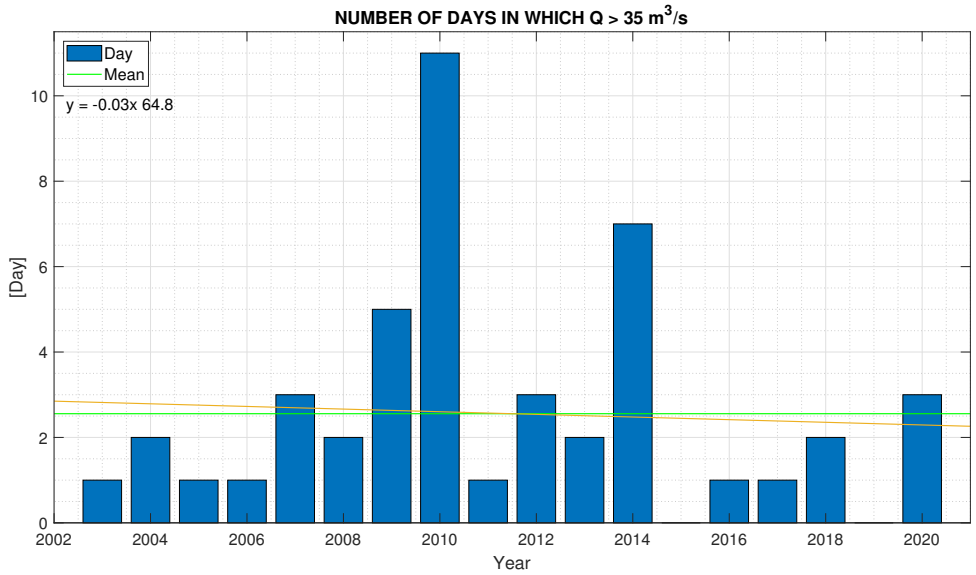


Figure 4.36: Number of days with $Q > 35m^3/s$

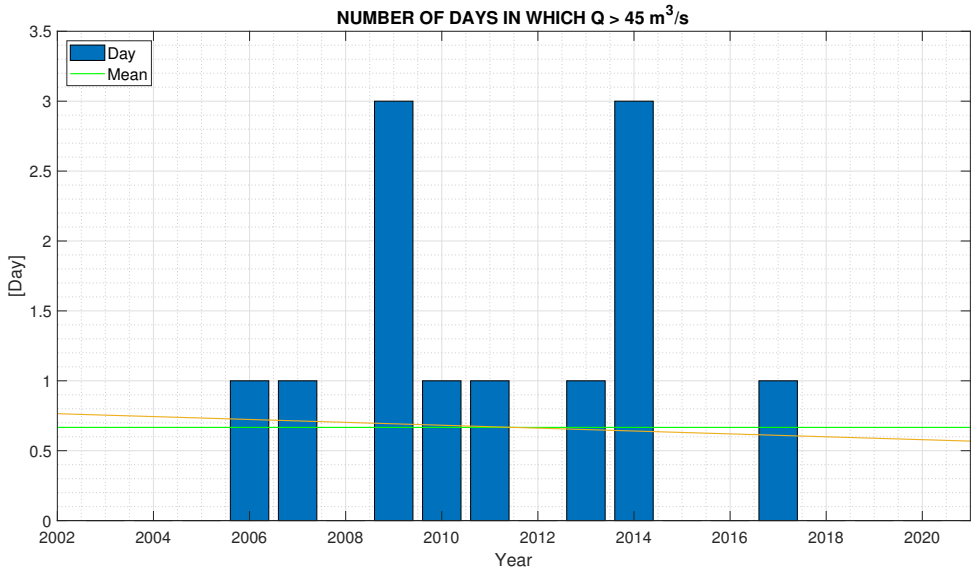


Figure 4.37: Number of days with $Q > 45m^3/s$

4.2 Observed vs MOLOCH comparison

From the results of the hydrological simulations forced with observed weather data coming from ARPA and MNW weather stations, and MOLOCH meteorological model forecasts in the period 2013 to 2020, it is possible to perform a comparison of the hydro-meteorological variables in order to evaluate the performance of MOLOCH model. In the following subsections are shown the results of the comparison of the meteorological forcings (temperature, solar radiation, relative humidity, wind speed, and precipitation) and the simulated hydrological variables (soil moisture, potential evapotranspiration, and discharge), in terms of the statistical tools and indexes described in Section 3.7. Furthermore, this is done at an hourly and daily scale, and considering daily maxima and minima. Additionally, comparison at seasonal scale is also performed for solar radiation, precipitation and discharge. Then, at the end, further comments concerning the comparison and the used forecasted data from MOLOCH are done.

4.2.1 Temperature

From Fig. 4.38 to Fig. 4.41 are shown the scatter plots comparing observed and forecasted hourly, mean daily, daily minimum and daily maximum temperature data. It is possible to observe a very good agreement between both datasets, reflected by means of the coefficient of determination which is larger or equal than 0.96 in all four figures. Nevertheless, by observing the sign of the mean error, the linear regression under the diagonal in the left-hand scatter plot, and the larger amount of data points under the 0 value in the right-hand scatter plot, it is clear that MOLOCH model consistently underestimates the temperature. Moreover, for the hourly and daily scale, from the statistical indexes which give some feedback about the error committed by MOLOCH model, it can be said that – on average- it is not large and the magnitude of the error decreases when averaging the temperature over each day.

Concerning the mean error, the trend of the forecasts can also be observed in Fig. 4.42 and Fig. 4.43. In both cases it is seen a left-shifted distribution meaning underestimation. At hourly scale, the mean error is mainly found within the range of $-3\,°C$ to $-1\,°C$, followed by the range $-1\,°C$ to $1\,°C$ which is the one containing the best score of the index. Similarly at daily scale, the mean error is mainly found within $-2.5\,°C$ to $-0.5\,°C$, also, the percentages of the ranges containing positive values of temperature are very small, meaning that overestimation happens but with less frequency. Finally, these values tell us that, although underestimation is observed, the errors do not have a large magnitude and are usually concentrated in the left side of the best score.

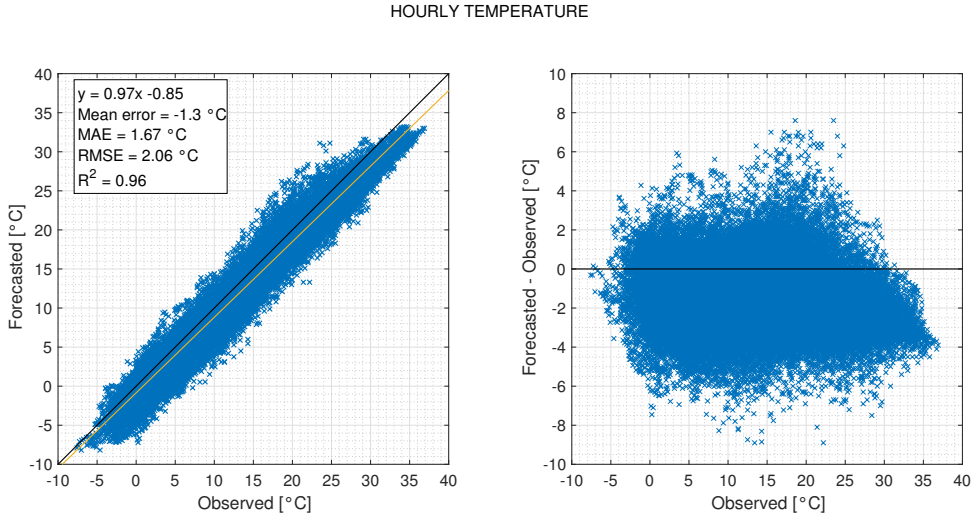


Figure 4.38: Hourly temperature

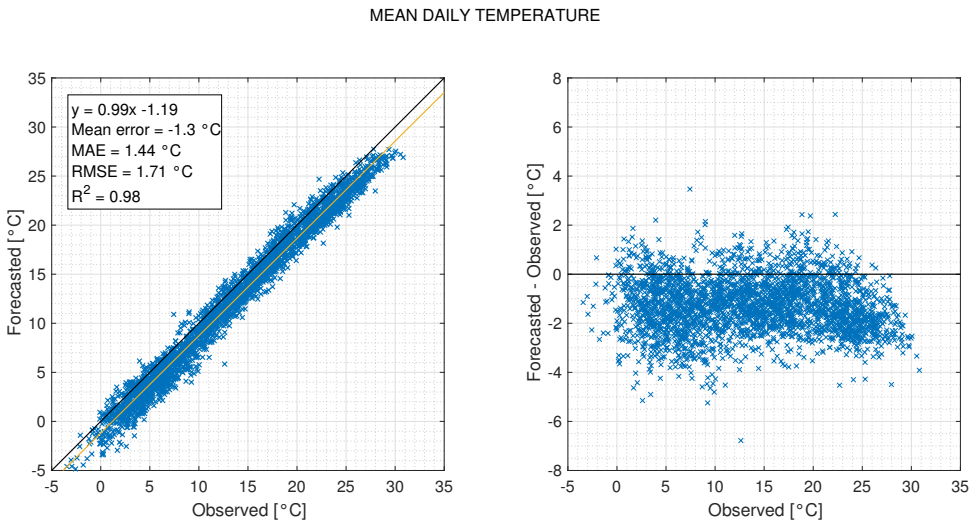


Figure 4.39: Mean daily temperature

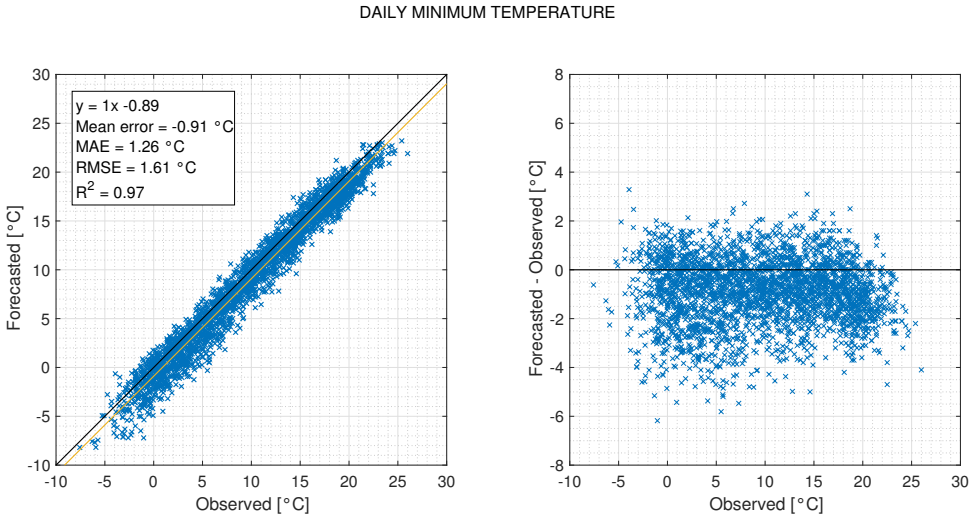


Figure 4.40: Daily minimum temperature

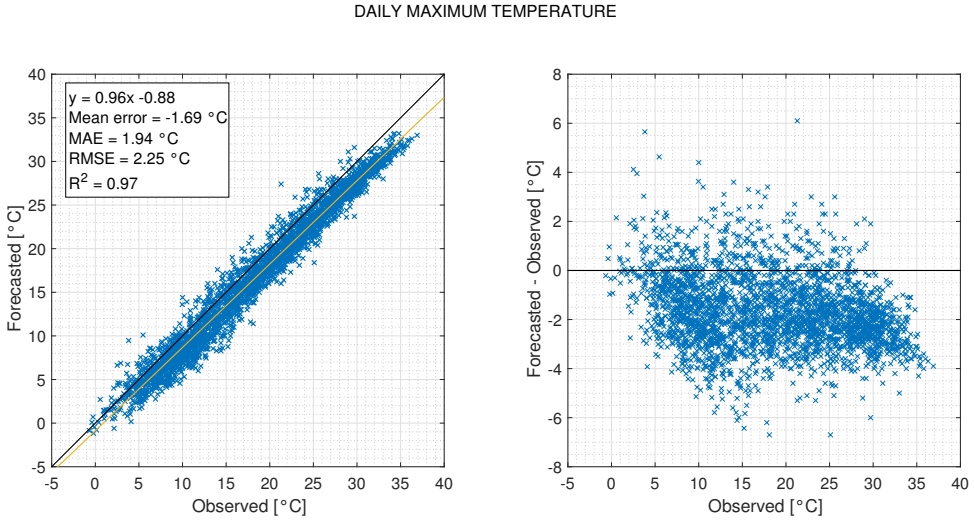


Figure 4.41: Daily maximum temperature

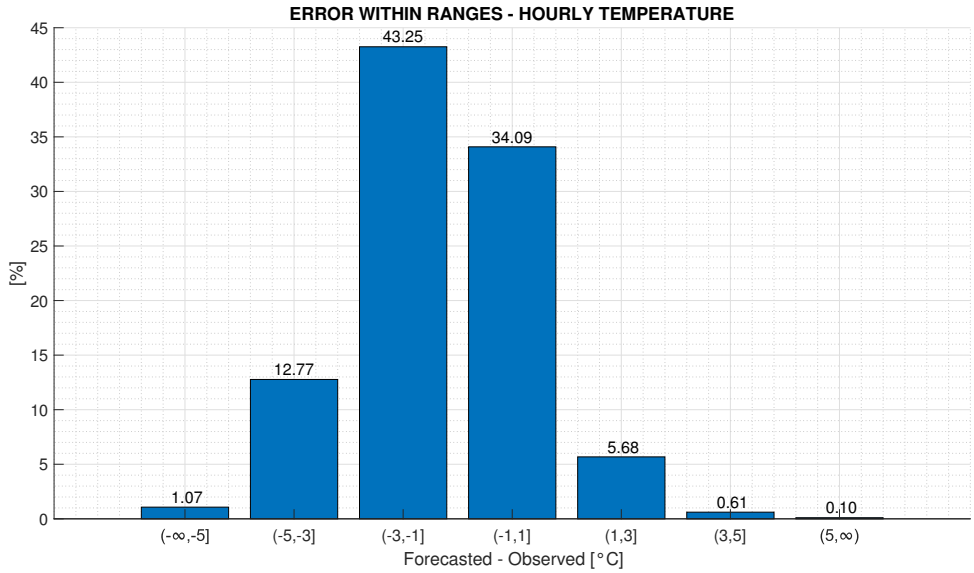


Figure 4.42: Error within ranges - Hourly temperature

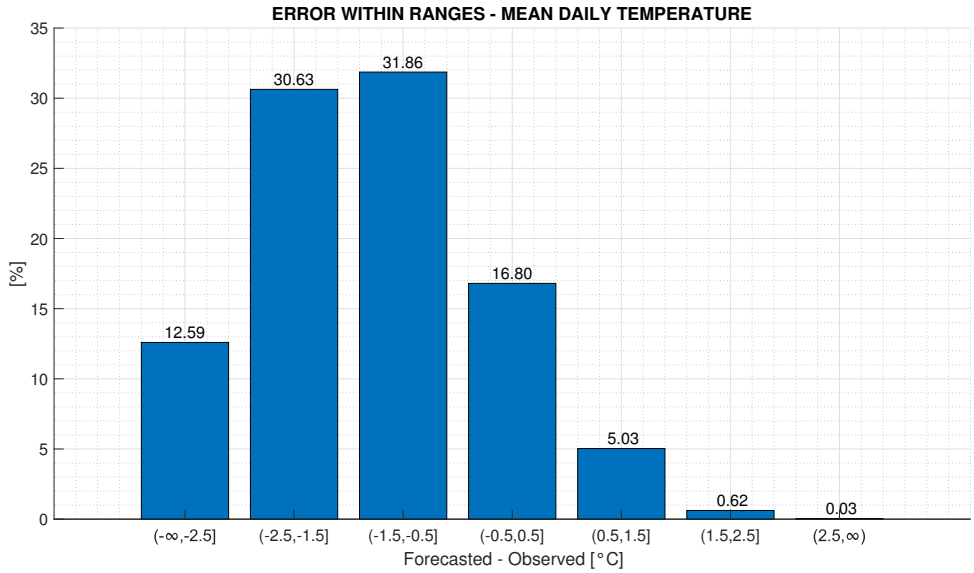


Figure 4.43: Error within ranges - Mean daily temperature

4.2.2 Solar radiation

Considerations must be taken concerning solar radiation, particularly, in the MNW dataset the solar radiation from 18:00 to 7:00 of the next day is discarded in order to homogenize the information since some measure problems were identified among the weather stations. This affects the spatialized observed information for the reason that, especially in summer, after 18:00 there is still incoming solar radiation which however is not considered. Consequently, in addition to the hydrological simulation from 2013 to 2020, two other situations were deemed: first, filtering the results of the simulation from 8:00 to 17:00 in each day, and second, a further simulation for the same period of time but considering ARPA observed dataset only. The following graphs (Fig. 4.44 to Fig. 4.46) represent the hourly, mean daily, and daily maximum solar radiation for the three above-described situations:

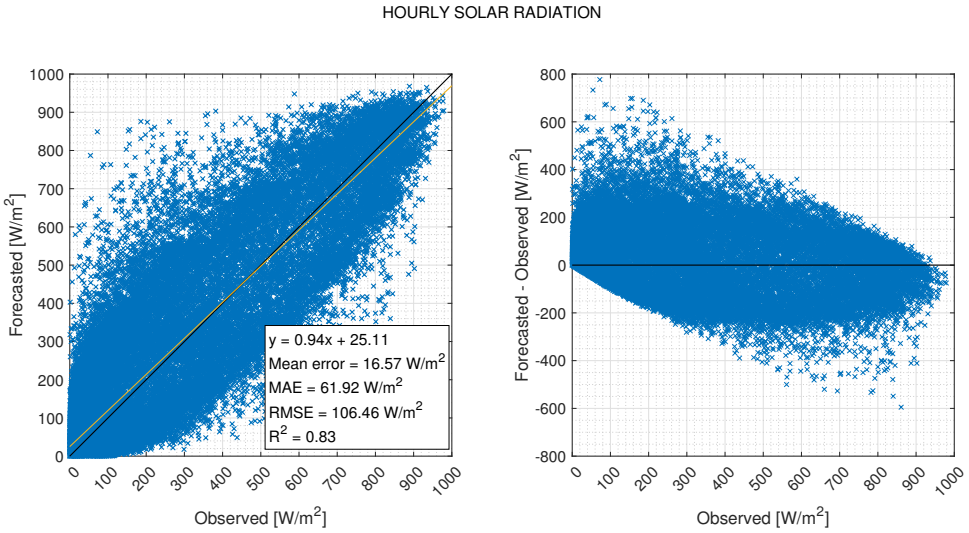


Figure 4.44: Hourly solar radiation

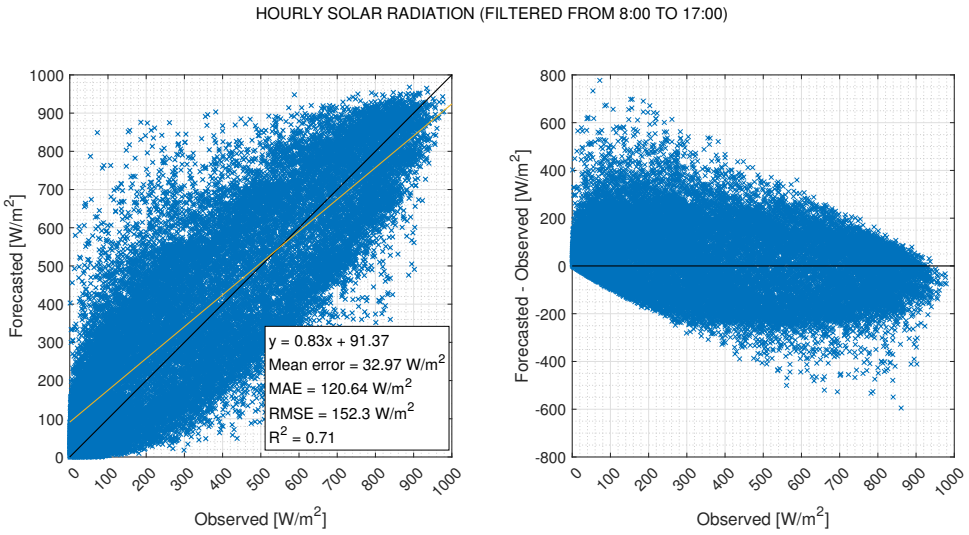


Figure 4.45: Hourly solar radiation - filtered

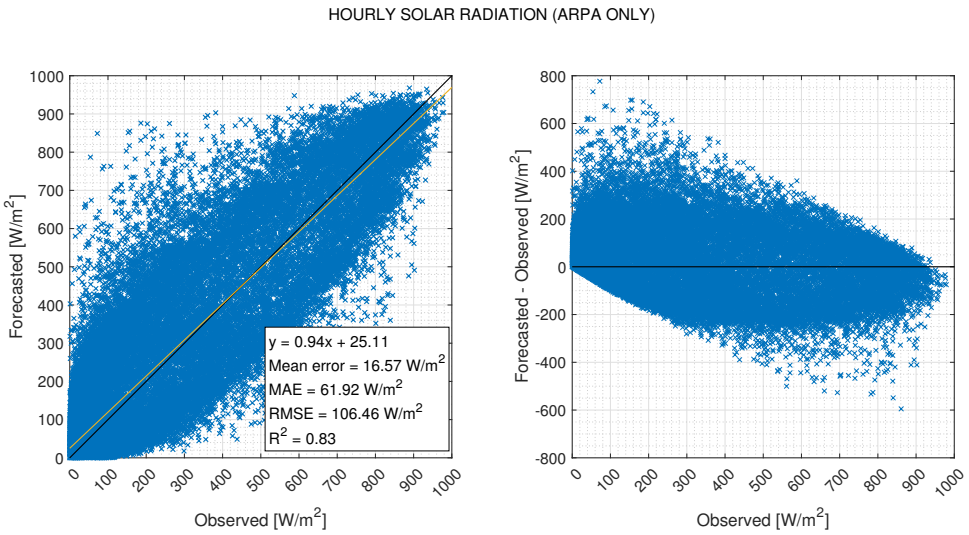


Figure 4.46: Hourly solar radiation - only ARPA

For the hourly solar radiation, observing the coefficient of determination, a good agreement between forecasted and observed data is evidenced, also, it is verified a tendency of overestimation and underestimation for low and high values, respectively. The mean error and the magnitude of the average error are small compared with the order of magnitude of the solar radiation values.

In respect of the three situations, in the filtered dataset there is a loss of correspondence since the R^2 coefficient is reduced up to 0.7. On the other hand, the simulation containing only ARPA information has the largest value of R^2 (0.83) and the lowest ME, MAE, and RMSE; the mean error can be explained due to the fact that the point in which the tendency passes from over to underestimation is found around $500\ W/m^2$ which is the central value of the data and, therefore, errors with different directions of the estimation compensate each other, the additional indexes are not too different when compared with all observed dataset simulation, meaning that – overall – the predictability and the magnitude of the average error do not improve greatly when considering only ARPA information.

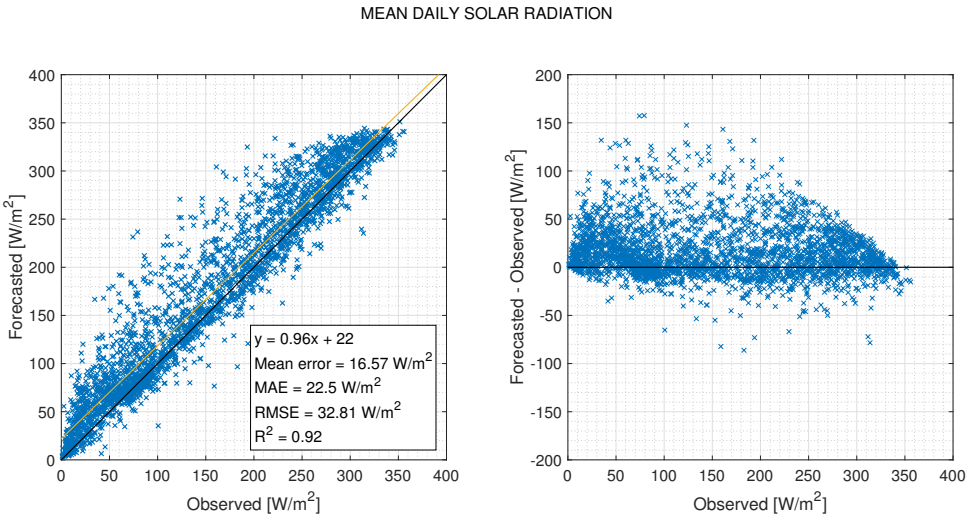


Figure 4.47: Daily mean solar radiation

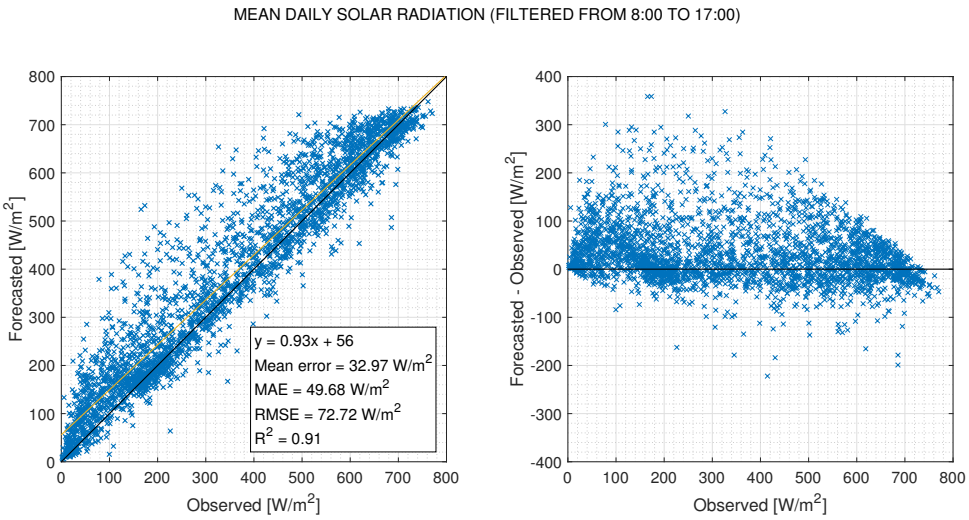


Figure 4.48: Daily mean solar radiation - filtered

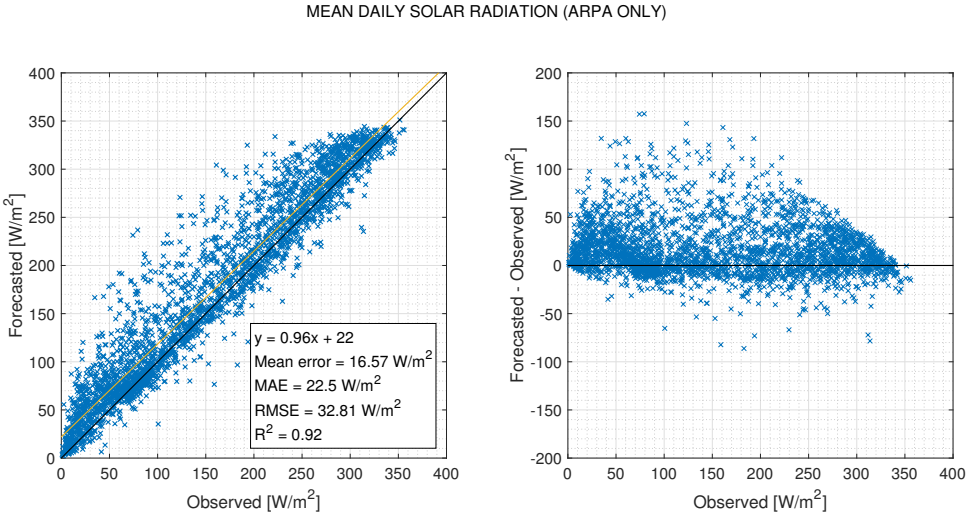


Figure 4.49: Daily mean solar radiation - only ARPA

In the case of mean daily solar radiation (Fig. 4.47 to Fig. 4.49), it is observed a good agreement with a coefficient of determination equal to 0.92 in all three cases, there is a tendency to overestimate, and the statistical indexes show low values compared with the order of magnitude of the variable. When comparing the three situations, similar comments to the hourly based case can be made: the simulation considering only ARPA information evidences lower mean error and magnitude of the mean error which, however, are not too different in comparison to the simulation forced with the entire dataset.

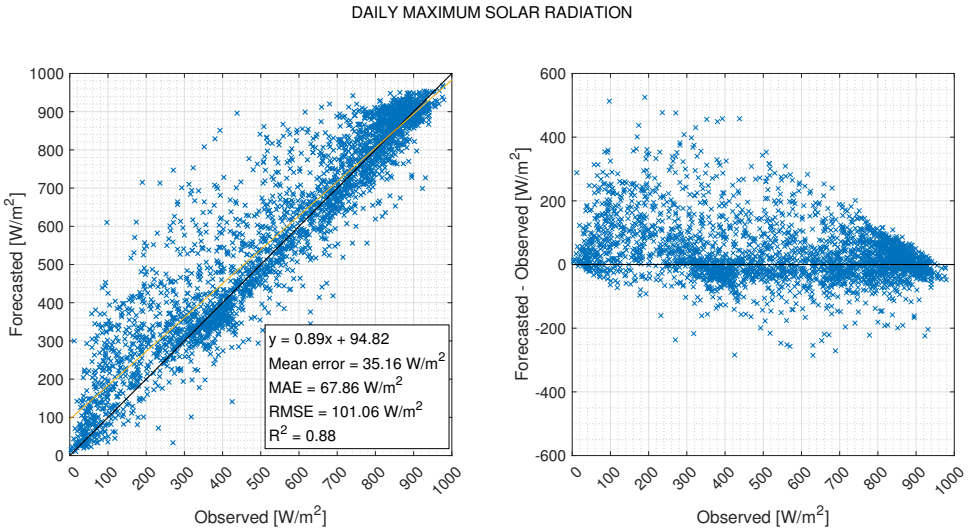


Figure 4.50: Daily maximum solar radiation

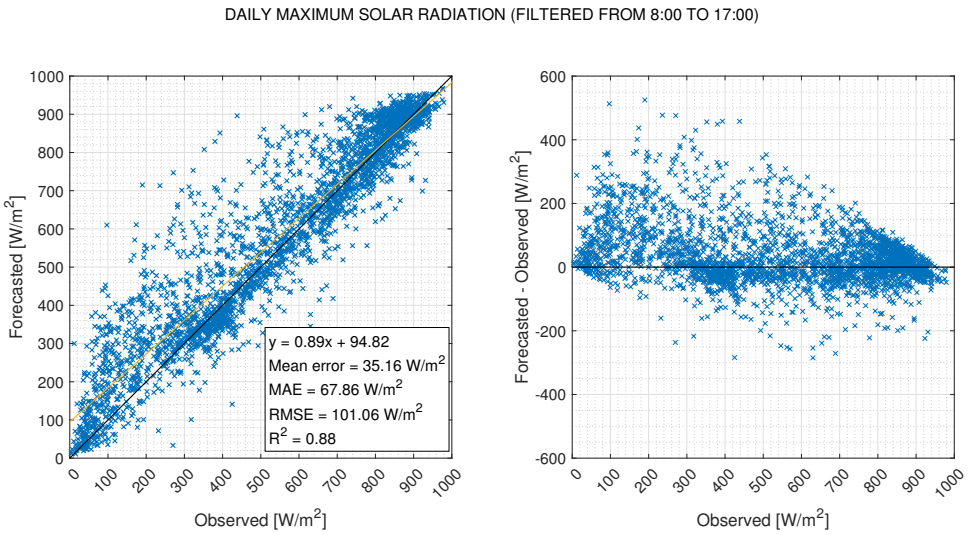


Figure 4.51: Daily maximum solar radiation - filtered

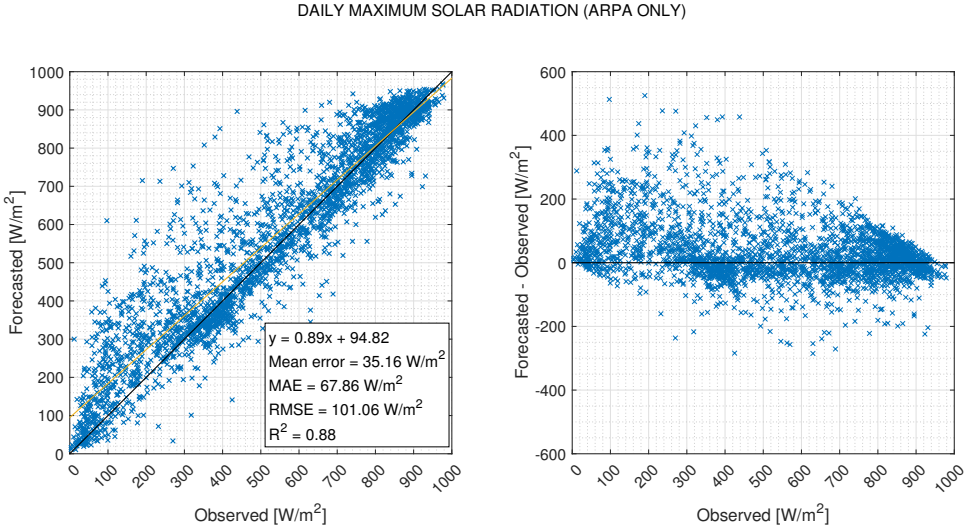


Figure 4.52: Daily maximum solar radiation - only ARPA

Concerning the daily maximum solar radiation (Fig. 4.50 to Fig. 4.52), the R^2 coefficient is around 0.89 which means a good agreement between the data, there is an overestimation trend which is reduced when increasing the values of the solar radiation where underestimation is also verified, additionally, the ME, MAE, and RMSE are low. Only ARPA simulation has lower mean error and magnitude of the average error, but the lowest coefficient of determination, whilst the other two cases are practically the same; this happens because maximum solar radiation is observed in the central hours of the day and, therefore, the filtering of the data does not influence the results.

An additional analysis that can be made is at the seasonal scale since solar radiation forecasts, especially for low values, might be affected by the presence of fog in the cold period of the year. Thus, in Fig. 4.53 and Fig. 4.54 are shown the hourly solar radiation discretized by season when using the entire dataset (from 2013 to 2020 including ARPA and MNW data):

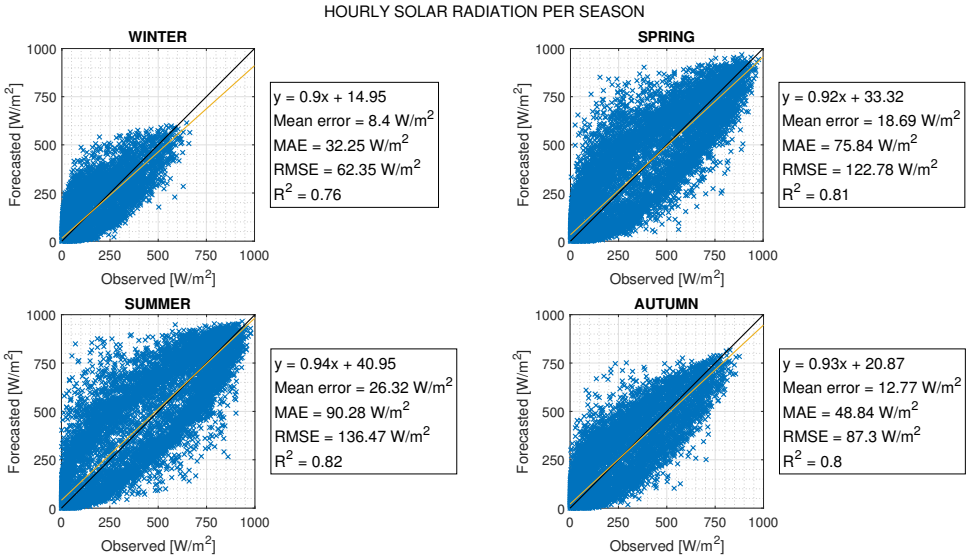


Figure 4.53: Hourly solar radiation per season: observed vs forecasted

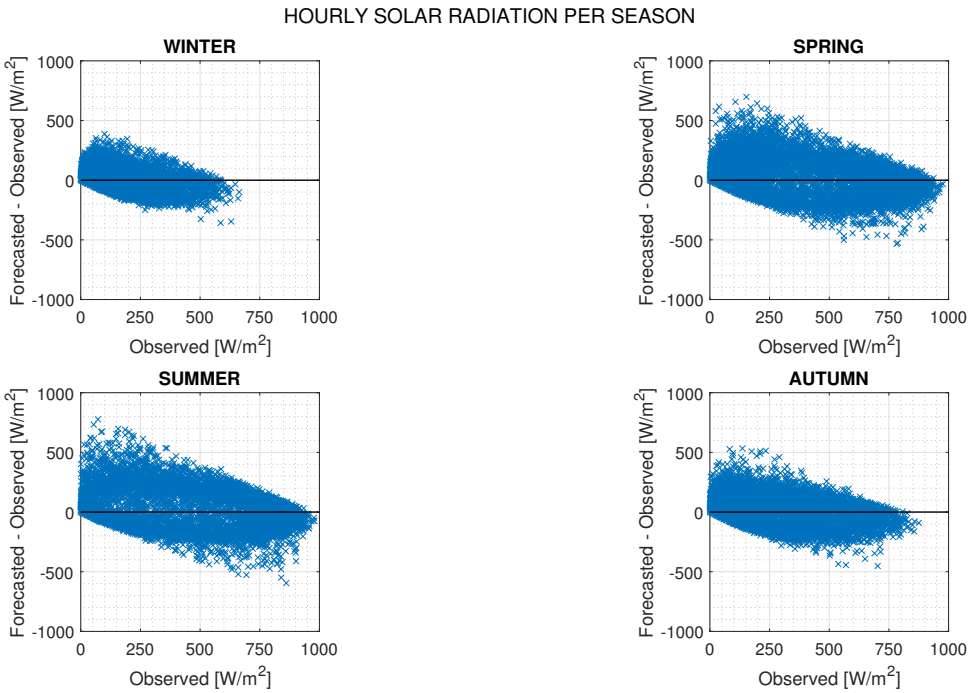


Figure 4.54: Hourly solar radiation per season: observed vs observed-forecasted

It is observed the under and overestimation trend for large and low values, respectively, for the four seasons. The ME, MAE, and RMSE are low when compared to the order of magnitude of the observed and forecasted values. Additionally, the correspondence of information is good and similar in spring, summer, and autumn, while the coefficient of determination is reduced in winter. This behaviour, as commented before, could be explained by the fact that forecasts are influenced by the presence of fog. The mean daily solar radiation are represented in Fig. 4.55 and Fig. 4.56.

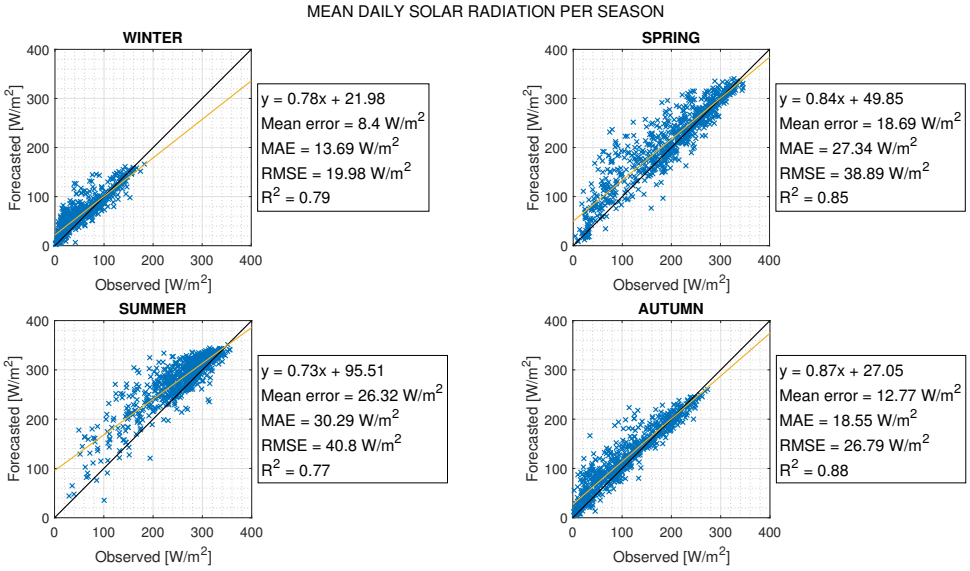


Figure 4.55: Mean daily solar radiation per season: observed vs forecasted

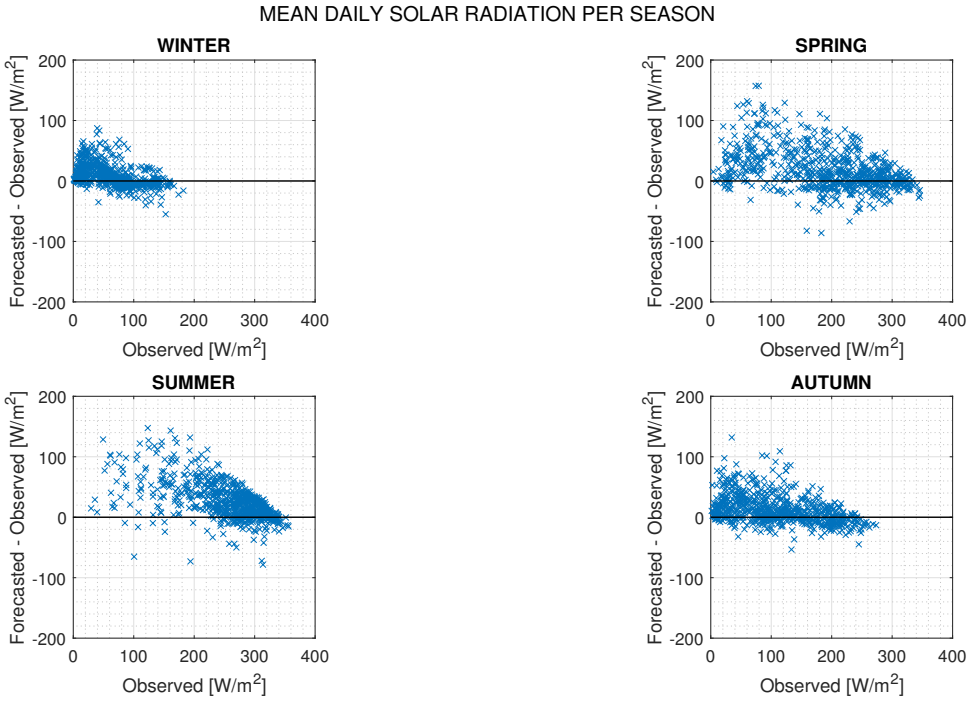


Figure 4.56: Mean daily solar radiation per season: observed vs observed-forecasted

In this case the correspondence of values in winter improves, while for summer decreases up to 0.77. Additionally, it is still clear the overestimation tendency for low observed values, which becomes more evident again in summer. On the other hand, improved behaviour is verified for spring and autumn. Finally, the daily maximum solar radiation are shown in Fig. 4.57 and Fig. 4.58.

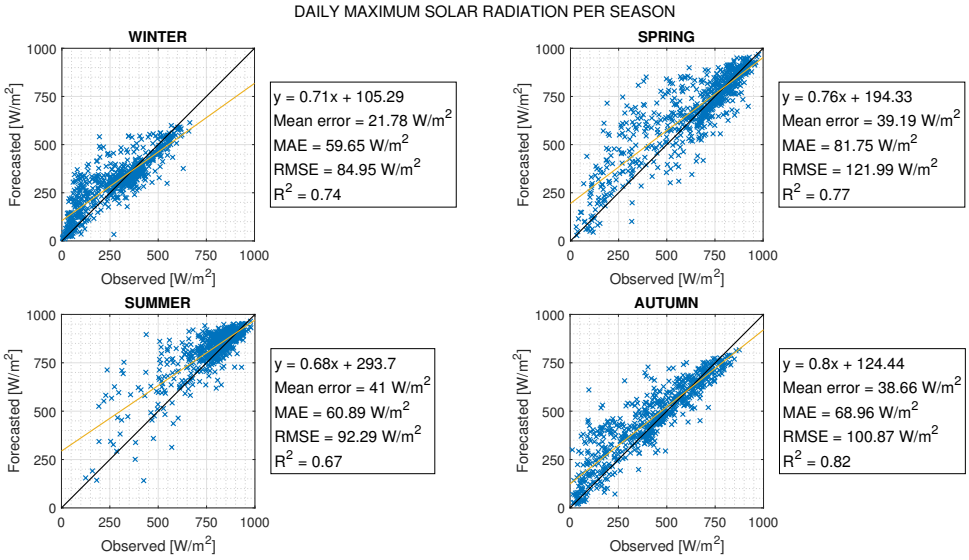


Figure 4.57: Daily maximum solar radiation per season: observed vs forecasted

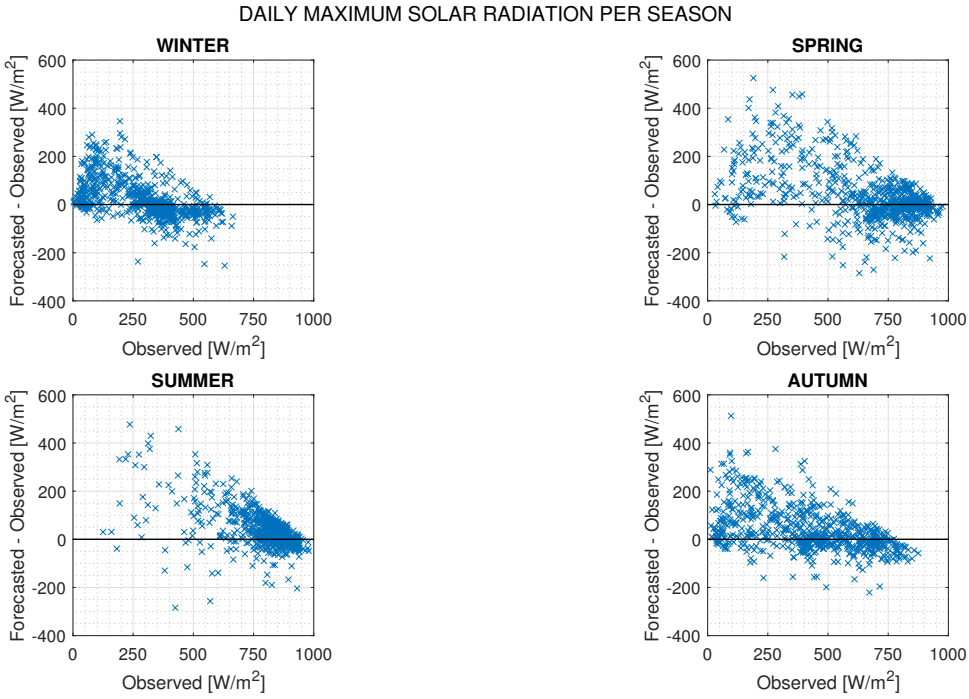


Figure 4.58: Daily maximum solar radiation per season: observed vs observed-forecasted

It is observed a decrease of the correspondence to a discreet value in the case of summer, while in winter there is a similar value of the R^2 coefficient when compared to the last two cases. Moreover, in spring a loss of correspondence is verified, while autumn remains with a good behaviour. Again, it is clear the overestimation trend when low values of incoming solar radiation are observed.

4.2.3 Relative humidity

From Fig. 4.59 to Fig. 4.62 are presented the scatter plots of hourly, mean daily, daily maximum and minimum relative humidity.

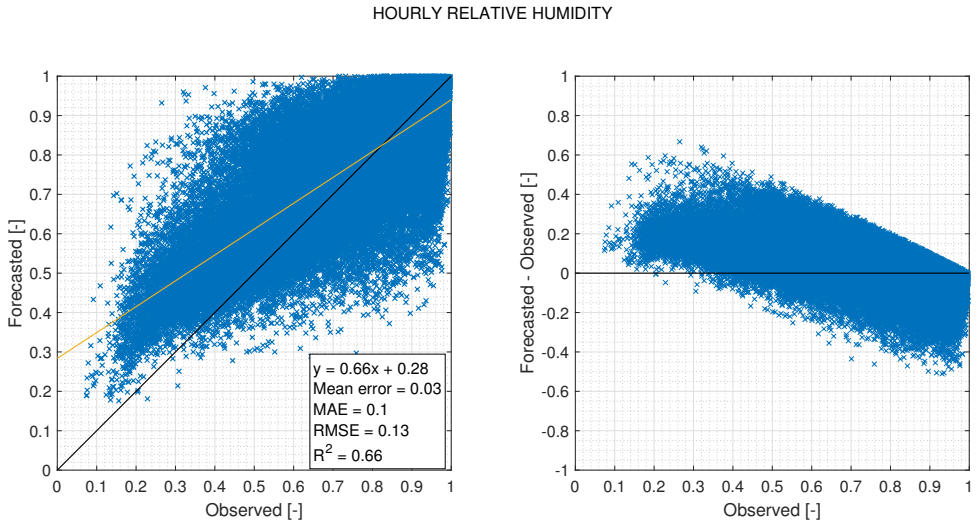


Figure 4.59: Hourly relative humidity

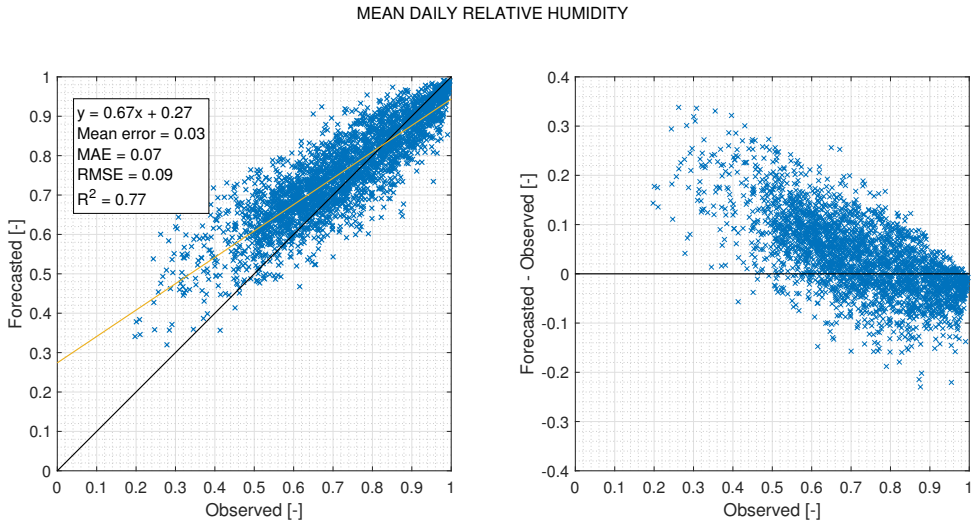


Figure 4.60: Mean daily relative humidity

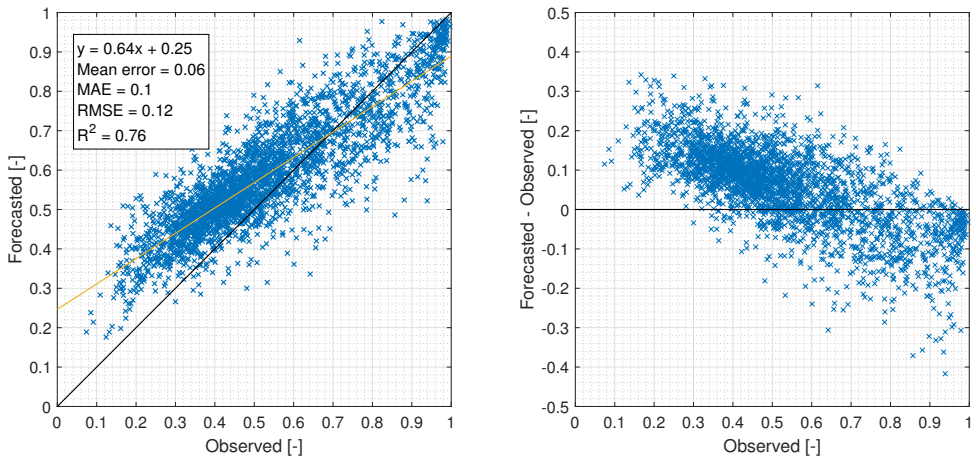


Figure 4.61: Daily minimum relative humidity

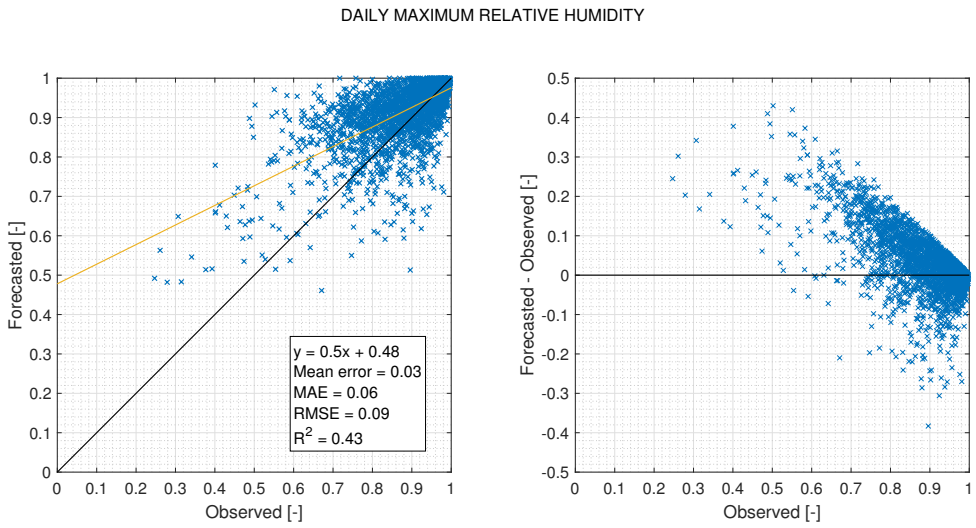


Figure 4.62: Daily maximum relative humidity

In all four cases there is a large overestimation tendency when low values of relative humidity are observed, while an underestimation trend for large values of the variable. The correspondence between observed and forecasted information is not particularly high since the coefficient of determination is 0.64 at hourly scale, then the predictability increases at a mean daily scale and in terms of daily minimum, with 0. 76 and 0.74, respectively, and finally, concerning the daily maximum the correspondence is very low with a value of 0.42. Moreover, in terms of the average error and its magnitude, it can be said that they are – overall – one order of magnitude smaller than the lowest admissible value of relative humidity (10 %).

It is important to highlight that this kind of behaviour was not expected for this variable and, therefore, it was decided to perform the same comparison when forcing the hydrological simulation with ARPA dataset only; this since, in general terms, it is a more homogeneous network of weather stations, considering that the MNW is a citizen scientists network in which the stations have different brands, and are located in different climatological zones (urban, sub-urban, rural, roof, etc.) and at different heights from the ground (often different from 2 metres). The results are presented in Fig. 4.63 to Fig. 4.66, where it is possible to observe a very small improvement of the correspondence between observed and forecasted relative humidity. Nevertheless, the great overestimation trend when low values of the variable are verified still remains, and the ME, MAE, and RMSE indexes get slightly worse than in the previous case.

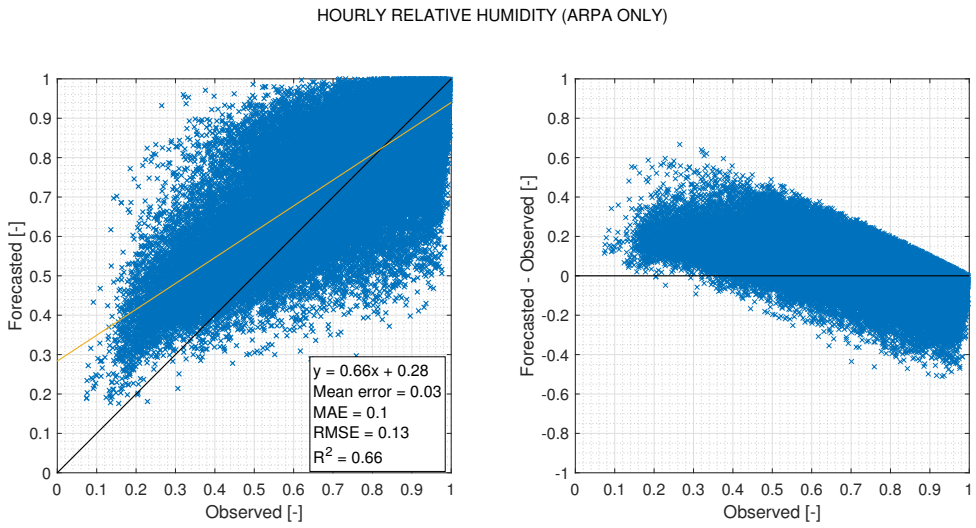


Figure 4.63: Hourly relative humidity - only ARPA

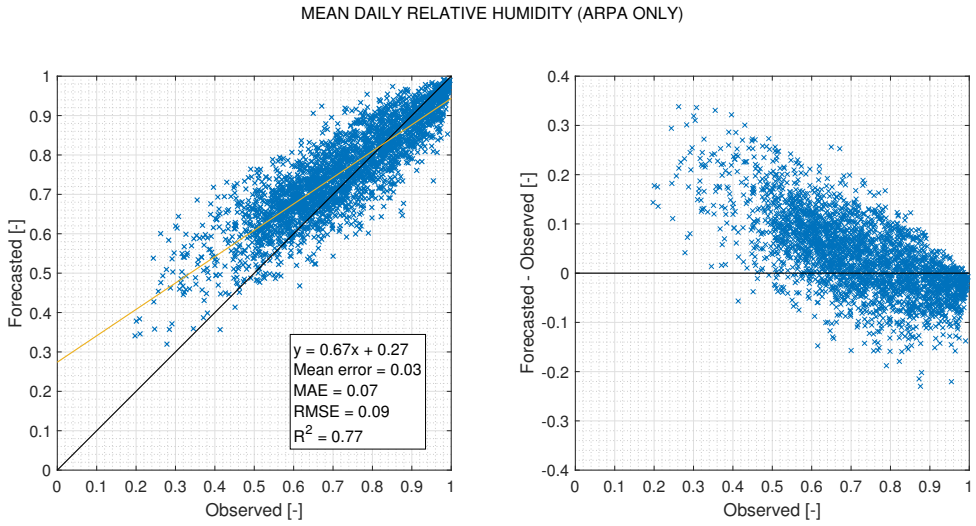


Figure 4.64: Mean daily relative humidity - only ARPA

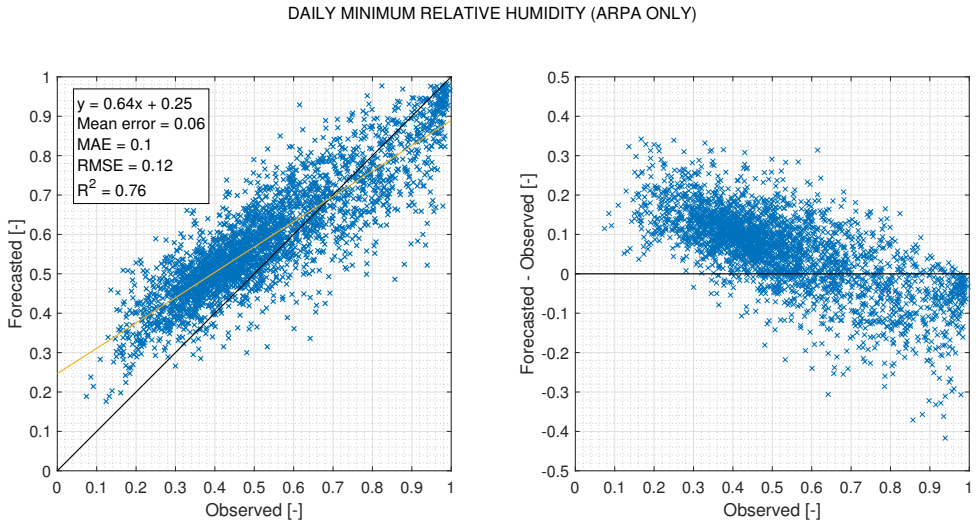


Figure 4.65: Daily minimum relative humidity - only ARPA

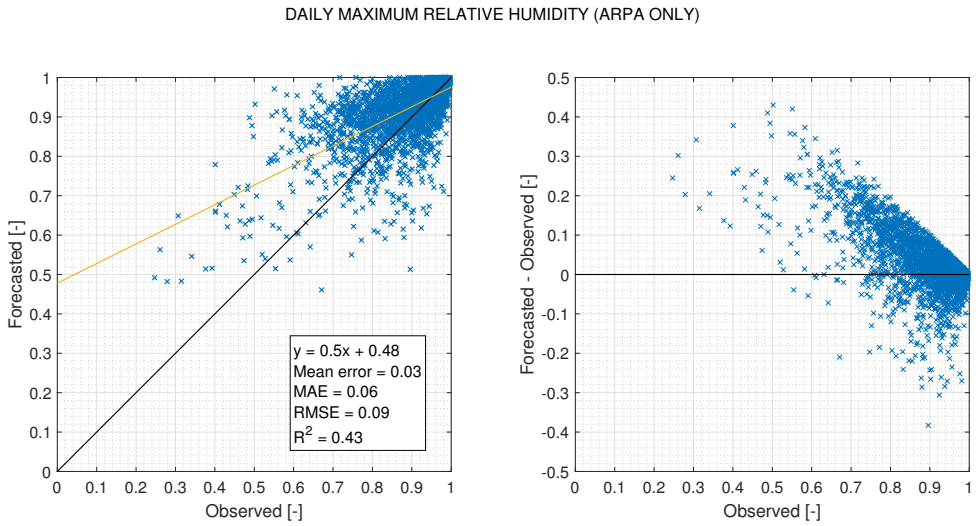


Figure 4.66: Daily maximum relative humidity - only ARPA

4.2.4 Wind speed

Scatter plots of wind speed are shown in Fig. 4.67 to Fig. 4.71. In the hourly based case there exists a very poor correspondence of the data since the R^2 coefficient is equal to 0.38, it is also observed an overestimation tendency especially for small values, which is also verified in Fig. 4.71 where a right-shifted distribution is evident, with the mean error mainly found in the range $-0.5\ m/s$ to $1.5\ m/s$. Concerning the daily mean and the daily maximum, it is again observed an overestimation trend which grows when increasing the wind speed values. There is better predictability of the data with a coefficient of determination equal to 0.56 in both cases, which is still low. Additionally, larger values of ME, MAE, and RMSE indexes are verified in the daily maximum case. Finally, over and underestimation for low and high values, respectively, are observed for the daily minimum wind speed, where there exists a very poor correspondence between observed and forecasted data. It is important to mention that weather stations acquiring wind speed information are not consistently at $10\ m$ from the ground – elevation at which the MOLOCH forecasts are reported -, and therefore, comparison of the data may not be accurate.

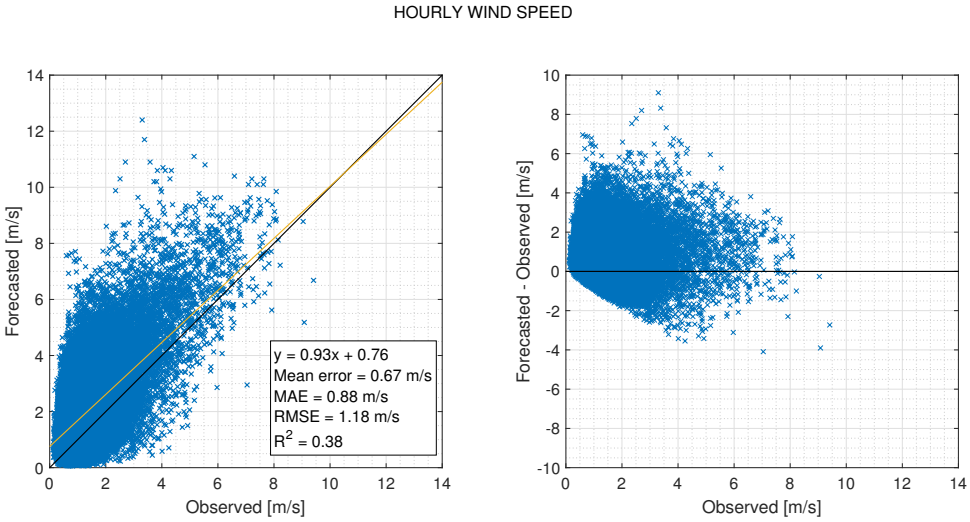


Figure 4.67: Hourly wind speed

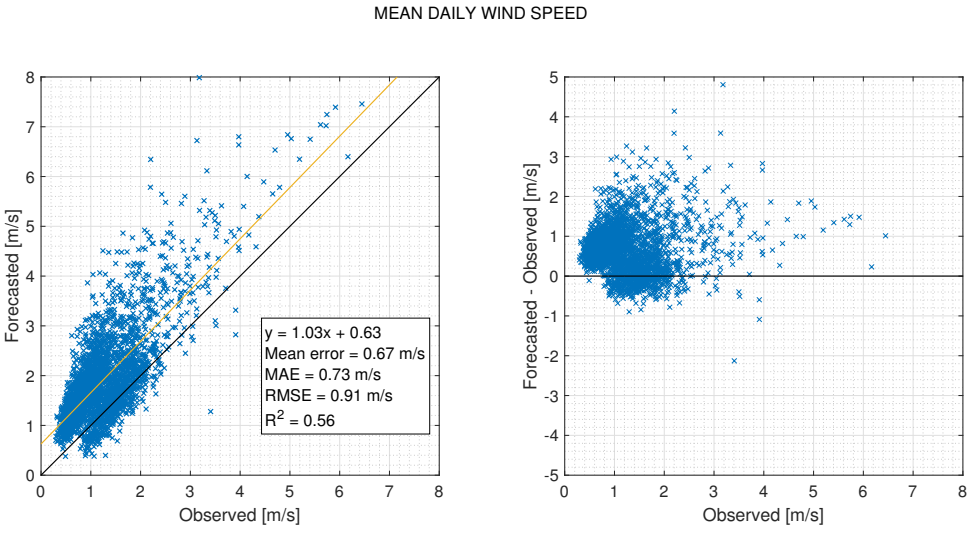


Figure 4.68: Mean daily wind speed

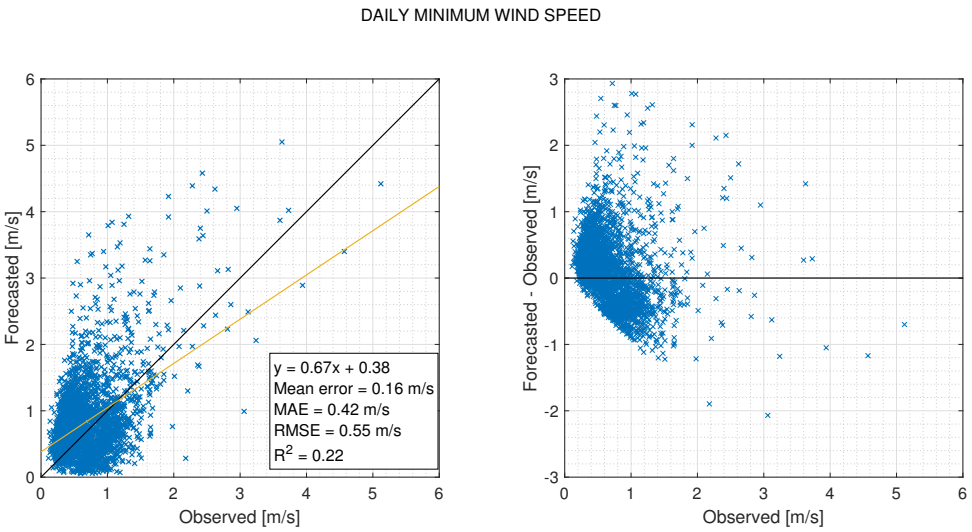


Figure 4.69: Daily minimum wind speed

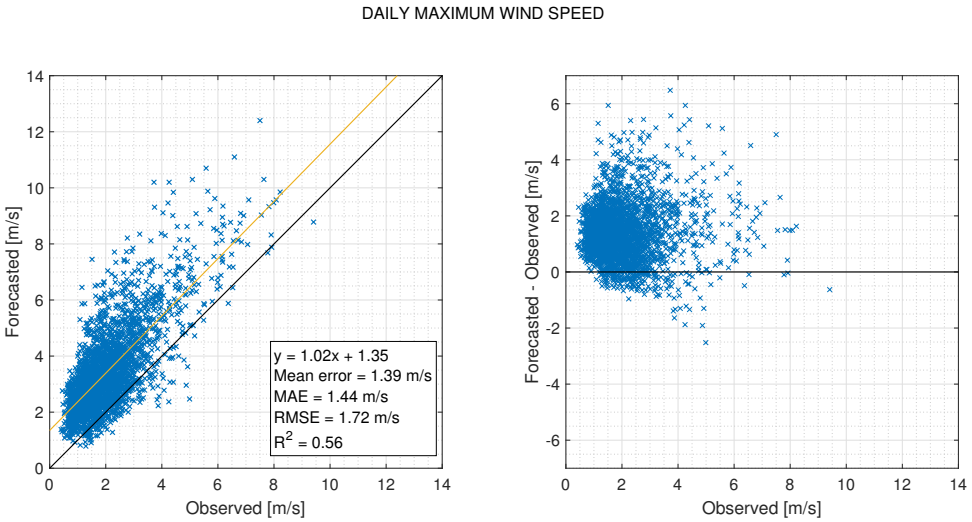


Figure 4.70: Daily maximum wind speed

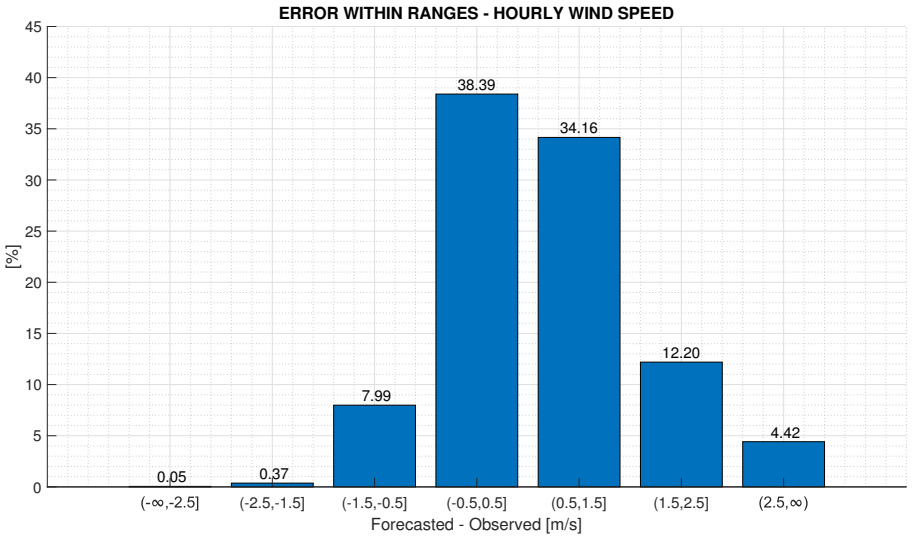


Figure 4.71: Error within ranges - Hourly wind speed

4.2.5 Precipitation

The daily and daily maximum precipitation are represented in Fig. 4.72 and Fig. 4.73, respectively.

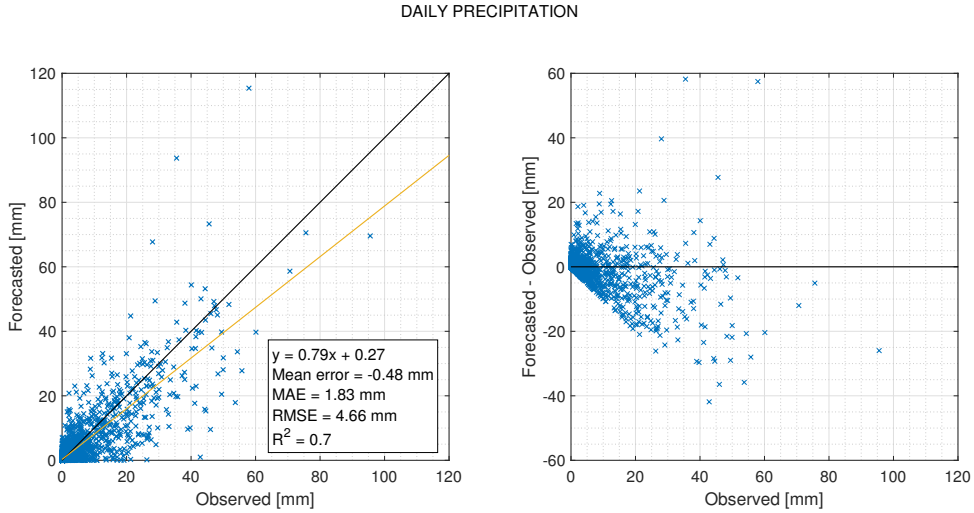


Figure 4.72: Daily precipitation

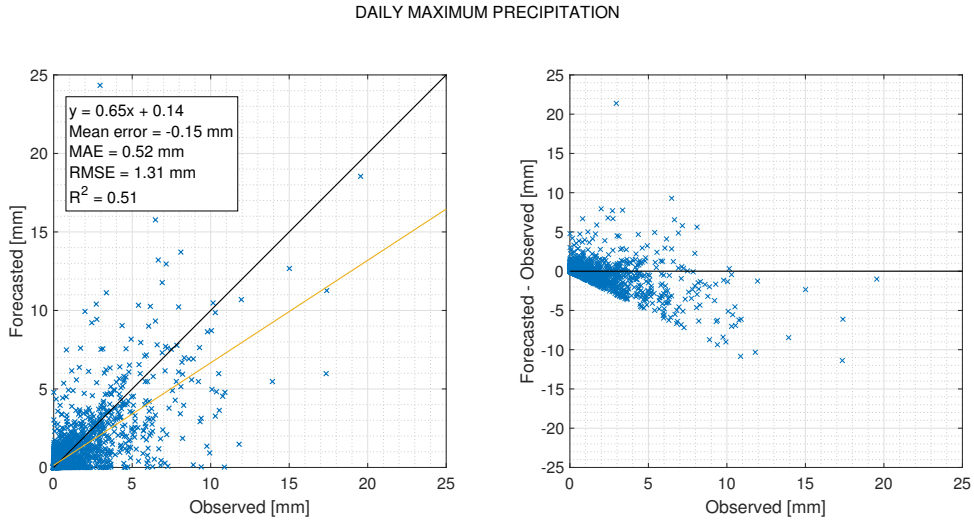


Figure 4.73: Daily maximum precipitation

In the first case there is a good correspondence ($R^2 = 0.70$) with an underestimation trend, especially for high precipitations, and an overestimation for very small values. The magnitude of the average error is less than 2 mm while the average error is near 0.5 mm. On the other hand, for the daily maximum, there is a loss in the correspondence since the coefficient of determination drops to 0.51, while the trend follows the same behaviour as the daily precipitation, it is, over and underestimation for low and high precipitations, respectively.

The underestimation tendency given by the meteorological model can also be observed in Fig. 4.74 where are represented the number of days in which a threshold has been overpassed in the case of observed and forecasted datasets. It can be seen the underestimation for the first five thresholds, and some overestimation for the last two, which however, are not common situations if we take into account that we are considering an 8-year studied period.

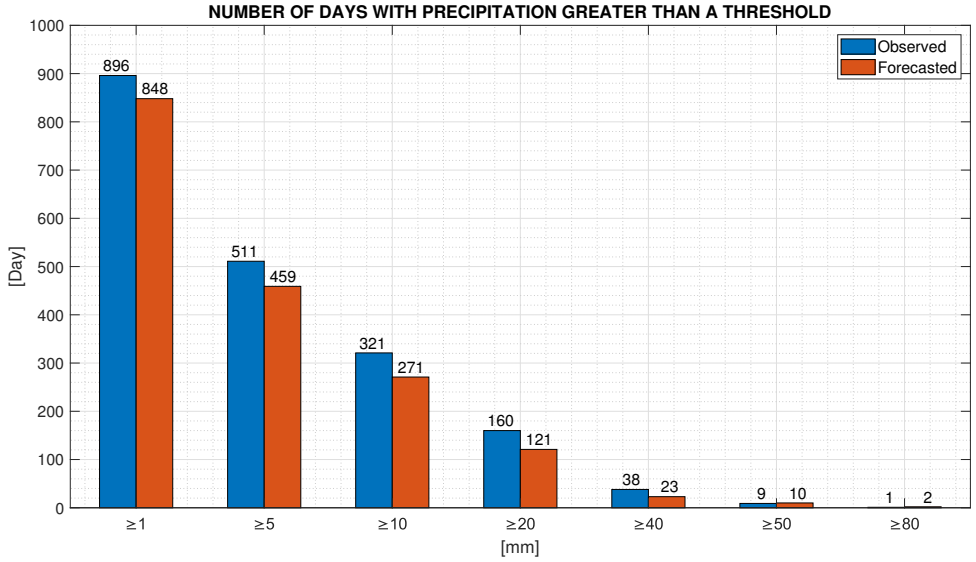


Figure 4.74: Number of days with precipitation greater than a threshold

The absolute error within ranges is shown in Fig. 4.75 where it is possible to observe that the error mainly ranges from 0 mm to 5 mm, meaning that the model do not perfectly predicts the observations, but the committed errors are – in the majority – low. This, bearing in mind that within the first range of absolute error are also the days without precipitation, which are more easily predictable and, consequently, increases the percentage of this range shown in the figure.

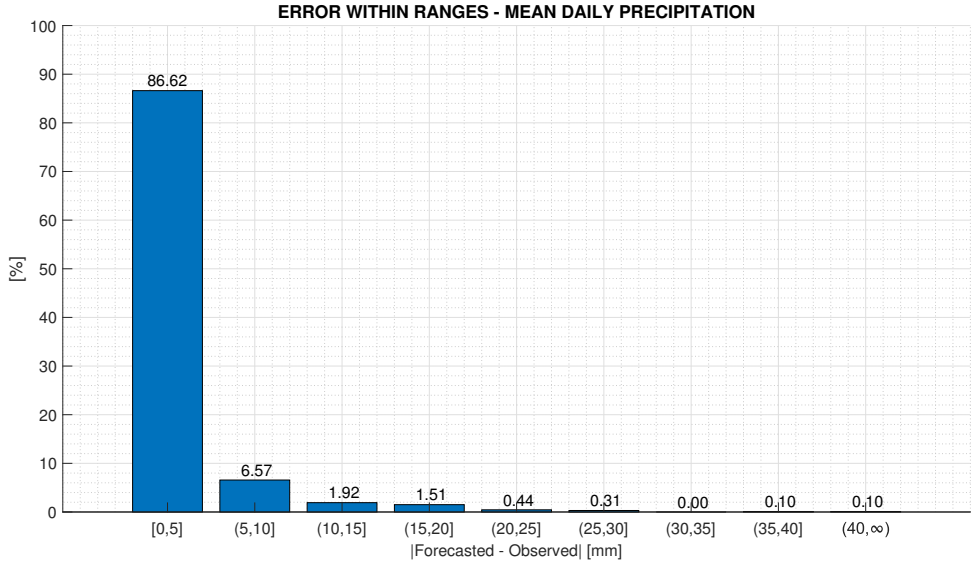


Figure 4.75: Error within ranges - Mean daily precipitation

Additionally, as stated at the beginning of the chapter, the same charts were obtained at a seasonal scale by considering winter from December to February, spring from March to May, summer from June to August, and autumn from September to November. This with the aim of verifying whether MOLOCH model behaves the same along the year or, conversely, its performance changes. Daily precipitation results are shown in Fig. 4.76 and Fig. 4.77:

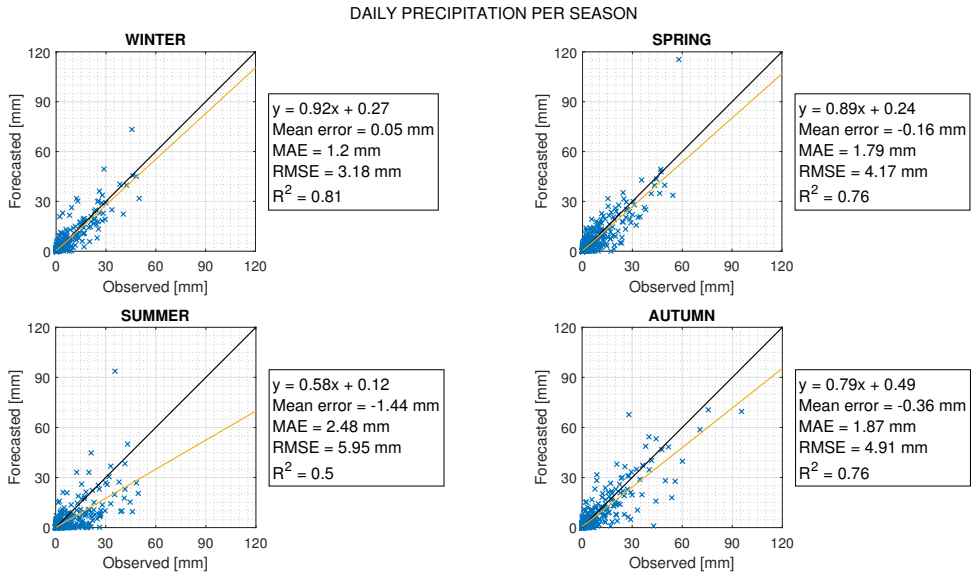


Figure 4.76: Daily precipitation per season: observed vs forecasted

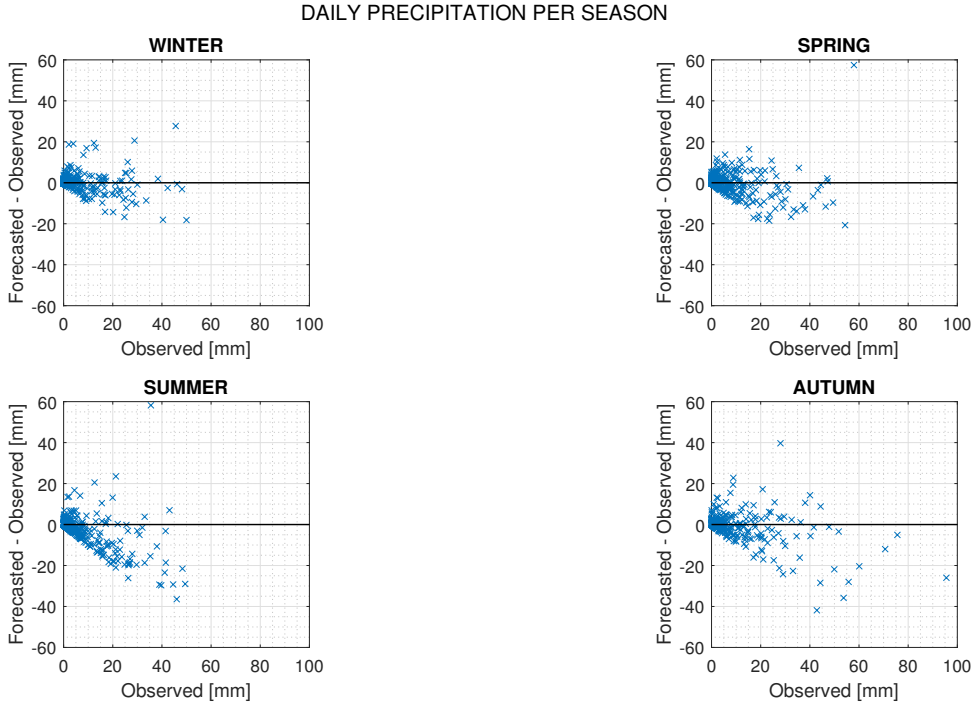


Figure 4.77: Daily precipitation per season: observed vs observed-forecasted

It is clear that the predictability is larger – when compared to the entire dataset – in winter, spring and autumn, whilst it gets worst in summer. The average error and its magnitude are better in winter and have a similar behaviour in spring and autumn, whereas in summer are found the largest values. Moreover, the underestimation tendency for large values of precipitation still remains, increasing in the following order: winter, spring, autumn, and summer, in correspondence with ME values. These results, especially in summer, can be explained by understanding the precipitation dynamics along the year described in the climatological characterization performed in the previous section. It has been demonstrated that the area of study has a bi-modal regime of precipitation with rainy periods in spring and autumn where a similar behaviour is present in the scatter plots, in winter less precipitation is verified and, consequently, it is expectable the forecast to be more accurate, and finally, in summer convective precipitations with very high intensities at the end of the day are common. The latter strongly influences the correspondence of values in summer since it is a type of phenomenon difficult to predict.

In the case of daily maximum precipitation, the behaviour is very similar but with less correspondence of the observed and forecasted values. Results are shown in Fig. 4.78 and Fig. 4.79.

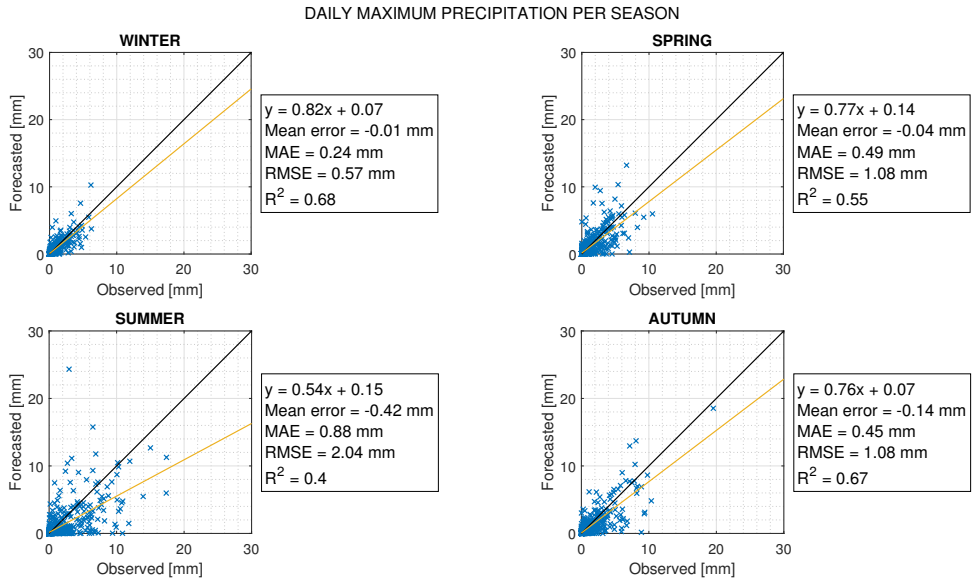


Figure 4.78: Daily maximum precipitation per season: observed vs forecasted

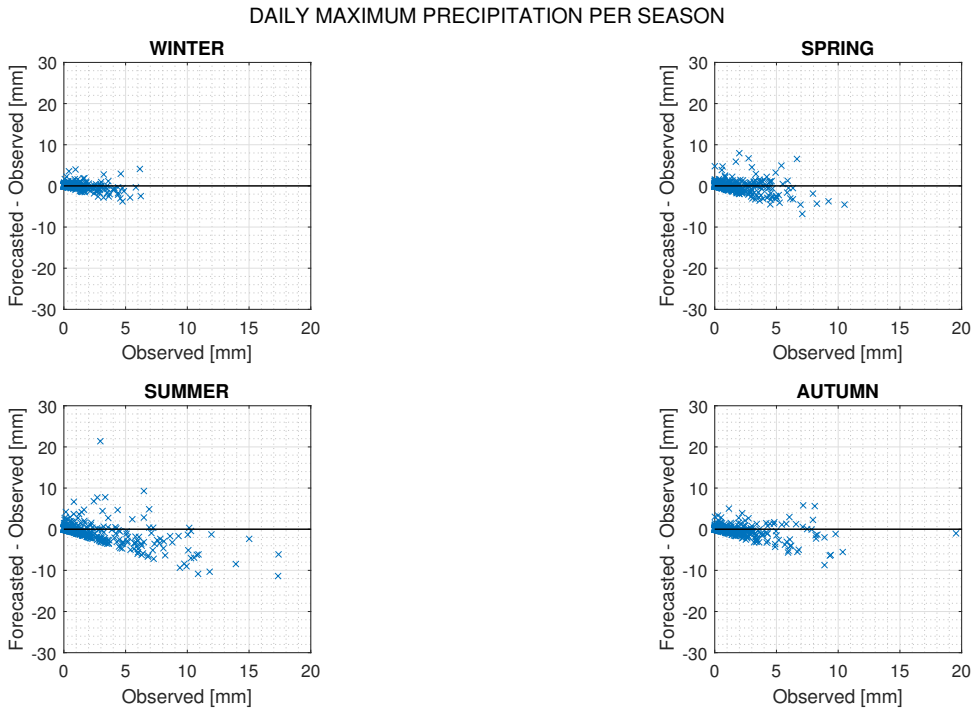


Figure 4.79: Daily maximum precipitation per season: observed vs observed-forecasted

Concerning the exceeding of a threshold (Fig. 4.80), the strong underestimation is evident in summer, whereas good accordance is verified in the other seasons with a slight overestimation in the smallest thresholds. Finally, the mean absolute error in all four seasons mainly ranges from $0\,mm$ to $5\,mm$ (Fig. 4.81) and same comments as in the entire dataset apply.

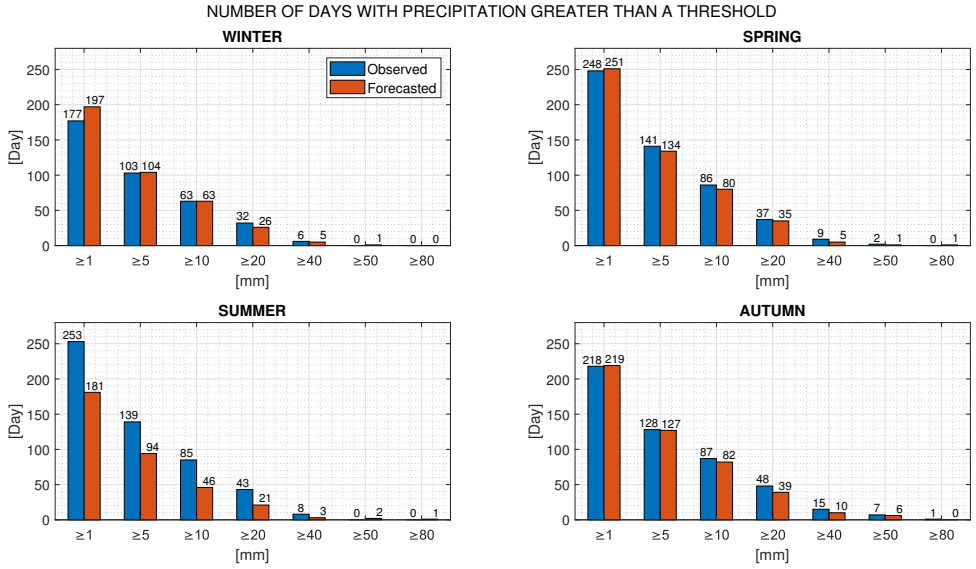


Figure 4.80: Number of days with precipitation greater than a threshold per season

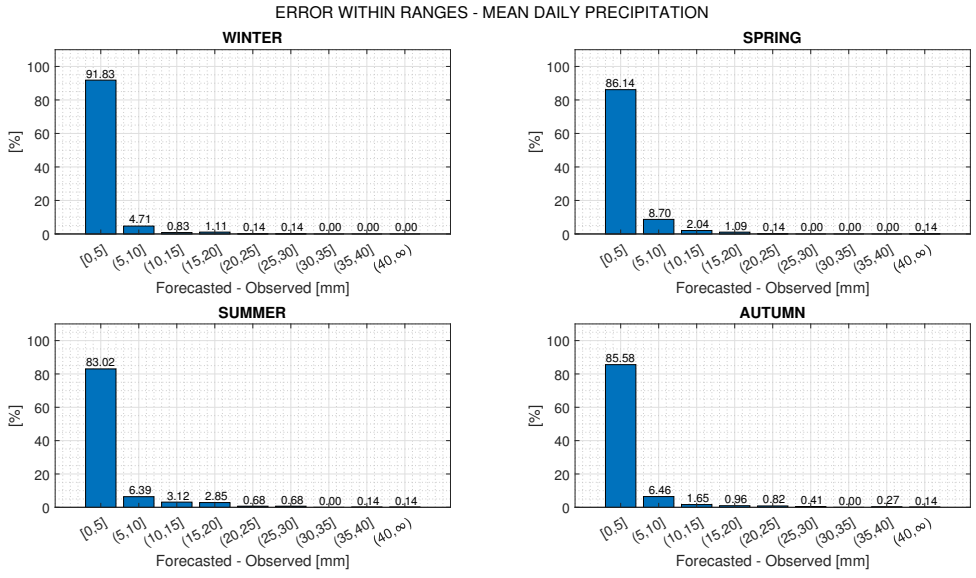


Figure 4.81: Error within ranges - Mean daily precipitation per season

4.2.6 Soil moisture

In respect of the simulated hydrological variables, soil moisture comparison has been also performed at hourly and daily scale, including daily maximum and minimum (Fig. 4.82 to Fig. 4.85). This, bearing in mind that comparison is assessed between simulated data from forecasted and observed information, and not between simulated data from observed information and the direct MOLOCH soil moisture forecast. It can be seen that the coefficient of determination indicates a good agreement of the data and is practically the same in all four cases, the same happens for ME, MAE, and RMSE indexes. It is also evident that there is a consistent underestimation trend which may be explained by the underestimation of precipitation that MOLOCH models presents. Nevertheless, the reasons why this behaviour is verified are illustrated in the performed sensitivity analysis. Furthermore, it must be taken into account that in the present study are compared the simulated soil moisture when the hydrological model is forced with observed and forecasted data. This, in spite of the fact that MOLOCH model performs soil moisture forecasts too.

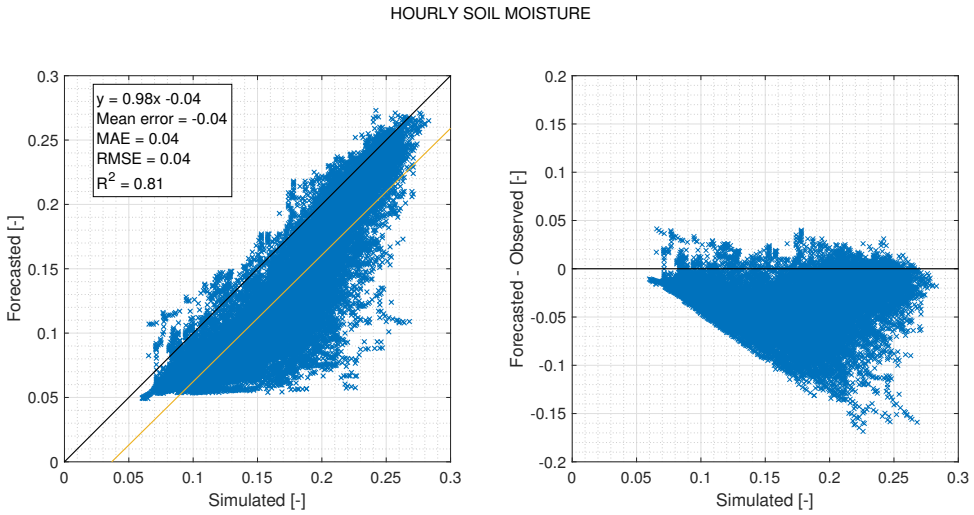


Figure 4.82: Hourly soil moisture

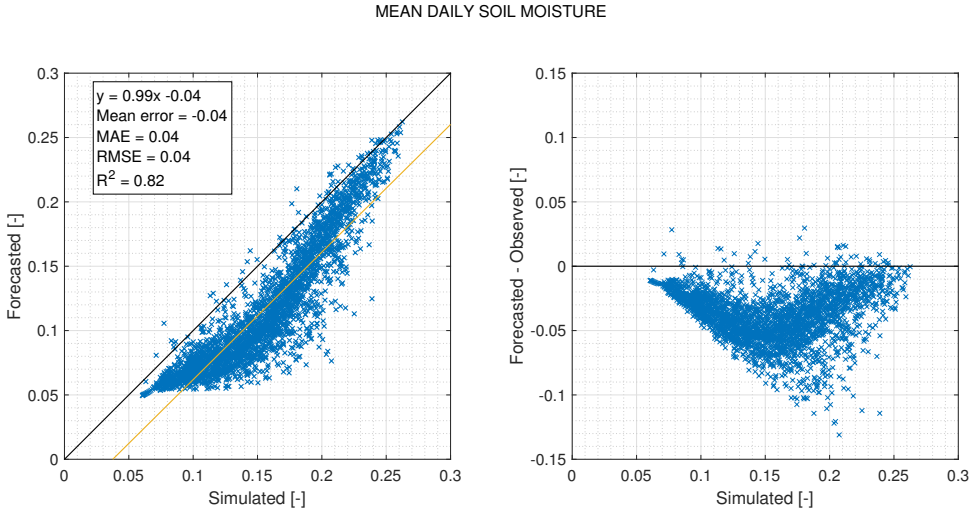


Figure 4.83: Mean daily soil moisture

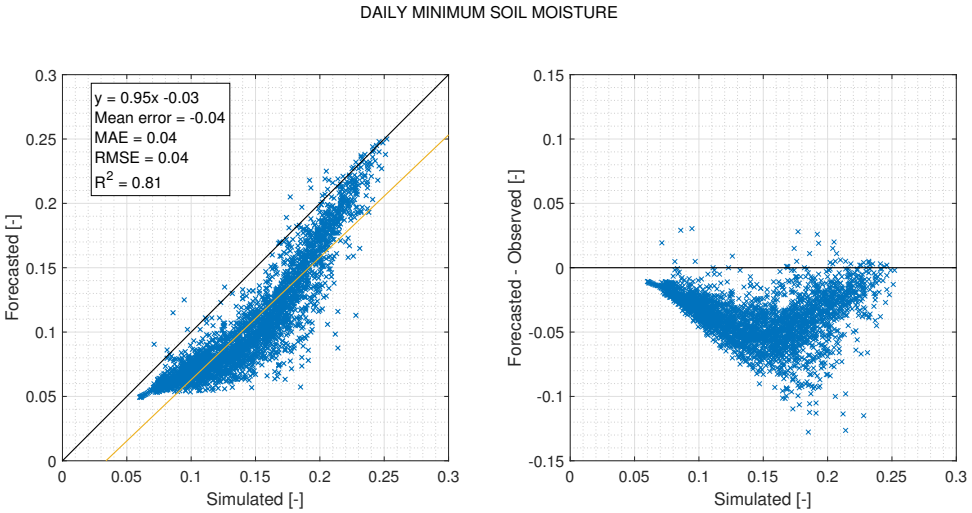


Figure 4.84: Daily minimum soil moisture

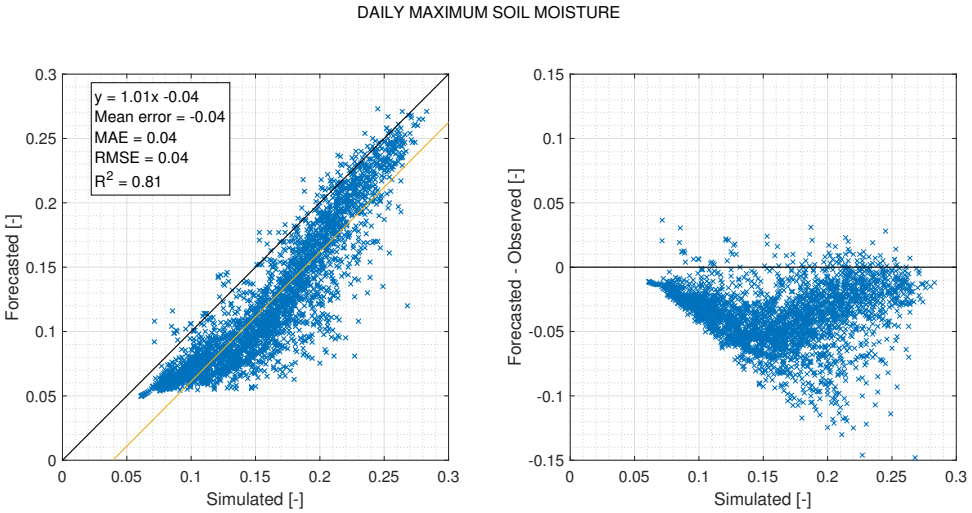


Figure 4.85: Daily maximum soil moisture

4.2.7 Potential evapotranspiration

In the case of potential evapotranspiration, the daily and daily maximum where considered (Fig. 4.86 and Fig. 4.87). It is clear that a strong overestimation by part of the forecast is verified, whilst a good agreement with the linear regression is observed (R^2 larger or equal than 0.9 in both cases). Indeed, the statistical indexes are large when compared with the order of magnitude of the potential evapotranspiration observed values – represented in the axis of the graphs. Further comments are given in the sensitivity analysis.

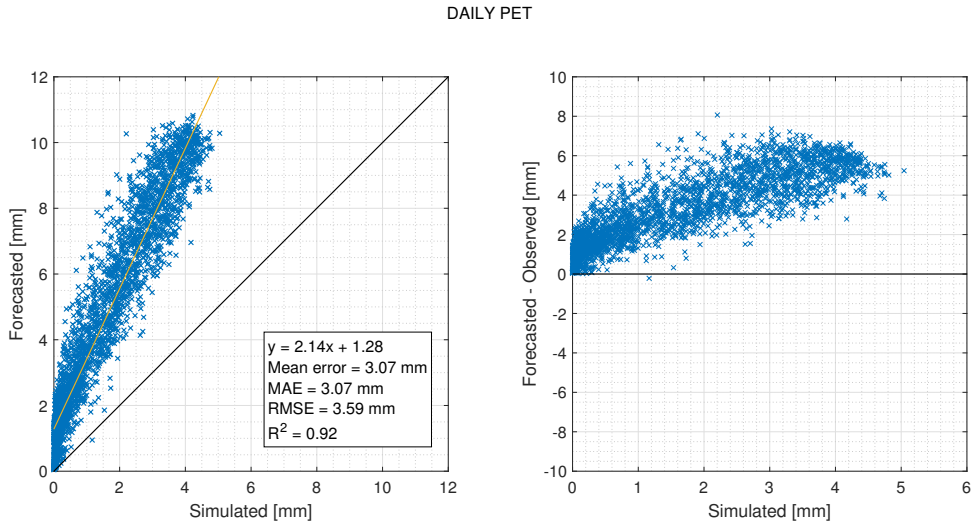


Figure 4.86: Daily potential evapotranspiration

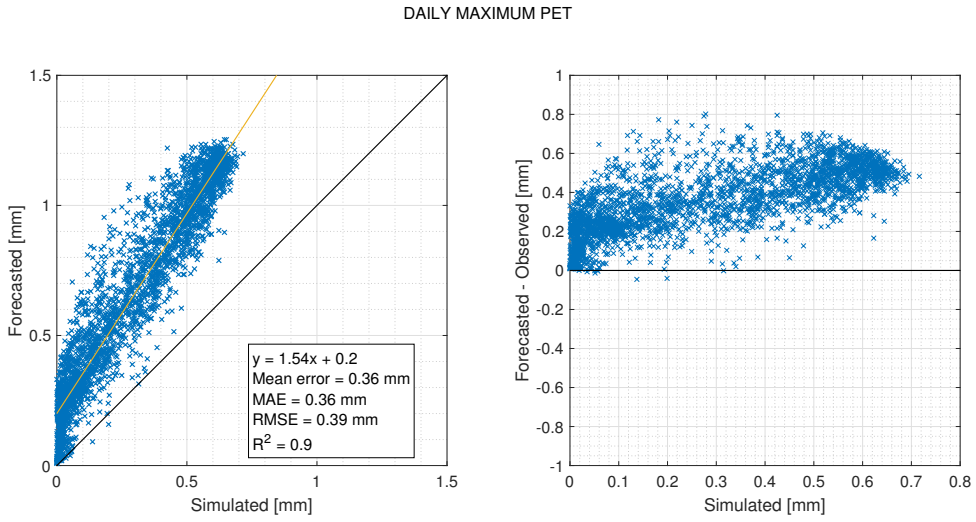


Figure 4.87: Daily maximum potential evapotranspiration

4.2.8 Discharge

The discharge at Bovisio station has also been considered. The hourly, mean daily and daily maximum are presented in Fig. 4.88 to Fig. 4.90. It can be seen that the correspondence of the data is sufficient at hourly scale and in terms of daily maximum, while it increases to a discreet value in the mean daily case. The related error indexes are similar at hourly and daily scales, but they increase in magnitude when considering the maximum values. Additionally, there is an underestimation trend especially for large discharges. This leads to big differences between the observation and forecasts, described – particularly - by the large RMSE in the daily maximum chart. The latter may be explained by the fact that flowrate is strongly conditioned by the amount of precipitation which – as already seen – is also underestimated for large values (a more extensive analysis is given in the sensitivity analysis).

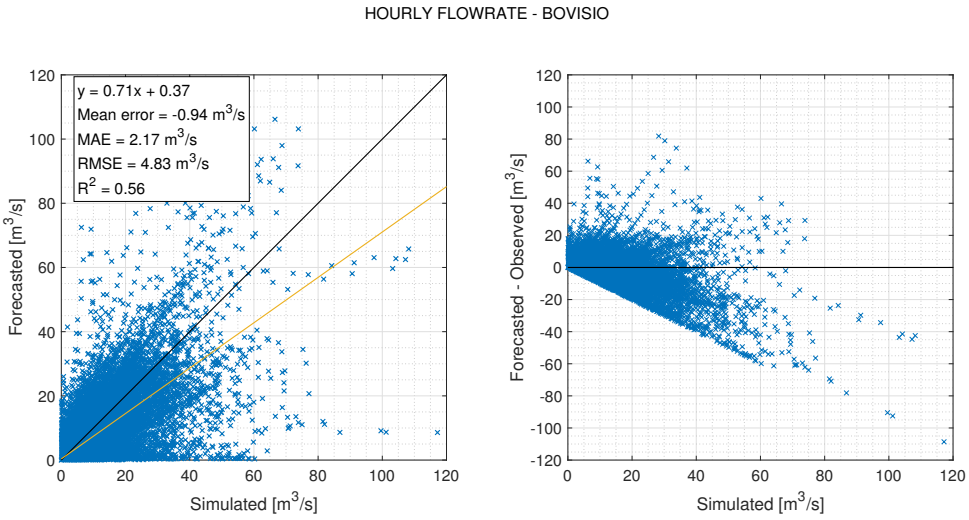


Figure 4.88: Hourly flowrate

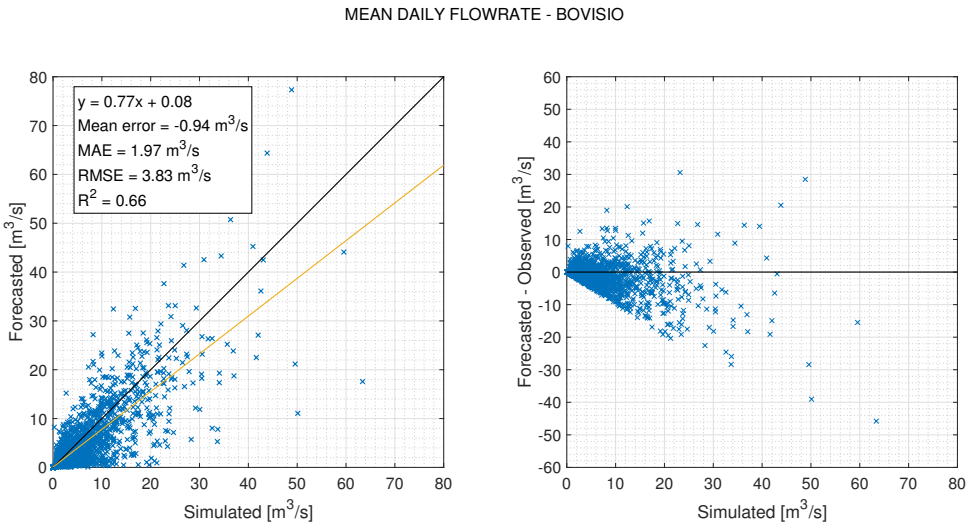


Figure 4.89: Mean daily flowrate

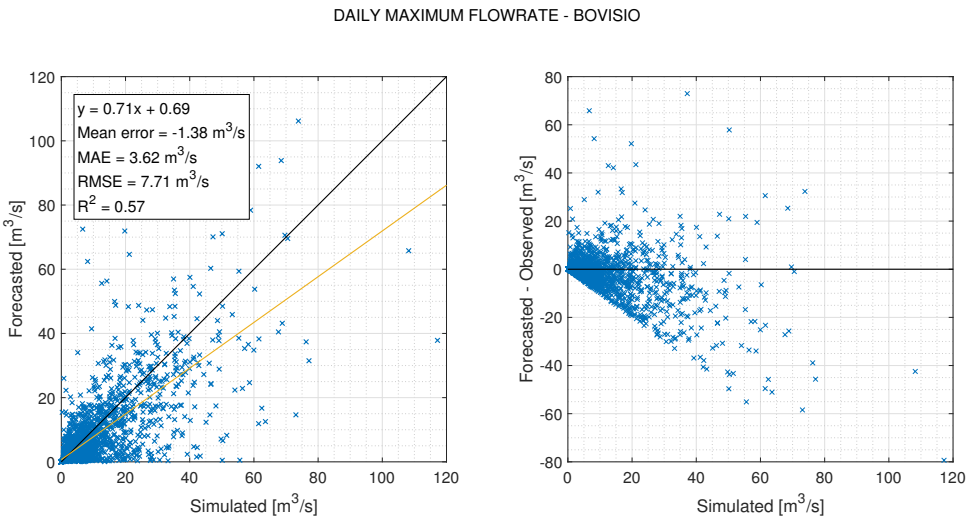


Figure 4.90: Daily maximum flowrate

Concerning the overpassing of a threshold, in Fig. 4.91 are presented the number of days in which the flowrate is greater or equal than the discharges related to usual alert values. It can be seen the underestimation of the flowrate in all three thresholds, with the reduction of the number of days which is coherent with the frequency with which discharges of that magnitude are verified. Additionally, it is not observed a particular trend of the difference between observed and forecasted number of days when increasing the threshold, this since the forecasts represent 78.2 %, 67.6 %, and 76.0 % of the observed values for each threshold in ascendent order.

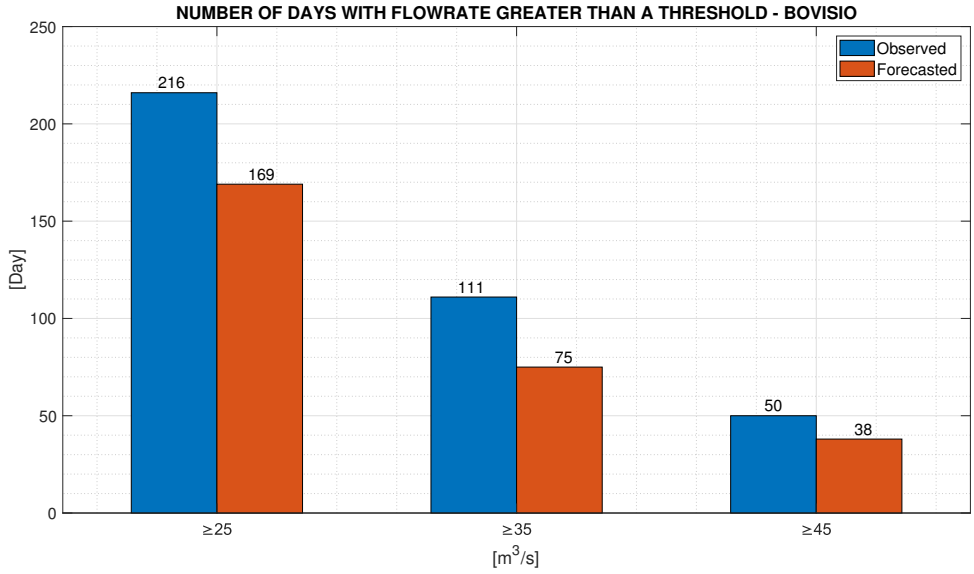


Figure 4.91: Number of days with flowrate greater than a threshold

Flowrates greater than 25 m^3/s were also considered (Fig. 4.92), nevertheless, it is verified a very bad agreement of the data and very high errors. Hence, it is again shown that correspondence of data is low when considering large discharges and an underestimation tendency is observed.

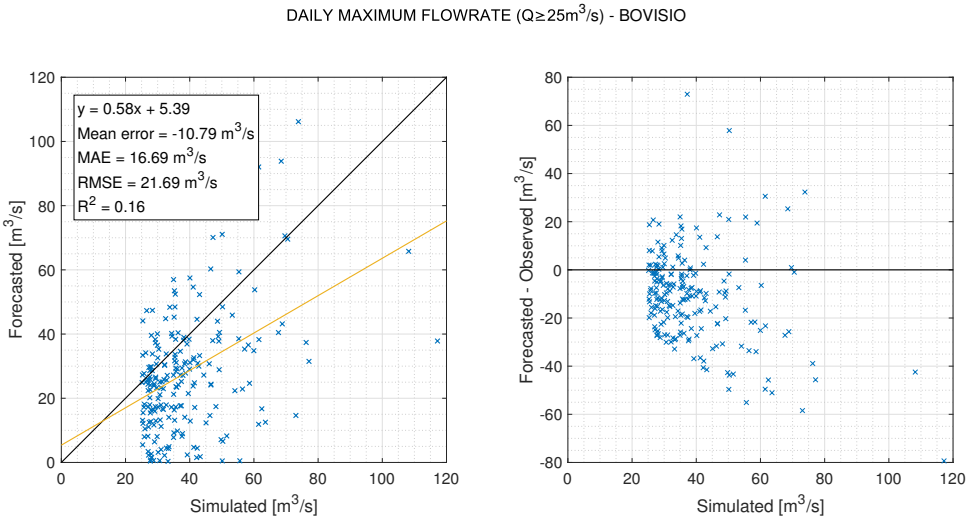


Figure 4.92: Daily maximum flowrate ($Q \geq 25\, m^3/s$)

Furthermore, plots were obtained at a seasonal scale (Fig. 4.93 to Fig. 4.99) where a behaviour similar to the one of precipitation is observed. Thus, the predictability of flowrate in winter is better than in the other seasons, followed by spring and autumn – with similar coefficients of determination -, and finally summer where the agreement is very low. Additionally, the underestimation tendency for hourly, mean daily, and daily maximum remains in all seasons, and – in general – errors are lower in winter – characterized for low precipitation heights - and spring – rainy period with less precipitation -, followed by spring – rainy period with larger precipitation -, and finally summer where – as already stated – convective precipitation phenomena affect the forecasts. Finally, concerning the surpassing of thresholds, there exists underestimation in all seasons, especially in summer.

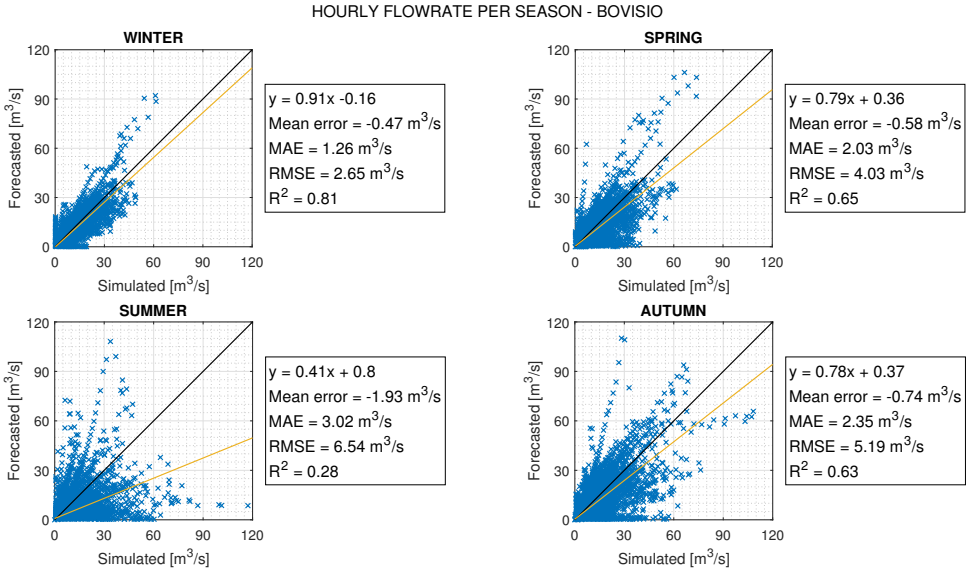


Figure 4.93: Hourly flowrate per season: simulated vs forecasted

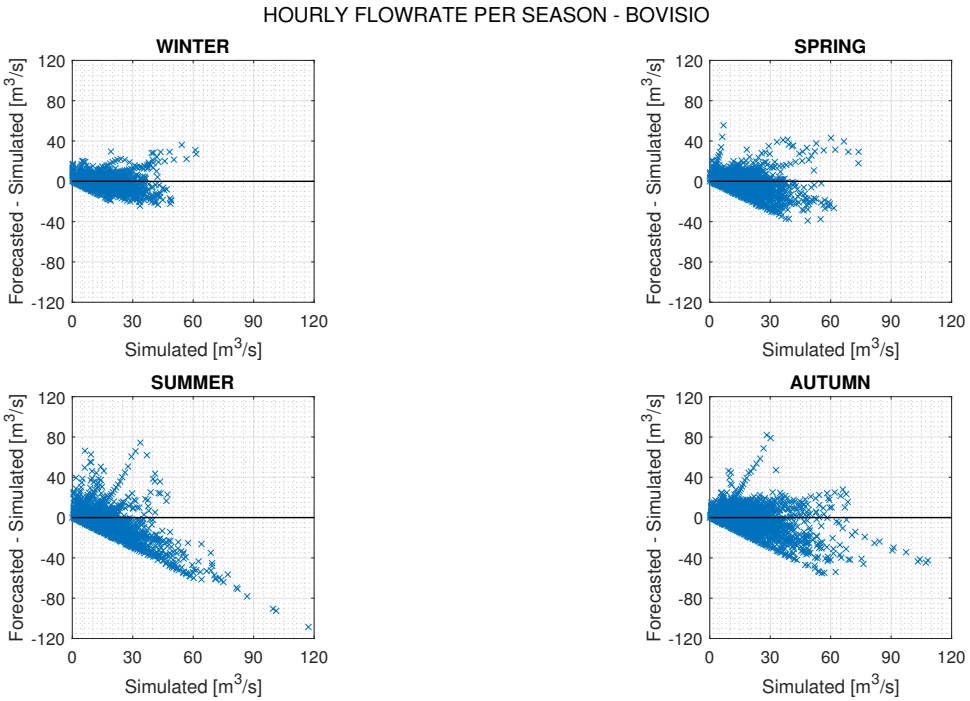


Figure 4.94: Hourly flowrate per season: simulated vs forecasted-simulated

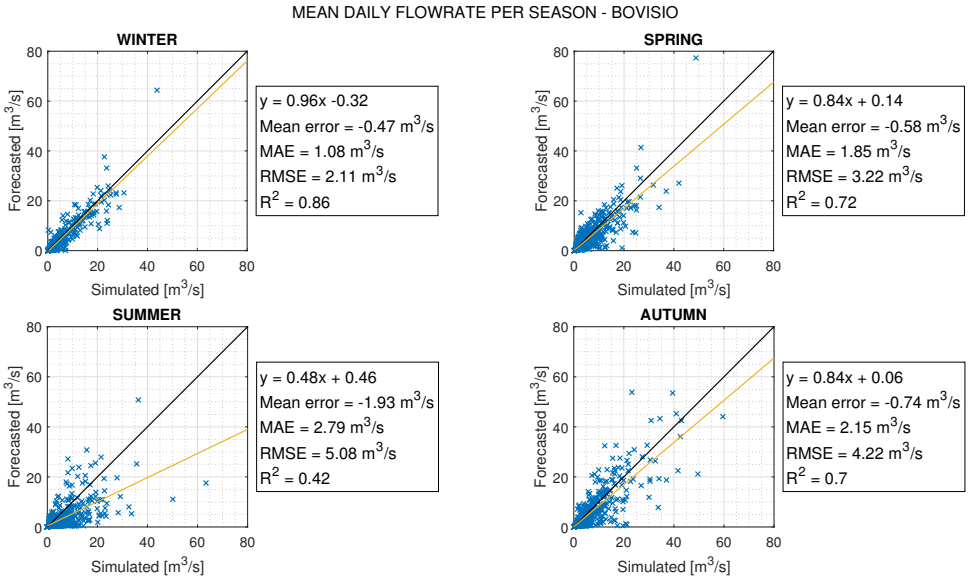


Figure 4.95: Mean daily flowrate per season: simulated vs forecasted

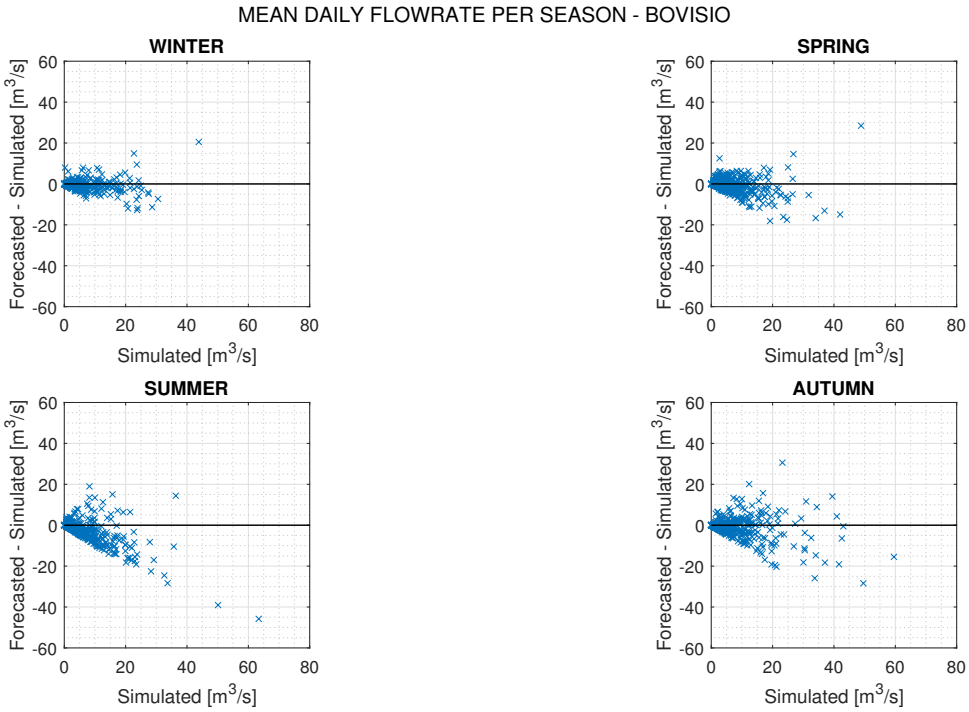


Figure 4.96: Mean daily flowrate per season: simulated vs forecasted-simulated

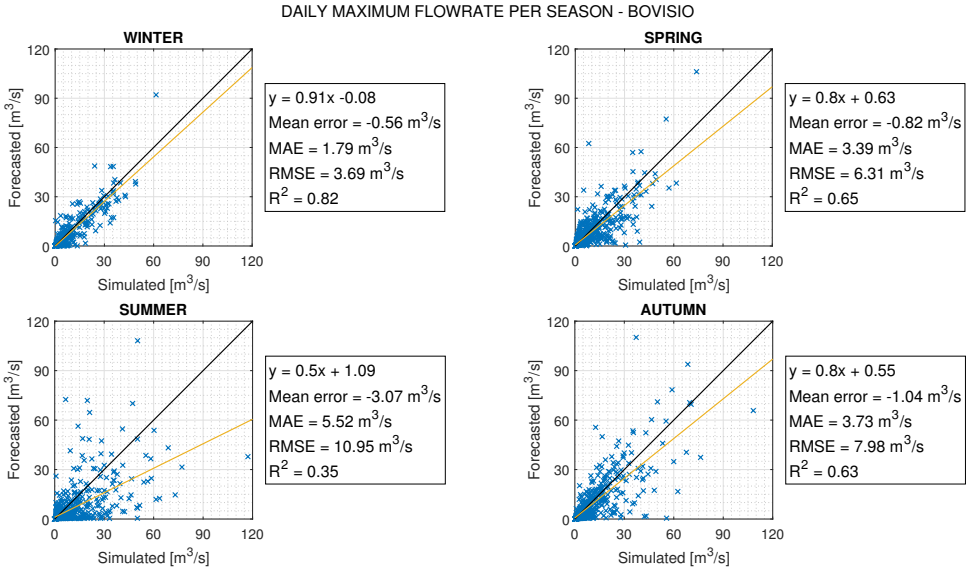


Figure 4.97: Daily maximum flowrate per season: simulated vs forecasted

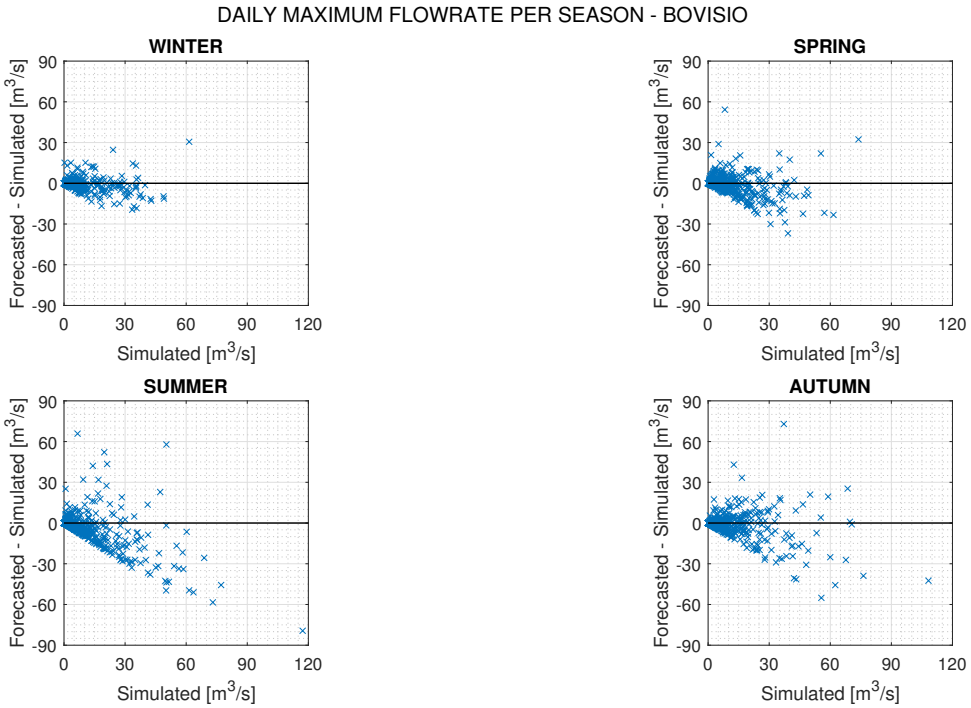


Figure 4.98: Daily maximum flowrate per season: simulated vs forecasted-simulated

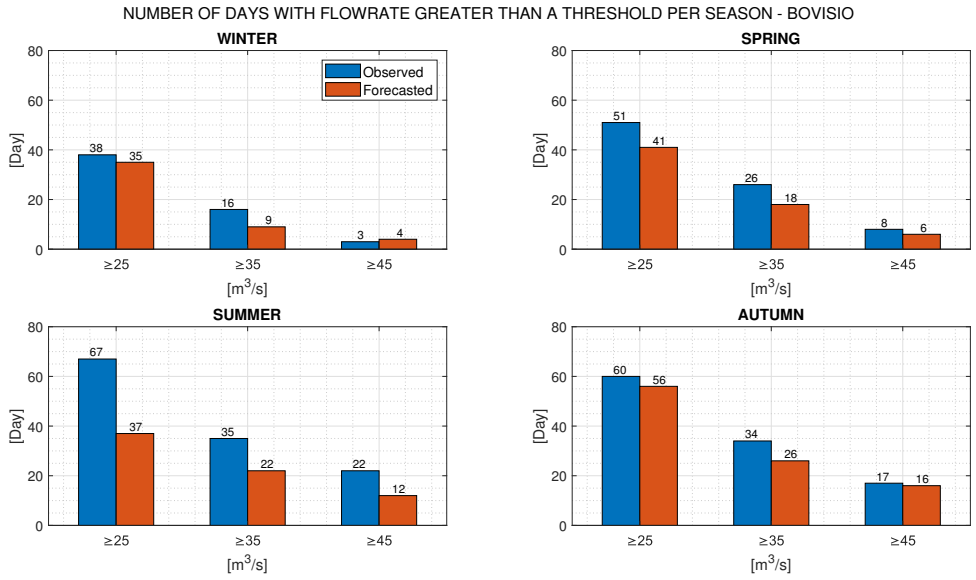


Figure 4.99: Number of days with flowrate greater than a threshold per season

4.2.9 Comparison overview

Once the comparison is performed for each variable, an overview of the performance of MOLOCH forecasts can be done. Concerning the meteorological variables, it has been shown that there is a very good agreement between forecasted and observed temperature, with an underestimation tendency by MOLOCH. Furthermore, there exists a good agreement of hourly, mean daily, and daily maximum solar radiation. In the first and third cases an overestimation and underestimation for low and high values is verified, respectively. On the other hand, in the second case an overestimation tendency is evident. Additionally, it has been shown that better statistical indexes are obtained when considering the simulation forced with the ARPA dataset only, which however, are not too different from the results when considering the entire dataset and, consequently, it does not reveal a great improvement of the predictability of solar radiation. Finally, at the seasonal scale, the lowest correspondence of information is

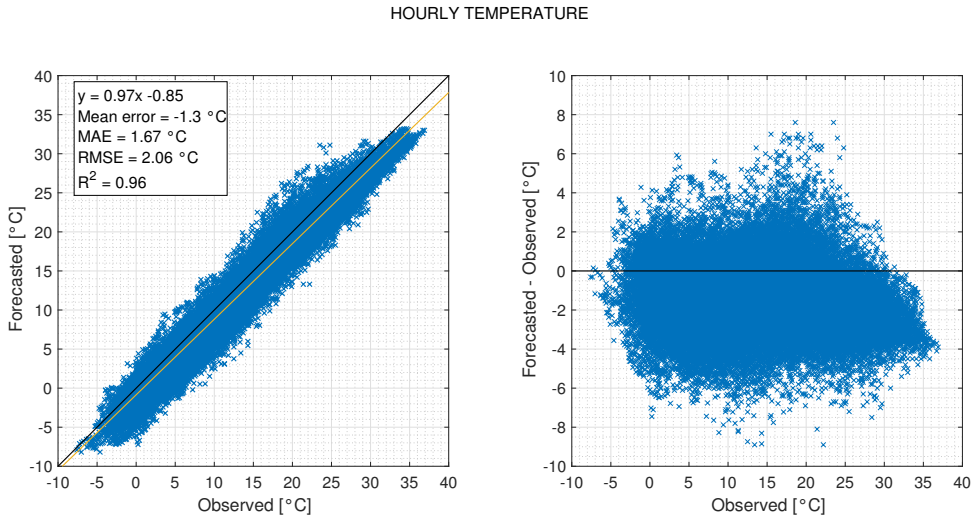
verified in winter in the hourly solar radiation, which may be a sign that the presence of fog affects the forecasts, especially for low values. Nevertheless, for mean daily and daily maximum solar radiation, it is summer where the coefficient of determination shows the lowest values and a clear overestimation trend for low values is verified. Also, the correspondence of the relative humidity data is not very high, and a large overestimation tendency is observed when low values are verified. Moreover, there is not a great improvement of the performance of the forecasts when compared with just the ARPA dataset. In addition, very low correspondence is verified for the wind speed variable. Overestimation is present in mean daily and daily maximum, and for low values when considering the hourly scale and the daily minimum where, at the same time, it is evidenced an underestimation trend for high wind velocity data. Furthermore, it must be taken into account that the present results may be affected by the heterogeneity of the measurement instruments' location, since MOLOCH forecasts are all reported at $10\,m$ of elevation from the ground. Therefore, it could be evaluated the possibility of applying a scale factor correction to the observed dataset according to the wind logarithmic equation to perform a better comparison. In any case, it must be remembered that wind speed values did not influence the results of the hydrological quantities since the potential evapotranspiration is assessed through Priestley-Taylor equation which does not consider the wind velocity component. In respect of precipitation, there exists an acceptable agreement of the forecasted and observed data. There is an overestimation trend especially for large values of precipitation, while small underestimations are present for small precipitation heights; the latter may be caused by instrumentation errors. Concerning the seasonal scale, the prediction of this variable is strongly affected by convective precipitation events in summer which reduce greatly the correspondence of data, increasing the underestimation tendency.

Regarding the hydrological variables it has been shown that in the case of soil moisture and potential evapotranspiration there exists a good correspondence of the information. Nevertheless, for the former it is observed a consistent underestimation, while for the latter it can be said that MOLOCH model strongly overestimates the data. Finally, a discreet agreement of the information is present in the mean daily discharge while a sufficient correspondence at the hourly scale is verified. There is an underestimation trend particularly high for large discharges, that may be explained by the observed behaviour of precipitation. Moreover, at the seasonal scale it is evident that the forecasts in summer fail in describing the observed data, while predictability gets better in autumn, spring, and especially in winter.

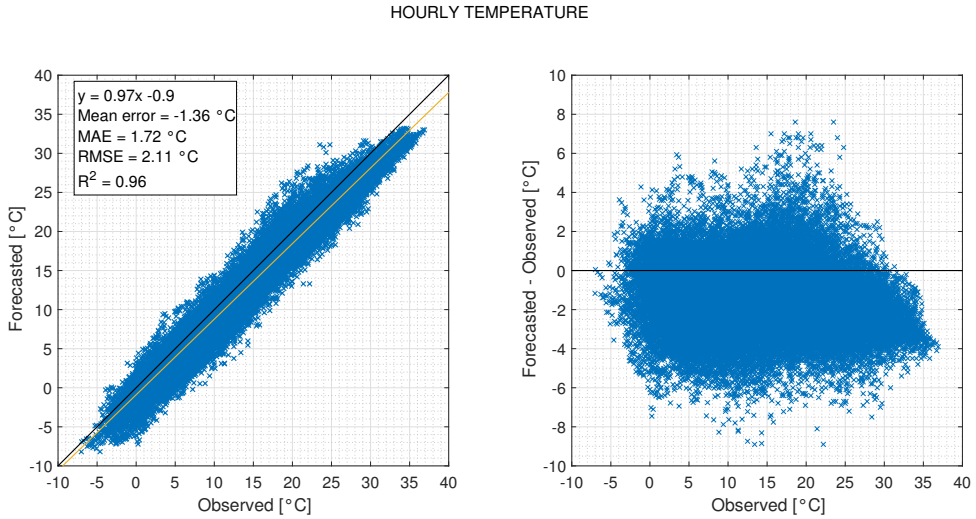
4.2.10 Additional comparisons

In the last subsections comparison between observed and forecasted data for the period from 2013 to 2020 and for every hydro-meteorological variable has been done. However, two further considerations can be made with the aim of improve the way in which the data is contrasted. In first place, it must be considered the fact that when performing forecasts there is an adjustment period in which the obtained results are not reliable. Thus, bearing in mind that MOLOCH forecasts have been taken from 04:00 UTC to 03:00 UTC of the next day, it is plausible to drop the first three hours of forecast related to the model spin-up and compare information from 07:00 UTC to 03:00 UTC of the next day. In second place, MOLOCH model was subjected to an upgrading and improvements, above all to terrain scheme, from December 2018, reason why the comparison could show better results if contrast of data is performed for the last two years of the studied period, it is, for 2019 and 2020. In the following charts (Fig. 4.100 to Fig. 4.114) results for hourly and daily scale of all variables are shown for the entire dataset (case A), the filtered one from 7 am to 3 am of the next day (case B), and the filtered for 2019 to 2020 (case C). Additional comments are also made.

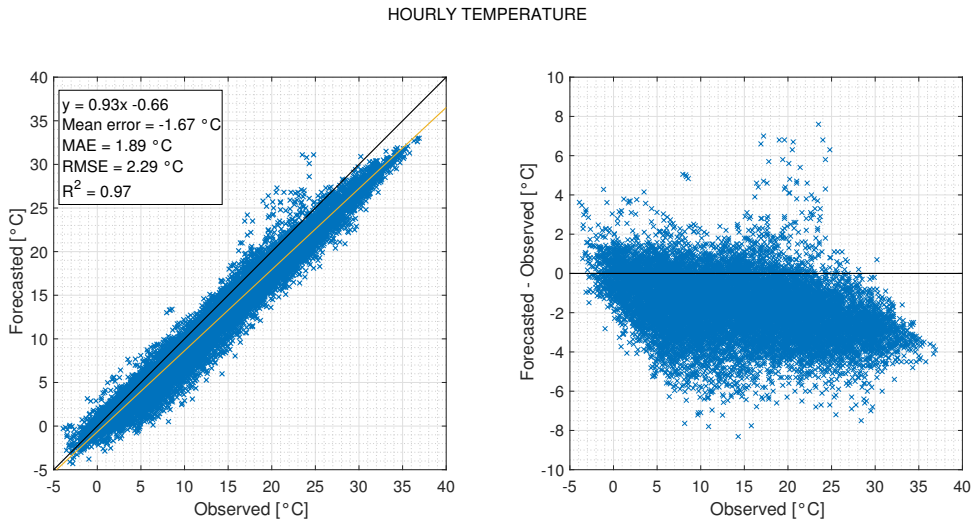
Hourly temperature (Fig. 4.100) is still cold bias, there is an improvement of correspondence in C, whilst the ME, MAE, and RMSE get slightly worse when compared with A. Moreover, similar behaviour is observed for the mean daily temperature (Fig. 4.101), with the difference that no improvement in terms of R^2 coefficient is verified.



(a) All data

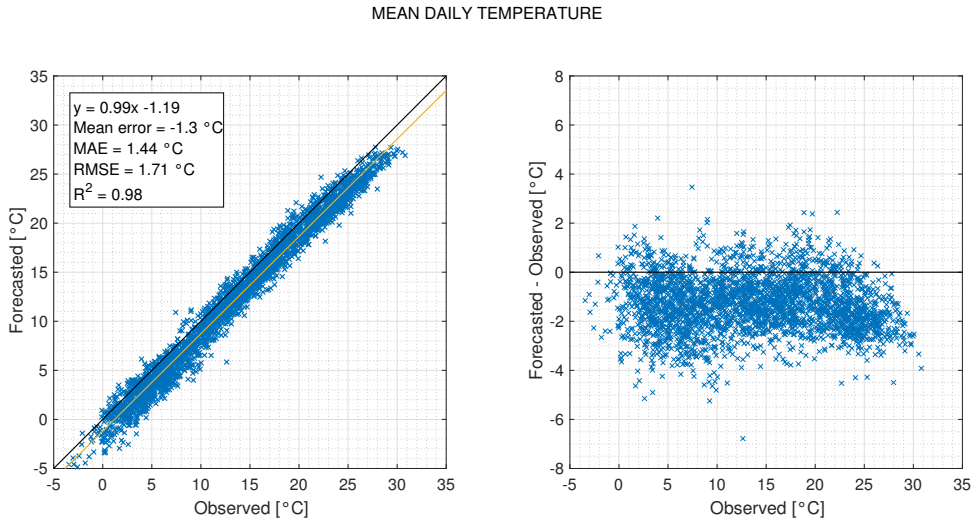


(b) Filtered from 7 am to 3 am

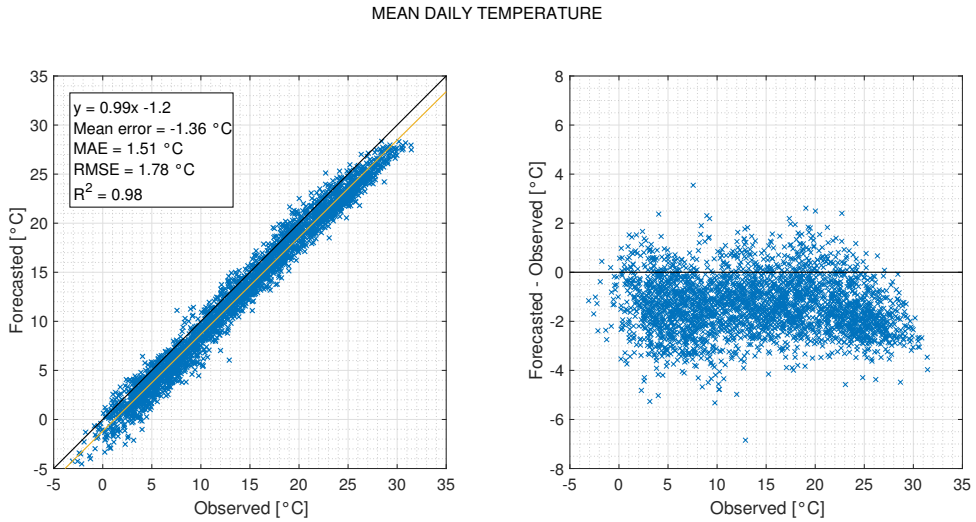


(c) Filtered from 7 am to 3 am (2019-2020)

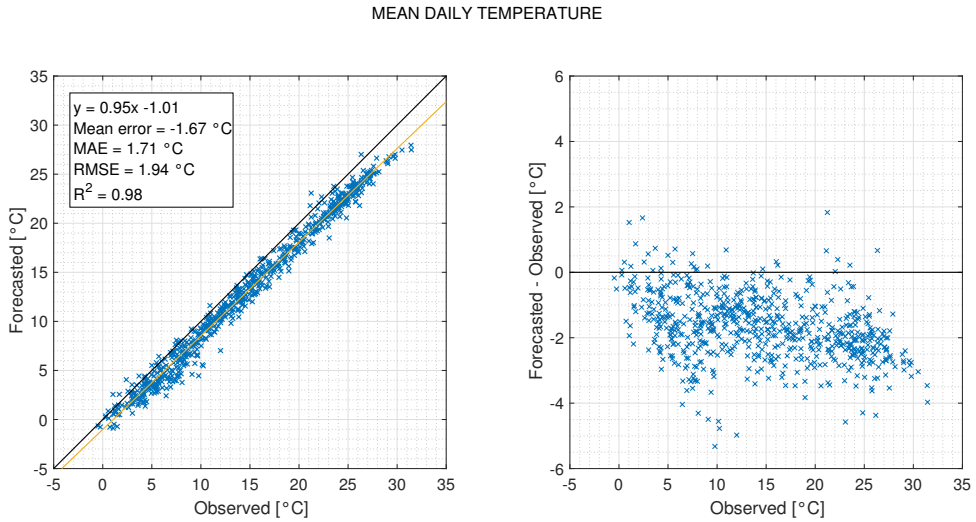
Figure 4.100: Hourly temperature with different dataset comparison



(a) All data



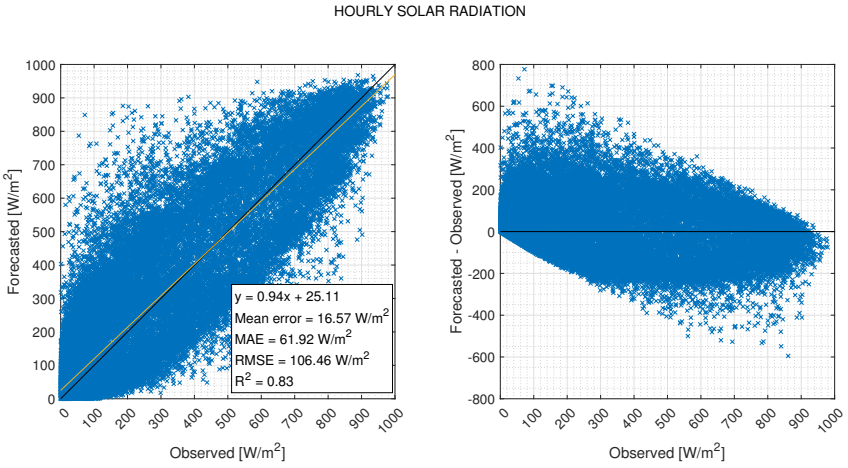
(b) Filtered from 7 am to 3 am



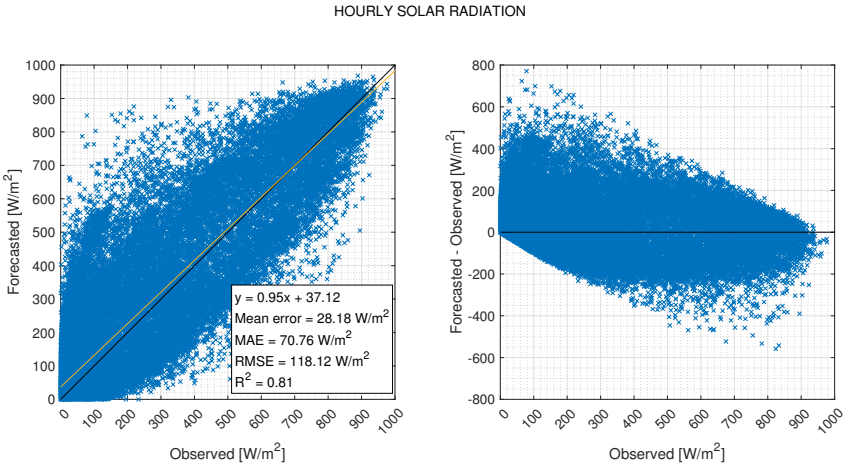
(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.101: Mean daily temperature with different dataset comparison

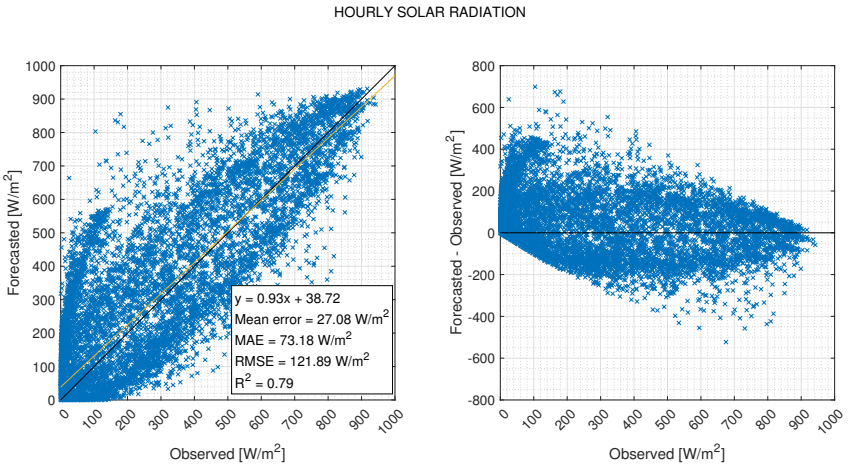
Hourly solar radiation (Fig. 4.102) shows an overestimation and an underestimation for low and large values, respectively. The coefficient of determination decreases from A to B, and from B to C. On the other hand, mean daily solar radiation (Fig. 4.103) presents the already described overestimation trend, and the correspondence gets better in C case. Moreover, the statistical indexes remain small in comparison with the order of magnitude of the values for both hourly and mean daily solar radiation.



(a) All data

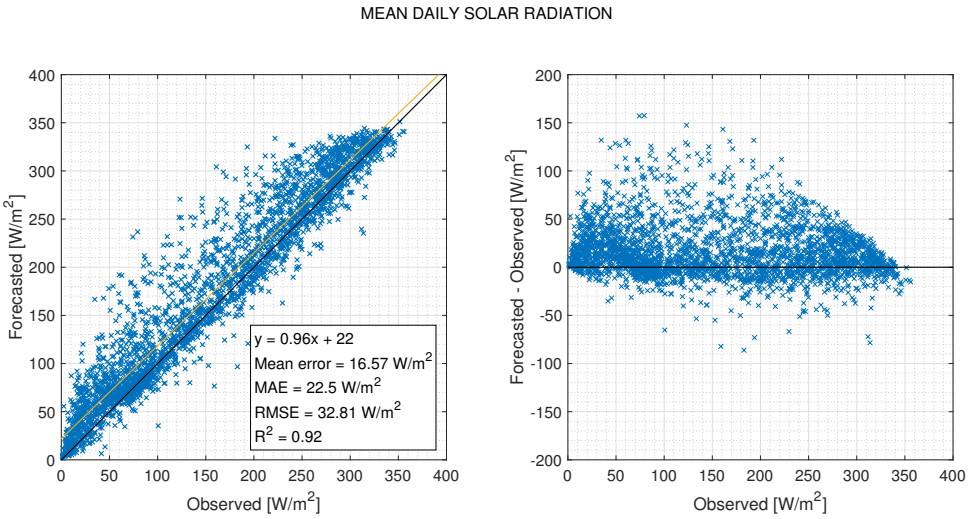


(b) Filtered from 7 am to 3 am

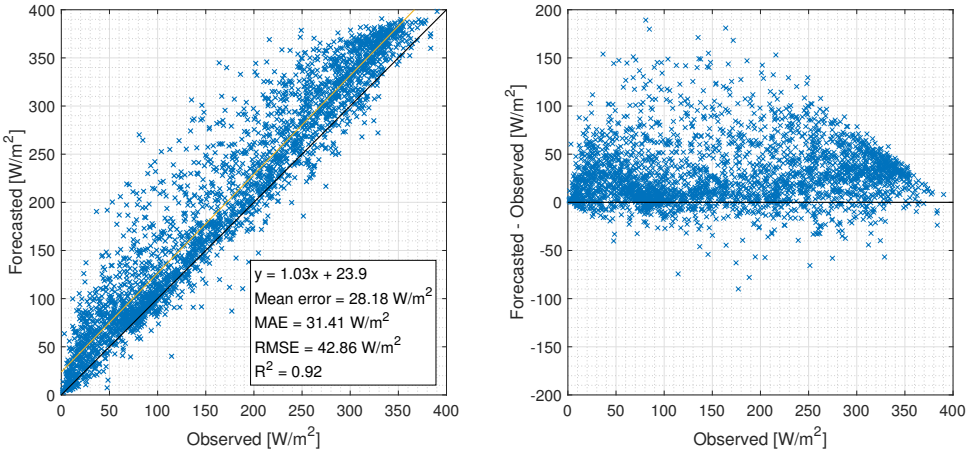


(c) Filtered from 7 am to 3 am (2019-2020)

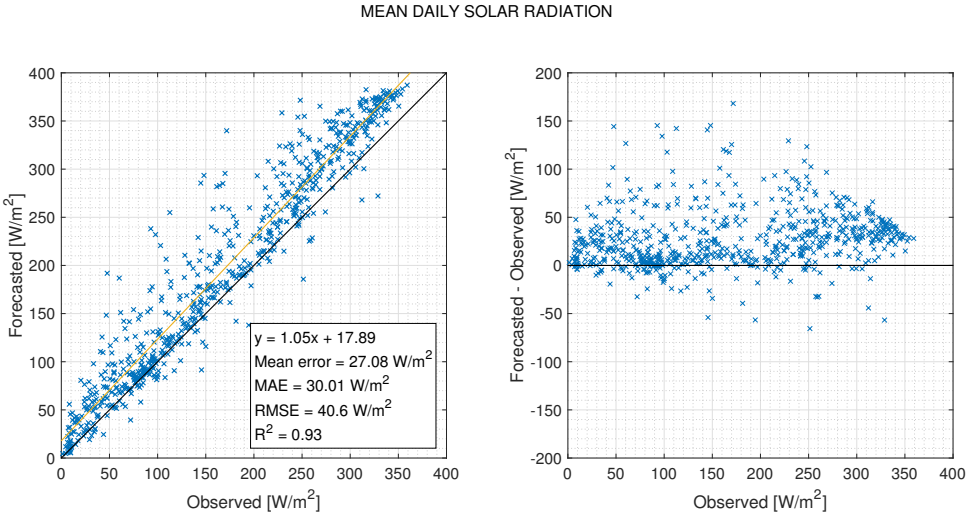
Figure 4.102: Hourly solar radiation with different dataset comparison



(a) All data



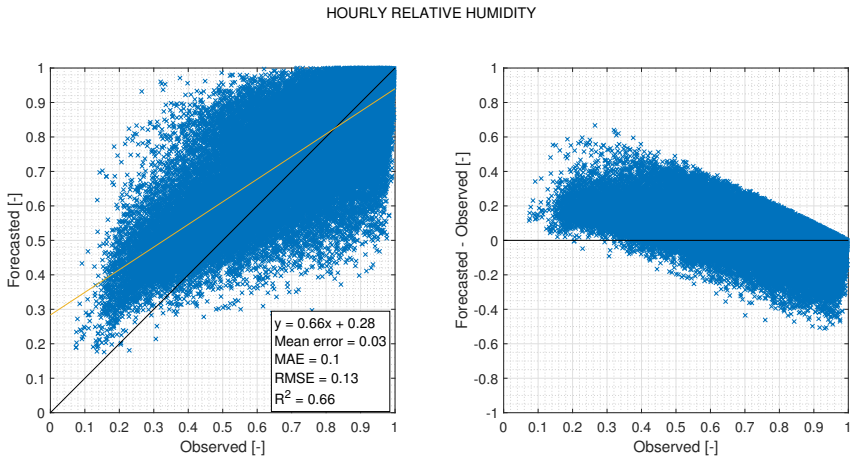
(b) Filtered from 7 am to 3 am



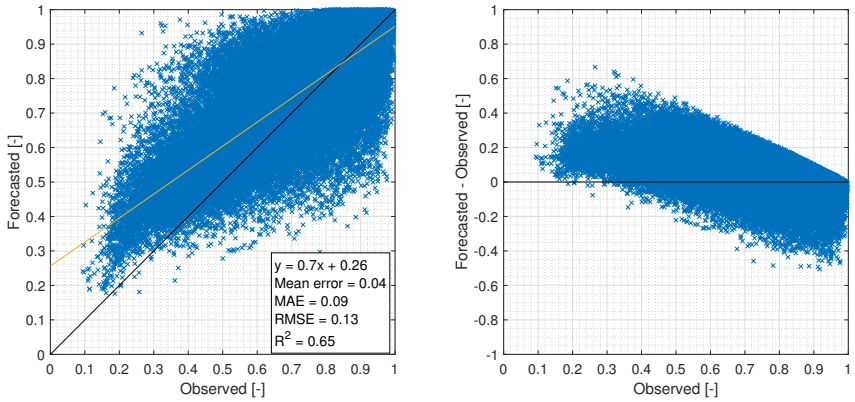
(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.103: Mean daily solar radiation with different dataset comparison

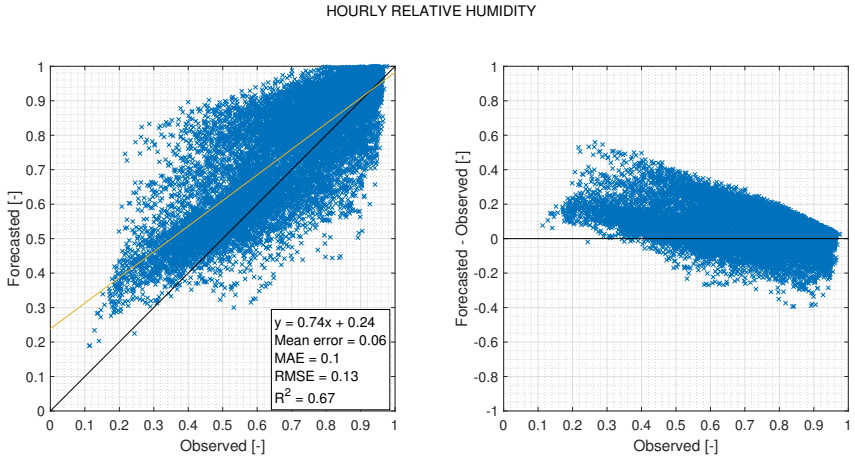
Correspondence is improved in hourly relative humidity case (Fig. 4.104). The trend of large overestimations when low values of the variable are verified is still present in all cases and, particularly in C, overestimation is also verified for larger values when compared with A and B. Moreover, statistical indexes are practically the same. Conversely, mean daily relative humidity (Fig. 4.105) predictability gets worse from A to B to C, the trend is the same as in the hourly based case, and statistical indexes are similar.



(a) All data

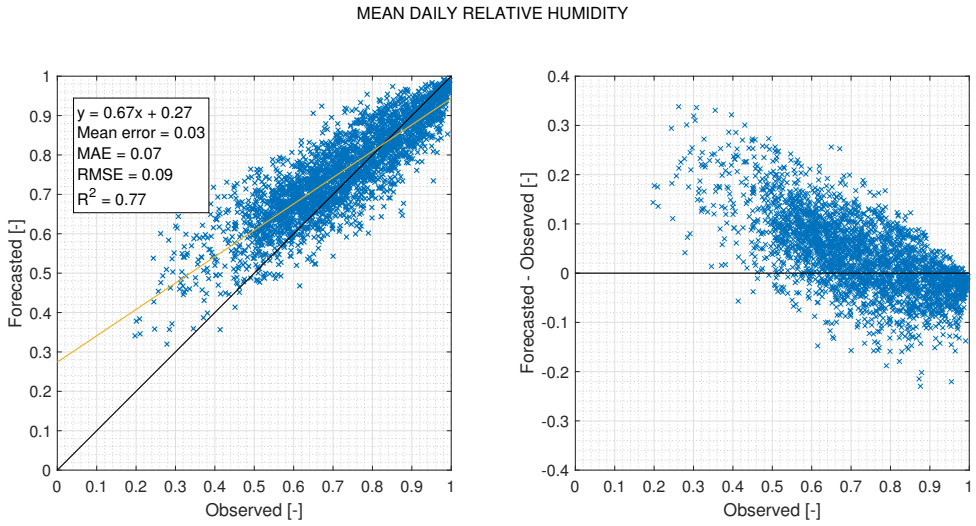


(b) Filtered from 7 am to 3 am

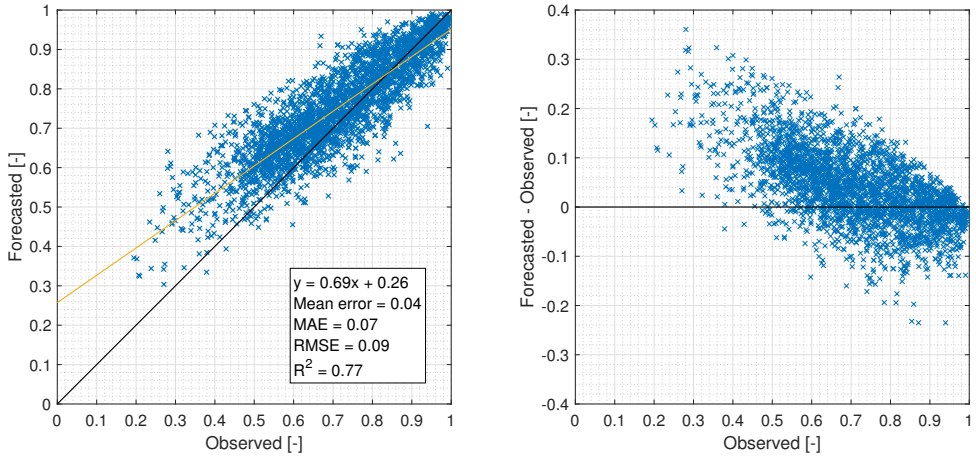


(c) Filtered from 7 am to 3 am (2019-2020)

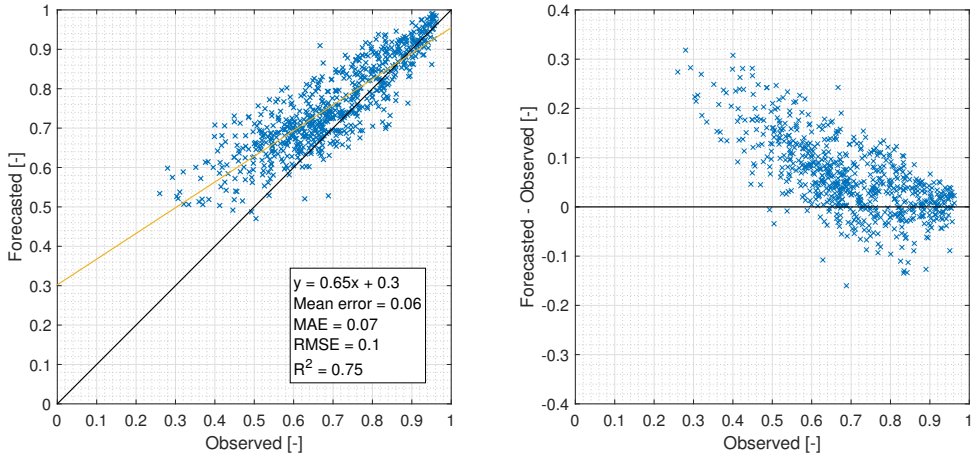
Figure 4.104: Hourly relative humidity with different dataset comparison



(a) All data



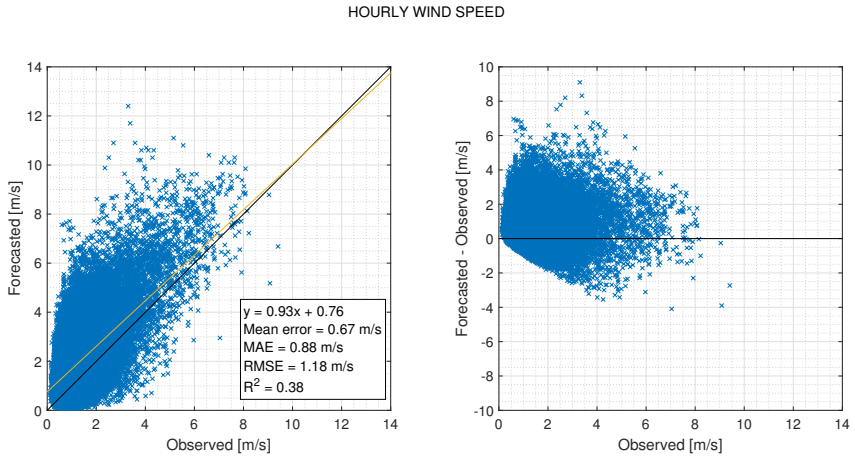
(b) Filtered from 7 am to 3 am



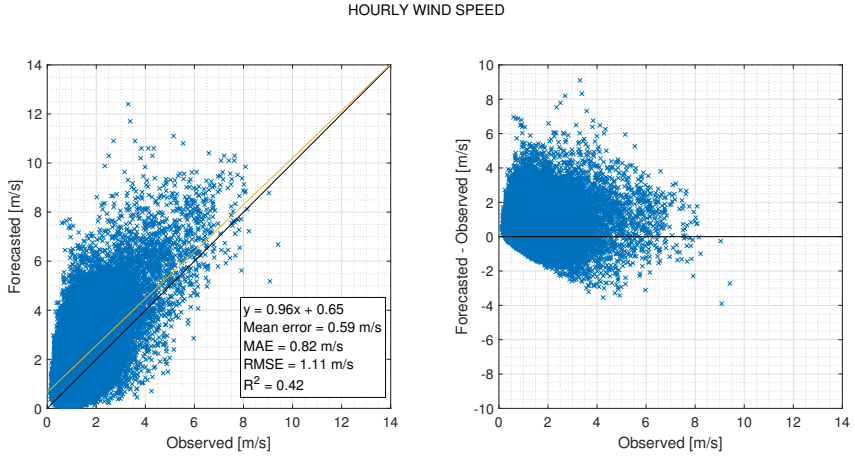
(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.105: Mean daily relative humidity with different dataset comparison

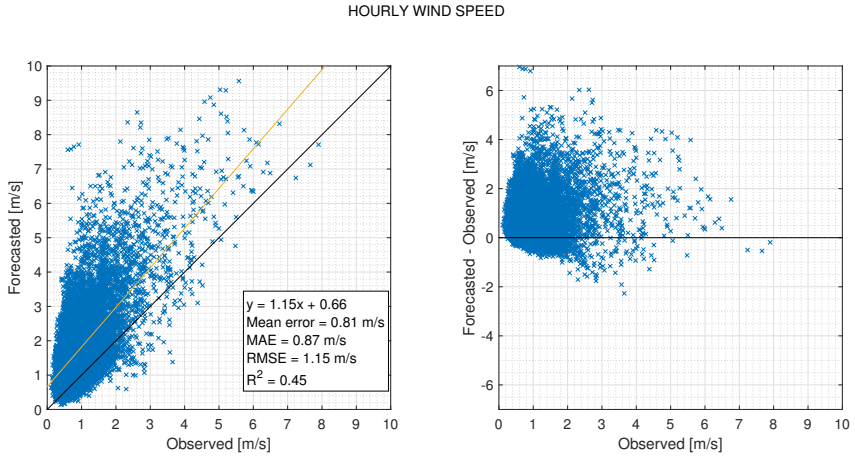
In hourly wind speed case (Fig. 4.106) it can be seen an improvement in the R^2 coefficient which, however, is very low. The trend also changes from over and underestimation for low and high values - respectively – in A, to a slight overestimation in B, to a consistent overestimation in C. On the other hand, the correspondence largely improves in case C for mean daily wind speed (Fig. 4.107), passing from R^2 values of 0.56 and 0.57 (in A and B, respectively) to 0.73. Moreover, the overestimation trend still remains, and it gets more evident in C. Nevertheless, it must be remembered that elevation of anemometers affects the observed information, and a correction coefficient could be evaluated.



(a) All data

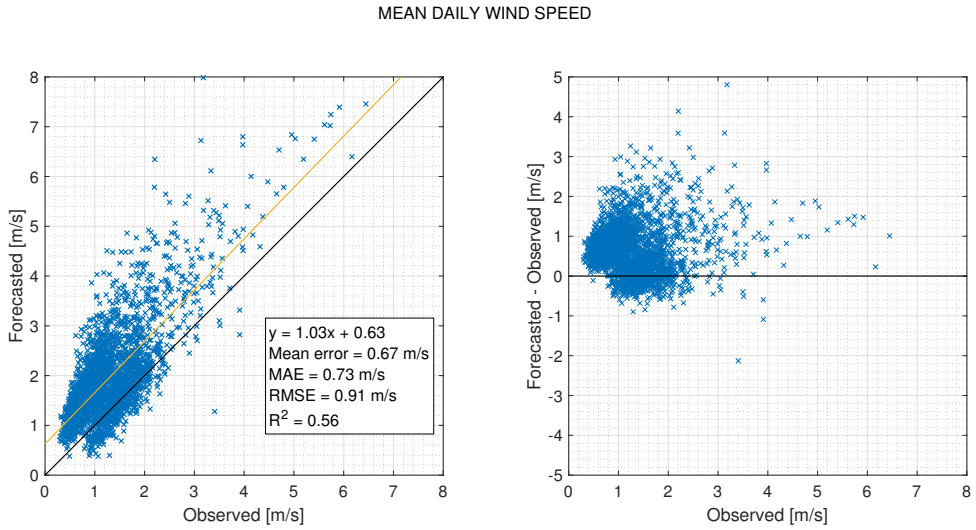


(b) Filtered from 7 am to 3 am

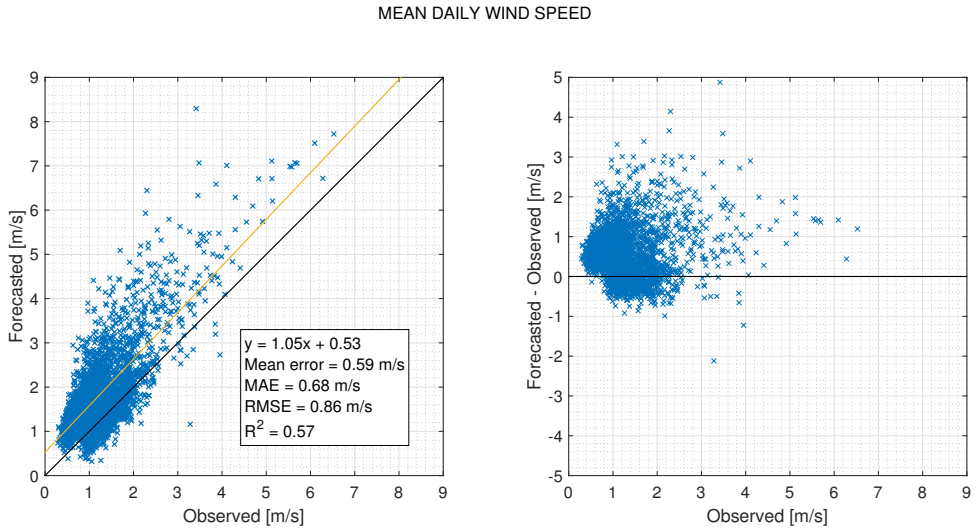


(c) Filtered from 7 am to 3 am (2019-2020)

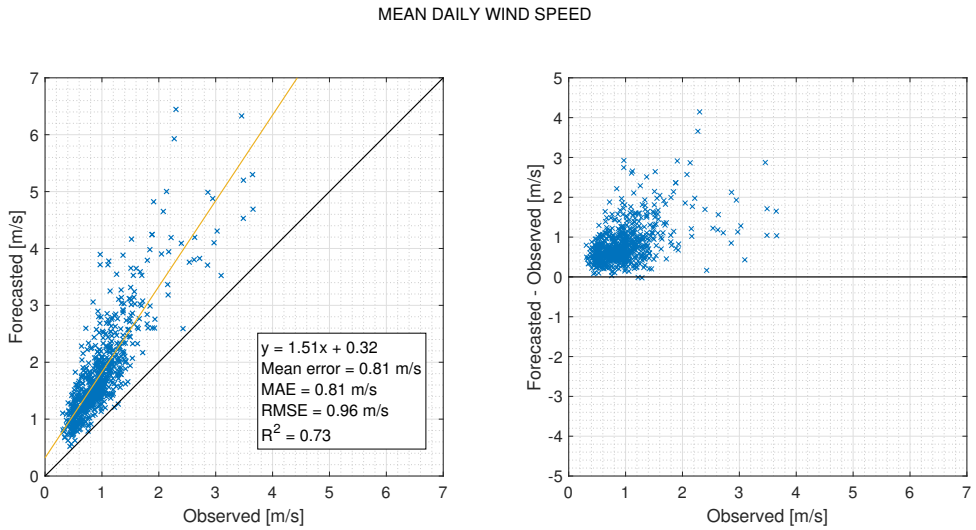
Figure 4.106: Hourly wind speed with different dataset comparison



(a) All data



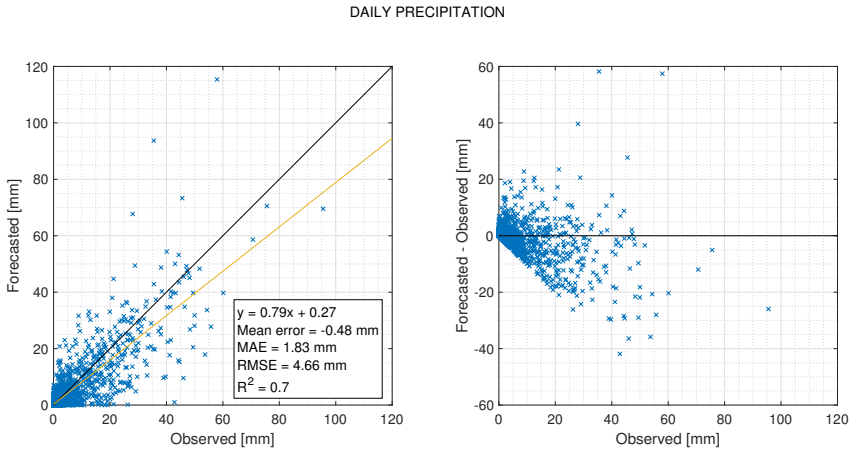
(b) Filtered from 7 am to 3 am



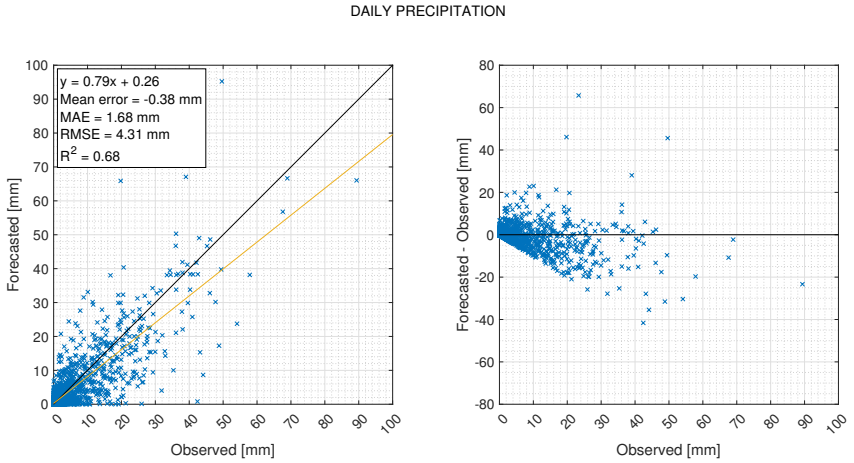
(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.107: Mean daily wind speed with different dataset comparison

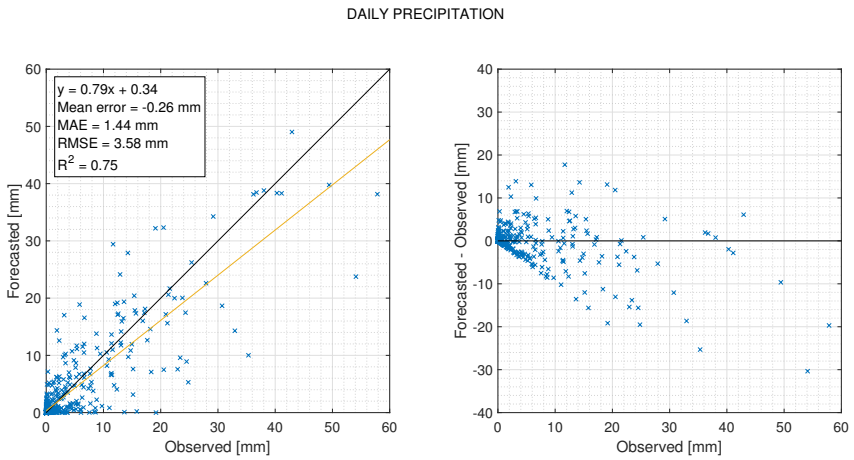
Underestimation tendency is present in all three cases for daily precipitation (Fig. 4.108). However, the coefficient of determination gets worse from A to B, and then improves to the largest value (0.75) in C case. Moreover, statistical indexes display the same behaviour.



(a) All data



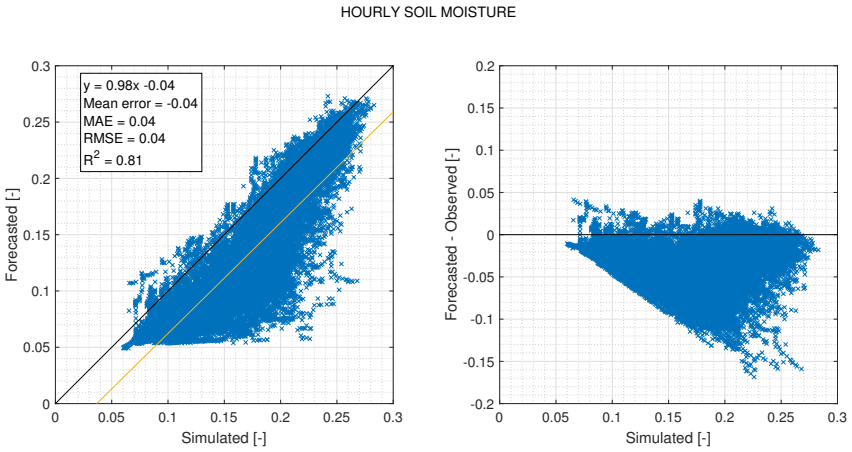
(b) Filtered from 7 am to 3 am



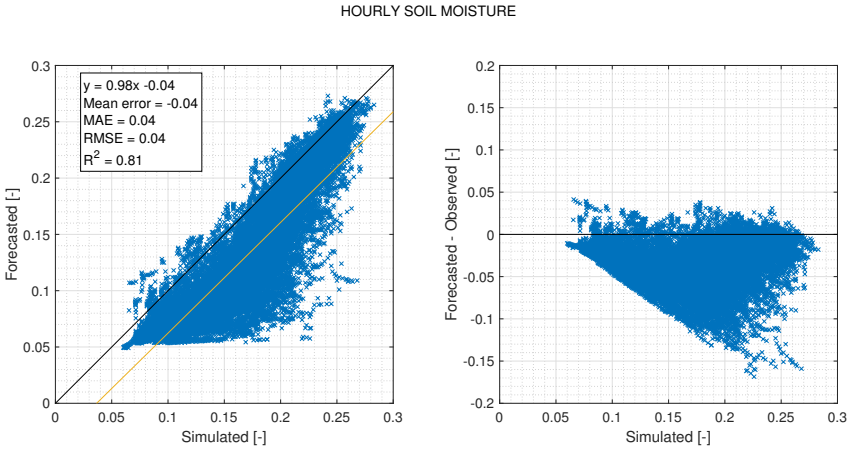
(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.108: Daily precipitation with different dataset comparison

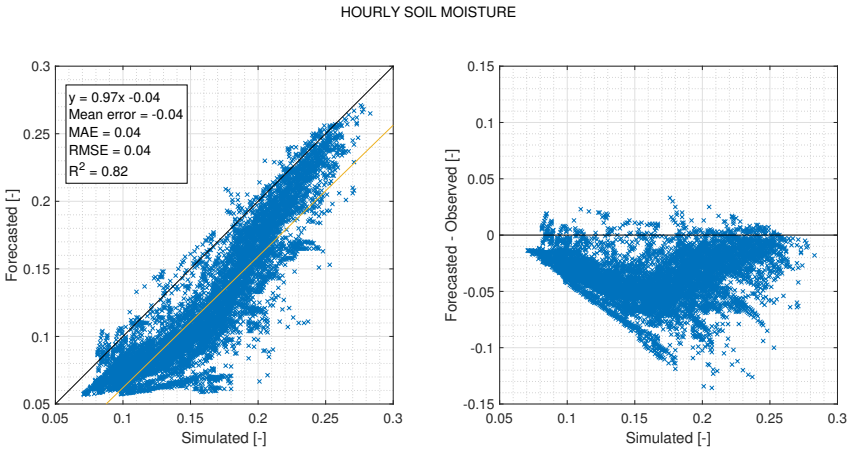
For the hourly soil moisture case (Fig. 4.109) a very slight improvement of the predictability is observed in C, while the ME, MAE, and RMSE are the same in all cases. The consistent underestimation remains too. Additionally, the exact same behaviour is verified for the mean daily case (Fig. 4.110).



(a) All data

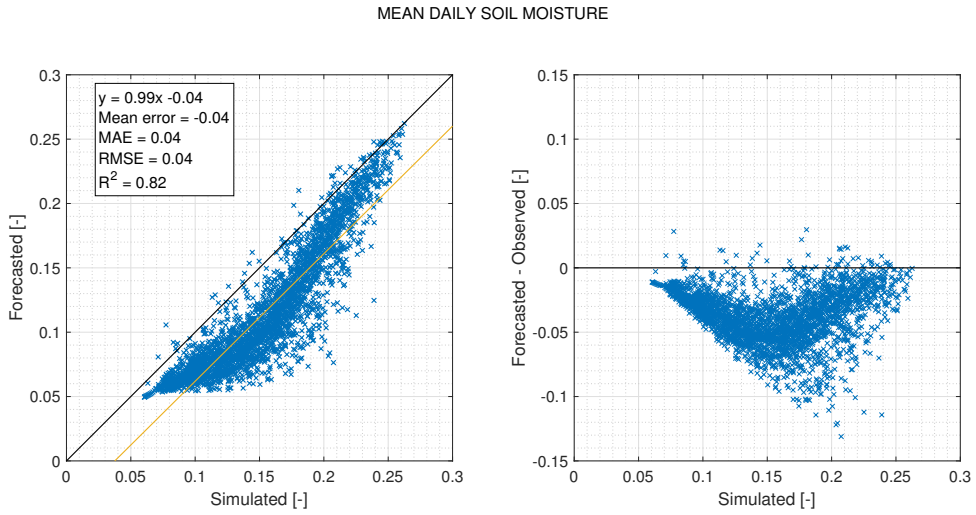


(b) Filtered from 7 am to 3 am

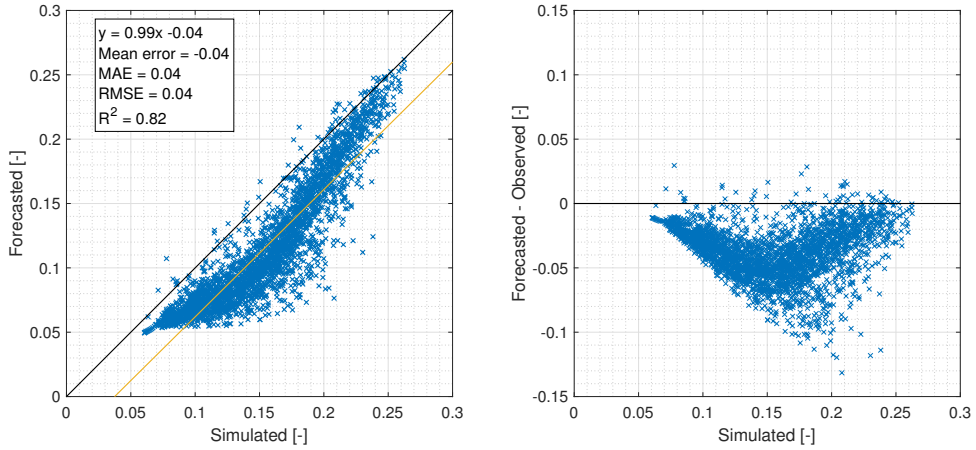


(c) Filtered from 7 am to 3 am (2019-2020)

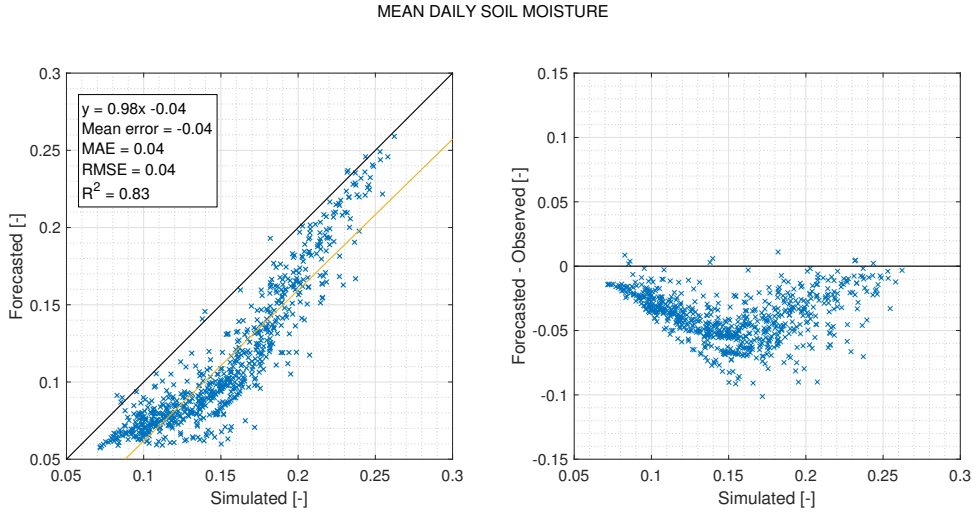
Figure 4.109: Hourly soil moisture with different dataset comparison



(a) All data



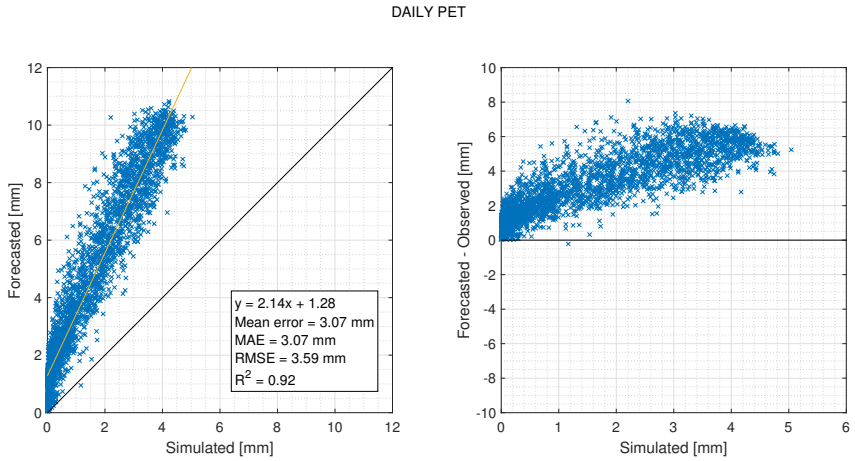
(b) Filtered from 7 am to 3 am



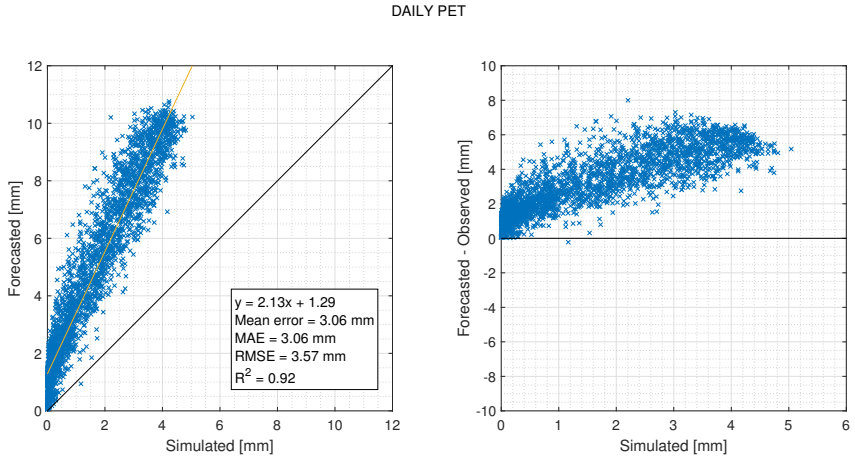
(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.110: Mean daily soil moisture with different dataset comparison

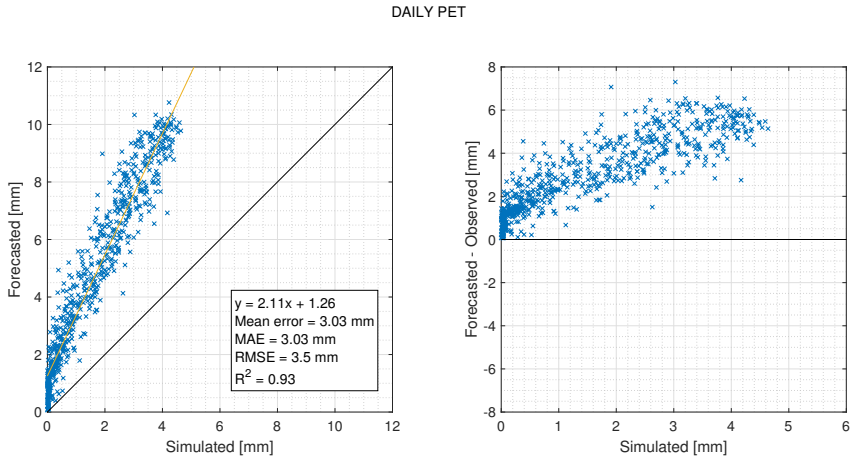
For both hourly and mean daily potential evapotranspiration (Fig. 4.111 and Fig. 4.112) it is observed that the strong overestimation is still present in all cases, and the coefficient of determination improves by one percentual point in C.



(a) All data

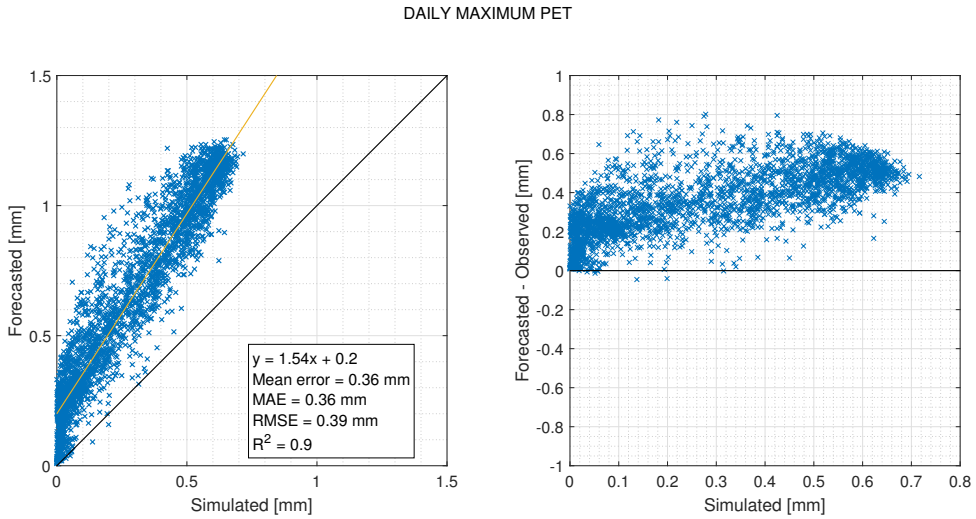


(b) Filtered from 7 am to 3 am

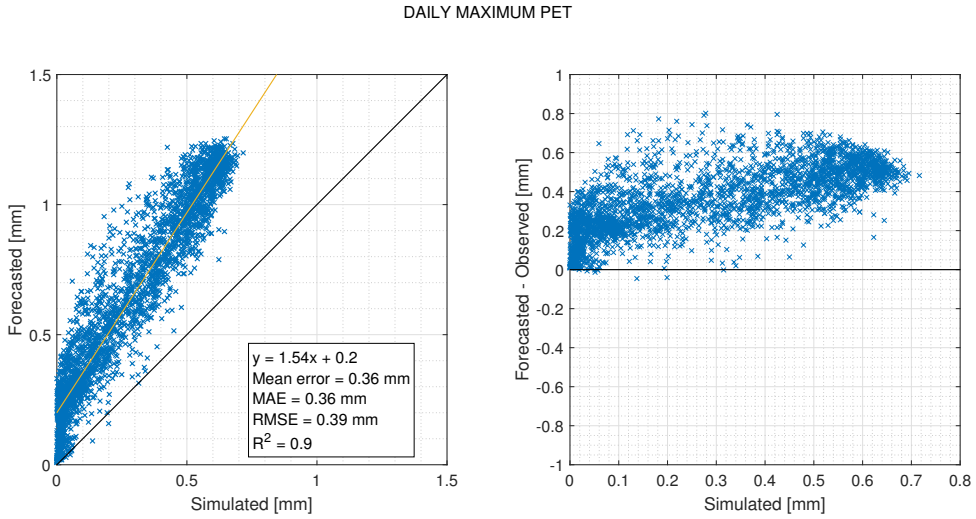


(c) Filtered from 7 am to 3 am (2019-2020)

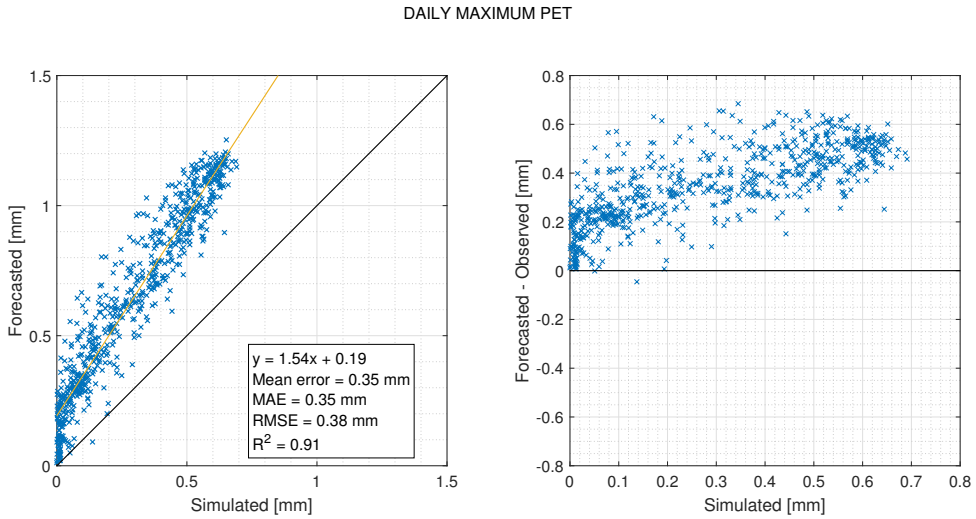
Figure 4.111: Hourly potential evapotranspiration with different dataset comparison



(a) All data



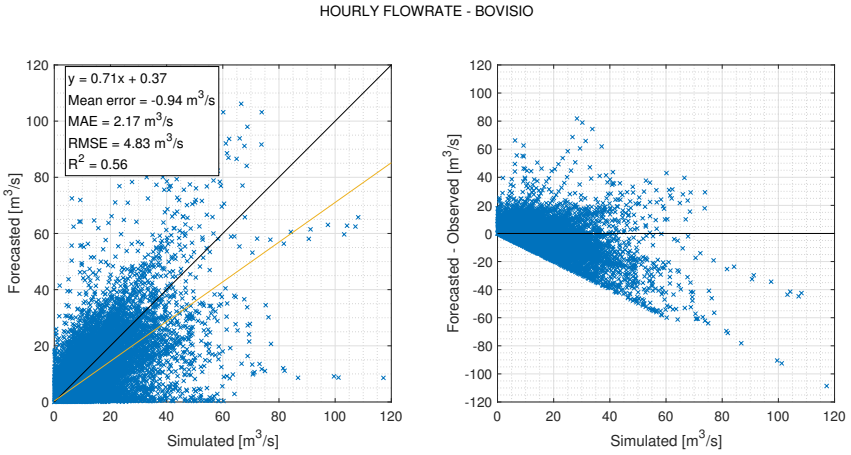
(b) Filtered from 7 am to 3 am



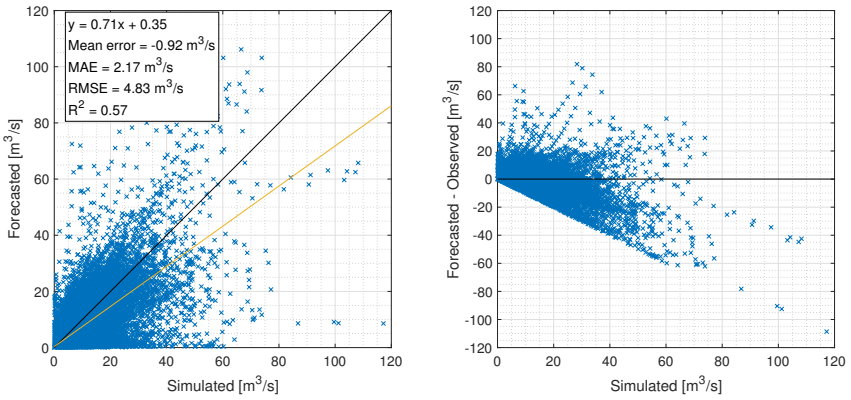
(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.112: Mean daily potential evapotranspiration with different dataset comparison

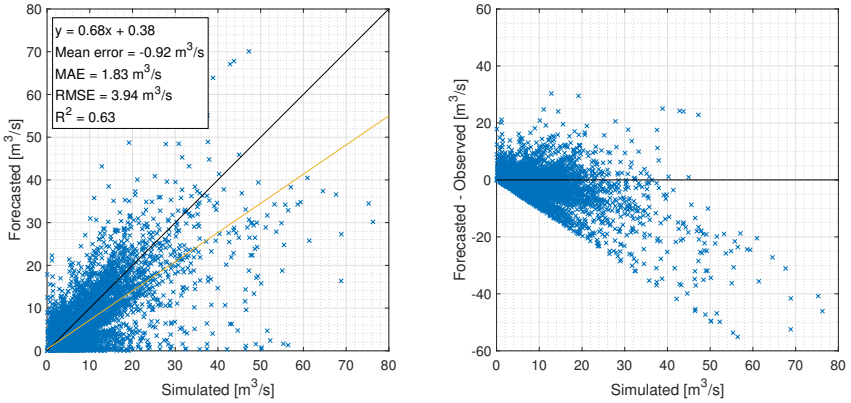
Finally, for the hourly discharge (Fig. 4.113) the agreement improves from A to B to C, but values of R^2 are still discreet. On the other hand, in terms of mean daily discharge (Fig. 4.114) the correspondence is equal in A and B cases (0.66), while it considerably improves to 0.74 when considering only 2019 and 2020. It is also evident that for hourly and mean daily discharges the underestimation trend – especially for large values – is still present.



(a) All data

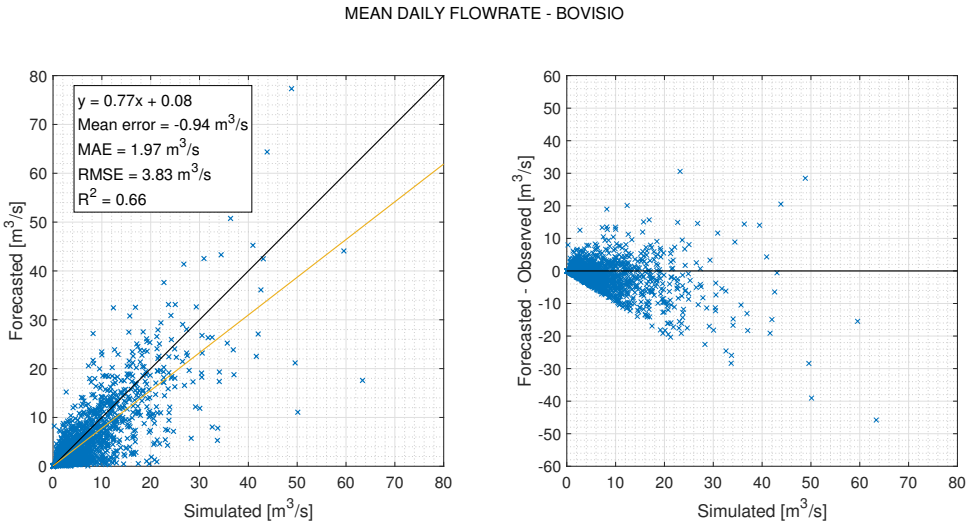


(b) Filtered from 7 am to 3 am

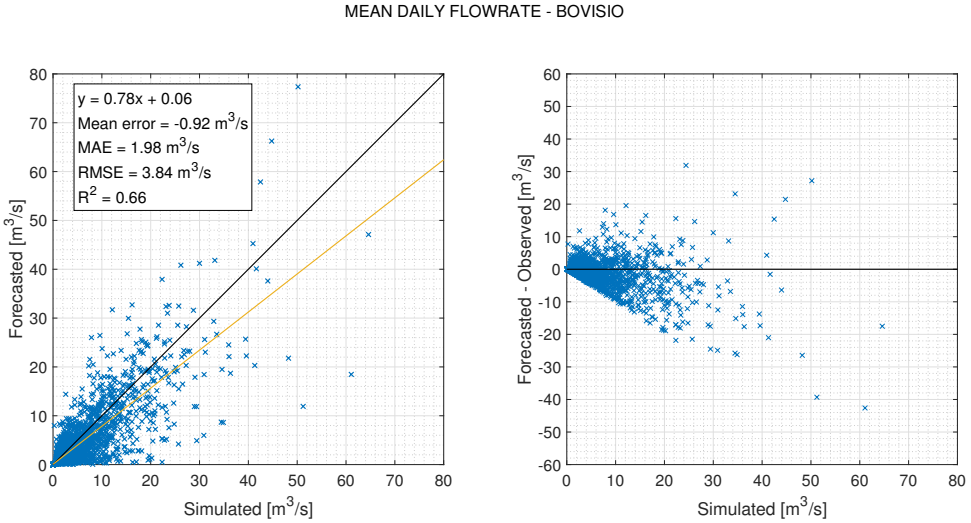


(c) Filtered from 7 am to 3 am (2019-2020)

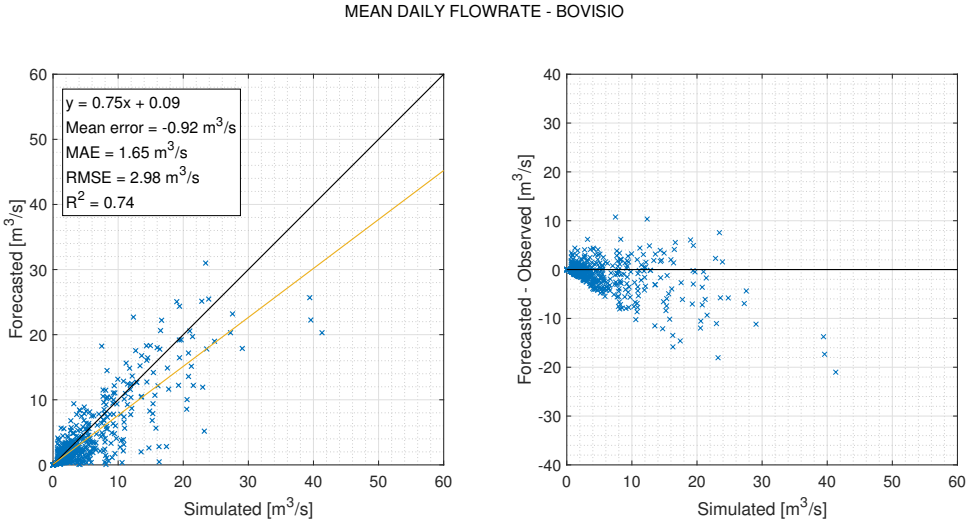
Figure 4.113: Hourly discharge at Bovisio with different dataset comparison



(a) All data



(b) Filtered from 7 am to 3 am



(c) Filtered from 7 am to 3 am (2019-2020)

Figure 4.114: Mean daily discharge at Bovisio with different dataset comparison

Hence, as a brief overview it can be said that the agreement improves in most of the cases when using the filtered dataset for 2019 to 2020, being the exception the hourly solar radiation and the mean daily relative humidity. Furthermore, the over or underestimation trends are conserved in all but hourly wind speed cases. Finally, it should be considered that for C case there is less amount of data analysed and, consequently, results may not be as robust as the ones shown for 2013 to 2020 period.

4.3 Sensitivity analysis

To define which one of the meteorological forcings influences the most the simulated hydrological variables (soil moisture, potential evapotranspiration, and discharge), as already mentioned, it is performed a sensitivity analysis by means of the one-factor-at-time (OAT) methodology [40] [41]. In the following charts (Fig. 4.115 to Fig. 4.122) are presented four scatter plots for each hydrological variable at daily scale, each one representing the results of the simulation obtained by forcing the model with observed data except for the variable in the title (which is the one forecasted by MOLOCH). Notice that wind speed is not considered since it does not influences the estimation of the hydrological quantities in the present study.

In first place, it can be seen that soil moisture (Fig. 4.115 and Fig. 4.116) is affected by solar radiation and, mainly, by precipitation forecasts. This, due to the fact that the coefficient of determination drops to 0.92 and 0.85 in the solar radiation and precipitation cases, respectively. Thus, this combined effect makes the forecast to have the behaviour shown in Fig. 4.83 in the last section, where the R^2 is 0.82. On the other hand, for temperature and relative humidity are obtained perfect forecasts of the variable. Hence, bearing in mind that the soil moisture is dynamically described by the water balance equation in FEST-WB model (Equation 3.6), which depends on precipitation, surface runoff and drainage fluxes, and evapotranspiration – which at the same time is strictly dependent on the potential evapotranspiration -, it is clear why forecasts of precipitation affects the soil moisture variable. Additionally, solar radiation is also present in Priestley-Taylor equation for the estimation of potential evapotranspiration and, as consequence, it also affects soil moisture forecasts.

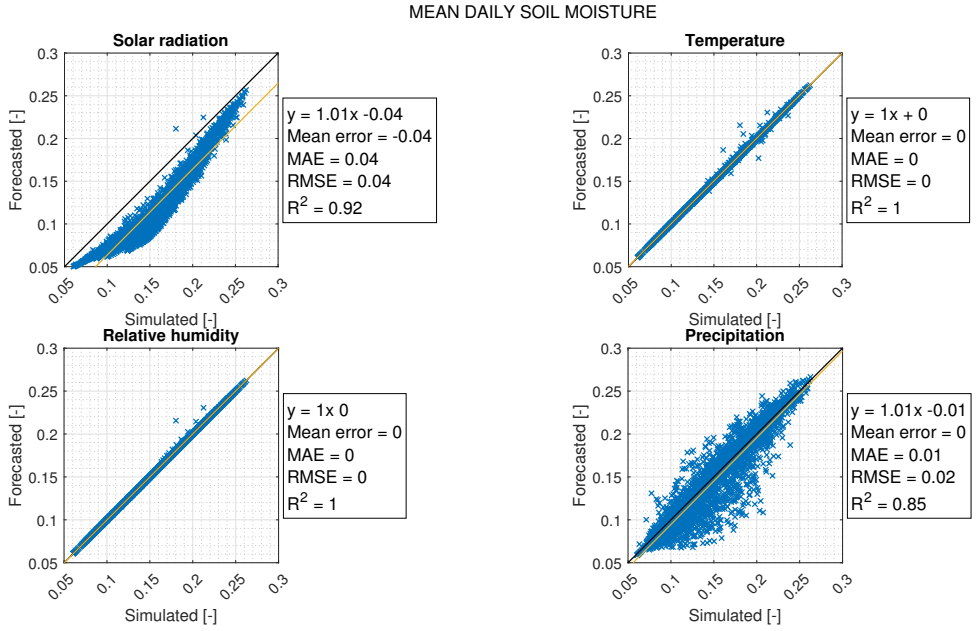


Figure 4.115: Mean daily soil moisture simulated vs forecasted sensitivity analysis

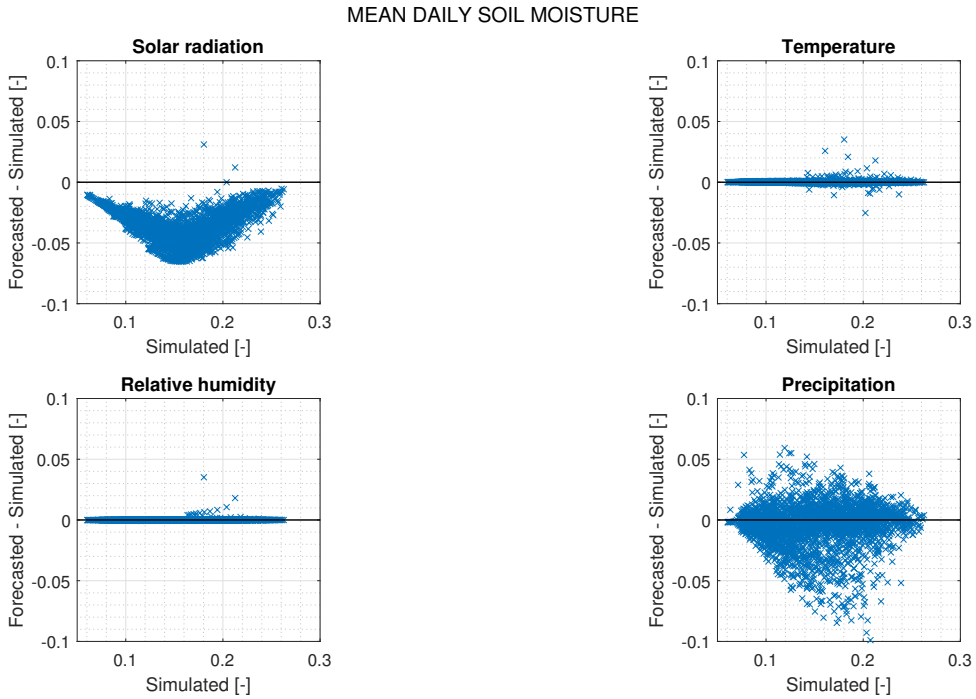


Figure 4.116: Mean daily soil moisture simulated vs forecasted-simulated sensitivity analysis

Results of potential evapotranspiration are shown in Fig. 4.117 and Fig. 4.118. It can be seen that, effectively, the solar radiation strongly influences the potential evapotranspiration while the additional meteorological forcings do not. Consequently, the overestimation trend verified for the solar radiation leads to the behaviour described in Fig. 4.86 where a strong overestimation is evident.

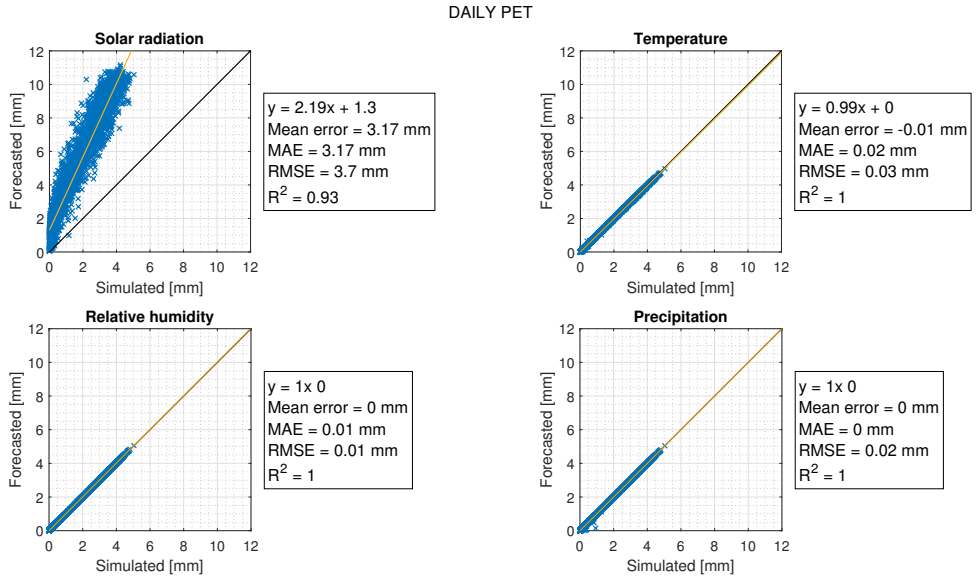


Figure 4.117: Mean daily potential evapotranspiration simulated vs forecasted sensitivity analysis

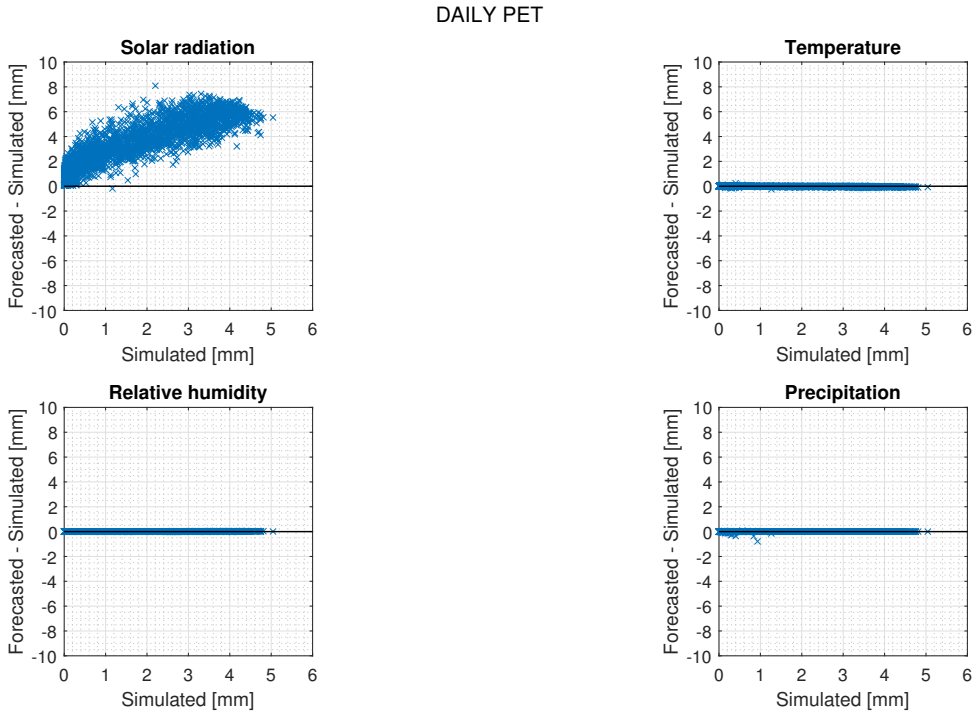


Figure 4.118: Mean daily potential evapotranspiration simulated vs forecasted-simulated sensitivity analysis

Concerning the discharge at Bovisio station, results are shown from Fig. 4.119 to Fig. 4.122. It is evident that the precipitation strongly influences the discharge forecasts leading to the underestimation observed in Fig. 4.89, while with the other meteorological forcings are obtained practically perfect forecasts. This results are also evident when considering discharges greater or equal than 25 m^3/s (Fig. 4.121 and Fig. 4.122), where the coefficient of determination decreases to 0.16 when the model is forced with forecasted precipitation.

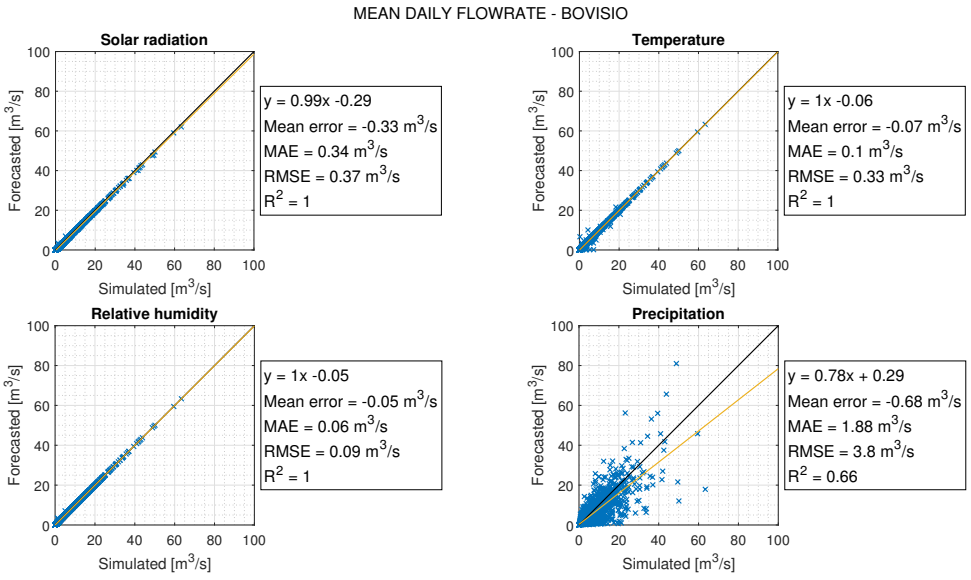


Figure 4.119: Mean daily discharge at Bovisio simulated vs forecasted sensitivity analysis

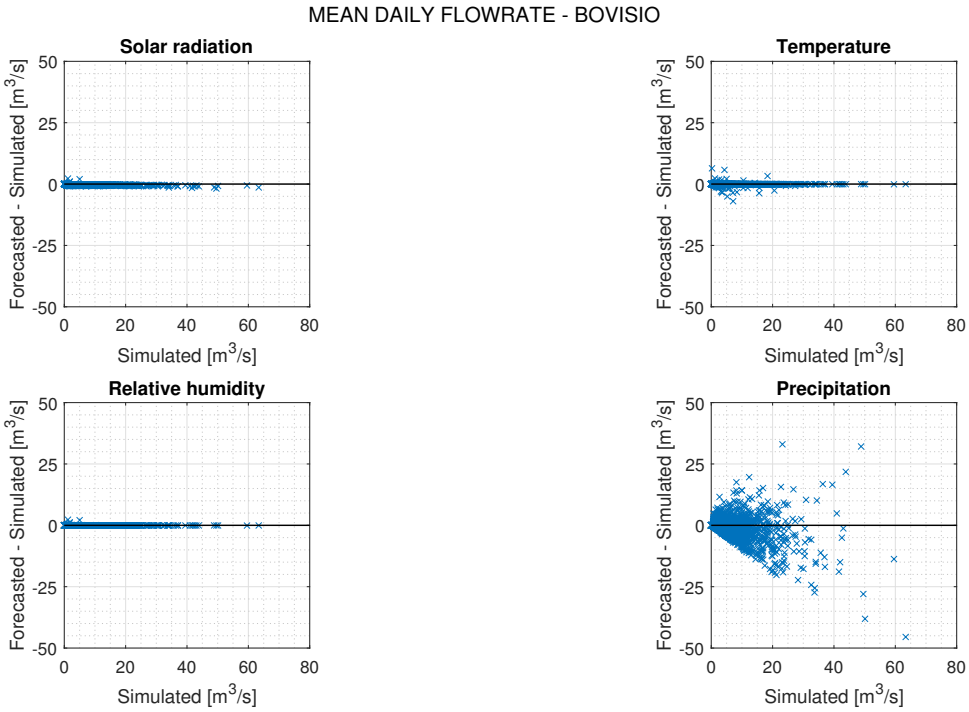


Figure 4.120: Mean daily discharge at Bovisio simulated vs forecasted-simulated sensitivity analysis

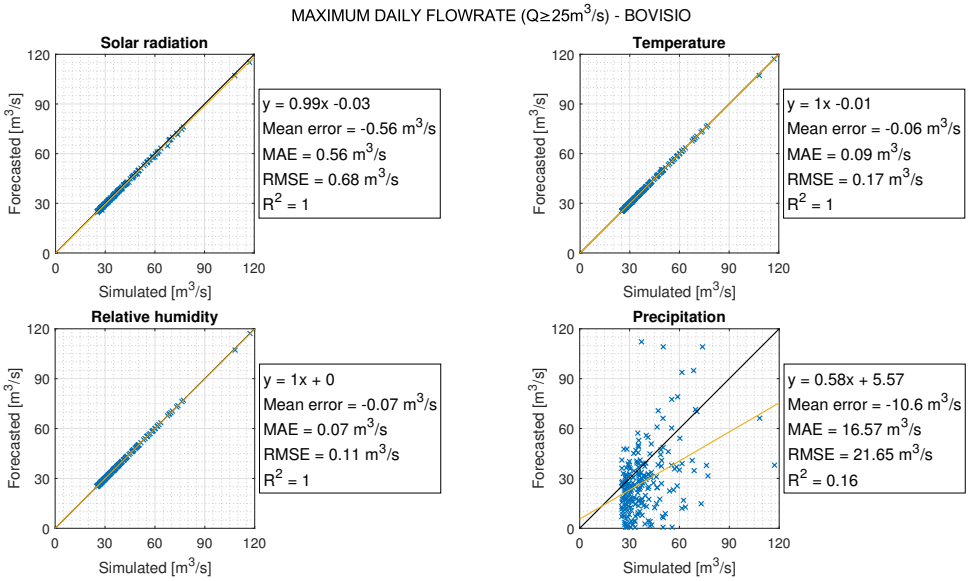


Figure 4.121: Mean daily discharge at Bovisio ($Q \geq 25\,m^3/s$) simulated vs forecasted sensitivity analysis

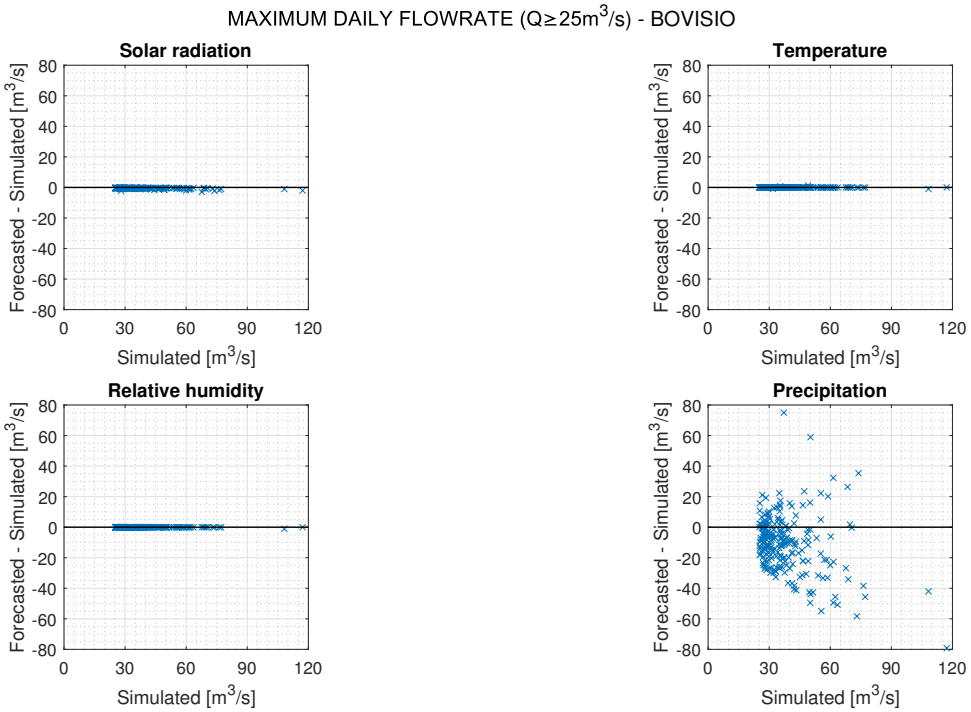


Figure 4.122: Mean daily discharge at Bovisio ($Q \geq 25m^3/s$) simulated vs forecasted-simulated sensitivity analysis

With this analysis, it can be said that solar radiation and precipitation are the variables influencing the most the prediction of the hydrological variables. Thus, the principal errors committed by MOLOCH model are the underestimation of precipitation – especially for large observed values – and the overestimation of solar radiation – which depends on the cloudiness index and the coverage -, which lead to the underestimation of the soil moisture, the strong overestimation of the potential evapotranspiration, and most important, to the underestimation of the discharge.

Chapter 5

Conclusions

In the present study a comparison analysis between observed weather data and MOLOCH forecasts, and between simulated hydrological variables obtained when forcing FEST-WB model with observed and forecasted meteorological information were performed. Concerning the meteorological variables, it was found good agreement with respect to a linear regression for temperature and solar radiation, while relative humidity of the air and precipitation have a discreet correspondence, and wind speed presents high dispersion at the hourly scale and sufficient agreement at daily scale. The solar radiation was also analysed by seasons and the lower correspondence of data was verified in winter in the hourly case, which may be explained by the presence of fog in the cold periods affecting the forecasts, however, in the mean daily and daily maximum cases the lowest agreement was present for summer with an overestimation for low observed values. Moreover, daily precipitation was also analysed at the seasonal scale, and it was found poor predictability in summer where convective precipitation is common. Conversely, better results were verified for spring and autumn – which are the rainy periods of the year corresponding to frontal episodes – and especially in winter where less precipitation is present. In respect of the simulated hydrological variables, good correspondence was verified for soil moisture and potential evapotranspiration, while discreet and sufficient agreement is present in daily and hourly flowrates, respectively. In addition, analysis at seasonal scale of discharge was also performed and it was found similar behaviour as the one of precipitation. Furthermore, comparison was also made with different forecasted datasets: the complete one, filtered from 07:00 am to 03:00 am of the next day, and filtered for the period 2019 to 2020. It was found that the agreement slightly improved in the last case, but it must be considered that less information is contained in that dataset.

Additionally, to evaluate the performance of MOLOCH model when using its forecasts as input of FEST-WB model to simulate discharges, the one-factor-at-time (OAT) methodology was applied. It was shown that solar radiation and precipitation influence the most the simulation of hydrological variables and, for instance, strong overestimation of potential evapotranspiration and underestimation of soil moisture and especially of discharge are observed.

Thus, bearing in mind that the present study is particularly interested in discharge forecasting for ungauged basins using forecasted meteorological information as input of a hydrological model, it can be concluded that – in the area of study - the underestimation of precipitation – especially large in summer – and the overestimation of solar radiation with respect to their observed values given by MOLOCH model, are the main reasons why discreet correspondence with an underestimation trend of discharge is verified. Nevertheless, it must be highlighted the good performance in terms of temperature and relative humidity, and the fact that, although solar radiation influences the discharge forecasting, it shows a high level of correspondence of the information. Therefore, it can be said that taking the first 24 hours of forecasts given by MOLOCH leads to a discharge estimation that is not sufficiently accurate, since the error committed – specifically for large flowrates - is not negligible. However, this does not mean that MOLOCH model is not suitable for the given task, instead, it must be remembered that it is a high spatial detail meteorological model with the scope of represent the convective phenomena.

Further comments are that it is worth to remember that wind speed did not affect the simulation of hydrological variables in the present study, and that problems concerning the positioning of measuring instruments could be considered by the estimation of a correction factor function of the wind logarithmic

equation. Furthermore, MOLOCH has its own forecast of soil moisture so comparison could be accomplished with simulations coming directly from observed weather data. In addition, it must be noticed that the area of study has a surface smaller than 2000 km^2 and, as a consequence, the possibility that MOLOCH meteorological forecasts may show better performance when compared with observed data in basins with different sizes should be evaluated. Finally, the present study considered only the Seveso-Olona-Lambro river basins in the period 2013 to 2020 which is characterized for the presence of urban areas, a mountainous part and a plain, therefore, further studies in basins with different characteristics – for example where snow dynamics play an important role – could be considered in order to clarify the robustness of the results here presented.

Finally, in respect of future developments, from a hydrological and civil protection point of view, actions should be evaluated in order to improve the predictability - especially for precipitation - by part of the meteorological model, or for example, by processing the forecasted data with some suitable bias correction methodologies before using it to force the FEST-WB hydrological model.

References

[1] Jun Erik Rentschler and Melda Salhab. 1.47 billion people face flood risk worldwide: for over a third, it could be devastating, 2020. Available online at: https://blogs.worldbank.org/climatechange/147-billion-people-face-flood-risk-worldwide-over-third-it-could-be-devastating, last accessed 2021-06-10.

[2] WMO. *Atlas of Mortality and Economic Losses From Weather, Climate and Water Extremes*. Number 1123. 2014.

[3] M. Verbunt, M. Zappa, J. Gurtz, and P. Kaufmann. Verification of a coupled hydrometeorological modelling approach for alpine tributaries in the Rhine basin. *Journal of Hydrology*, 324:224–238, 2006.

[4] J. Bartholmes and E. Todini. Coupling meteorological and hydrological models for flood forecasting. *Hydrology and Earth System Sciences*, 9(4):333–346, 2005.

[5] B. T. Gouweleeuw, J. Thielen, G. Franchello, A. P. J. De Roo, and R. Buizza. Flood forecasting using medium-range probabilistic weather prediction. *Hydrology and Earth System Sciences*, 9(4):365–380, 2005.

[6] Y. He, F. Wetterhall, H. L. Cloke, F. Pappenberger, M. Wilson, and J. Freer. Tracking the uncertainty in flood alerts driven by grand ensemble weather predictions. *Meteorol*, 16:91–101, 2009.

[7] Jinyin Ye, Yuehong Shao, and Zhijia Li. Flood Forecasting Based on TIGGE Precipitation Ensemble Forecast. *Advances in Meteorology*, 2016:9, 2016.

[8] F. Wetterhall, Y. He, H. Cloke, and F. Pappenberger. Effects of temporal resolution of input precipitation on the performance of hydrological forecasting. *Advances in Geosciences*, 29:21–25, 2011.

[9] Y. Xuan, I. D. Cluckie, and Y. Wang. Uncertainty analysis of hydrological ensemble forecasts in a distributed model utilising short-range rainfall prediction. *Hydrology and Earth System Sciences*, 13:293–303, 2009.

[10] Francesco Silvestro and Nicola Rebora. Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting chain. *Journal of Hydrology*, 519:1052–1067, 2014.

[11] J. E. Reynolds, S. Halldin, J. Seibert, C. Y. Xu, and T Grabs. Flood prediction using parameters calibrated on limited discharge data and uncertain rainfall scenarios. *Hydrological Sciences Journal*, 65(9):14, 2020.

[12] J. Seibert and K. J. Beven. Gauging the ungauged basin : how many discharge measurements are needed? *Hydrology and Earth System Sciences*, 13:883–892, 2009.

[13] E. Ward, W. Buytaert, L. Peaver, and H. Wheater. Evaluation of precipitation products over complex mountainous terrain: A water resources perspective. *Advances in Water Resources*, 34:1222–1231, 2011.

[14] Singaiah Chintalapudi, Hatim O. Sharif, and Hongjie Xie. Sensitivity of Distributed Hydrologic Simulations to Ground and Satellite Based Rainfall Products. *Water*, 6:1221–1245, 2014.

[15] Xue Yang, Jan Magnusson, Jonathan Rizzi, and Chong-Yu Xu. Runoff prediction in ungauged catchments in Norway : comparison of regionalization approaches. *Hydrology Research*, 49.2:487–505, 2018.

[16] I. Chawla and P. P. Mujumdar. Evaluating rainfall datasets to reconstruct floods in data-sparse Himalayan region. *Journal of Hydrology*, 588(125090):15, 2020.

[17] Amor V. M. Ines and James W. Hansen. Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and Forest Meteorology*, 138:44–53, 2006.

[18] D. Cane, S. Ghigo, D. Rabuffetti, and M. Milelli. Real-time flood forecasting coupling different postprocessing techniques of precipitation forecast ensembles with a distributed hydrological model. The case study of may 2008 flood in western Piemonte, Italy. *Natural Hazards and Earth System Science*, 13:211–220, 2013.

[19] L. Crochemore, M. H. Ramos, and F. Pappenberger. Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts. *Hydrology and Earth System Sciences*, 20:3601–3618, 2016.

[20] Jack Mcdonnell, Keith Lambkin, Rowan Fealy, Deirdre Hennessy, Laurence Shalloo, and Caroline Brophy. Verification and bias correction of ECMWF forecasts for Irish weather stations to evaluate their potential usefulness in grass growth modelling. *Meteorological Applications*, 25:292–301, 2018.

[21] Patricio Velasquez, Martina Messmer, and Christoph C Raible. A new bias-correction method for precipitation over complex terrain suitable for different climate states. *Geoscientific Model Development*, 131(Preprint):1–27, 2019.

[22] ARPA. Fiumi sicuri in Lombardia: obiettivi e azioni, 2018.

[23] Sergio Malcevschi. La Rete Ecologica della Provincia di Milano. In *La rete ecologica della provincia di Milano*, chapter Lo stato a. FrancoAngeli, Milano, 7 edition, 2007.

[24] Lisa Sacchi. Linee guida per interventi di ingegneria naturalistica lungo i corsi d'acqua. In *Linee guida per interventi di ingegneria naturalistica lungo i corsi d'acqua*, chapter Peculiarit. Angelo Guerini e Associati, Milano, 1 edition, 2003.

[25] Autorità di Bacino del Fiume Po. Linee generali di assetto idrogeologico e quadro degli interventi bacino dell'Olona, 2015.

[26] ARPA. Fiume Olona, 2018. Available online at: https://www.regione.lombardia.it/wps/portal/istituzionale/HP/DettaglioRedazionale/servizi-e-informazioni/Enti-e-Operatori/territorio/interventi-per-l-assetto-idrogeologico/fiumi-sicuri/interventi-assetto-idrogeologico-fiume-olona/interventi-assetto-idrogeologico-fiume-olona, last accessed 2021-05-19.

[27] Settore Monitoraggi Ambientali, Centro Regionale Qualità delle Acque, U.O. Monitoraggio Acque - Macroarea 2, U.O. Centro Regionale Laghi e Monitoraggio Biologico Acque Superficiali, and U.O. Risorse Idriche - Programmazione e coordinamento. Stato delle acque superficiali bacino dei fiumi Lambro e Olona, 2015.

[28] Città Metropolitana di Milano. Corso e analisi del torrente Seveso. Technical report, 2005.

[29] ARPA. Torrente Seveso, 2018. Available online at: https://www.regione.lombardia.it/wps/portal/istituzionale/HP/DettaglioRedazionale/servizi-e-informazioni/Enti-e-Operatori/territorio/interventi-per-l-assetto-idrogeologico/fiumi-sicuri/interventi-assetto-idrogeologico-torrente-seveso/interventi-assetto-idrogeologico-torrente-seveso, last accessed 2021-05-19.

[30] Alessandro Balducci, Mariella Borasio, Alberto Magnaghi, Sergio Malcevschi, and Marco Prusicki. Scenari strategici di valorizzazione delle risorse idriche per la riqualificazione del sistema ambientale e territoriale del bacino del Seveso. Technical report, Milano, 2001.

[31] ARPA. Fiume Lambro, 2018. Available online at: https://www.regione.lombardia.it/wps/portal/istituzionale/HP/DettaglioRedazionale/servizi-e-informazioni/Enti-e-Operatori/territorio/interventi-per-l-assetto-idrogeologico/fiumi-sicuri/interventi-assetto-idrogeologico-fiume-lambro/interventi-assetto-idrogeologico-fiume-lambro, last accessed 2021-05-19.

[32] Autorità di Bacino del Fiume Po. Linee generali di assetto idrogeologico e quadro degli interventi. Bacino del Lambro. Technical report, 2001.

[33] CNR-ISAC. Short description of the MOLOCH model (CNR-ISAC), 2020. Available online at: https://www.isac.cnr.it/dinamica/projects/forecasts/moloch_short_description_2012.htm, last accessed 2021-05-25.

[34] Gabriele Lombardi, Alessandro Ceppi, Giovanni Ravazzani, Silvio Davolio, and Marco Mancini. From deterministic to probabilistic forecasts: The 'shift-target' approach in the milan urban area (Northern Italy). *Geosciences (Switzerland)*, 8(5):1–14, 2018.

[35] Alessandro Ceppi. *Real time flood forecasting coupling meteorological and hydrological models*. Ph.d. thesis, Politecnico di Milano, 2011.

[36] Gabriele Lombardi. *La potenzialità di previsioni probabilistiche idro-meteorologiche sull'area urbana di Milano*. PhD thesis, Politecnico di Milano, 2016.

[37] D. Rabuffetti, G. Ravazzani, C. Corbari, and M. Mancini. Verification of operational Quantitative Discharge Forecast (QDF) for a regional warning system - The AMPHORE case studies in the upper Po River. *Natural Hazards and Earth System Science*, 8(1):161–173, 2008.

[38] G Ravazzani, D Rabuffetti, C Corbari, and M Mancini. Validation of FEST-WB, a continuous water balance distributed model for flood simulation. *31° Convegno Nazionale di Idraulica e Costruzioni Idrauliche*, page 9, 2008.

[39] Jason Mercer. Penman-Monteith and Priestley-Taylor Evaporation, 2018.

[40] A. Ceppi, G. Ravazzani, A. Salandin, D. Rabuffetti, A. Montani, E. Borgonovo, and M. Mancini. Effects of temperature on flood forecasting: Analysis of an operative case study in Alpine basins. *Natural Hazards and Earth System Science*, 13(4):1051–1062, 2013.

[41] U. Stein and P. Alpert. Factor Separation in Numerical Simulations. *Journal of the atmospheric sciences*, 50(14):2107–2115, 1992.

[42] Météo Frannce, Purdue University, NOAA/NWS Hydrometeorological Predcition center, Harold Brooks, Barb Brown, Environment Canada, Deutscher Wetterdienst, Beth Ebert, Chris Ferro, ECMWF, Johannes Jenkner, Ian Jolliffe, University of Munich, Tieh-Yong Koh, Finnish Meteorological Institute, Paul Robber, David Stephenson, The Met Office, and Research en Prévision Munérique. Forecast verification, 2015. Available online at: https://www.cawcr.gov.au/projects/verification/#FAQs, last accessed 2021-06-01.

[43] Daniel S Wilks. *Statistical methods in the atmospheric sciences*, volume 91. Elsevier Inc., 2 edition, 2006.

Appendix A - Post-processing of MOLOCH data

To create a unique homogeneous file readable by FEST-WB model a post-processing of the downloaded information was performed by means of the Climate Data Operator (CDO) program. In the following is presented a Bash script example for processing information of MOLOCH forecasts in which the following operations are made:

- Merging of grib2 files.
- Converting grib2 files into netcdf.
- Reduce the area of interest by selecting a grib box with specific coordinates.
- Calculate wind speed for u and v components.
- Set time references.
- Remove x and y variables since latitude and longitude are already present in the netcdf file.
- Remove variables that are not used in the hydrological simulation.

```
#!/bin/bash

#Merge all grib2 files in one single file
cdo cat *.grib2 unico_file.grib2

#Convert grib2 file into netcdf file
wgrib2 unico_file.grib2 -netcdf unico_file.nc

#Select this grib box with these coordinates
cdo sellonlatbox,8,12,44,47 unico_file.nc ritaglio.nc

#Calculate the wind speed for u and v components
ncap2 -O -s "windspeed=sqrt(UGRD_10maboveground^2+VGRD_10maboveground^2)"
    ritaglio.nc  ritaglio_vento.nc

#Set the time references
ncap2 -O -s "time@units=\"seconds since 1970-01-01 00:00:00\""
    ritaglio_vento.nc moloch_time.nc

#Remove x and y variables, since latitudine and longitude information are
    already present in the netcdf file
ncks -C -O -x -v x,y moloch_time.nc moloch_allvar.nc

#Remove all the variables we do not use for hydrological simulations
cdo delname,ASNOW_surface,PRMSL_surface,PRMSL_meansealevel,TCDC_surface,
    WEASD_surface,UGRD_50maboveground,UGRD_80maboveground,
    UGRD_100maboveground,VGRD_50maboveground,VGRD_80maboveground,
    VGRD_100maboveground,LAPR_surface,DPT_2maboveground,
    TSOIL_0D03underground moloch_allvar.nc moloch_netcdf_"$data_0b"03.nc
```

List of abbreviations

AMC	Antecedent Moisture Condition
ARPA	Regional Environmental Protection Agency
BOLAM	Meteorological model developed at CNR-ISAC
CDO	Climate Data Operator
CNR-ISAC	National Research Council of Italy - Institute of Atmospheric Sciences and Climate
CSNO	North-West Overflow Channel (Canale Scolmatore di Nort-Ovest)
ECMWF	European Centre for Medium-Range Weather Forecasts
FEST-WB	Flash-flood Event-based Spatially-distributed rainfall-runoff Transformation - Water Balance model
GCM	Global Circulation Model
GFS	Global Forecast System
IDW	Inverse Distance Weighting
MAE	Mean absolute error
ME	Mean error
MNW	Meteorological Network
MOLOCH	Meteorological model developed at CNR-ISAC
m a.s.l.	Metres above sea level
NCEP	National Centers of Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NWPs	Numerical Weather Predictions
OAT	One-factor-at-time
PET	Potential evapotranspiration
RMSE	Root mean square error
SCS-CN	Soil Conservation Service - Curve Number
SOL	Seveso-Olona-Lambro
TMPA	Multisatellite Precipitation Analysis
UTC	Universal Time Coordinated
WRF	Weather Research Forecasting

Acknowledgments

I would like to give thanks to CNR-ISAC for allowing us to use their meteorological forecasts in the present study and for their feedback and suggestions when partial results were presented.

Additionally, I express thanks to professors Alessandro Ceppi and Giovanni Ravazzani who let me get deeper into the hydrology world and, at the same time, to start an interest for the meteorology. I especially thank Alessandro for being of constant help, for his incredible willingness to resolve my doubts, and for showing me a proper way of accompanying a thesis' elaboration.

On the other hand, I express my gratitude to Universidad Nacional de Colombia and Politecnico di Milano for providing me the opportunity of making the master's degree. Both institutions will be fondly remembered as my academic homes.

Finally, I would like to extend warm thanks to all people who helped me to arrive here and enjoy this wonderful experience. To Loren for being an excellent friend and for showing me how to have a good time during the university experience beyond the academy. To my aunt Carmocha, my grandma Chela, my grandparents Julio and Consuelo, who were fundamental in my arrival and stay. And finally, to my parents Claudia and Mauricio, my sister Laura, and my niece Amélie, who belong to my nucleus, who are a constant support, and who encourage me to believe and to be confident in myself to conclude in the best way this journey.