Politecnico di Torino

Master's Degree in Environmental and Land Engineering

Climate Change

A.Y. 2023/2024

Exam session October/2024

Assessment of Seasonal Forecasts of Meteorological Variables and Irrigation Requirements of Maize in Pianura Padana

Thesis advisor:

Stefania Tamea

Candidate:

Stefania Oliva

Abstract

The impact of climate change on agriculture represents a crucial challenge for food and water security in the coming decades. In this context, seasonal forecasts are an essential tool for adaptation in the agricultural sector, offering valuable information for medium-term decision-making. This study, therefore, aims to analyse the uncertainty and to assess the consistency of seasonal forecasts of meteorological variables such as maximum temperature, minimum temperature, and precipitation in the Po River Basin, with a particular focus on the Pianura Padana area, where the agricultural production is concentrated. Moreover, the seasonal forecasts of irrigation requirements of maize in the area of interest will be estimated to assess the propagation of uncertainty to dependent variables.

Medium-term climate forecasts rely on slow, complex changes in Earth's components. Climate models, initialized with data on the system's state, predict climate evolution over weeks or months. "Ensemble Members" represent possible evolutions, indicating the probability of different future scenarios.

In this thesis, products from 2018 to 2022 from the ECMWF forecast centre (SEAS5 system) and those generated by two multi-model ensembles (MME and MMES) only in 2022 will be examined. The goal is to assess the variability of the Ensemble Members over time and space by comparing different initialization times (including those from March to June), various lead times (from March to September), and the differences associated with the spatial grid cells at a 1°x1° horizontal resolution. Subsequently, the errors generated by ECMWF, MME, and MMES forecasts for the entire Po Basin area will be analysed, comparing them with data from the ERA-5 climate reanalysis dataset at a 0.25°x0.25° resolution. A more detailed analysis will also be conducted on two grid cells in the Cuneo province, comparing ECMWF data both with ERA-5 data and with data downloaded from ARPA Piemonte, in order to highlight differences between forecasts in grid cells at elevations below 400m and those at elevations above 400m. Finally, using a model for estimating irrigation requirements, these needs can be calculated based on both observed and forecasted data, then compared to assess their quality.

The results show that during the summer period, forecasts for precipitation and temperature have narrower ranges of variability compared to the spring period, with a corresponding lower error range. This pattern is particularly evident for temperatures, much less so for precipitation. In general, the errors associated with precipitation are significantly higher than those related to temperatures. Specifically, errors related to ten-day cumulative precipitation show magnitudes similar to the forecasts themselves, while the average errors over three intervals within a month can be very small.

Moreover, minimum temperatures appeared to have larger errors than maximum temperatures, despite having narrower confidence intervals. However, through the estimation of bias using hindcast data and the subsequent bias correction of real-time data, it emerged that minimum temperatures showed a significant increase in accuracy, with very low error values.

From the detailed analysis of the two grid cells, it was observed that the ECMWF products are more accurate when compared to ERA5 products than to the observed data from ARPA Piemonte.

Comparing the single ECMWF model with the two multi-model ensembles, it was found that the multi-model ensemble with the largest number of members exhibited both lower errors and a greater degree of uniformity among the ensemble members for temperature, while no significant differences were found among the three models for precipitation.

Lastly, preliminary results on irrigation requirements highlight how the accuracy of meteorological forecasts can significantly impact the dependent variable, thus consequently affecting the planning and management of water resources in the study area, in particular for achieving a more resilient agriculture.

Contents

1. INTRODUCTION

1.1 Climate Change and Agriculture

The impact of climate change on agriculture represents a crucial challenge for food and water security in the coming decades. According to the Synthesis Report of the latest Climate Change Assessment Report[1] , between 2011 and 2020, the average surface temperature of the Earth increased by 1.1°C compared to the end of the 19th century, marking an unprecedented rate of warming in the past two thousand years.

The Mediterranean is a climate change hotspot, where rising temperatures are expected to affect water availability, food security, and ecosystem health [2]. These changes could lead to significant reductions and, in some cases, failures in the quantity and quality of crops, with potential consequences for agricultural production and the economy at both regional and global levels [3].

An analysis of long-term climate scenarios projected for Italian territory [4] by the CMCC research center (for the worst-case scenario RCP8.5) forecasts a 5°C increase in average temperatures by 2100 compared to the beginning of the century, along with a decrease in annual rainfall and rainy days, and an increase in high-intensity events. This will result in more drought periods and a higher risk of extreme events. Risk is defined as the interaction between hazard, exposure to damage, and the vulnerability of the affected elements, including adaptation capacity and sensitivity.

While southern regions show lower resilience levels, even northern regions are no longer immune or better prepared to face these risks. These risks involve water resources and are higher during the summer months. As a result, Italy has been classified by the Organisation for Economic Co-operation and Development (OCSE) as a country experiencing medium-high water stress. Agriculture is one of the sectors most vulnerable to the impacts of climate change in the country compared to other European regions. Although climate adaptation is an inherent characteristic of the primary sector, the scope, uncertainty, and rapidity of ongoing and expected climate changes make it necessary to enhance its adaptive capacity. This is essential not only to mitigate impacts but also to seize opportunities presented by new climate conditions. Currently, there are no available national indicators to assess the effects of climate change on agriculture [5].

Crops respond to the projected rise in average temperatures by altering the length of the growing season, advancing phenological stages, and potentially shifting cultivation areas toward higher

latitudes and altitudes, where more favourable growing conditions may be found. In the future, Italy is expected to see a decline in the productivity of spring-summer crops, especially if not irrigated. Additionally, a shift in the seasonal peak of rainfall will result in the arrival of nutrients at times when high or low temperatures may limit phytoplankton organisms, leading to a further reduction in primary productivity [6].

In Pianura Padana, climate change has had a significant impact, reducing water reserves in the Po basin and drastically altering precipitation and temperature patterns. This has affected the hydrological cycle and the availability of water for agricultural irrigation. For example, the 2018 summer heatwave showed a substantial deviation in rainfall compared to the 1981-2010 period, causing severe crop damage [7], an issue echoed during the 2022 drought emergency, which saw low precipitation and above-average temperatures, severely impacting agricultural production [8]. In the Piedmont region, using a baseline reference period (2000-2015) and a reporting period (2016-2019), it has been estimated that 1.82 million hectares of land have been degraded. Degradation refers to the reduction in the soil's capacity to be biologically productive. Specifically, changes in three sub-indicators are assessed: land cover and its changes over time, soil productivity, and organic carbon content. This phenomenon is closely linked to the impacts of climate change [5].

The most significant economic losses occur in the country's infrastructure and networks due to the worsening of landslides and in agriculture and tourism, both during the summer and winter seasons. Climate change will require significant investments in the future, offering Italy an opportunity to focus on sustainable development, recognized by the European Green Deal as the only growth model for the future. Adaptation represents both a necessity and an opportunity to respond to the effects of climate change.

1.2 Objective of the thesis

With a view to contributing in the short term to the implementation of adaptation strategies, this thesis aims to address this issue by analysing the uncertainty and assessing the consistency of seasonal forecasts of climate variables such as temperature and precipitation in the Po River basin, with particular focus on the Pianura Padana area, where the main agricultural production areas are concentrated. In the second part of the thesis, seasonal forecasts of the irrigation requirements for maize crops in the area of interest will be estimated.

To evaluate the quality of seasonal forecasts, it is necessary to compare the predicted data with observed data. However, to date, no observational data exist regarding irrigation requirements; therefore, it is necessary to study the quality of the meteorological variables used to assess irrigation needs accordingly. Specifically, the accuracy and uncertainty of the minimum temperature, maximum temperature, and precipitation variables from the ECMWF seasonal forecast center and two multi-model ensembles will be analysed against ERA-5 data for the entire study area. Additionally, a more detailed analysis will be conducted on two grid cells in the territory of the province of Cuneo, comparing ECMWF data with both ERA-5 data and data downloaded from ARPA Piemonte, in order to highlight differences between forecasts in cells at altitudes below 400m and those at altitudes above 400m. Subsequently, the observed and predicted irrigation requirements will be estimated in order to assess the propagation of error.

2. STUDY AREA

Figure 1 shows the map of the area analysed, which was created with Qgis by reprojecting the various layers on the WGS84/UTM reference system zone 32N EPSG 32632. In particular, the delimitation of the Po basin, the province of Cuneo and the Digital Elevation Model (DEM) of the area of interest are represented. The latter is used to distinguish the area with altitudes lower than 400m from the mountainous areas with altitudes higher than 400m. The area with an altitude of less than 400m, included in the polygon of the Po River basin, will be extracted later to roughly identify the Pianura Padana.
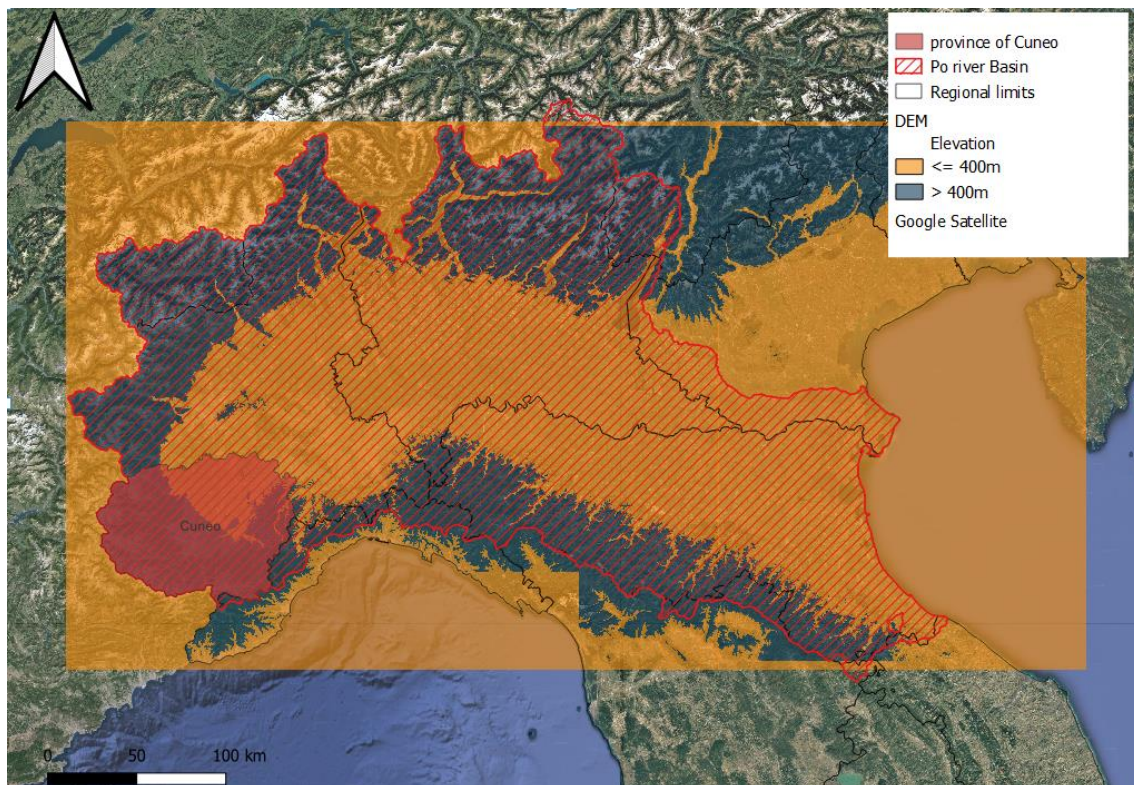
Figure 1. *Study area*

The Pianura Padana is an alluvial plain bordered by the Italian Alps and Prealps to the north and west, the Northern Apennines to the south, and the Upper Adriatic to the east, covering parts of the regions of Piedmont, Lombardy, Emilia-Romagna, Veneto, and Friuli-Venezia Giulia [9].

This area is the centre of major agricultural activities and can be further divided into the Upper Plain, Lower Plain, and Springs, each characterized by different features.

The Upper Plain includes altitudes between 200 and 400 meters, with coarse, highly permeable, and dry soil, resulting in less developed agriculture. The Lower Plain lies below 100 meters, with soil composed mainly of clays and sands, making it more fertile and suitable for intensive farming.

In the area between the Upper and Lower Plains, the more permeable soils meet the impermeable ones, leading to the phenomenon known as the resurgent springs. In more detail, water from the Upper Plain infiltrates until it encounters an impermeable layer, then flows until it can resurface from the water table, creating these natural springs[9].
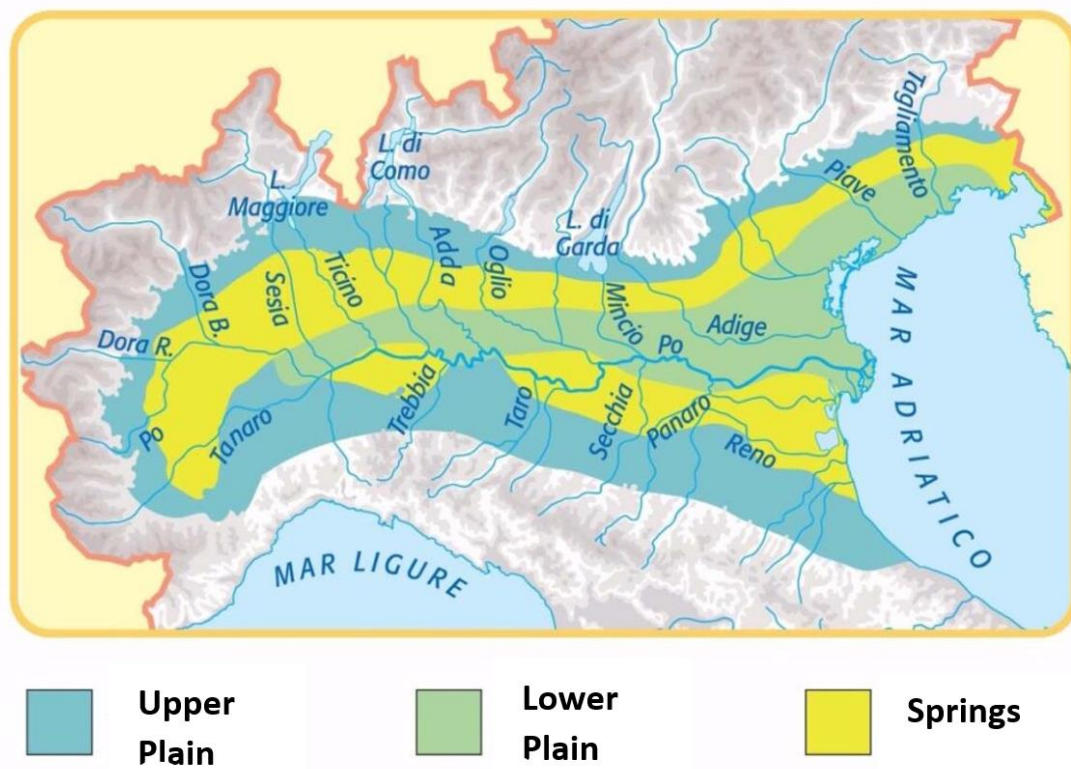

Figure 2. Pianura Padana characterization[10]

2.1 Climatic characterization of the study area

Before focusing on the climate of the area of interest, it is important to explore the general climatic framework at the national level. Italy, in fact, exhibits significant climatic variability due to several factors, including [11]:

- Its latitude span from 36°N to 47°N, which results in different solar angles received. This causes temperature differences across the territory, given other equal conditions;

- The presence of mountain ranges such as the Alps and the Apennines. These elevations influence the climate in various ways, including temperature variation, the amount of precipitation, and especially their interaction with prevailing winds, impacting the air currents over the peninsula. For instance, the Alps protect the Pianura Padana and the Veneto Plain from cold northern currents, while the Apennines limit the influence of moist western air to the Tyrrhenian side;

- The presence of the Mediterranean Sea, which exerts a strong moderating effect on the national climate, especially in coastal areas, and interacts with the air masses crossing it, encouraging the development of low-pressure systems.

To categorise the Italian climate, it was chosen to use Köppen and Geiger's 1936 classification. This is based on both temperature and precipitation on an annual and monthly time scale in order to generate six main groups: A (tropical), B (arid), C (temperate), D (cold), and E (polar). A second letter describes the climate based on seasonal precipitation, while a third letter indicates the level of warmth [12].

Italy falls within the temperate zone of the Northern Hemisphere. In particular, northern Italy exhibits a humid temperate climate with hot summers (Cfa), while the centre-south is dominated by a Mediterranean climate with dry summers (Csa).

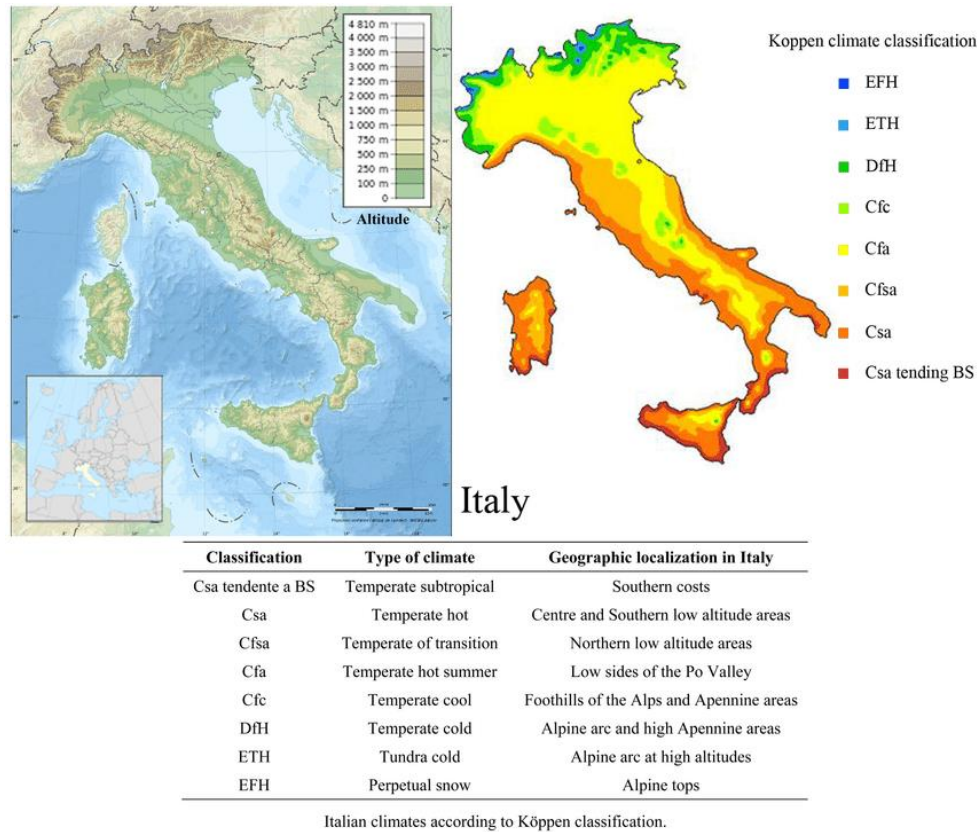A detailed representation is shown in Figure 3:

Classification	Type of climate	Geographic localization in Italy
Csa tendente a BS	Temperate subtropical	Southern costs
Csa	Temperate hot	Centre and Southern low altitude areas
Cfsa	Temperate of transition	Northern low altitude areas
Cfa	Temperate hot summer	Low sides of the Po Valley
Cfc	Temperate cool	Foothills of the Alps and Apennine areas
DfH	Temperate cold	Alpine arc and high Apennine areas
ETH	Tundra cold	Alpine arc at high altitudes
EFH	Perpetual snow	Alpine tops

Italian climates according to Köppen classification.

Figure 3. *Classification of the Italian climate according to Köppen-Geiger*

Furthermore, it is possible to distinguish climatic regions in accordance with the morphology of the territory. Thus, the morphological subdivision of Italy proposed by Guzzetti and Reinchenbach into 8 major topographic provinces is shown: 1. Alpine Mountain System, 2. North Italian Plain, 3. Alpine-Apennine Transition Zone, 4. Apennine Mountain System, 5. Tyrrhenian Borderland, 6. Adriatic Borderland, 7. Sicily, 8. Sardinia[13] (figure 4).
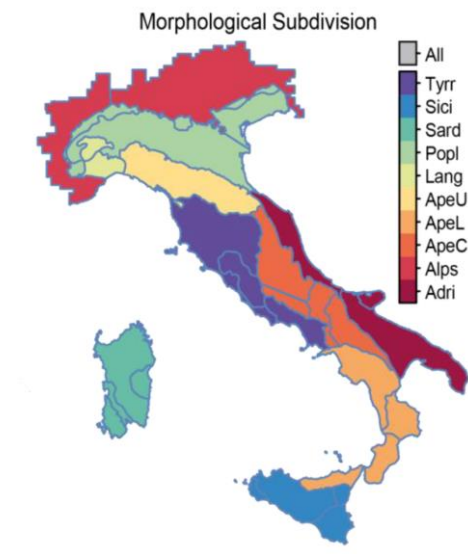
Figure 4. Morphological subdivision of Italy proposed by Guzzetti and Reinchenbach[14]

After this overview, it can be asserted that the landscape distinguishing the Po Basin is morphologically complex, characterized by the Alpine arc to the west and north, and by the Pianura Padana. This results in significant climatic and consequently hydrological variability, influencing seasonal processes.

The climate of the Alpine region is determined by its altitude, being classified as cold temperate and transitioning to nival (snow-dominated) above 2,700-2,800 meters. The Alps and Prealps are subject to substantial precipitation, reaching up to 3,000 mm per year in areas most exposed to cold polar currents and warm currents from Africa.

The climate of the Pianura Padana on the other hand, is marked by strong seasonal variation, with summer temperatures often exceeding 30°C and winter temperatures frequently falling below freezing. Precipitation is relatively modest, ranging between 600 and 800 mm per year, with a higher incidence in autumn and spring. Summer is also marked by frequent thunderstorms.

According to the Köppen and Geiger classification, the climate in this area can be classified into:

- Nival altitude climate (EFH): Found at the highest peaks of the Alps, covered by perpetual snow, generally at altitudes above 3,500 meters;

- Cold tundra altitude climate (ETH): Present in the Alpine arc at altitudes above 2,000-2,200 meters a.s.l. and on some Apennine peaks. This climate is characterized by very cold night and winter temperatures and mostly summer precipitation. The landscape gradually shifts from high-altitude grasslands to glaciers;

- Cold temperate altitude climate (DfH): Found in the Alpine arc below 2,000-2,200 meters a.s.l., including mid-altitude valleys. This climate is characterized by long, harsh winters (with an average temperature in the coldest month below -3°C) and slightly dry conditions, with predominantly summer precipitation;

- Temperate climate with hot summer (Cfa): The maximum altitude generally reaches 400 meters in upper Piedmont. This area experiences two peaks of precipitation, one in spring and one in autumn, and two periods of minimal precipitation, one in winter (usually in January) and one in summer (July or August). Summer tends to be wetter than winter in areas north of the Po, while in the sub-Apennine lands and in the province of Cuneo, the difference evens out;

- Temperate climate with mild summer (Cfb): This climate is typical of the hills and plateaus of Piedmont above approximately 400 meters a.s.l. in the north and 500 meters a.s.l. in the south, the Ligurian and Tuscan-Emilian Apennines, and the lower Alpine and Prealpine slopes, excluding the more internal ones;

- Cool temperate climate (Cfc): This climate is typical of high-altitude Prealpine and Apennine areas. Precipitation can be substantial, especially during intermediate seasons, but also abundant during the summer.

2.2 Agriculture of the area

According to ISTAT data, the Pianura Padana is one of the most productive agricultural regions in Italy, with a wide variety of crops, including cereals, potatoes, sugar beets, and fodder, as well as fruit trees, contributing about 40% of Italy's agricultural GDP. In particular, the province of Cuneo is a significant example of this productivity.

In the second part of the thesis, as previously mentioned, we will focus on estimating the irrigation requirements for maize cultivation in this area. Maize, originally from tropical regions, requires temperatures of at least 12°C for sowing and is at risk of irreversible damage below 5°C. Temperatures above 33°C cause stress to the plant, especially under water-scarce conditions. Maize needs abundant water, with regular rainfall or artificial

irrigation, which is why its cultivation is concentrated in areas with a moderate climate and sufficient water resources, such as the Pianura Padana. Here, the eastern part has benefited from summer rainfall, while in the central and western areas, irrigation has always been essential. Therefore, maize cultivation is limited in hilly areas, reduced in central Italy, and almost absent in southern Italy [15].

In the following figure, the distribution of maize cultivation in the north-Italy is illustrated. The data was downloaded from the Crop Calendar Dataset provided by the Center of Sustainability and the Global Environment at a spatial resolution of 5' in NetCDF format. The data corresponds to the year 2000, updated to 2010[16].
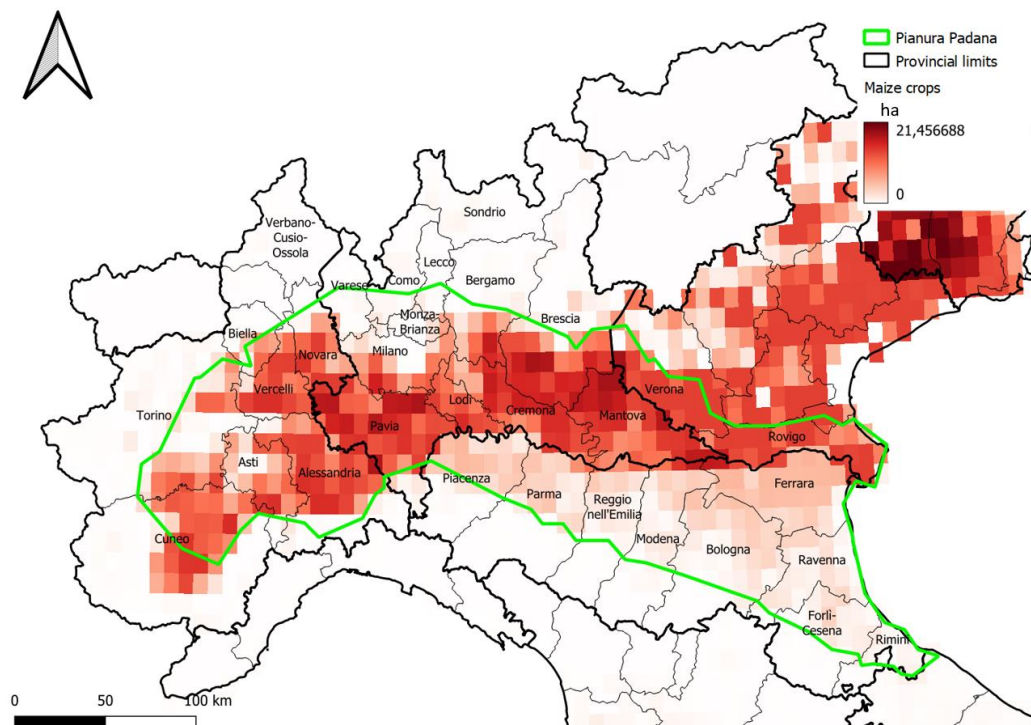

Figure 5. Distribution of maize crops in Pianura Padana

According to the ISTAT report on agricultural crops in 2021, between 2010 and 2020 the cultivated area of maize in Italy decreased from 26.7 % to 20.1 % for a total of 603,000 hectares, remaining however the second most cultivated cereal in Italy after durum wheat[17].

In particular, using the dataset of the *7° Censimento dell'agricoltura* by ISTAT [18], it was possible to determine, for each province, the area of agricultural land used for maize cultivation within the study area. The results are summarized in the following table.

The total area designated for maize cultivation within the green polygon was estimated to be approximately 438,183 hectares.

Table 1. *Area of agricultural land used for maize cultivation*

Region							
Piedmont		*Lombardy*		*Emilia-Romagna*		*Veneto*	
Province	*Hectars*	*Province*	*Hectars*	*Province*	*Hectars*	*Province*	*Hectars*
Alessandria	21,075.00	Bergamo	11,840.00	Bologna	8,787.00	Verona	26,670.00
Asti	8,128.00	Brescia	44,103.00	Ferrara	26,734.00	Vicenza	15,606.00
Biella	1,635.00	Como	2,094.00	Forlì-Cesena	887.00	Padova	31,986.00
Cuneo	44,903.00	Cremona	35,527.00	Modena	6,899.00	Treviso	20,952.00
Novara	8,962.00	Lecco	999.00	Parma	4,908.00	Venezia	31,091.00
Torino	52,419.00	Lodi	16,902.00	Piacenza	12,393.00	Rovigo	30,895.00
Verbano-Cusio-Ossola	129.00	Mantova	38,020.00	Ravenna	5,232.00	Belluno	1,523.00
Vercelli	10,467.00	Milano	16,128.00	Reggio nell'Emilia	5,724.00		
		Monza-Brianza	2,346.00	Rimini	263.00		
		Pavia	19,696.00				
		Sondrio	908.00				
		Varese	1,293.00				
Tot	147,718.00	**Tot**	189,856.00	**Tot**	71,827.00	**Tot**	158,723.00

3. INPUT DATA

As previously mentioned in the thesis objective (chapter 1.2); in order to calculate seasonal forecasts of irrigation requirements and their quality, it is necessary to start with the variables of precipitation, maximum temperature, and minimum temperature.

In this chapter, the concept of seasonal forecasting, the existing models and those used for the current study, and the input data used to compare and evaluate the quality of the forecast data will be analysed.

3.1 Seasonal forecast introduction

Seasonal forecasts, which provide a medium-term view of climatic variations, are an essential tool for adaptation in the agricultural sector, offering valuable information for medium-term decisions such as sowing planning, selection of the most suitable crop varieties, fertilisation scheduling and irrigation management [19], achieving significant economic savings as a result of targeted actions taken with the help of seasonal forecasts [20].

They are based on predictable changes in the slowly evolving and interacting components of the planet, providing a medium-term view of climate variations over weeks or months in the Earth system.

For example, ocean surface temperatures (SST) tend to vary gradually over weeks or months. Since the ocean influences the overlying atmosphere, changes in its properties, such as temperature, can alter atmospheric conditions both locally and remotely. These changes from 'normal' atmospheric conditions are the basis for medium to long-term forecasts, such as seasonal forecasts [21]. However, to fully grasp the complex and non-linear dynamics that govern these interactions, it is necessary to rely on advanced climate models.

Indeed, climate models are an essential tool for predicting climate evolution over longer periods, overcoming the limitations of traditional weather forecasting, which is less effective on extended time scales.

Even though the most intense interactions between ocean surface temperatures (SST) and seasonal atmospheric climate occur mainly in the tropics, seasonal forecasts are still very useful for indicating likely climate conditions in Europe.[22]

Due to the chaotic nature of the atmosphere, which generates large uncertainties, and the complex and non-linear interactions between the various elements, these models are used to provide more accurate predictions in the medium-long term. They begin with data that capture the system's initial state at the forecast's starting point and are subsequently employed to simulate its progression over time by a distribution of results [23]. The models, in fact, are composed of 'Ensemble Members', which individually are complete descriptions of the evolution of climate variables over time, but which collectively, indicate the probability of a series of future weather scenarios [24].

These models integrate a wide range of components of the Earth system, including the various layers of the atmosphere, ocean, sea ice and land surface. The uncertainties associated with these models can be distinguished mainly into two categories:

- Uncertainties related to the approximate knowledge of initial conditions: these uncertainties are often handled through the use of several ensemble members, each starting from slightly different initial conditions than the current 'recorded' state;

- Uncertainties associated with model selection and its relative approximations: these arise from the inherent limitations in faithfully representing the complexity of the Earth system.

These sources of uncertainty reflect the complexity of the Earth system and affect the accuracy of the models' seasonal forecasts.

In the context of these seasonal forecasts, it is important to familiarize with some specific technical terms used in the drafting of the paper, such as:

- Ensemble Member: as previously mentioned, it is a single prediction simulation within a group of predictions made by a model. In figure 6, there is an example of ensemble members. The ensemble correspond to the total distribution of results defining the probability of a given event;
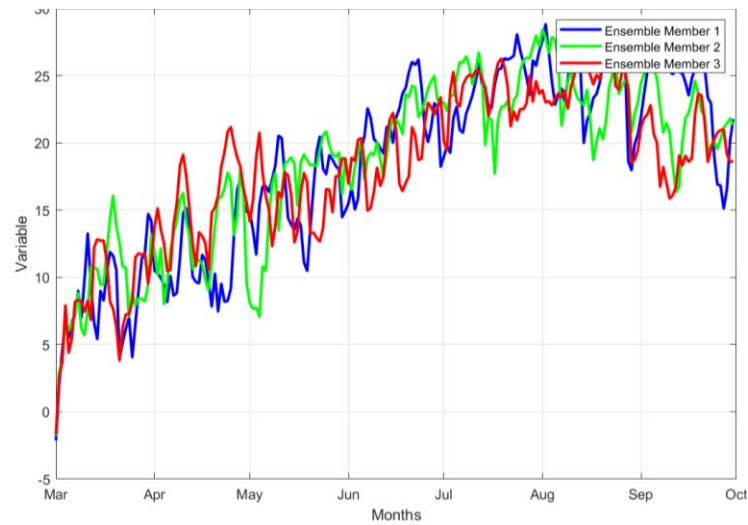
Figure 6. *Example of ensemble Members*

- Starting time: initial point from which the climate prediction model starts. This is the time reference point for the initialisation of the model and determines the start of the forecast period;

- Lead time: time interval between the moment a forecast is issued and the period to which the forecast refers. When we speak of lead time 0, we therefore refer to the month in which the forecast is initialised. When we speak of lead time 1, we refer to the next month after initialisation and so on (example: if the month from which the forecast starts is March, lead time=2 corresponds to the month of May);

- Forecast Range: time interval for which a forecast is provided. In other words, it represents the period covered by the forecast itself.
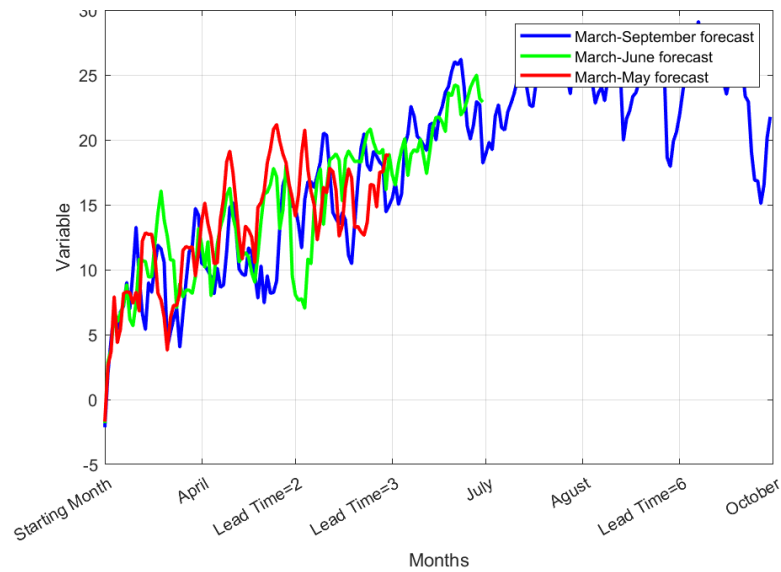
Figure 7. *Key words in the context of seasonal forecasts*

During the drafting of the paper, we will frequently compare both the results of forecasts initialized in a given month on different lead times and forecasts made with a single target forecast month but initialised in different months, in order to identify the best skills in the various cases, if any.

3.1.1 Available models

In 2006, the World Meteorological Organisation (WMO) launched an initiative that is an essential pillar of the WMO Global Data Processing and Forecasting System (GDPFS). Specifically, the WMO Global Producing Centres for Long-Range Forecasts (GPCs-LRF), responsible for the production of global seasonal forecasts, and WMO Standard Verification System for Long-Range Forecasts (SVSLRF), a set of standard verification measures, were designated[25].

There are officially 14 global production centres for long-term forecasting (GPC-LRF) identified by the WMO. Some of the most important are listed below:

1. ECMWF (European Centre for Medium-Range Weather Forecasts) – Europe;
2. JMA (Japan Meteorological Agency) – Japan;
3. UKMO (UK Met Office) - United kingdom;
4. Météo-France – France;
5. CMCC (Centro Euro-Mediterraneo sui Cambiamenti Climatici) – Italy;
6. DWD (Deutscher Wetterdienst)- Germany.

Each production centre issues forecasts at a certain frequency, spatial and temporal resolution, with a different time horizon among centres and different variables available. In addition, the different models also have different ways of creating Ensemble Members, which will be analysed in more detail in the following sections.

A convenient way to quantify the effect of the approximations from the individual models is to combine the results obtained, developed, initialised and run independently to create a multi-model system.

As pointed out in the research "Temperature and precipitation seasonal forecasts over the Mediterranean region: added value compared to simple forecasting methods"[26], during the summer period, precipitation and temperature forecasts obtained by averaging different models show a significant correlation with observations, outperforming many forecasts made by individual systems, thus being more accurate and reliable. In particular, temperature forecasts prove more accurate and consistent than precipitation forecasts. However, verification methods must be implemented in order to assess the quality of seasonal forecasts as a function of the area and period considered.

3.1.2 Hindcasts and Real time forecast

Seasonal forecasts, as previously mentioned, start from the current observed state of the system, which has then evolved over several months through model integration. These initial observations, however, are imprecise and therefore contain errors that grow during model integration, reaching magnitudes comparable to those of the predictable signals. These errors can be categorised into:

- casual: which are taken into account through the use of ensemble members;
- systematic: which, if determined, could allow these to be corrected in the predictions themselves.

Systematic errors can be determined by comparing retrospective forecasts (hindcasts) with observations. Hindcasts are performed on past events using historical data, which are used to set initial conditions to be realistic. During the simulation, historical data up to the point of interest are used to calibrate and validate the model. Finally, systematic differences between model and observation are identified as biases in order to assess the model's accuracy by pointing out discrepancies between simulation and observed data that can then be removed in order to improve the model's real time prediction results. This is then done through bias correction of future predictions[27].

Each forecast system that is part of the Copernicus Climate Change Service (C3S) system is characterised by a hincast period and relative ensemble size and a real time forecast period and relative ensemble size as shown in the table. In addition, several techniques are used to sample the uncertainty of the initial conditions, and thus to construct the ensemble members:

- 'burst': the initial conditions of all ensemble members refer to the same start date but with slightly perturbed states;

- 'lagged': the initial conditions of ensemble members refer to different start dates with sufficiently small perturbations [27].

Table 2. *Seasonal forecasting centres and characteristics of each*

Feature / Forecast Centre	Hindcast period	Hindcast ensemble size	Hindcast initial condition	Real time forecast period	Real time forecast ensemble size	Forecast initial condition
ECMWF	1981-2016	25 members	1st of month	2017-2024	51 members	1st of month
Météo-France (MF)	1993-2018	12 members	last and penultimate Thursday of previous month,1st of month	2017-2024	25 members	last and penultimate Thursday of previous month, 1st of month
Deutscher Wetterdienst (DWD)	1993-2019	30 members	1st of month	Nov 2018-2024	50 members	1st of month
The Met Office (UKMO)	1993-2016	7 members/ start date	1st, 9th, 17th, 25th of month	2017-2024	4 members/ day	each day of month
Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC)	1993-2016	40 members	1st of month	Nov 2018-2024	50 members	1st of month

3.2 Prediction models used for the study

For the final paper, the C3S dataset provided by Copernicus Climate Change Service will be used. This is a multi-system forecasting service that collects, processes and combines daily data produced by forecast centres from various European countries at a spatial resolution of 1 degree by 1 degree on individual levels. The entities involved are The Met Office(UKMO), European Center Medium Weather Forecast(ECMWF), Euro-Mediterranean Centre on Climate Change (CMCC), Deutscher Wetterdienst (DWD), Météo-France(MF), NCEP, the Japan Meteorological Agency (JMA) and Environment and Climate Change Canada (ECCC)[28].

More specifically, the real-time products from 2018 to 2022 of the ECMWF forecast centre SEAS5 system and those constructed by two multi-model MME ensembles on 2022 alone will be examined in order to compare the uncertainties associated with the three.

Various initialisation times, including those from March to June, will be examined in order to analyse the predictive capabilities of the meteorological variables at a maximum lead time of seven months. These variables will then be used to reconstruct crop evapotranspiration and to estimate the irrigation requirements from May to September, which is the growing season for maize crops in the Pianura Padana. The analysis will be based on daily forecasts of maximum and minimum temperature and cumulative precipitation at daily, ten-day and monthly intervals. The results will then be compared and evaluated with observational data to highlight associated uncertainties. It is expected that accuracy and reliability will increase with time series monitoring, as indicated in the study "Seasonal climate forecast can inform the European agricultural sector well in advance of harvesting". This study, which examines wheat cultivation in different European regions, also shows a progressive loss of memory as the number of forecast months increases and a greater predictive capacity as the spatial resolution increases [19]. The products obtained will then be evaluated with the ERA5 reanalysis and ARPA Piemonte observation products for two specific grid cells at a resolution of 0.25°x0.25°. The reference period studied will be from 2018 to 2022, thus considering very different years in terms of temperature and precipitation and therefore impacts on agriculture.

3.2.1 ECMWF SEAS5- How ensemble members are constructed

The thesis work will therefore focus in abundance on the evaluation of the variables predicted by the fifth forecast system (SEAS5) launched by the single ECMWF centre in

2017 to produce real-time forecasts. For this reason, it is important to fully understand how the ensemble members of this model are constructed.

Seasonal forecasts are constructed through the use of numerical models that combine hydrodynamic equations, which describe the evolution of the planet's main components (atmosphere and ocean), taking into account a wide range of parameterizations[29].

SEAS5 implemented some updates to the previous 'system 4', which covered the period from 2011 to 2017. In particular, it is worth emphasising the importance of the initial conditions that need consistency in order to avoid forecast invalidity. As demonstrated by Andrej Ceglar's study [30], seasonal climate forecast initialised with realistic land-surface achieves better skill than climatological land-surface initialization. The fundamental issue with using realistic conditions is the different horizontal resolution between reanalysis data and real-time data. This issue was significant in System 4, but was resolved in System 5 through the offline recalculation of the initial conditions of the land surface, at the required resolution and with modified forcing factors[31].

In addition, the initial atmospheric conditions for the construction of the 25 ensemble members of the retrospective forecasts (1981-2016) come from the ERA-Interim dataset, while for the 51 members of the real-time forecast (2017 onwards) from ECMWF operational analyses, through the combination of perturbations of the SST and atmospheric initial conditions and the activation of stochastic physics.

SEAS5 includes prognostic ozone and requires initial ozone conditions. However, the ERA-Interim ozone data are not entirely reliable because they are affected by changes in satellite instruments over time, causing errors in stratospheric temperatures. For this reason, instead of using the "problematic" data, SEAS5 uses a seasonal climatology (a seasonally variable average ozone profile) derived from a specific model that is "forced" with ERA-Interim data. This avoids introducing initial errors while maintaining a realistic ozone evolution during the forecast.

Ensemble member 0 is initialized with unperturbed atmospheric initial conditions. The initial conditions for all other ensemble members have perturbations applied to certain fields (including upper-air fields and a limited set of soil moisture, soil temperature, snow, sea ice temperature, and surface temperature) to account for the uncertainty in the initial atmospheric state due to observation system limitations. Small initial errors can lead to significant later errors due to the chaotic nature of the system; therefore, it is necessary to

consider a set of ensemble members. Perturbations are applied to both the atmospheric initial conditions and a limited set of land surface variables, such as soil moisture and snow temperature. This methodology ensures that the ensemble explores a realistic range of possible future evolutions, improving the robustness of the forecasts.

Since the perturbations applied to the real-time forecast were not available for previous years in the hindcast period, the 2015 perturbations were retroactively applied to all years of historical re-forecasts (hindcasts) to preserve consistency[29].

This ensures that the entire re-forecast dataset is consistent, allowing for more reliable comparisons between current and historical forecasts.

Finally, it is important to highlight the significance of hindcast data. Seasonal forecast models themselves contain errors, also known as "biases," which can vary depending on the region, lead time of the forecast, or area of interest. In this context, hindcast data are an excellent tool for estimating biases, which are then expected to have the same magnitude in real-time forecasts. As demonstrated in the article "Skill decreases in real-time seasonal climate prediction due to decadal variability," real-time forecasts may exhibit lower predictive skill compared to hindcast-based forecasts. This could be due to the difficulty in adequately representing the decadal variability of ocean-atmosphere teleconnections [32]. Although this thesis will not include a direct comparison between the predictive capabilities of the two types of forecasts, it is reasonable to assume that the real-time data analysed may have gaps or limitations.

3.2.2 Multi-Model Ensembles

For a better interpretation and evaluation of the seasonal forecasts, it was decided to also consider blended forecast systems. This analysis will be carried out only for the year 2022, as no other years of forecasts shared by all four forecasting centers are available on Climate Data Store for the 2018-2022 period [28].

More specifically, the two multi-model ensembles are defined as follows:

- Multi-Model Ensemble (MME): This includes all available Ensemble Members from the forecasting centres Météo-France (MF), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Euro-Mediterranean Centre on

Climate Change (CMCC), and The Deutscher Wetterdienst (DWD), for a total of 202 members;

- Multi-Model Ensemble Small (MMES): This includes only ten members from each forecasting centre (MF, ECMWF, CMCC, DWD), randomly selected from all available members for each centre, for a total of 40 available members.

The following table provides the technical details of each forecasting centre considered, such as the forecasting system and the number of ensemble members taken from each to build the two multi-ensembles.

Table 3. Forecasting centres considered for the construction of MME and MMES

Forecast Centre \ Feature	Prediction System	Institution	Ensemble size for MME	Ensemble size for MMES
MF	System8	Météo-France	51 members	10 members
ECMWF	SEAS5	European Centre for Medium Range Weather Forecasts	51 members	10 members
CMCC	SPS3.5	Centro Euro-Mediterraneo sui Cmabiamenti Climatici	50 members	10 members
DWD	GCFS 2.1	Deutscher Wetterdienst	50 members	10 members
MME			202 members	
MMES				40 members

The chosen forecasting centres exhibit technical differences in model construction, including forecast initial conditions, time range forecasts, and the original resolution of the models.

As previously explained, there are different techniques for generating forecast members, such as the 'burst' and 'lagged' methods. The ECMWF, DWD, and CMCC members use the 'burst' approach, meaning they are all initialized on the same date (specifically, the first of each month) with slightly perturbed initial conditions to account for observational uncertainty. MF, on the other hand, follows a 'lagged' approach, where members are initialized on different dates but with more similar starting conditions. Specifically, 25 members are initialized on the second-to-last Thursday of the month before the forecast begins, another 25 on the last Thursday of that month, and one member on the first day of the forecast month.

Regarding the time range, ECMWF covers a forecast period of up to 215 days, MF covers seven calendar months, while DWD and CMCC cover six calendar months. Therefore, since the forecast period analysed in this thesis spans from March to September (7 calendar months), for the forecast with a March initialization month, only the period from March to August will be analysed for the two Multi-Model Ensembles, as forecasts from DWD and CMCC are not available.

Finally, the original resolution of the individual models varies, with some even having better resolution than that of the C3S Climate Change Service. Nonetheless, Copernicus Climate Data Store provides the data at the same spatial resolution of 1°x1° to ensure comparability and data uniformity.

3.3 Reanalysis and Observation data

To assess the forecast error, it is essential to compare forecast data with reanalysis and/or observation data.

For this thesis, the decision was made to compare ECMWF, MME, and MMES data for the entire Pianura Padana with the ERA-5 climate reanalysis dataset developed by the European Centre for Medium-Range Weather Forecasts (ECMWF). For a more detailed analysis of two cells in the Cuneo area, ECMWF data will be compared both with ERA-5 data and with data downloaded from ARPA Piemonte.

This choice was made because obtaining observation data for the entire studied area might have been challenging. By proceeding this way, the aim is to ensure temporal consistency, as ERA-5 data are available daily for the entire considered period and area.

3.3.1 ERA-5 data

Specifically, ERA-5 data were downloaded from Climate Data Store at single levels, with an hourly temporal resolution and a spatial resolution of 0.25°x0.25°, in NetCDF format. The variables considered were maximum and minimum temperatures and cumulative precipitation, for the period from March to September for the years 2018 to 2022. To ensure a comparison between forecast data and ERA-5 data, the hourly maximum and minimum temperature data were averaged daily, while daily cumulative precipitation was extracted from the total cumulative precipitation, and subsequently, the ten-day or monthly totals were calculated depending on the type of analysis conducted.

Furthermore, given the resolution differences between forecasts and reanalysis data (1° vs 0.25°), a downscaling process was performed by replication and averaging at the edges. Specifically, to preserve valuable information from the higher-resolution data, seasonal forecast data from various models at 1°x1° resolution were converted to 0.25°x0.25° resolution, reflecting the original values from the 1° cell onto the higher-resolution cell. At the borders between two or four 1° cells, due to the lack of exact correspondence between 0.25° grid cells and a single value, the average of the adjacent cells was calculated, as shown in the example in Figure 8.
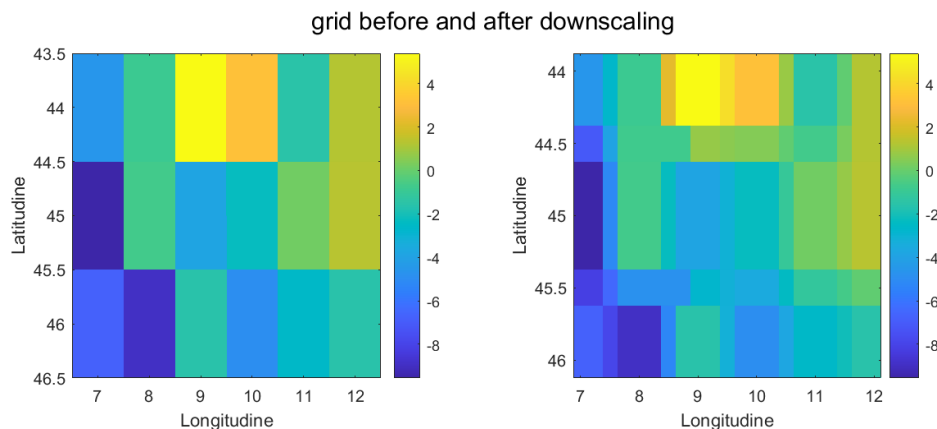
Figure 8. *Downscaling for replication from 1° resolution grid to 0.25° resolution grid*

It should be noted, however, that this operation has resulted in the analysis losing the yellow band that is not shared between the highest resolution grid (red grid) and the lowest resolution grid (grey grid) as shown in the figure 9. This operation did not significantly compromise the study, as we are more interested in the lowland areas, where most of the

crops are concentrated. The yellow band, in fact, mainly concerns areas with altitudes greater than 400m.
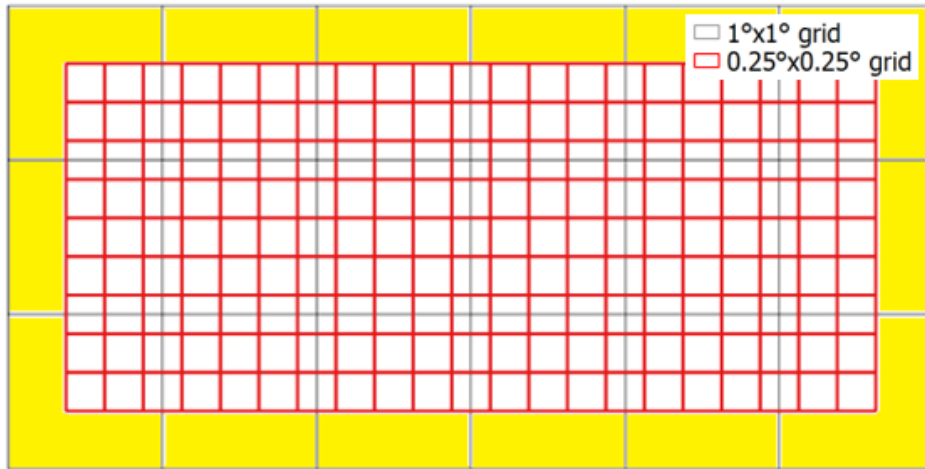
Figure 9. *0.1°x0.1° resolution grid vs 0.25°x0.25° resolution grid*

3.3.2 ARPA Piemonte data

As previously mentioned, for two grid cells with a resolution of 0.25°x0.25° in the territory of the province of Cuneo, a more accurate analysis of the quality of the forecast data was desired. For this reason, for the cells at latitude 44.5°N and longitude 7.25°E (cell 1) and latitude 44.75°N and longitude 7.75°E (cell 2) respectively, the forecast data of the ECMWF centre were compared with both the ERA5 reanalysis data and the average observation data of four meteorological stations belonging to the two cells (two for each cell) equipped with rain gauge and thermometer.
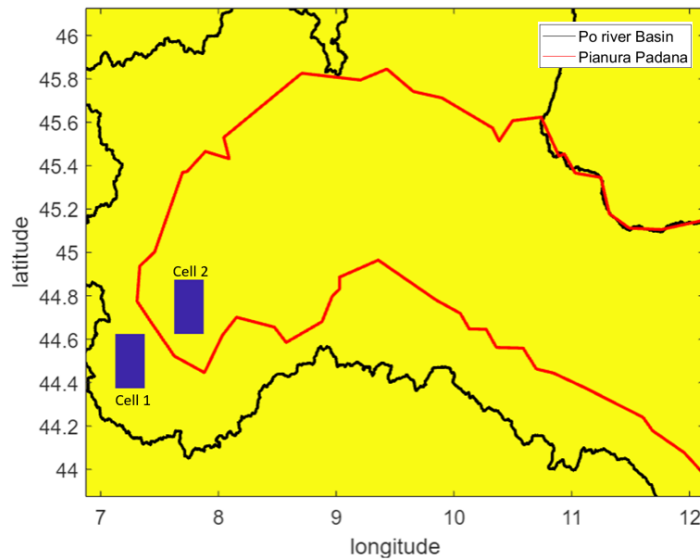
Figure 10. *Considered cells for a punctual analysis*

These cells were chosen to compare the forecasting behaviour in a predominantly mountainous area with that in a lowland area.

The meteorological stations considered for cell 1, i.e. in the area at altitudes greater than 400m, are those of:

- Brossasco: longitude=7.36°E, latitude=44.56°N, altitude=580m;
- San Damiano Macra: longitude=7.25°E, latitude=44.48°N, altitude=1095m.

For cell 2, on the other hand, at altitudes lower than 400m, data recorded by the meteorological stations of:

- Bra: longitude=7.85°E, latitude=44.69°N, altitude=290m;

- Marene: longitude=7.73°E, latitude=44.66°N, altitude=310m.

Subsequently, the average altitude of the two cells was found using the Qgis tool 'Raster Layer Statistics' and then compared with the average altitude of the two corresponding stations:

- $Altitude_{Cell1}$=1290.37m > Average altitude of the two weather stations = 837.5m;
- $Altitude_{Cell2}$=268.59m < Average altitude of the two weather stations = 300m.

In the case of cell number 1, the difference is much greater than 100m. Therefore, it was appropriate to consider the linear relationship between altitude and temperature to 'adjust' the temperatures recorded by the two stations. In the end, the following procedure was used to obtain a single temperature observation value:

$$T_{1obs} = T_{Brossasco} - b * (z_{cella} - z_{Brossasco})$$

$$T_{2obs} = T_{SanDamianoMacra} - b * (z_{cella} - z_{SanDamianoMacra})$$

$$T_{obs} = \frac{(T_{1obs} + T_{2obs})}{2}$$

where b=6.5°C/km, corresponding to the temperature gradient used to compare values at different altitudes in the troposphere [33].

This was done for the maximum, minimum and mean temperatures of cell 1.

For cell 2, on the other hand, given a difference between the average altitude of the cell and the average of the altitudes of the two stations of less than 100m, the average of the values of the variables recorded by the stations was taken directly as:

$$T_{obs\_cella2} = \frac{(T_{Bra} + T_{Marene})}{2}$$

Finally, precipitation in both cells was found as the average of the precipitation values recorded by the two rain gauges for each cell.

$$P_{obs\_cella1} = \frac{(P_{Brossasco} + P_{SanDamiano\text{Macra}})}{2}$$

$$P_{obs\_cella2} = \frac{(P_{Bra} + P_{Marene})}{2}$$

4. DATA ANALYSIS

As previously described, forecasting models provide predictions through the use of ensemble members. In this chapter, we will analyse the data provided by the single ECMWF model for the years 2018 to 2022, as well as the forecasts composed from the union of the MME and MMES for the year 2022. The aim is to highlight the variability, any trends and/or anomalies present in the models and to identify differences between single and multi-model ensembles.

Specifically, the analysis will cover both temporal and spatial aspects, seeking significant spatial differences at specific lead times and starting months, as well as temporal trends by considering the spatial average over time of the temperature and precipitation variables. In this chapter, for the spatial resolution analysis, we will use the original resolution of the data downloaded from Climate Data Store (1° x 1°), for a total of 18 cells considered.

4.1ECMWF analysis

4.1.1 Temperature

- *Ensemble members representation*

To graphically visualize the 51 ensemble members of the ECMWF model and analyse their variability and trends, the average of the 18 cells considered in this study was calculated. This reduced the matrix from four dimensions (longitude, latitude, ensemble members and time) to two dimensions (ensemble members and time). Subsequently, the average of all ensemble members was calculated, considering all starting months for each year from 2018 to 2022. In the following figures, all ensemble members of the maximum and minimum temperature variables have been represented, with starting months of March, April, May, and June, and target forecasts extending to September, along with the respective ensemble means for the years 2018, 2019, 2020, 2021, and 2022.
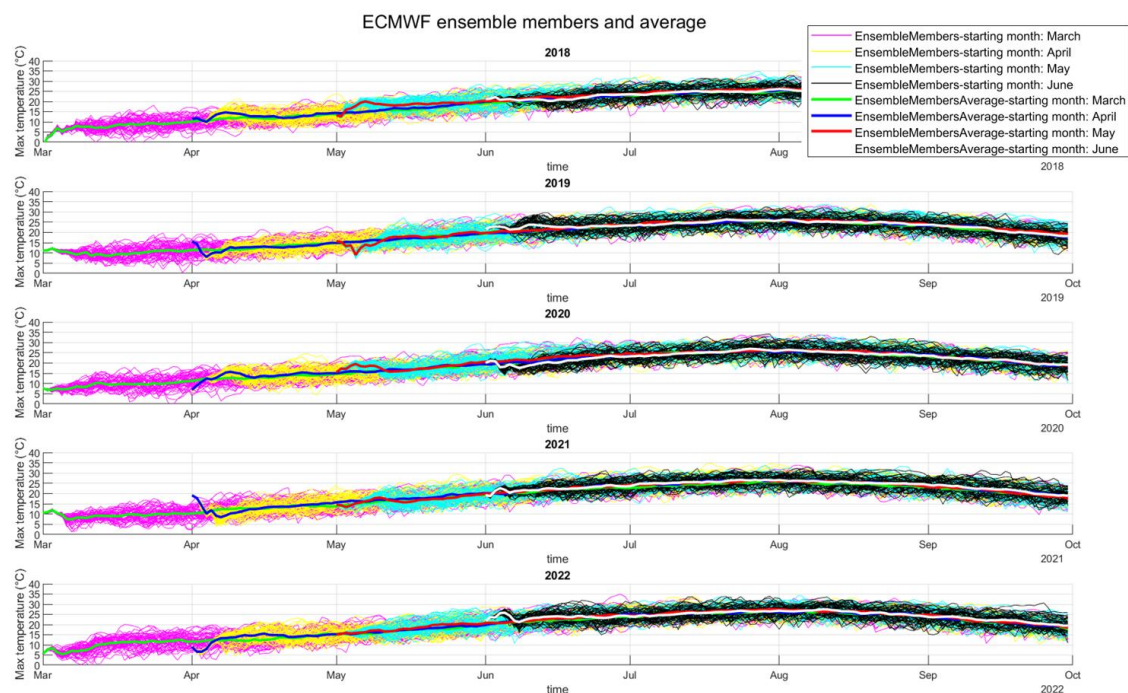
Figure 11. *Ensemble members representation and their average considering four months of initialisation and five years for Maximum temperature*
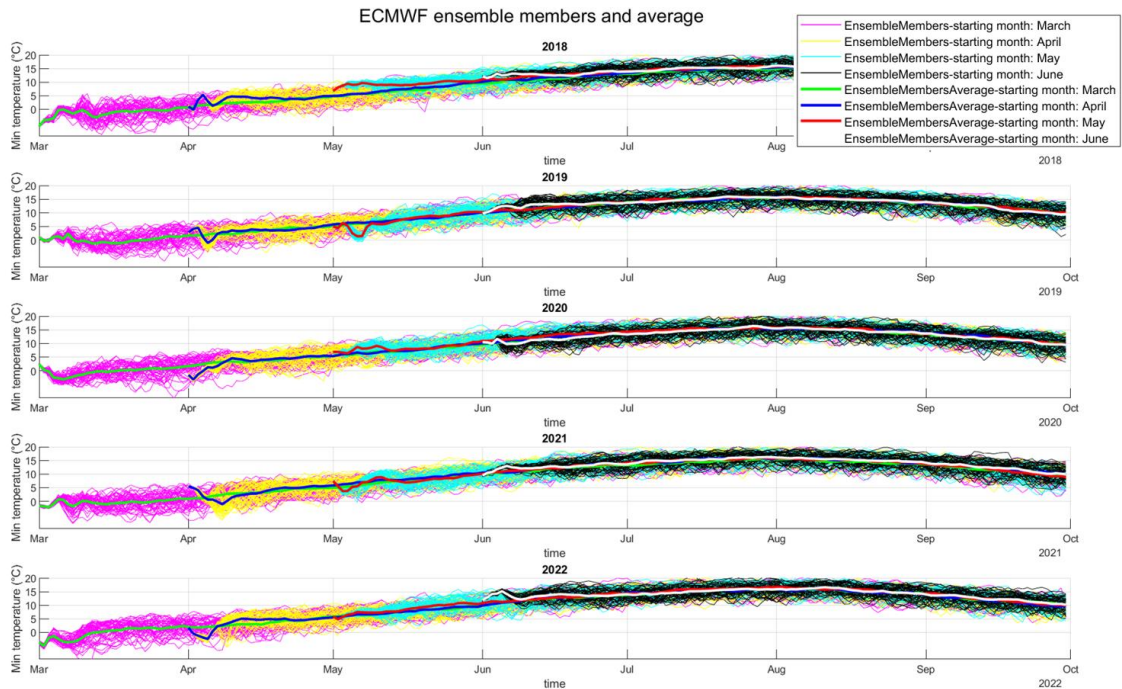
Figure 12. *Ensemble members representation and their average considering four months of initialisation and five years for Minimum temperature*

These graphs are of fundamental importance, as they provide valuable information. It is evident that at each starting month, the ensemble members exhibit considerable agreement, showing a highly variable average behaviour across the first days. However, in certain cases, such as with maximum temperatures in various years with an initialization month of April, anomalies are observed, characterized by strongly ascending or descending trends in the early days of the forecast. This, as previously discussed in Section 3.2.1, may be attributed to the method of constructing the ensemble members, which could introduce initial errors.

For each forecast, even after 10-15 days, the ensemble members provide highly variable information, but eventually converge to a more linear average trend. It is evident that the averages of forecasts initialized in March, April, May, and June for the months of July, August, and September are very similar, with a stable and increasing trend in July and a decreasing trend in August and September, showing no significant fluctuations. This behaviour is due to the loss of memory of the initial conditions in the forecasting models as the lead time increases.

This phenomenon is also noticeable when comparing different years. Comparing the same starting month over the years, the forecast values appear to be distinct throughout the lead

time equal to zero with forecasted temperatures increasing over the years. Conversely, from lead time equal to 1, particularly in July, August and September, the forecasts seem to converge and become very similar. Therefore, to compare different years, forecasts with a lead time of 0 will be used, as this period highlights the greatest differences between years. To compare different lead times, a single representative year (2018) will be referenced.

Initial analysis suggests that forecasts are more accurate at the beginning of the forecast period. However, this will be examined in more detail in the chapter comparing with reanalysis data.

Finally, it is noted that maximum and minimum temperatures exhibit similar behaviour. Specifically, if maximum temperatures increase or decrease, minimum temperatures follow the same trend. This is particularly evident in the early days of the forecast, where predictions show more detailed behaviour. Therefore, in subsequent temperature data analyses, focus will be mainly on maximum temperatures, keeping in mind that minimum temperatures show a symmetric behaviour.

- *Comparison with same starting month and different lead times*

In this paragraph, we aim to examine in detail the behaviour of ensemble members as lead time increases. To this purpose, the monthly average for each member was calculated, thus obtaining 51 representative values for each forecast month to statistically study the data distribution.

Specifically, the data distribution for the year 2018 was analysed, with the initialization month being June, for two distinct lead times (zero and two). In Figure 13, the monthly averages of each ensemble from June to September are plotted (Figure 13.a), along with the empirical distribution functions (EDFs) of forecasts initiated in June with lead time 0 and lead time 2 (Figure 13.b).
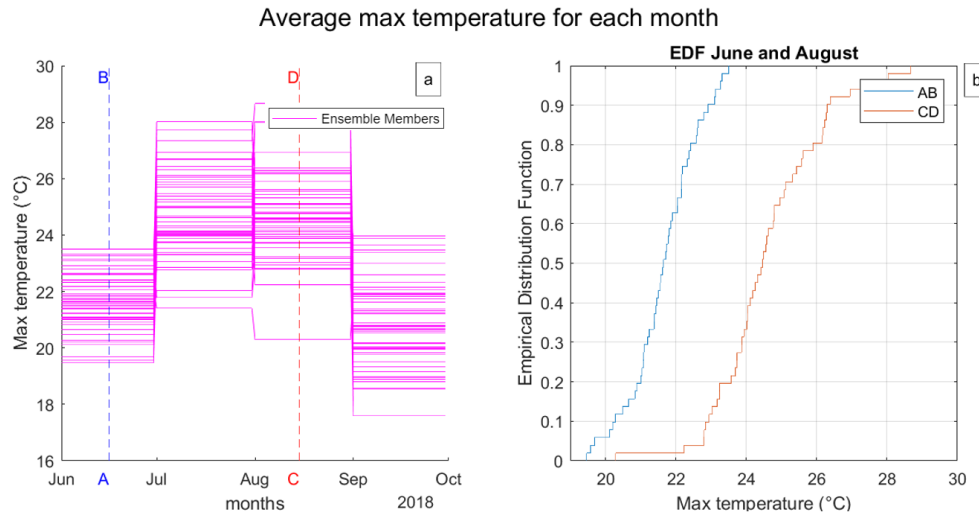
Figure 13. *Monthly mean of the Ensemble Members and EDFs for lead times 0 and 2 with starting month: June*

In accordance with the previous paragraph, it is observed that the distribution with lead time 2 has a wider range of values, with maximum temperatures ranging from 20 to 29 °C. The EDF of AB shows a steeper slope compared to the curve of CD, which instead has longer tails, indicating greater data dispersion. However, analysing August (CD), it is noted that the curve still shows a steeper slope in a more central range. In other words, it can be concluded that the 10th and 90th percentiles are still far from the maximum and minimum extremes of the covered value range.

Nonetheless, from Figure 13.a, it emerges that this behaviour is not generalizable a priori for lead times greater than zero. In July, for example, the data are distributed more homogeneously, and the EDF would therefore have a much less steep slope. In this regard, in the next paragraph, the confidence intervals of the forecasts will be analysed.

Continuing the analysis of the AB and CD curves, the histograms of the data distributions provided by the forecasts were compared with theoretical distributions (Figure 14) to identify the one that best fits and summarizes the characteristics with few parameters (mean, variance, etc.). For the analysis, the normal, log-normal, and Gumbel distributions were compared using the Quantile-Quantile plot and the probability density function.
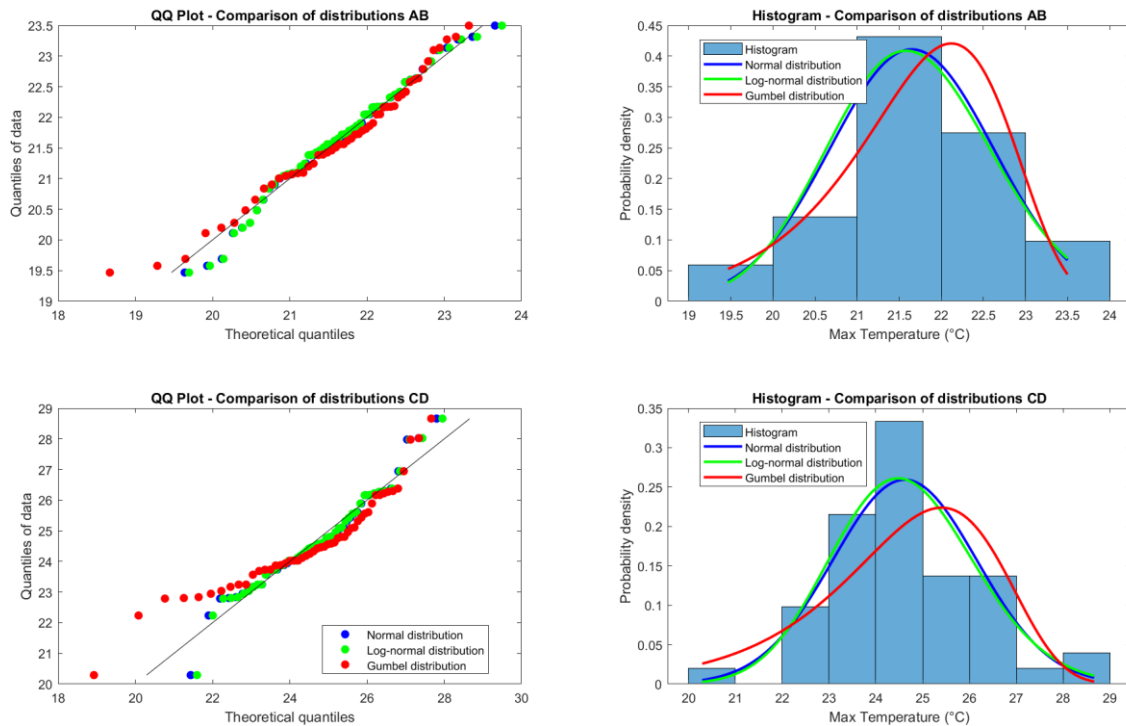
Figure 14. *Statistical analysis: QQ plot comparing the theoretical distribution with the real data, with histograms showing the empirical distribution of the data against the expected distribution for lead times 0 and 2*

Visually, it is immediately noticeable that in both cases, the best distributions are the normal and log-normal ones. For the CD case, the data are well aligned along the diagonal straight line only in the central part, while showing more pronounced deviations at the beginning and end of the distribution, indicating possible outliers. In the case of AB, however, the data fairly accurately follow the theoretical normal and log-normal distributions. For practical purposes, considering that the two curves do not significantly differ from each other, the normal distribution is chosen as the theoretical representative of the data. Thus, it is concluded that the data, both with lead time 0 and with lead time greater than 0, exhibit symmetric behaviour, so their mean well represents the centre of the distribution. Additionally, a fundamental parameter of the data set is the standard deviation, which provides important information about the data dispersion relative to the mean.

With this premise, in the figure 15, we wanted to compare a single forecast month (July), varying the starting months (March, April, May, June) using only the normal distributions.
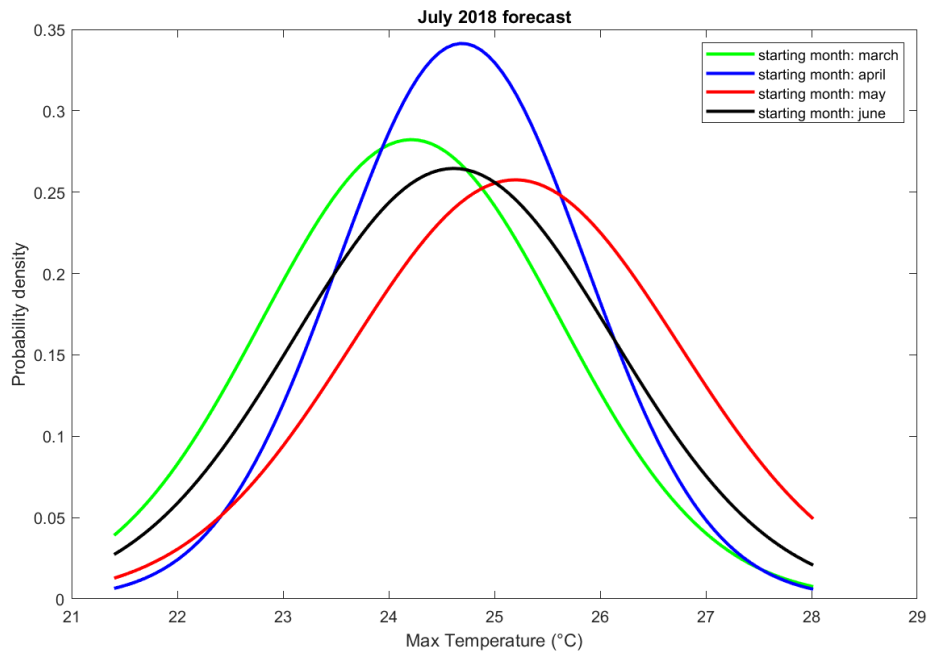
Figure 15. *Comparison of normal distributions for the July forecast, using four different initialization months*

Table 4 - *Standard deviation of July forecast comparing different starting months*

Starting month	Standard deviation(°C)
March	1.41
April	1.17
May	1.54
June	1.5

From the figure 15 and the table 4, it is evident that the lowest standard deviation is associated with the forecast for July with the starting month of April, while the highest standard deviation pertains to the starting months of May and June. It seems, therefore, that the confidence interval tends to narrow with lead times greater than 3, i.e., with forecasts made three months in advance. Conversely, forecasts made one or two months in advance show higher variability.

To further analyse this behaviour, in the next paragraph, the confidence intervals of forecasts with different starting months and forecast months will be examined in more detail.

- *Comparison with different starting months and different lead times*

Confidence intervals (CIs) are fundamental tools in statistics, used to estimate the range within which a population parameter is likely to lie, based on a sample of data.[34]

In the present case, the goal is to calculate the 90% confidence interval (from the 5th to the 95th percentile) for the predicted data, within which the observed value is expected to fall.

In Figure 16 and 17, four initialization months are represented with forecasts from one to seven months for March, from one to six months for April, and up to four months for June for both maximum temperature and minimum temperature.
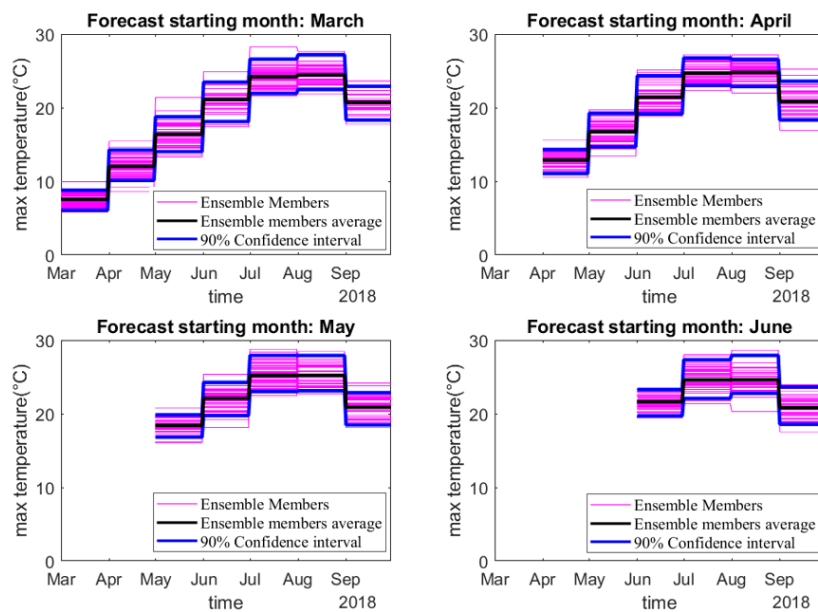
Figure 16. *Monthly average of ensemble members, along with the ensemble mean, including a 90% confidence interval for maximum temperatures*
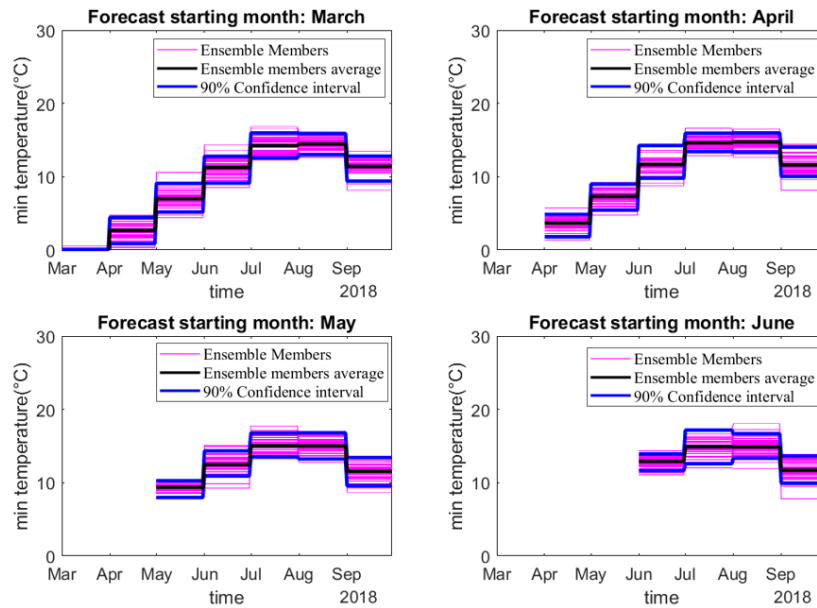
Figure 17. *Monthly average of ensemble members, along with the ensemble mean, including a 90% confidence interval for minimum temperatures*

The narrowest confidence intervals in all four cases for each variable are generally associated with a lead time of zero. Subsequently, these intervals tend to widen from lead time 1, then narrow again from lead time 3 onwards.

As can be seen when comparing the two variables, minimum temperatures tend to show narrower confidence intervals compared to maximum temperatures with the ensemble members much closer to each other and therefore more consistent.

In Figures 19 and 20, the analysis focuses on the forecast period from June to September through the representation of boxplots, means, and standard deviations. Boxplots are useful tools for visualizing data distribution, representing the central 50% of the data within the coloured rectangle, with the central red line indicating the median and the "whiskers" indicating dispersion. These plots provide 50% confidence intervals, including additional information such as the median and the presence of any outliers.
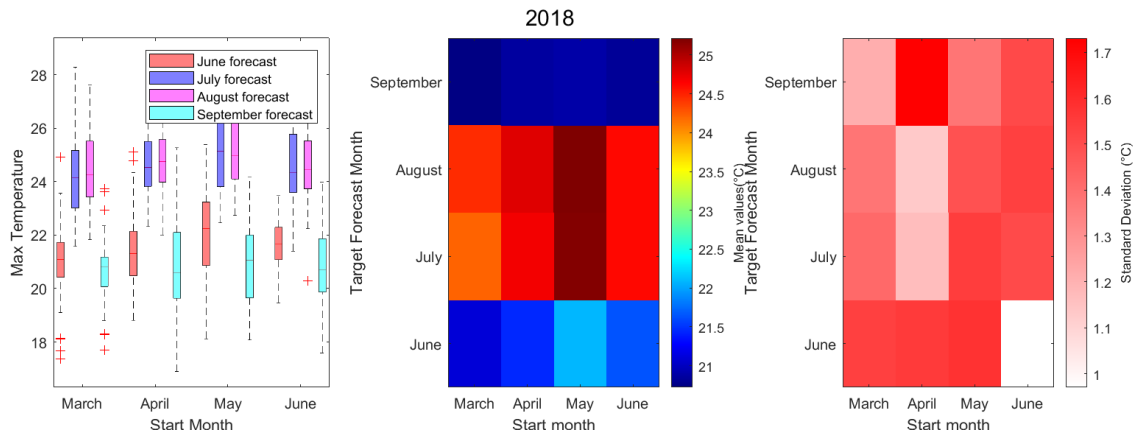
Figure 18. *Representation of boxplots, monthly means, and their respective standard deviations, comparing four different initialization months and four forecast months for maximum temperature*
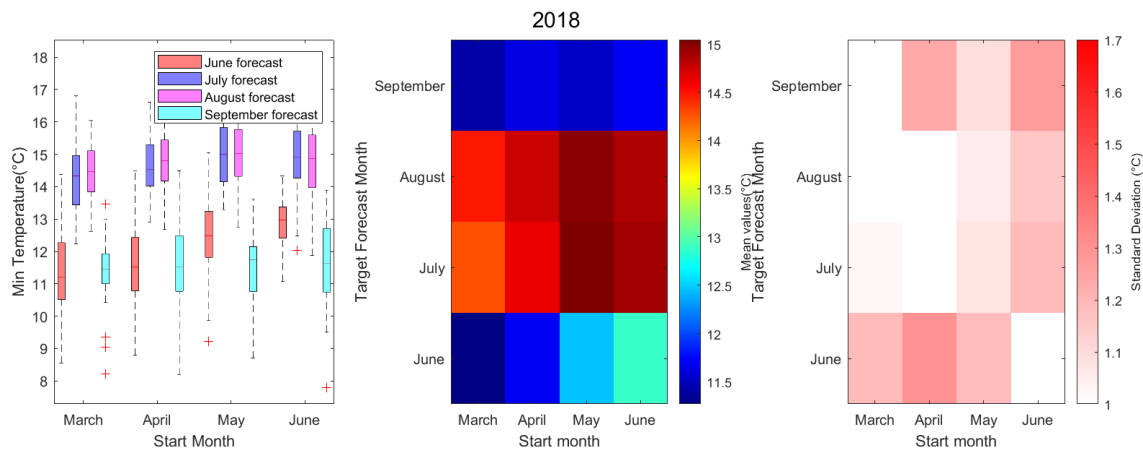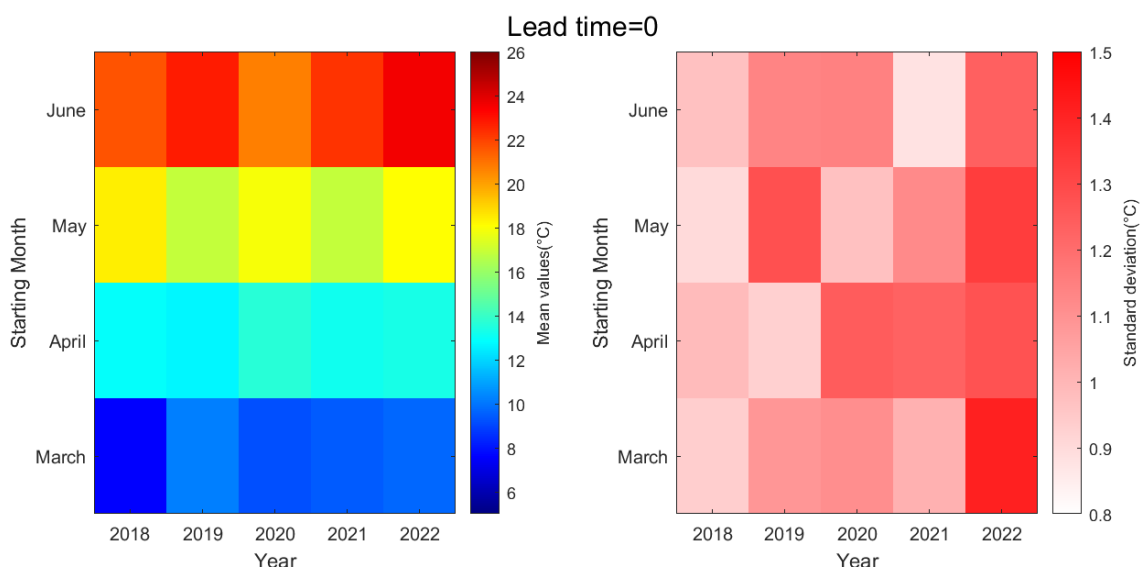

Figure 19. *Representation of boxplots, monthly means, and their respective standard deviations, comparing four different initialization months and four forecast months for minimum temperature*

From the analysis of the standard deviation, it is observed that the forecast from March for June (lead time 3) shows a higher standard deviation compared to the forecasts from July onwards, indicating a narrowing of the interval. Forecasts for July and August with the initialization month of June show a dispersion about one degree higher than that for June itself, confirming that the months with the highest standard deviation correspond to lead times 1 and 2.

Moreover, the narrower confidence intervals for minimum temperatures correspond to much lower standard deviation values compared to maximum temperatures. This

indicates less variability among the ensemble members, but it does not necessarily mean they are more accurate. This aspect will be analysed in Chapter 5, where the forecast errors will be evaluated.

In conclusion, the variation of the confidence intervals with lead time highlights greater variability in short-term forecasts (lead times 1 and 2), while longer lead time forecasts show less dispersion, indicating greater reliability over the considered period.

- *Comparison among different years*

This paragraph will focus on the behaviour of forecasts over the five years analysed, using a lead time of zero for each initialization month. During this time frame, as discussed previously, the values across the years show significant variation.

Considering four starting months (March, April, May, June) and their respective lead times of zero (e.g., forecast for March with initialization month March, forecast for April with initialization month April), the differences in both monthly mean values and standard deviation of the ensemble members were evaluated. These two parameters, as demonstrated earlier, are sufficiently representative of the data distributions for the various months.

Figure 20. *Representation of monthly means and their respective standard deviations, comparing four different initialization months with lead time 0 from 2018 to 2022*

From figure 20, which displays the mean and the standard deviation values of the ensemble members averaged monthly, it can be observed that the mean and dispersion of the predicted maximum temperature tend to increase over the years, with higher values in 2022. A high standard deviation signifies greater dispersion of the data, while a low standard deviation indicates that the data are closer to the mean. In this context, the mean is a good representative of the data overall.

In conclusion, the graphs reveal that the ensemble members in 2022 displayed significantly greater variability compared to the previous four years. This increase in variability can be attributed to several factors. First, climate change has made the climate more unpredictable. As global temperatures rise, we are witnessing a rise in extreme events like heatwaves, droughts, and intense rainfall. These phenomena contribute to greater uncertainty in seasonal forecasts and, consequently, affect the standard deviation. Moreover, climate models depend heavily on historical data to make predictions. However, with the ongoing changes in the climate, the current conditions often diverge significantly from past data. This discrepancy reduces the accuracy of these models, leading to increased uncertainty and variability in the forecasts they produce.

More detailed analyses for the year 2022 will be presented in the subsequent paragraphs, comparing different models and examining forecasts for specific spatial grid cells.

- *Spatial representation*

As a final temperature analysis, in the following figures, the spatial distribution of maximum and minimum temperatures is represented. Because the downloaded variables are depicted as four-dimensional matrices, in this case, the mean was calculated both for the 51 ensemble members and the daily data, in order to determine the monthly average value for all members. This resulted in temperature matrices with dimensions of 3x6 (latitude x longitude). June was chosen as the starting month for the forecast, with a lead time of 1 month. In other words, the forecast period for July was extracted from the data matrix. To better visualize the area of interest, the shapefile of the Po basin and the boundary of the region with altitudes below 400 m were included. The latter representation is an approximate reconstruction, obtained using the "polygons" tool in QGIS, with the aid of the previously shown Digital Elevation Model. Three years (2018, 2020, 2022) were compared to provide an indication of the differences in space among the forecasts for the years considered, thus deepening the results found in the previous paragraph.
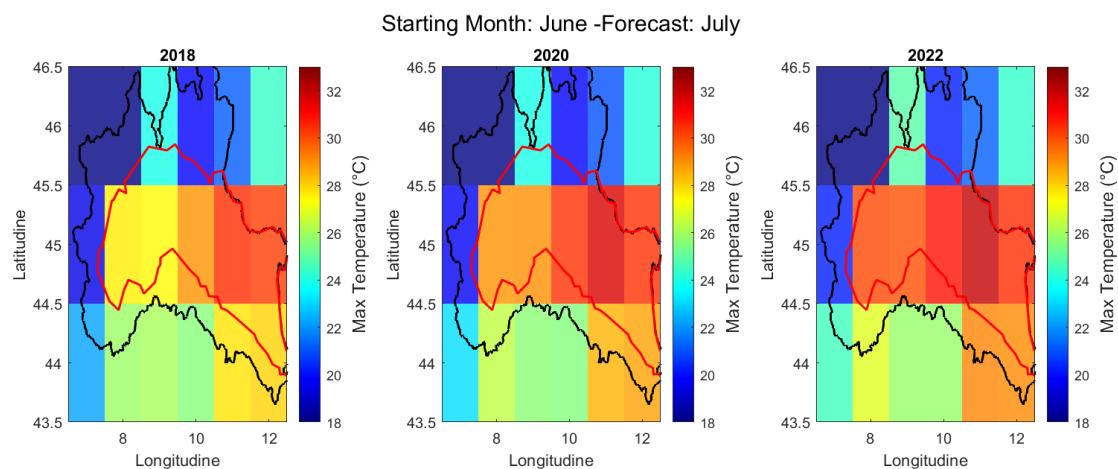
Figure 21. *Spatial distribution of the monthly mean Maximum temperature with a starting month of June and a forecast month of July*
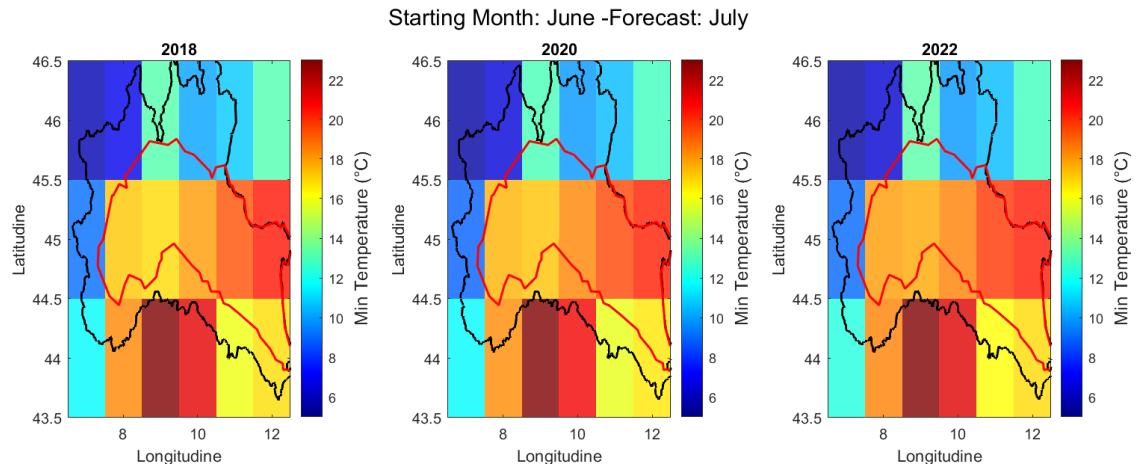
Figure 22. *Spatial distribution of the monthly mean Minimum temperature with a starting month of June and a forecast month of July*

A significant difference in temperatures between the cells in the northwest and those in the plain area is immediately noticeable. For both maximum and minimum temperatures, there is a discrepancy of about 10°C.

Specifically, it is noteworthy that the cells at latitude 44 and longitude from 7 to 10 show little difference between maximum and minimum temperatures. This could be related to the thermo-regulatory effect of the sea, which, with its great thermal capacity, changes its temperature more slowly than the land, resulting in an exchange of heat with the latter. In other words, during the day the sea, being colder, absorbs heat, while at night it releases it, thus smoothing out the high temperature variations that would occur on land.

Conversely, in the cells at latitude 45 and longitude 11 and 12, the temperatures remain high despite the apparent end of the Pianura Padana in that area, because there is still a plain area (the Venetian Plain).

It is also observed that, over the three years, there has been an increase in the forecasted temperatures in the central cells. For maximum temperatures, an increase from 27°C to over 30°C has been noted, while for minimum temperatures, an increase from 17°C to 18/19°C has been recorded.

Regarding the spatial distribution of the data dispersion, the following figure analyses the standard deviation of the different maximum temperature cells.
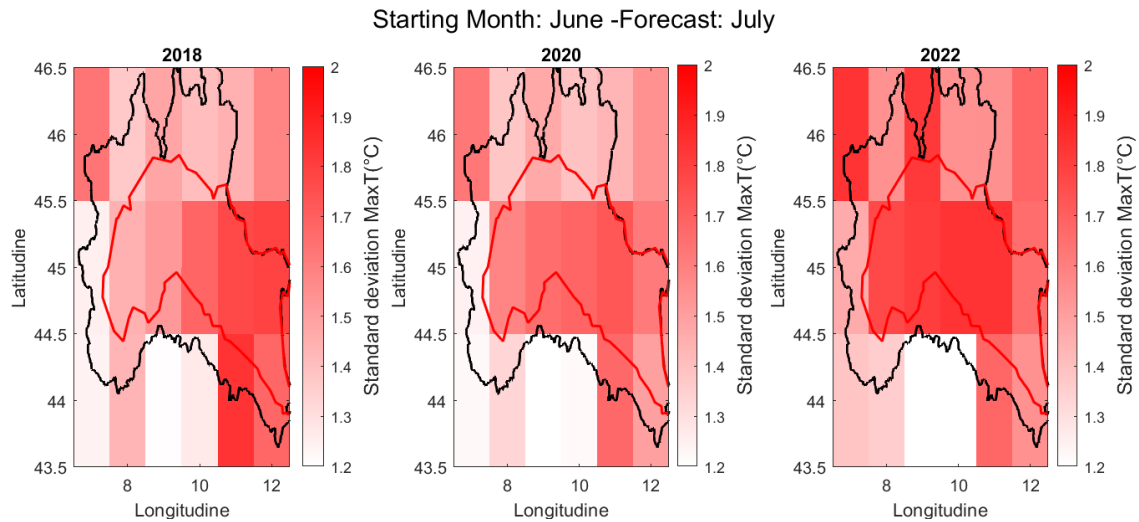
Figure 23. *Spatial distribution of the monthly mean standard deviation of maximum temperature, with a starting month of June and a forecast month of July*

It can be observed that the increase in standard deviation over the years, as discussed in the previous paragraph, is mainly attributable to the greater dispersion in the central and northwestern regions. In the cells at latitude 44°, there is an increase, although the values remain relatively low. This phenomenon is likely due to the thermoregulatory effect of the sea on the surrounding areas, which helps to reduce temperature variability.

4.1.2 Precipitation

- *Comparison between temperature and precipitation*

To provide an initial indication of the different variability between temperatures and monthly cumulative precipitation, it was decided to plot both variables, spatially averaged, on the same graph using two distinct y-axes. This choice allows for a visual comparison of the two-time series, highlighting the differences in their respective variability behaviours. The use of two separate y-axes is essential to maintain the appropriate scale for each variable, given that the units of measurement and magnitudes of variations can differ significantly. This approach facilitates data interpretation and offers a clearer understanding of the climatic dynamics under examination.
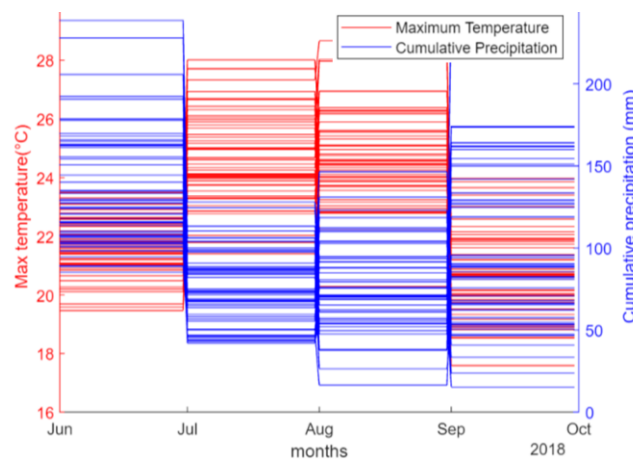
Figure 24. Comparison of the ensemble members' distribution of monthly cumulative precipitation and average maximum temperatures, with a starting month of June

From Figure 24, it is observed that higher temperatures correspond to lower precipitation, especially in the months of July and August, during which we therefore expect higher irrigation requirements. Regarding data variability, there is a significantly greater dispersion in precipitation compared to temperatures. Consequently, it was decided to further analyse this aspect using the coefficient of variance (CV). The coefficient of variation is a useful tool for comparing the standard deviations of two variables with very different scales, such as temperature and precipitation in our case. This allows for the comparison of relative variability with respect to their respective means, regardless of their units of measurement. A higher CV indicates greater relative variability. The CV is expressed as:

$$CV = \frac{\sigma}{\mu}$$

Where σ represents the standard deviation and μ the mean. The following table compares σ, μ, and CV for the forecasts of maximum temperature and cumulative precipitation in June and August, with initialization in June.

Table 5. *Precipitation and Max temperature CV*

	Cumulative precipitation			Max Temperature	
Parameters	*Lead time 0*	*Lead time 2*	*Parameters*	*Lead time 0*	*Lead time 2*
σ (mm)	35.53	26.39	σ (°C)	0.97	1.53
μ (mm)	136.95	77.28	μ (°C)	21.65	24.6
CV(-)	0.25	0.34	CV(-)	0.04	0.06

 As can be seen from the results in Table 5, the coefficient of variation (CV) values for monthly cumulative precipitation are about an order of magnitude higher than the CV values for temperatures. This suggests that precipitation data is more dispersed around the mean compared to temperature data. Additionally, high CV values can indicate a certain degree of forecast uncertainty.

Given this higher variability, the next paragraphs will analyse the behaviour of the ensemble members with both ten-day cumulative data (for a detailed analysis of short-term variability) and monthly cumulative data (for a more general and stable view of long-term trends and to facilitate the application of statistical models).

- *Representation of the ensemble members*

As was done for temperature, the following graphs will show the behaviour of the 51 ensemble members over time. Specifically, the distribution for the year 2018 has been analysed, with the assumption and subsequent demonstration that the ensemble members behave similarly in other years as well. Given that the initial downloaded data were cumulative over the whole time, daily relative precipitation was first calculated and then re-accumulated into ten-day intervals. Spatial averaging was then performed along latitude and longitude.
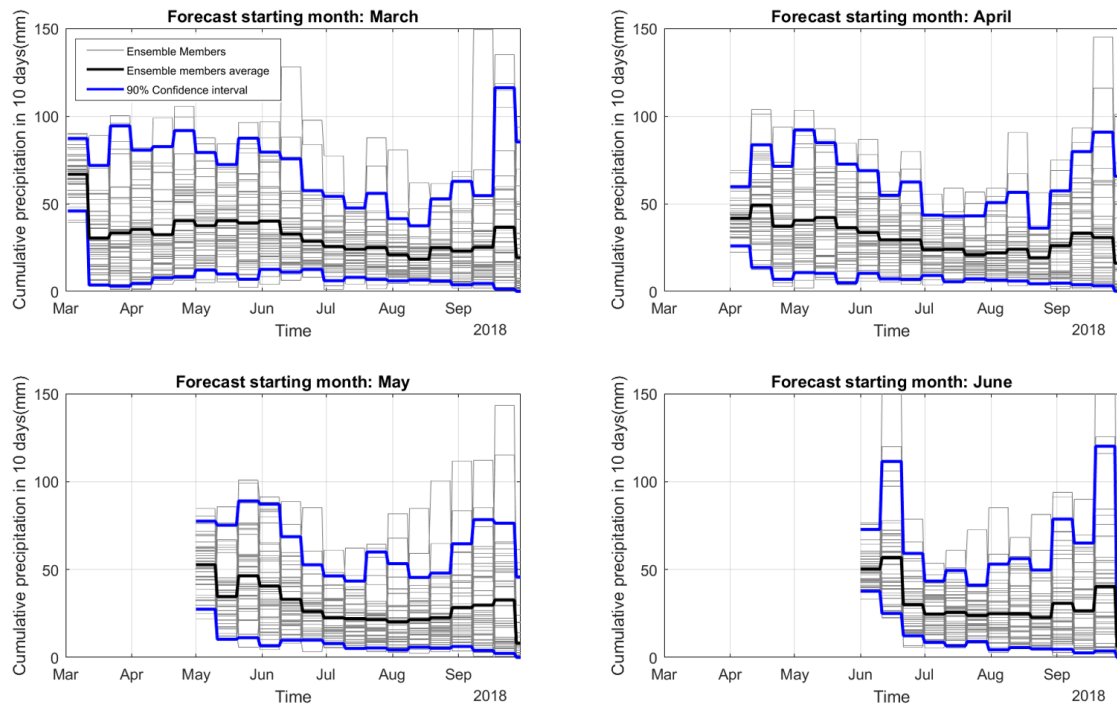
Figure 25. *Monthly cumulative precipitation of ensemble members, along with the ensemble mean, including a 90% confidence interval*

It is immediately evident that the ensemble members for precipitation show greater consistency and lower variability in the first interval considered, indicating that the forecasts from different members are very similar at the beginning. It could suggest higher probability in the accuracy of these initial estimates. However, starting from the second time interval, there is a noticeable trend towards greater dispersion among the ensemble members. This increasing variability, on the other hand,

reflects a rise in uncertainty. This behaviour may be attributed to various factors, including the intrinsic complexity of meteorological phenomena that increases over time, making accurate forecasts more challenging. This phenomenon highlights the importance of considering both the initial consistency of ensemble members and their subsequent dispersion for a comprehensive assessment of weather forecasts. Moreover, this dispersion across different time intervals results in a smoother average trend, as extreme fluctuations of individual members tend to offset each other.

Additionally, attention is drawn to the differences in variability between the spring and summer months. Generally, the confidence intervals from April to May are wider, indicating much more variable forecasts. The variability in precipitation patterns is

largely influenced by the transitions between winter and summer climates, characterized by cold fronts and extreme events which complicate predictions, resulting in wider uncertainty intervals. In contrast, summer months tend to be more stable due to consistent factors like solar heating and high-pressure systems, which allow for more accurate precipitation forecasts with narrower confidence intervals[35] .

In order to compare different years, it was decided to plot only the averages of precipitation for different starting months (March, April, May, June). This approach allowed for an understanding of whether there were significant differences in the various lead times across the years or if the results were relatively consistent.
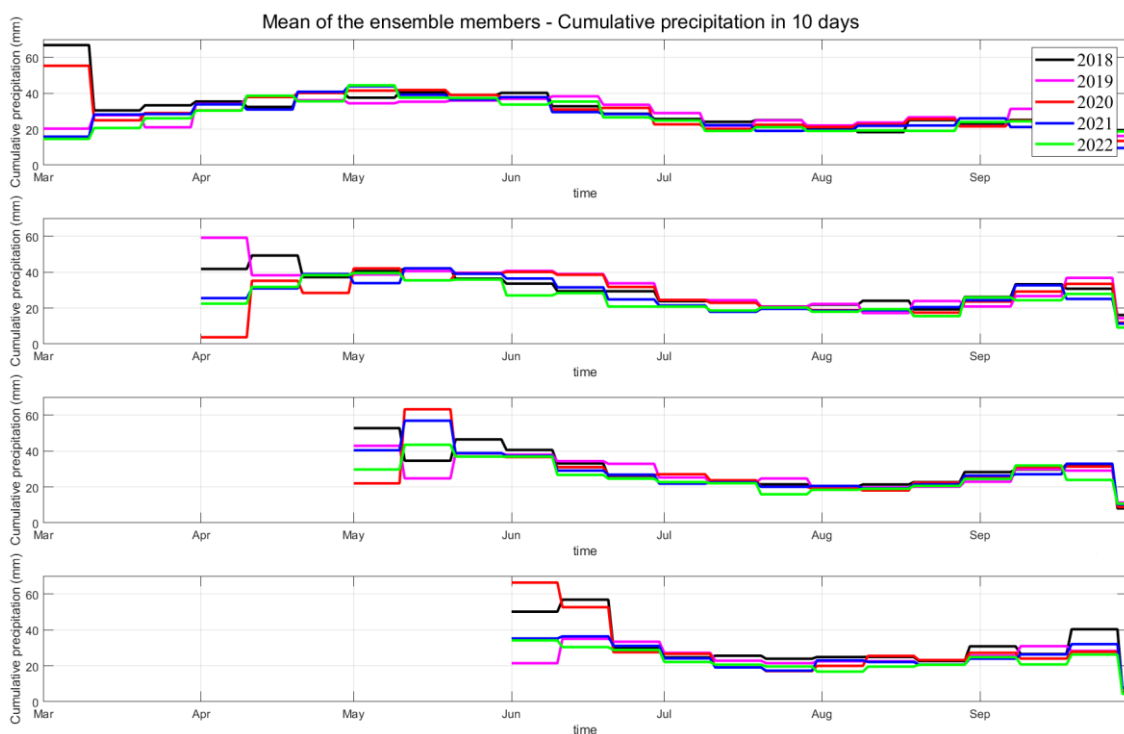
Figure 26. *Ensemble members average representation for four months of initialization and five years for 10 days interval cumulative precipitation*

It was found that the overall trend of the averages is very similar across years, with more pronounced differences particularly in the first and second-time intervals. However, the greater discrepancies in the initial intervals suggest that initial conditions can vary significantly from year to year, affecting short-term forecasts. These initial variations, however, tend to diminish in subsequent periods, resulting in a less significant impact over the long term.

- *Comparison with same starting month and different lead times*

In detail, regarding the behaviour of ten-day cumulative precipitation, we aim to depict the empirical distribution function (EDF) for three different ten-day intervals starting in June: the first, second, and fourth intervals. This approach allows for the analysis of how precipitation is distributed within each time interval and for comparing the variability among these periods. Depicting the EDF for these specific intervals helps to better understand the short-term dynamics of precipitation and how they evolve over time, highlighting the influence of initial conditions. Additionally, this analysis can reveal how precipitation changes throughout the month and beyond, providing valuable insights for improving weather forecasts and managing water resources. By comparing these intervals, it is possible to observe whether precipitation trends remain consistent or vary significantly, offering a more comprehensive picture of the behaviour of cumulative precipitation in the early intervals of the forecast.
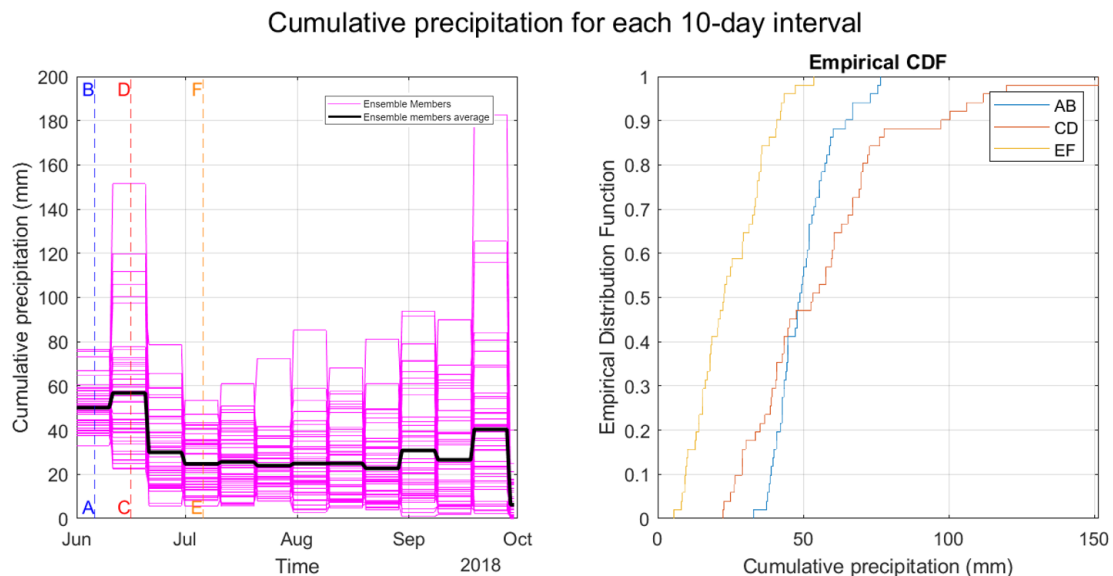
Figure 27. Distribution of cumulative precipitation at 10-day intervals for ensemble members and EDFs, with lead intervals of 0, 1, and 3, initialized in June

By examining the provided graphs, we can draw several conclusions about the behaviour of ten-day cumulative precipitation and their empirical distribution. Each vertical line on the right graph represents the EDF for specific ten-day intervals from the left graph:

- AB (Blue): First ten-day interval
- CD (Orange): Second ten-day interval
- EF (Yellow): Fourth ten-day interval

The greater consistency in the first interval (narrower and closer lines) suggests that initial precipitation conditions are more predictable and less variable, while the increasing dispersion in subsequent intervals reflects growing uncertainty over time.

Having demonstrated that variability between different intervals is most noticeable at lead time 0, while later months show more consistent interval averages, it was decided to analyse monthly cumulative data distributions. This choice was made because larger intervals can reduce noise and make the data more suitable for certain statistical models. Given that the data are limited or highly variable, aggregating over longer periods can improve the robustness of the analysis. Therefore, the aim is to find theoretical distributions that best approximate the available data. In this case, the analysis will focus on forecasts with an April starting month to evaluate predictions both in spring (lead times 0-AB and 2-CD) and in summer (lead time 3-EF) (figure 28).
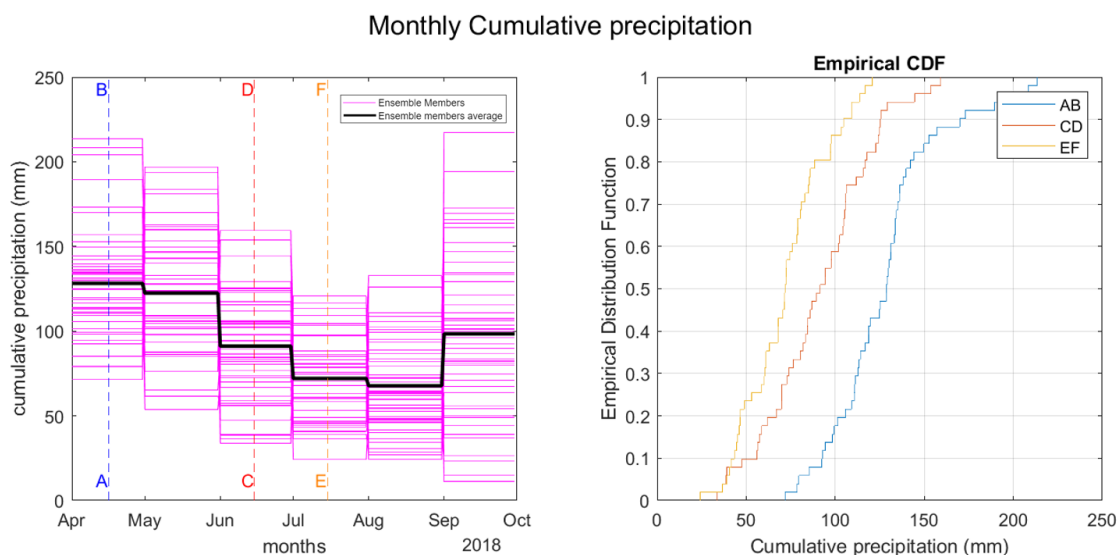
Figure 28. *Distribution of monthly cumulative precipitation for ensemble members and EDFs, with lead times of 0, 2 and 3, initialized in June*
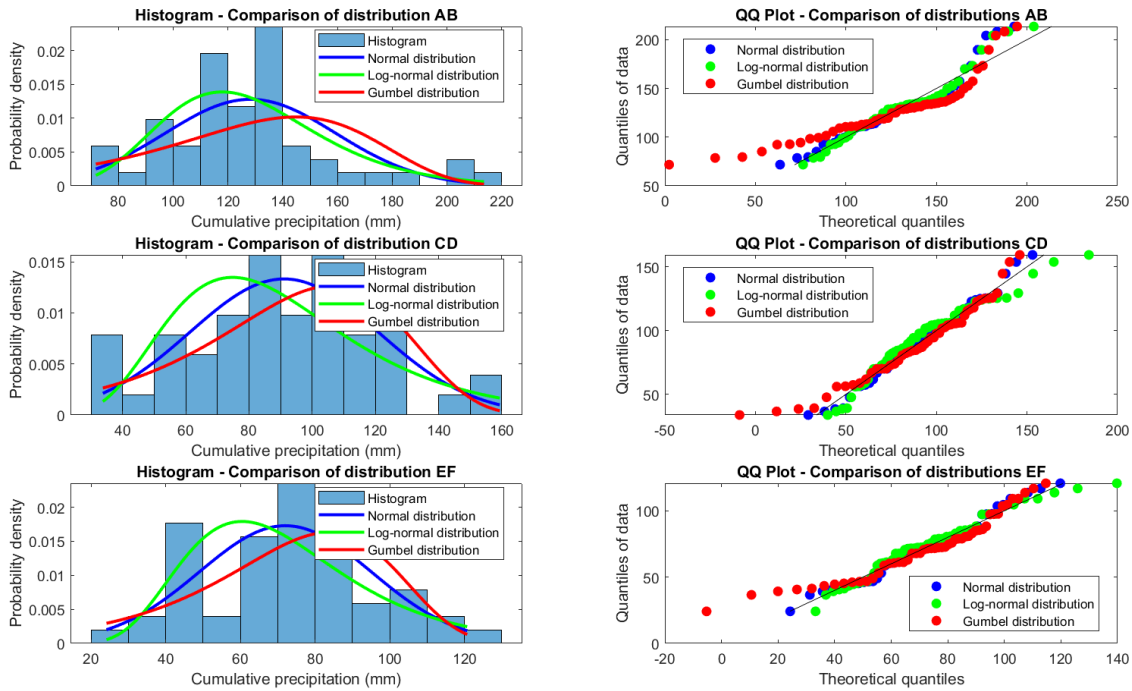
Figure 29. *Statistical analysis: QQ plot comparing the theoretical distribution with the real data, with histograms showing the empirical distribution of the data against the expected distribution for lead times 0, 2 and 3*

From the analysed graphs, it is evident that the most variable distribution is that of the first forecast month. This increased variability is due to two main factors:

- Differences in Dispersion Across Intervals: The second and third ten-day intervals show significantly wider dispersion compared to the first interval of the month.

- Seasonal Influence: The first forecast month falls in spring, a season characterized by high uncertainty in weather predictions. Seasonal transitions between winter and spring introduce greater variability in precipitation, contributing to a broader and less predictable distribution.

For longer lead times, it is observed that the distribution that best describes the monthly cumulative precipitation is the normal distribution. This can be explained by the Central Limit Theorem (CLT), which states that "the means of a random sample of size, n, from a population with mean, μ, and variance, σ^2, distribute normally with mean, μ, and variance, σ^2/n"[36].

In summary, using monthly cumulative data rather than ten-day intervals offers significant advantages for representing precipitation distributions. Monthly cumulative

data tend to follow a normal distribution, as predicted by the CLT, which allows for clearer representation through mean and standard deviation parameters. However, it is important to consider that monthly cumulative data may obscure finer variations within the month, reducing the ability to analyse specific short-term variability. Therefore, while the normal distribution can provide a good approximation for monthly cumulates, it is essential to be aware of potential differences and finer variability that may exist within the monthly periods.

- *Comparison with different starting months and different lead times*
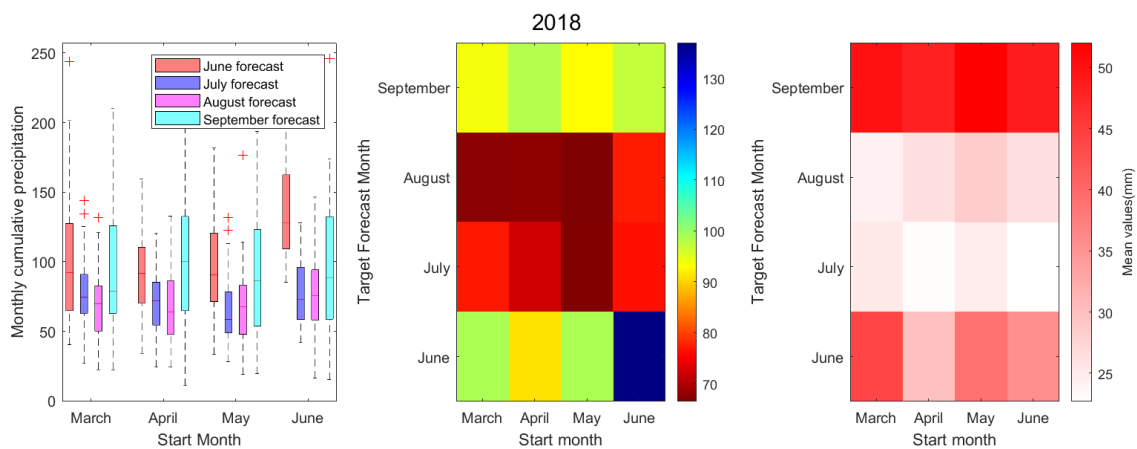
Figure 30. *Representation of boxplots, means of ensemble members, and their respective standard deviations, comparing four different initialization months and four forecast months for monthly cumulative precipitation*

In the graph, it is observed that the standard deviation of the forecasts does not follow a linear trend or a clear pattern when comparing different initialization months with the same target forecast month. In other words, there is no evident relationship between the month in which the forecasts are made and the variability of these forecasts for a given target month. However, a recurring feature can be noted: July and August exhibit lower standard deviations compared to other months. This suggests greater stability and predictability of weather conditions during these summer periods. In other words, forecasts for summer months tend to be less variable and therefore more reliable compared to those for other months, such as those in spring or fall, which may be influenced by seasonal transitions and more variable weather conditions.

- *Comparison among different years*

To compare different years, it was decided to compare the forecasts of monthly accumulated precipitation with different starting months and a lead time of 0 in order to observe, where present, differences in the mean and standard deviation of the values provided by ECMWF. The choice of a lead time equal to zero, as for temperatures, was made because in this time period the greatest differences between the different years are observed.
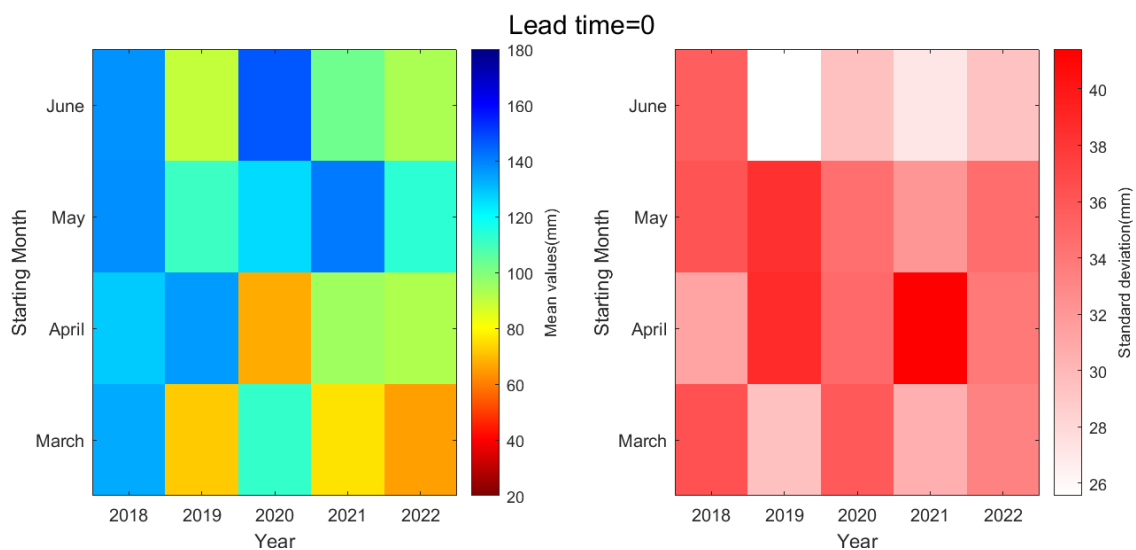

Figure 31. *Representation of means of ensemble members and their respective standard deviations for monthly cumulative precipitation, comparing four different initialization months with lead time 0 from 2018 to 2022*

In the graph on the left, there is a downward trend in the forecasts of average cumulative precipitation over the years. This decrease in the forecasted precipitation, combined with the increasing temperature forecasts (discussed in section 4.1.1), implies a rise in irrigation needs. The increase in temperatures exacerbates evapotranspiration, requiring more irrigation to compensate for the reduced natural water availability.

The graph on the right shows the standard deviations of precipitation forecasts with a lead time of 0 months for all initialization months. The standard deviations mostly range between 32 and 36 mm in most years. However, there are anomalies in 2019, where the standard deviations are either higher or lower than the overall average of other years. These "outlier" values suggest exceptional variability in precipitation forecasts for that year, indicating possible extreme meteorological events or greater uncertainties in climate predictions.

- *Spatial representation*

Finally, this section will focus on spatial distribution. As with temperatures, we will represent the averages of the 51 ensemble members, but instead of using monthly averages, we will use the cumulative values for each month, averaged over the 51 ensemble members. The analysis will be based on the initialization month of June with a lead time of 1 month (forecast for July) for three different years: 2018, 2020, and 2022.
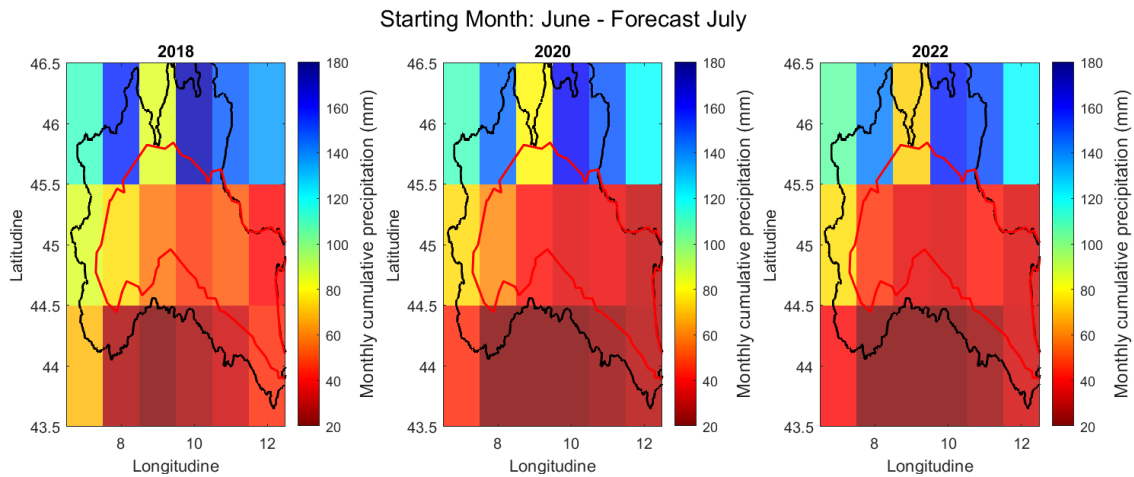

Figure 32. Spatial distribution of the mean of the Ensemble members for monthly cumulative precipitation with a starting month of June and a forecast month of July

A significant difference is immediately noticeable between the cells at latitude 46 and those at latitudes 44 and 45. The values differ markedly, with cumulative precipitation reaching up to 180 mm in higher altitude areas and much lower values in the flatlands and cells including the sea. Focusing on the areas below 400 meters elevation (within the red polygon), the seasonal forecasting model shows a significant decrease in precipitation from 2018 to 2022. However, it should be noted that, as previous studies have indicated, precipitation values can be highly variable and sometimes unreliable.
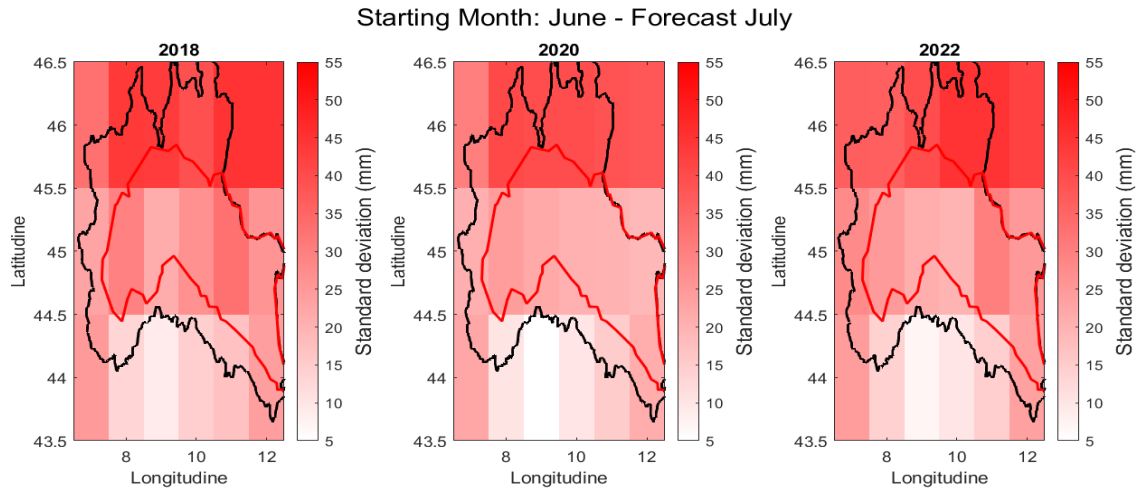
Figure 33. *Spatial distribution of the standard deviation of the Ensemble members for monthly cumulative precipitation with a starting month of June and a forecast month of July*

This graph illustrates that, although average cumulative precipitation appears to be higher in the cells at higher latitudes, the dispersion of the data also increases. Unlike temperature, where the standard deviation has shown an increasing trend over the years, no particular trend is observed in the precipitation data over time.

Therefore, in the chapter comparing with reanalysis and observational data, a more detailed focus will be placed on two specific cells within the Cuneo area.

4.2 MME-MMES-ECMWF comparison

As various studies have demonstrated, multi-model ensembles generally offer higher skills compared to single models [37] [26]. Therefore, in this sub-chapter, the behaviours of the ensemble members in the three different cases (MME, MMES, ECMWF), described in more detail in Chapter 3.2.2, will be analysed and compared in order to evaluate the differences in terms of variability. The accuracy and reliability of the various models will be analysed in Chapter 5.

As previously mentioned, the multi-ensembles also include forecasts from the single CMCC model, which have a maximum time horizon of 183 days. For this reason, a comparison between MME and MMES with ECMWF for the September forecasts initialized in March cannot be made.

More specifically, the mean values of the ensemble members and the standard deviation with various initialization months and lead times for the year 2022 alone will be compared, considering that the models have a different number of members (MME=202, MMES=40, ECMWF=51).

4.2.1 Temperature

- *Comparison with different starting months and different lead times*

In the following graphs, box plots of the spatially averaged maximum temperatures will be depicted, with initialization months of March, April, May, and June and target forecasts from March to September. Additionally, the monthly means of the ensemble members with target forecasts from June to September and initialization months from March to June will be shown. Finally, the standard deviations will be compared. It is important to reiterate that for MME and MMES, the values found for September are not anomalies, but rather null values due to the absence of forecast data from the multi-ensembles in this case.
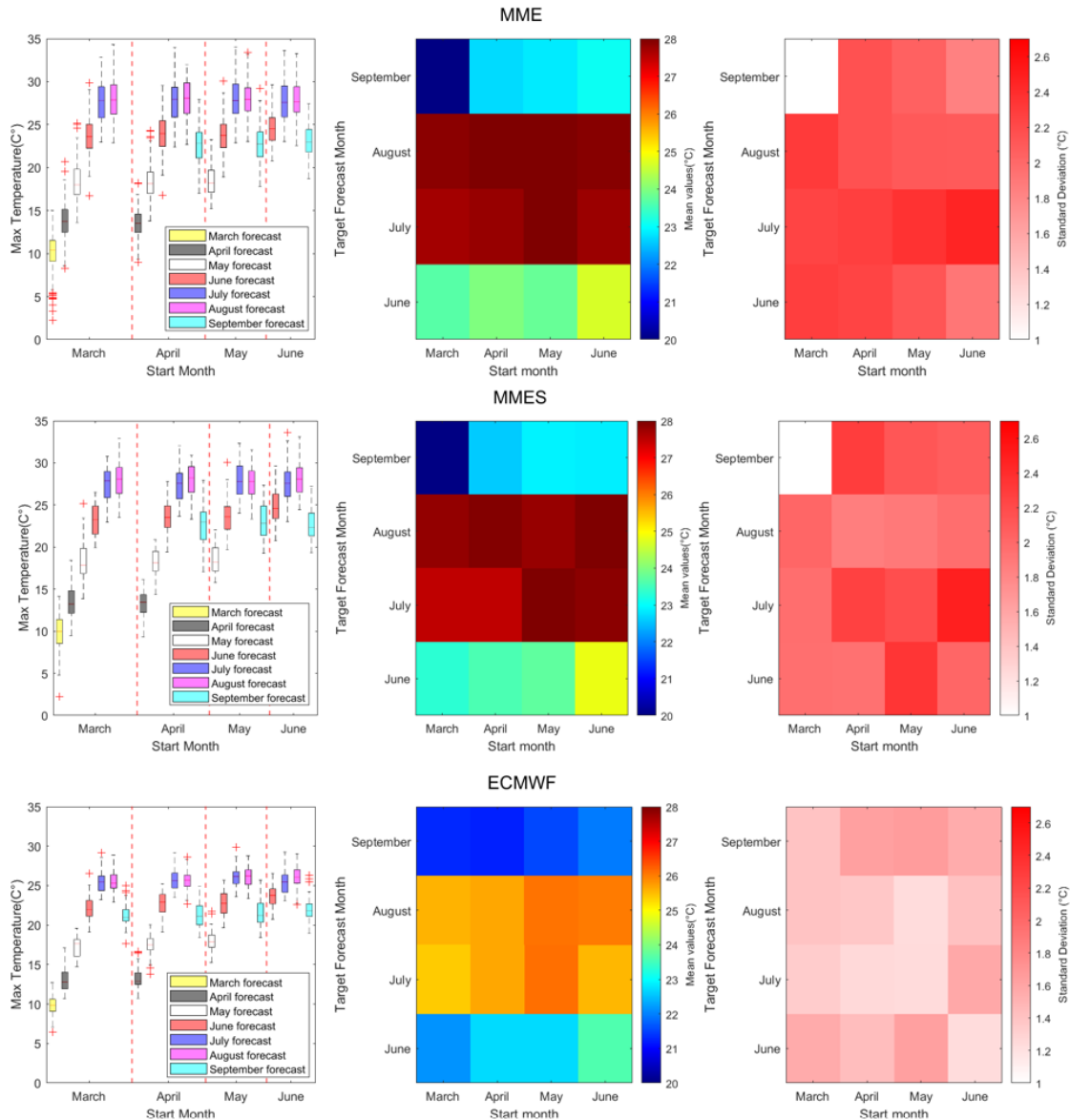
Figure 34. *Representation of boxplots, means of ensemble members, and their respective standard deviations, comparing four different initialization months and four forecast months for monthly mean Maximum temperature assessing ECMWF, MME and MMES behaviours*

The presented graphs reveal several significant trends related to the analysed forecast models. The mean values of the maximum temperatures predicted by the single ECMWF model are systematically lower than those obtained from the multi-model ensembles MME (Multi-Model Ensemble) and MMES (Multi-Model Seasonal Ensemble), likely due to the combination of multiple models that can capture and explore a wider range of scenarios, including the extreme ones. Moreover, the standard deviation for the MME and MMES models is significantly higher than for the single model. This can be explained by

the fact that the ECMWF ensemble, despite having 51 members, is made up of a single model, which means that the internal variability among the members is limited to initial perturbations and variations in the model's parameters. This tends to produce lower dispersion compared to the multi-model ensembles, where the members reflect not only initial perturbations but also structural differences between the various models.

By the way, there is a general consistency in the forecasts across different target months and various models, with a pattern observed in relation to the initialization months. For example, for the target forecast month of June, an increasing trend in predicted temperatures is observed as the lead time decreases. This means that forecasts made in June tend to show higher temperatures compared to those made in March, regardless of the model used.

Finally, the box plots show that as the lead time increases, there is greater symmetry in the box plots, particularly for the MME model. This might indicate greater stability in long-term forecasts. In the MME boxplot, there are also more outliers (red points) and a wider spread of maximum temperatures, reflecting that these ensembles sample a wider range of scenarios.

- *Spatial representation*

After comparing the temporal trends of the models, we will move into a more detailed spatial analysis. In the following figures, the mean of the ensembles and the standard deviation among the various grid cells for the July forecasts with an initialization month of June will be compared.
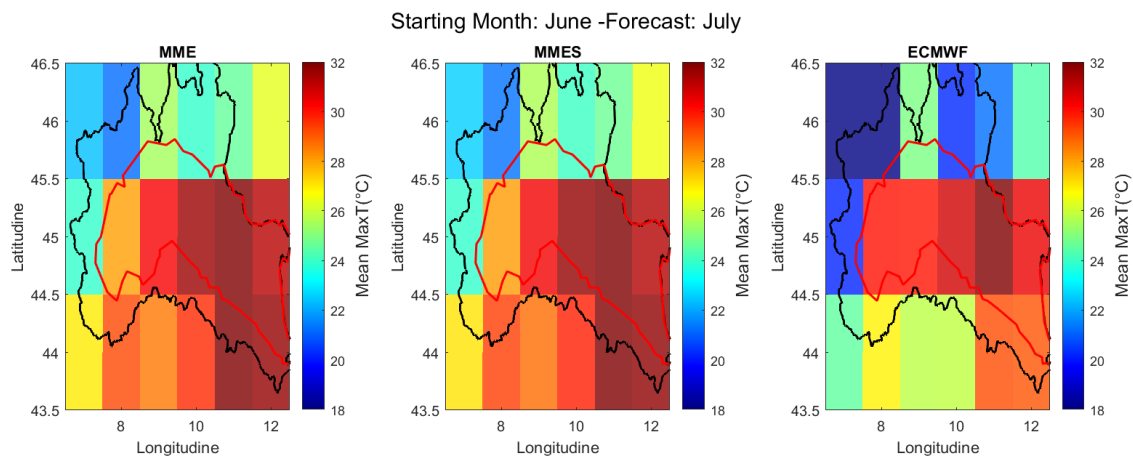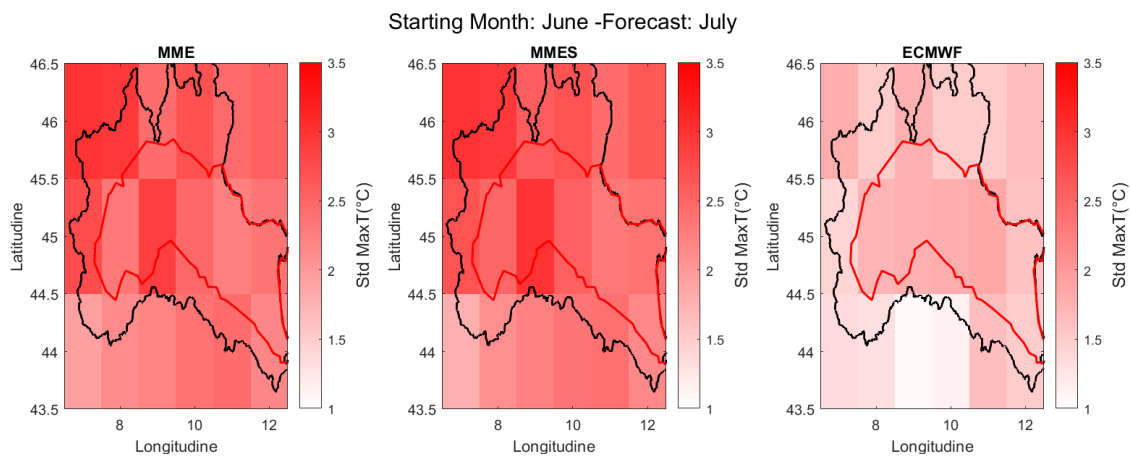

Figure 35. *Spatial distribution of the mean of the Ensemble members for monthly mean Maximum Temperature with a starting month of June and a forecast month of July assessing ECMWF, MME and MMES behaviours*

Figure 36. *Spatial distribution of the standard deviation of the Ensemble members for monthly mean Maximum Temperature with a starting month of June and a forecast month of July assessing ECMWF, MME and MMES behaviours*

From the analysis of the spatial distribution of temperatures, it emerges that the previously identified higher values are not uniformly distributed over the area of interest. In particular, the multi-model ensembles MME and MMES show higher temperatures

compared to the ECMWF model in the grid cells located at latitudes 46° and 44°. However, for the cells at latitude 45°, the situation is almost the opposite: for example, at longitude 8°, the ECMWF model predicts higher temperatures compared to the multi-model ensembles.

Regarding the standard deviation, the previously observed trend is confirmed: the MME and MMES models exhibit greater variability compared to the other model. Unlike the mean temperatures, this greater variability is uniformly distributed across all the grid cells. Additionally, all three models agree in showing higher standard deviations in the cells in the northwest, suggesting greater dispersion of forecasts in these areas.

4.2.2 Precipitation

- *Comparison with different starting months and different lead times*

Similarly to the temperatures, the same comparisons will be made for the monthly cumulative precipitation.
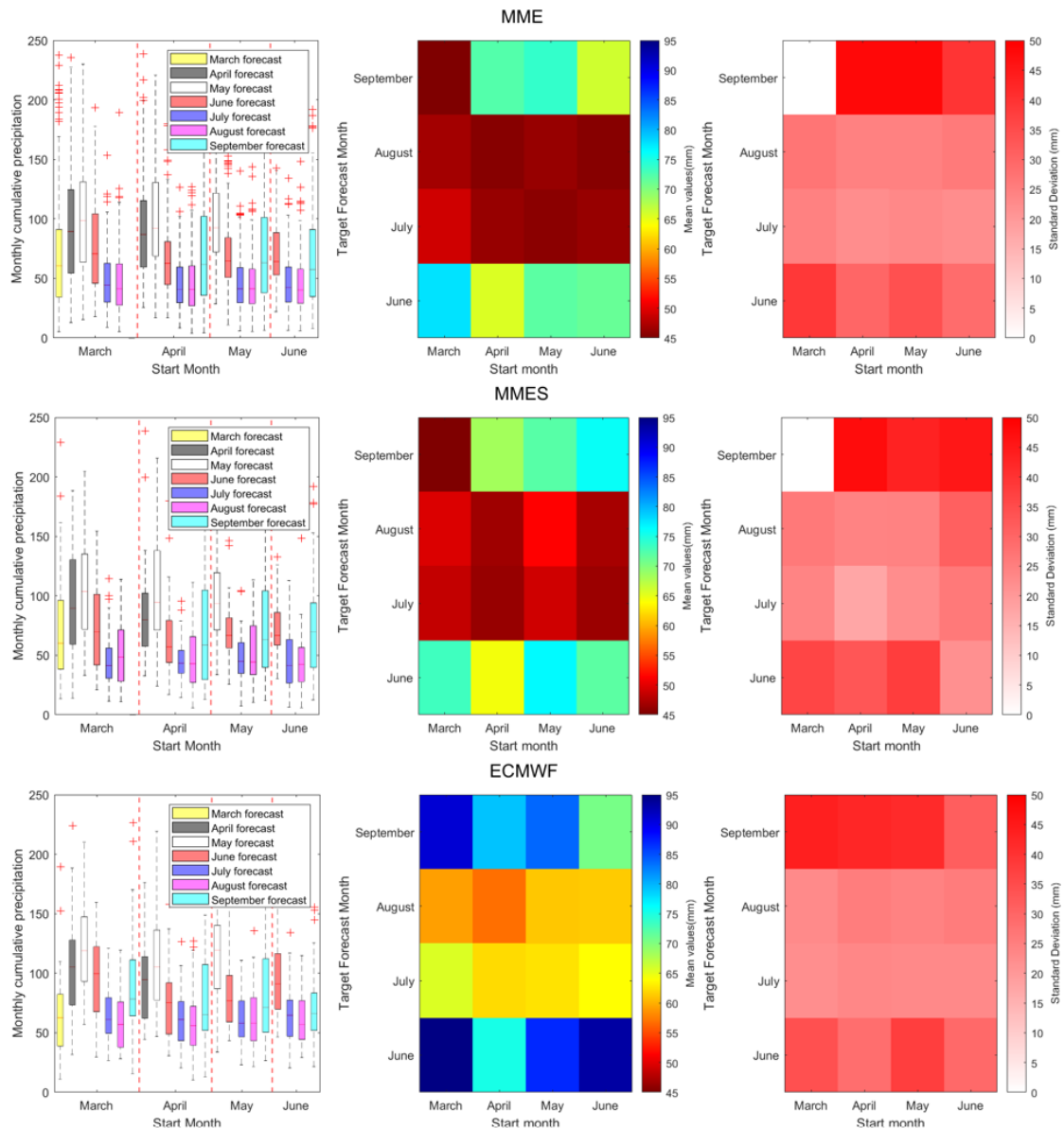
Figure 37. *Representation of boxplots, means of ensemble members, and their respective standard deviations, comparing four different initialization months and four forecast months for monthly cumulative precipitation assessing ECMWF, MME and MMES behaviours*

The three different models exhibit similar behaviour. In particular, the box plots show that the highest predicted precipitation occurs during the months of April and May, then

decreases during the summer months. Similarly, as discussed in Chapter 4.1.2 for the precipitation predicted by the multi-model ensembles, the spring months' box plots show the first and third quartiles farther apart, resulting in a wider interquartile range. It is noted that between the first and third quartile, 50% of the ensemble members are included; hence, the larger the range, the more dispersed the data. With regard to the standard deviation of the three models, in contrast to the temperatures, no significant differences are noted for the forecasts for the months of June to September

However, it should be noted that despite the similar trend, the values can vary significantly between models. In general, the ECMWF predicts higher average cumulative precipitation compared to the other two models. This, combined with lower predicted temperature values, will result in lower water needs compared to the other two models.

- *Spatial representation*

Finally, the spatial distribution of the average monthly cumulative precipitation forecasts is analysed among the different multi-ensembles, starting from the initialization month of June with the target forecast for July.
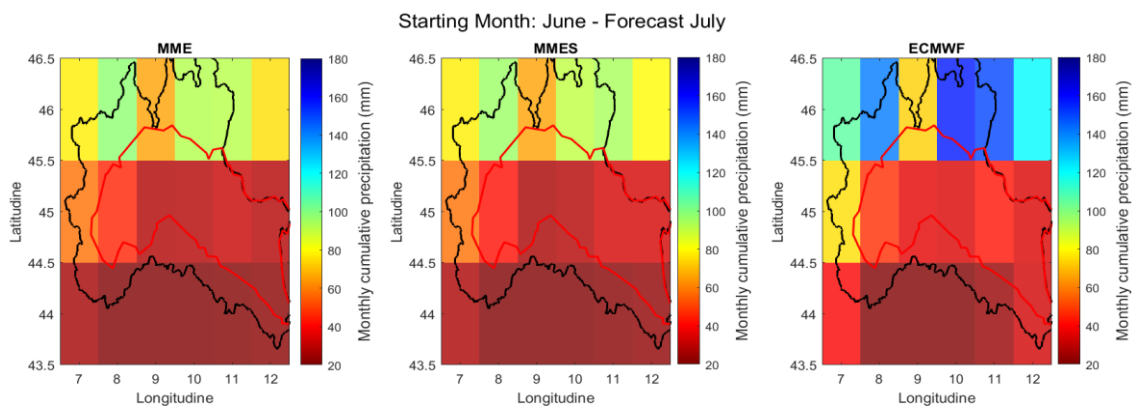

Figure 38. *Spatial distribution of the mean of the Ensemble members for monthly cumulative precipitation with a starting month of June and a forecast month of July assessing ECMWF, MME and MMES behaviours*
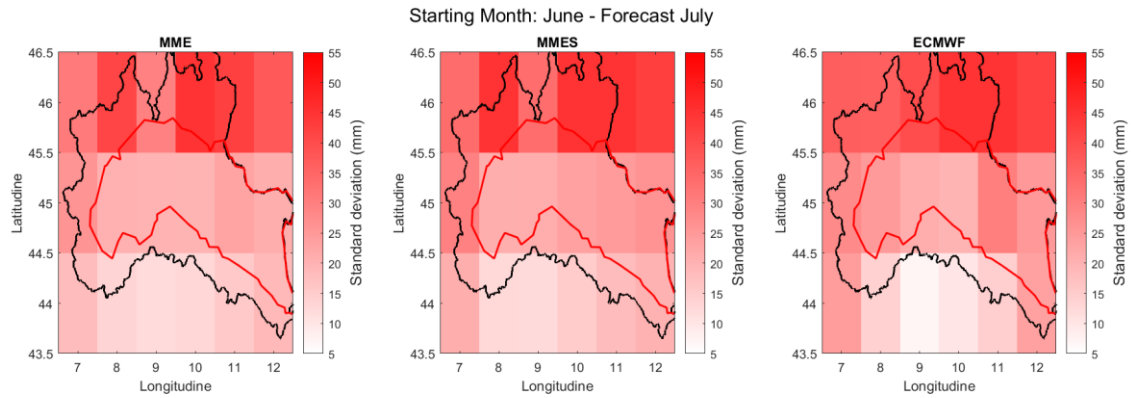
Figure 39. *Spatial distribution of the standard deviation of the Ensemble members for monthly cumulative precipitation with a starting month of June and a forecast month of July assessing ECMWF, MME and MMES behaviours*

In this case as well, for the averages of the ensemble members, an inversely proportional behaviour to the temperatures is observed: lower predicted temperatures correspond to higher predicted precipitation. For instance, the ECMWF model predicts higher values of monthly cumulative precipitation, especially for cells at latitude 46. The standard deviations, on the other hand, show higher values in cells at latitude 46, consistent with the temperatures, indicating greater variability in those regions.

4.3 Estimation of irrigation requirements

The aim of this section is to describe the methodology used to calculate the irrigation requirements of maize.

The irrigation requirements have been obtained through a numerical model (AGRHYD) developed by a PhD student at the Polytechnic University of Turin based on FAO's 56 paper[38].

After calculation of Extra-terrestrial solar radiation on top of the atmosphere, the model continues with computation of the crop evapotranspiration ET_C. The crop evapotranspiration ET_C is calculated as:

$$ET_C = k_C * ET_0;$$

where ET_0 is the reference evapotranspiration and k_C is the crop coefficient.

ET_0 is related to a reference surface (grass) with standardized features, such as a constant crop height of 0.12 m, a fixed surface resistance of $70 \frac{s}{m}$ and an albedo of 0.23[38]. ET_0 is obtained via the Hargreaves equation. Droogers and Allen found that the Hargreaves equation performed quite well under data-limited conditions, with acceptable accuracy compared to the more complex Penman-Monteith method. In their study, they highlighted that the Hargreaves method is particularly useful when resources for collecting data on wind speed and humidity are scarce[39].

The Penman-Monteith equation, in fact, requires radiation, air temperature, air humidity and wind speed data [38]. The Hargreaves equation, on the other hand, depends only on the daily maximum temperature(T_{max}), daily minimum temperature(T_{min}), mean daily temperature(T_{mean}) and the extraterrestrial solar radiation (which is a function of latitude and the day of the year) [40].

The full equation is:

$$ET_0 = k_h * (T_{mean} + 17.8) * (T_{max} - T_{min})^{0.5} * R_a;$$

where k_h is the Hargreaves constant, equal to 0.0023, that was derived empirically by Hargreaves and Samani through calibration against more complex ET equations like Penman-Monteith.

The crop coefficient k_c, on the other hand, depends specifically on the crop characteristics and changes among the different growth stages. For the maize we can distinguish four main stages, characterized by different duration and k_c

***Table 6.** Maize Crop Coefficient*

Growth stage	Duration	Kc
Initial stage	30 days	0.3
Crop development	40 days	0.3-1.2
Mid-season	50 days	1.2
Late season	30 days	1.2-0.5
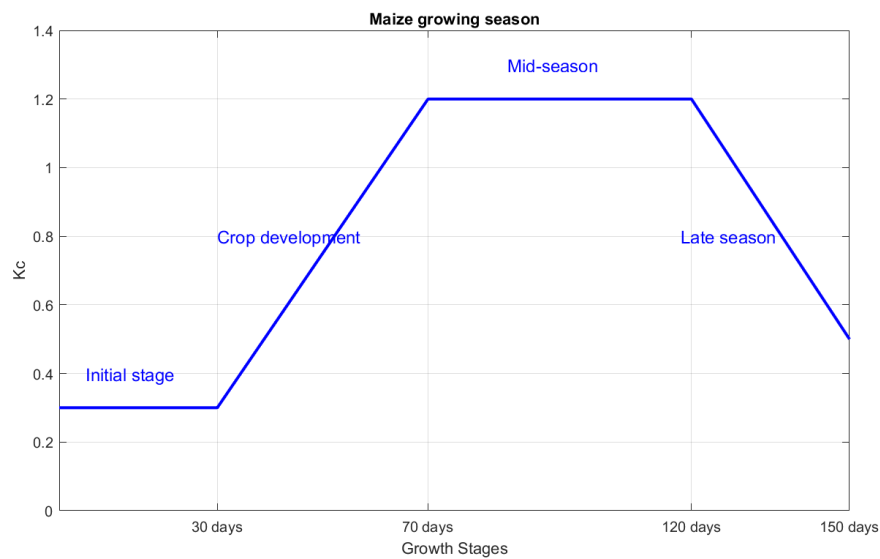
***Figure 40.** Maize growing season*

The maize growing season in the area of study (Pianura Padana) is within May to early October thus we assume the growing season May 15-Sep 30, so that the mid-season will correspond to the summer hottest months and the late season and consequently the harvesting period in October.

The process of calculating irrigation requirements starts with determining the rooting zone moisture (RM), which corresponds to the water content needed for the plant's evapotranspiration. This is calculated by taking the difference between the available water

capacity of the soil (AWC), which is the potential value for each cell of the grid, and the soil moisture for each day(assumed or observed for the first day of season only), then multiplying that by the daily rooting depth(d_{roots_daily}), to account for the portion of soil the plant can access for water uptake each day. In particular, the AWC is the amount of water which the soil can retain and make available to plants while the soil moisture is the actual amount of water present in the soil early in the beginning of the day. The first day soil moisture is assumed as the mean of AWC in Pianura Padana using the Harmonized World Soil Database version 2 [41].

$$RM = (AWC - Initial\ Soil\ Moisture) * d_{roots\_daily}.$$

Next, by subtracting the daily precipitation from the RM, we can determine the soil moisture state (SMS) and whether there is moisture deficit (negative values) and thus irrigation might be necessary. To this regard, the readily available water (RAW), which is the portion of AWC that the plant can use before experiencing stress, is set as the threshold beyond which (Moisture deficit >= RAW) irrigation is required to avoid water stress in the plant. We should note that this is a conservative point of viewtaking RAW as the threshold assuming that the plant should not enter in the stress state.

RAW is calculated by multiplying the AWC by the depletion factor (DF), a crop-specific value that indicates how much water can be depleted before the plant's growth is affected (for maize ~0.5).

$$SMS = RD - P_{daily};$$

$$RAW = AWC * DF.$$

When the SMS exceeds the RAW but is less than the actual evapotranspiration (ET_a), the plant enters water stress, and the irrigation requirement is equal to the deficit minus the RAW (the threshold). If the deficit surpasses the ET_a, the irrigation requirement equals the full evapotranspiration amount, since precipitation alone cannot meet the plant's water needs. This approach ensures that irrigation is applied only when necessary to prevent water stress and optimize plant growth.

- If 0 < SMS-RAW < ET_a → Irrigation Requirements= SMS -RAW.
- If SMS-RAW > 0 & SMS-RAW > ET_a → Irrigation Requirents=ET_a.

5. FORECAST ASSESSMENT: COMPARISON WITH REAL DATA

Up to this point, we have focused on the "isolated" behaviour of the ensemble members and their variability over time and space for both temperature and precipitation variables. However, to better assess the consistency and uncertainties related to the forecast data, it is essential to compare the data from the forecasting centres used with 'real' data.

Therefore, as explained in detail in Chapter 3.3, this study will use ERA5 data for comparison to evaluate the accuracy across the entire Po River basin. Additionally, for two specific cells in the Piedmont region, the analysis will also be supplemented with observed data from ARPA Piemonte meteorological stations.

In the following paragraphs, reference will often be made to "errors". In particular, we will distinguish:

-error: the difference between predicted and observed values

$$\varepsilon = PredictedValue - ObservedValue;$$

- absolute error: the difference in absolute value between predicted and observed values

$$\varepsilon_{abs} = |PredictedValue - ObservedValue|;$$

- relative error: the difference in absolute value between predicted and observed values divided by the observed value

$$\varepsilon_{rel} = \frac{|PredictedValue - ObservedValue|}{ObservedValue}.$$

5.1 ECMWF data vs ERA5 data

5.1.1 Temperature

- *Ensemble members average and ERA5 data representation*

In the following graphs, the ensemble members' means have been compared with ERA5 data (red line) for both maximum and minimum temperatures. This comparison was conducted for the initialization months of March, April, May, and June across the years 2018 to 2022. Since both ECMWF and ERA5 data are latitude- and longitude-dependent matrices, the spatial mean of all considered grid cells was computed.
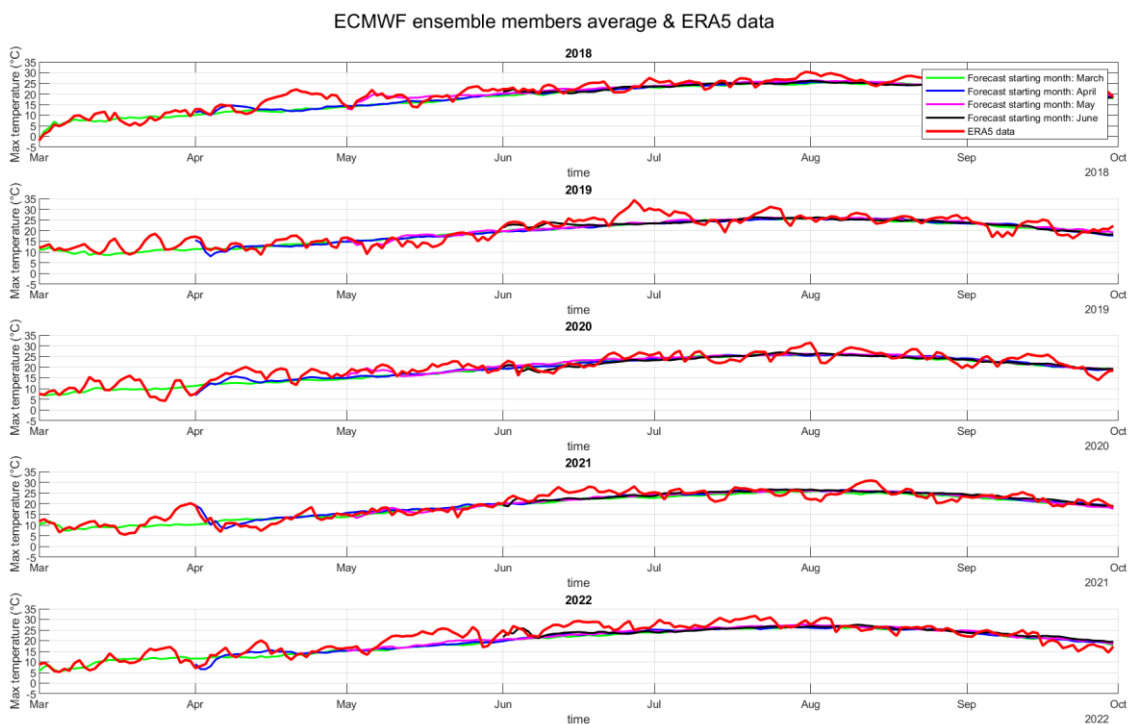
Figure 41. *Representation of ensemble member averages vs ERA5 considering four months of initialisation and five years for Maximum temperature*
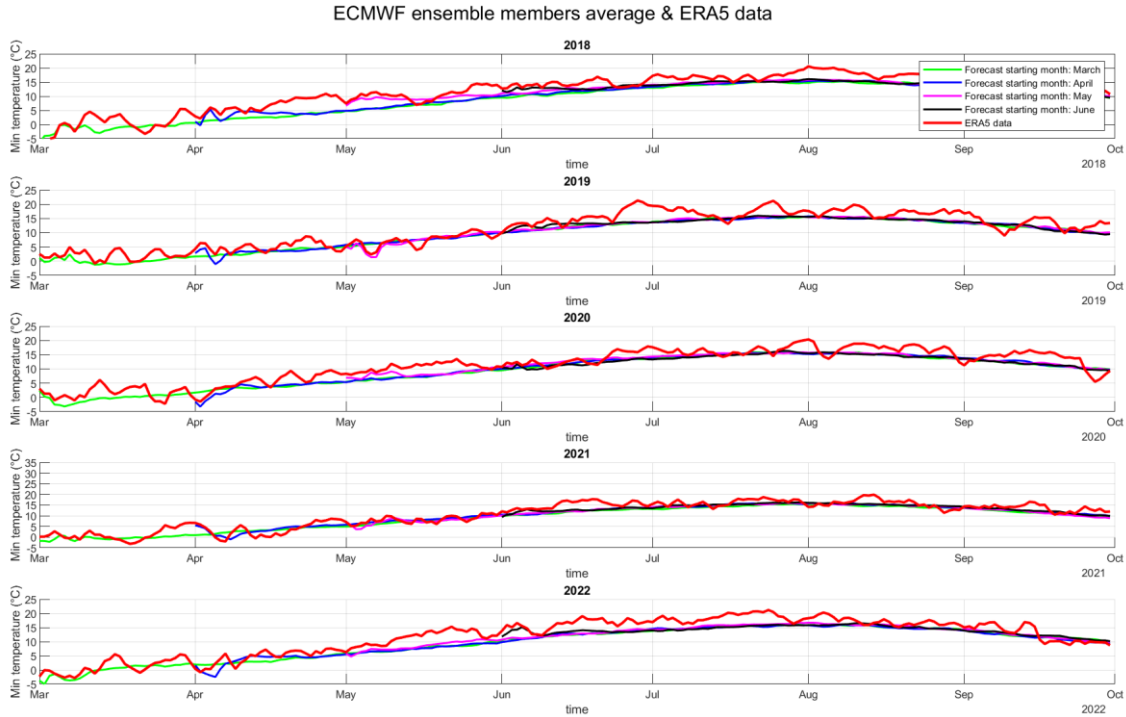
Figure 42. *Representation of ensemble member averages vs ERA5 considering four months of initialisation and five years for Minimum temperature*

In general, the seasonal forecasts seem to capture the observed temperature trends well, especially during the summer months. There is a general consistency in the seasonal temperature patterns, with an increase towards summer (June-July) and a decrease towards autumn (September). However, the accuracy of the forecasts can vary significantly depending on the initialization month, lead time, and the year considered, with a range of error that may be more substantial or less wide.

Discrepancies between the forecasts and observed data are also noticeable particularly during temperature peaks. These discrepancies might be due to anomalous climatic conditions or limitations in the forecasting models. Specifically, ECMWF data show strong coherence with ERA5 data, especially in the initial days of the forecast, with this coherence lasting up to approximately 10 days. As the forecast days increase, and thus the lead time, there is no consistent trend of increasing error margin as one might expect; instead, this varies depending on the initialization month. For instance, maximum temperature forecasts initialized in March 2021 with lead time 0 tend to have a larger error margin compared to those with lead time 2, initialized in the same month.

Moreover, as noted in Chapter 4, forecasts from July onwards tend to converge towards similar ensemble member means across different initialization months. This results in a

similar forecast error, regardless of the initialization month and thus the forecast considered.

Comparing the behaviour of maximum and minimum temperatures, there is a general underestimation of forecasts relative to the reanalysis data in both cases. The trend of both variables in ERA5 and ECMWF data is quite similar. However, it is noteworthy that despite the similar shape of the curves, the error recorded for minimum temperatures is slightly higher. Therefore, this aspect will be analysed more thoroughly in the following paragraphs.

Finally, it is worth highlighting that, when comparing the years, the forecasts for 2022 appear to be the least accurate. This year will be examined in greater detail in Chapter 5.2, where the single ECMWF model and the multi-ensemble MME and MMES will be compared.

- *Comparison with different starting months and different lead times*

In this section, the differences between maximum and minimum temperatures are examined in greater detail. To this end, the forecasts for the year 2018 have been chosen as a case study. The following graphs depict the 90% confidence intervals of the ensemble members, averaged monthly, which were analysed in Chapter 4. These confidence intervals are here compared with the monthly means of the ERA5 reanalysis data (red line).
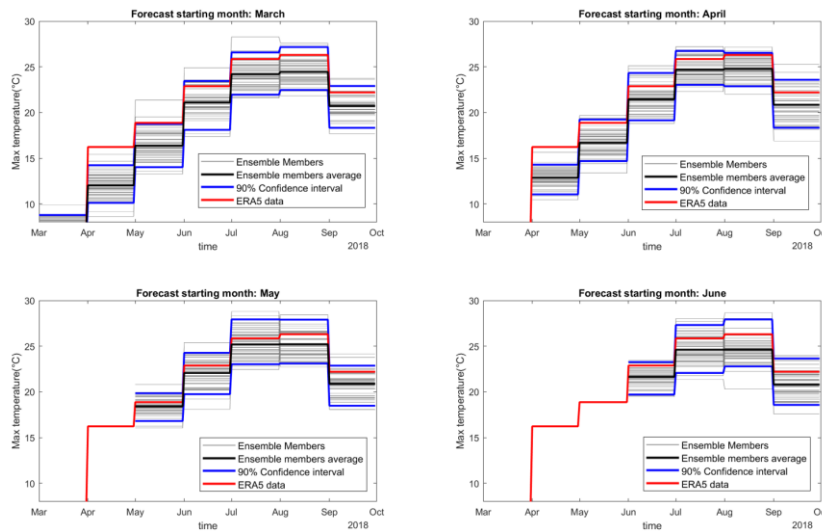

Figure 43. *Monthly average of ensemble members, along with the ensemble mean, including a 90% confidence interval vs ERA5 data for maximum temperatures*
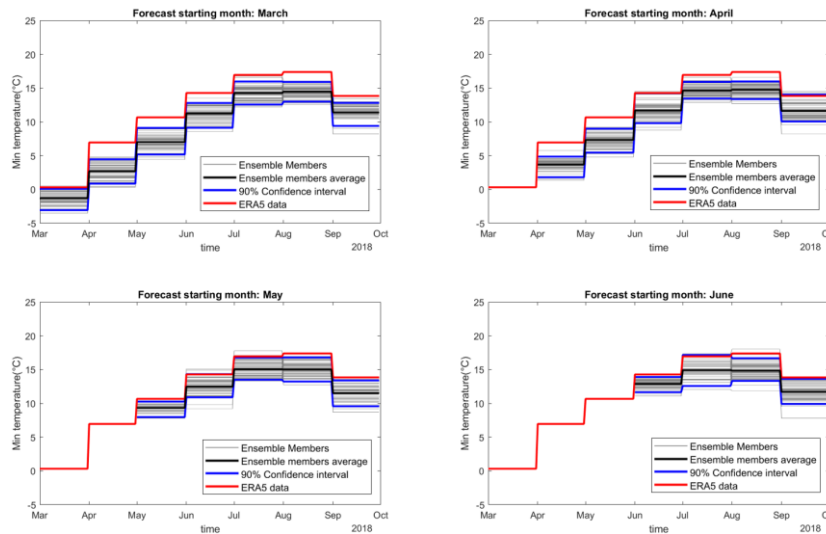
Figure 44. *Monthly average of ensemble members, along with the ensemble mean, including a 90% confidence interval vs ERA5 data for minimum temperatures*

In Chapter 4, it was highlighted that minimum temperatures showed narrower confidence intervals, hypothetically suggesting greater model certainty in estimating these temperatures. However, the graphs indicate that the reanalysis data rarely fall within these intervals, deviating from the ensemble members' means and, in many cases, not fitting within the 90% confidence intervals.

Conversely, for maximum temperatures, ERA5 data consistently fall within the considered intervals. However, it is important to note that these data exhibit greater variability, which results in wider confidence intervals and, consequently, a higher likelihood that the ERA5 data will fall within the 90% confidence intervals. Additionally, it is noted that the highest forecast errors, for both maximum and minimum temperatures, appear to occur in April and May.

To quantitatively assess the error, the following figures show histograms of normalized forecast errors initialized in four different months (from March to June) with reference to three different lead times (0, 2, 3). In this case, the errors are averaged monthly for all 51 ensemble members, focusing on specific forecast months. The error is defined here as the difference between the monthly mean of each ensemble member and the monthly mean of ERA5.
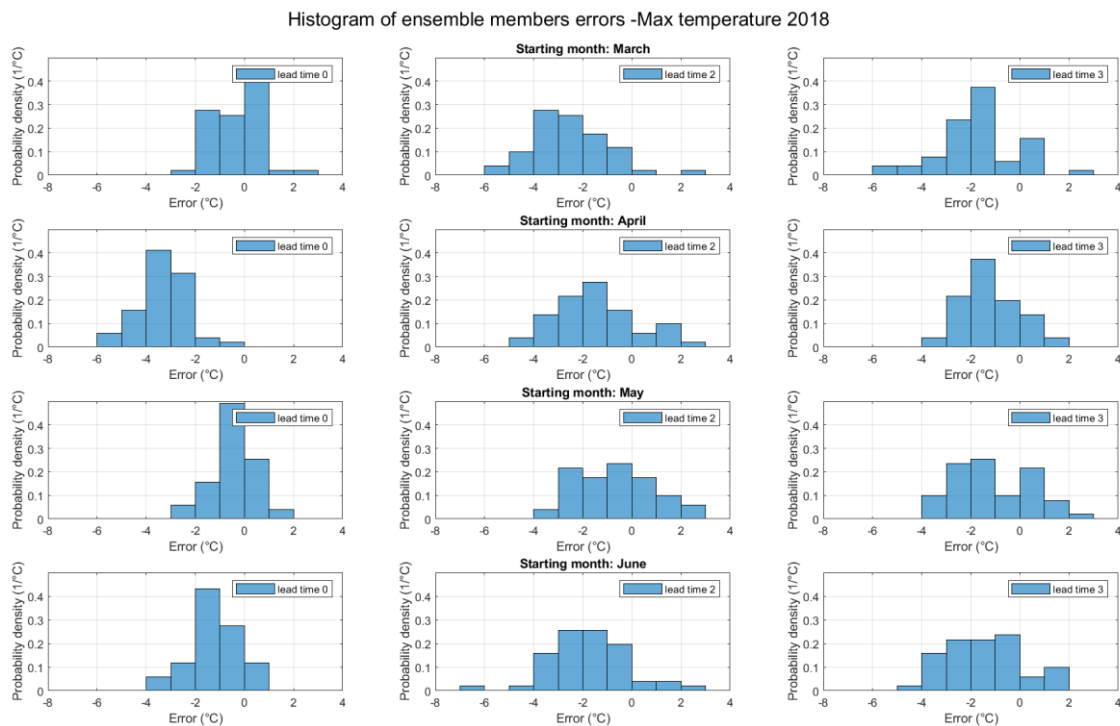

Figure 45. *Histograms of ensemble member errors, comparing four months of initialisation with lead times of forecast 0, 2 and 3 of Maximum temperature*
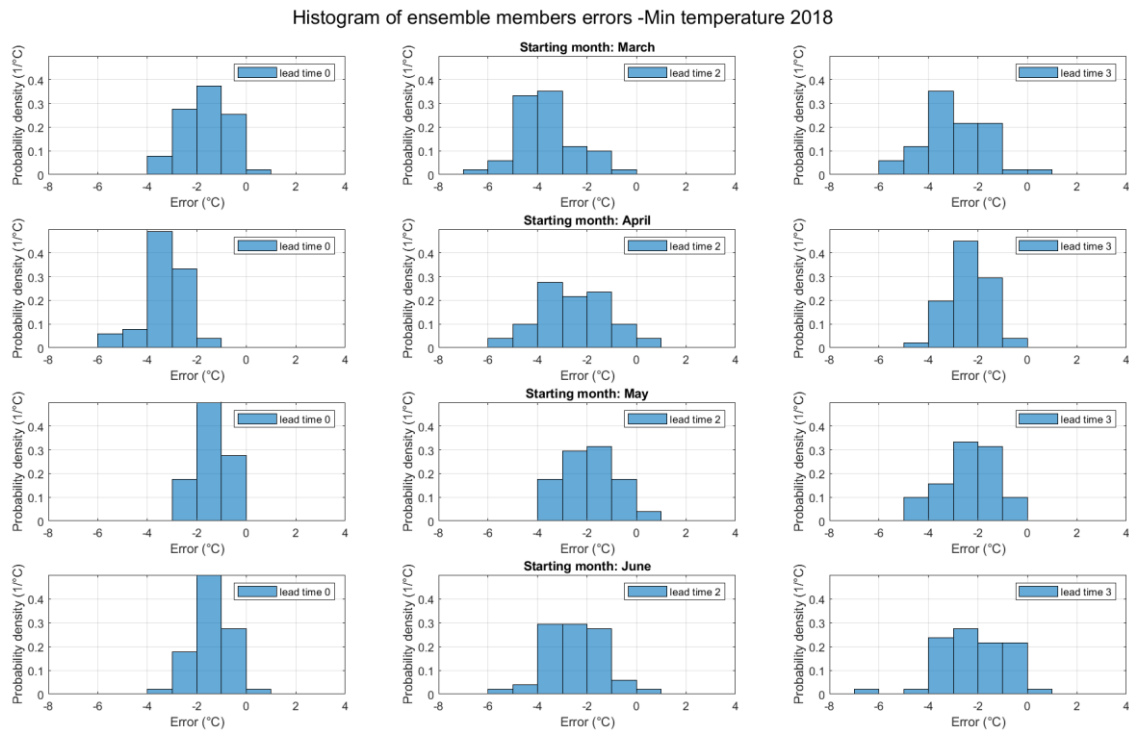
Figure 46. *Histograms of ensemble member errors, comparing four months of initialisation with lead times of forecast 0, 2 and 3 of Minimum temperature*

The graphs show that minimum temperatures have a narrower error range compared to maximum temperatures, as can be seen, for example, in the histograms with an initialization month of May and lead time 0. This results in generally slightly higher probability density values.

For both minimum and maximum temperatures at lead time 0, the histograms are fairly concentrated: most of the data falls around a central value, with few error values in the tails. Conversely, as lead time increases, the data becomes more dispersed.

Overall, most histograms exhibit approximate symmetry, with a slight tendency toward negitive skewness (a longer tail on the left). Based on this observation, in figures 47 and 48, the means of the ensemble members were calculated and then from these were subtracted the monthly mean value of ERA5. In this way, the monthly mean absolute errors were depicted for four different initialization months and four different target forecast months, in order to assess how much the monthly forecast errors generally deviate from zero (indicating a perfect forecast).
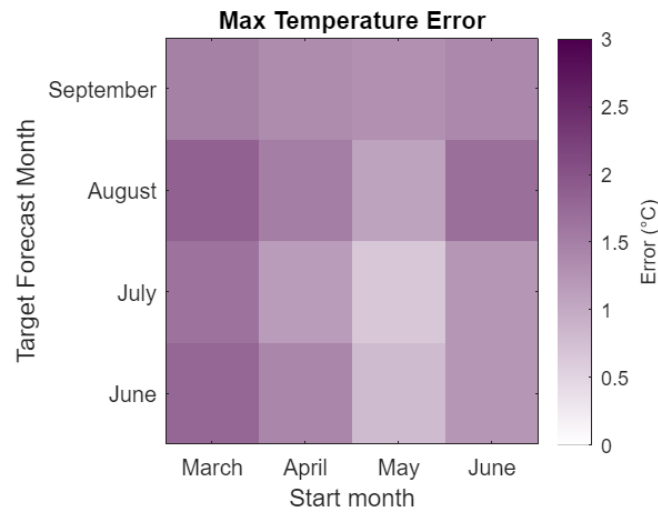
Figure 47. *Representation of monthly average absolute error comparing four different initialization months and four forecast months for Maximum temperature*
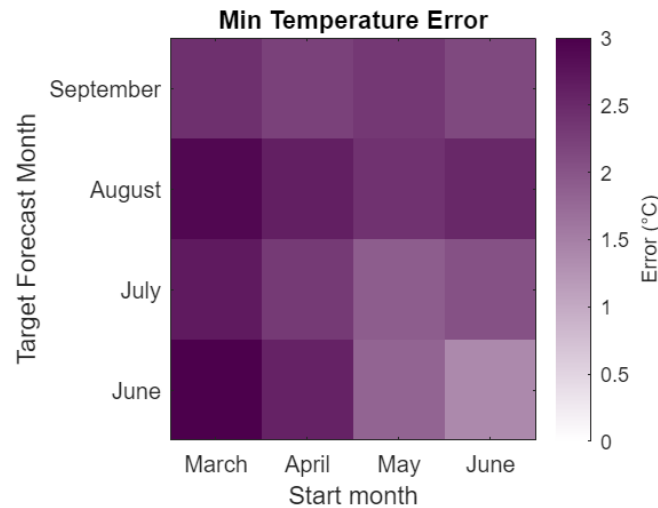

Figure 48. *Representation of monthly average absolute error comparing four different initialization months and four forecast months for Minimum temperature*

As shown in the figures above, the results reflect the previously made observations. Minimum temperatures exhibit a higher systematic mean error of about 1 degree compared to maximum temperatures.

The fact that minimum temperatures exhibit this behaviour, despite having narrower confidence intervals than maximum temperatures, might indicate that the forecasts for minimum temperatures are systematically biased or affected by a structural bias. The narrow confidence intervals suggest low variability in the forecasts, indicating greater

model confidence in estimating minimum temperatures. However, if the model has a systematic bias, this "confidence" might be misplaced, leading to larger errors when compared to reanalysis data.

- *Comparison among different years*

After analysing the behaviour of the sets of 51 members at specific time points, using the year 2018 as the reference year, this section will focus on the differences between the different years from 2018 to 2022, in order to extract possible error trends of the two variables.

The data downloaded from the Climate Data Store, as previously described, are daily data. In the previous section we treated the monthly averages of each ensemble member, but, in order not to lose important information on the temporal resolution, in the following figures we will compare the day-by-day ensemble averages with the daily ERA5 data using scatter plots for the entire duration of the forecast.
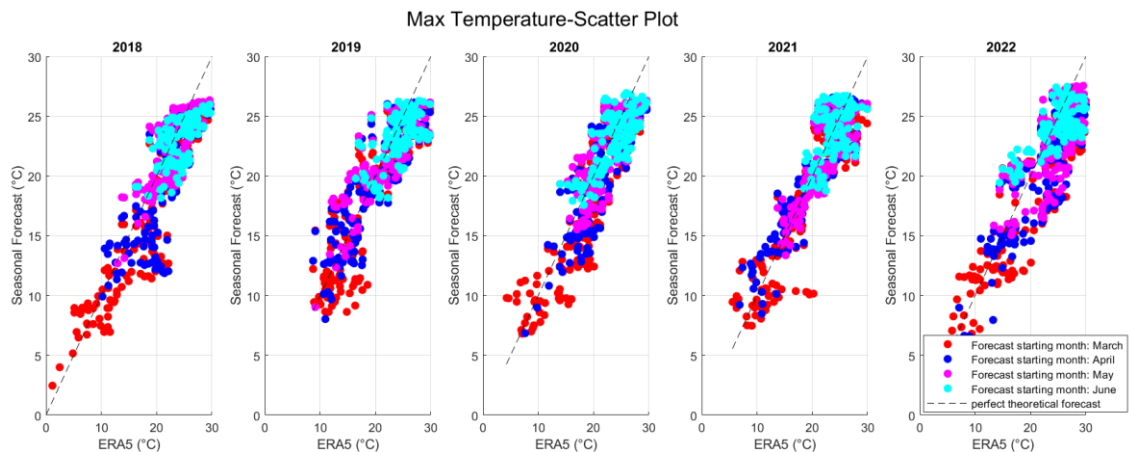

Figure 49. *Scatter plot of ensemble members average vs ERA5 data considering four months of initialisation and five years for Maximum temperature*
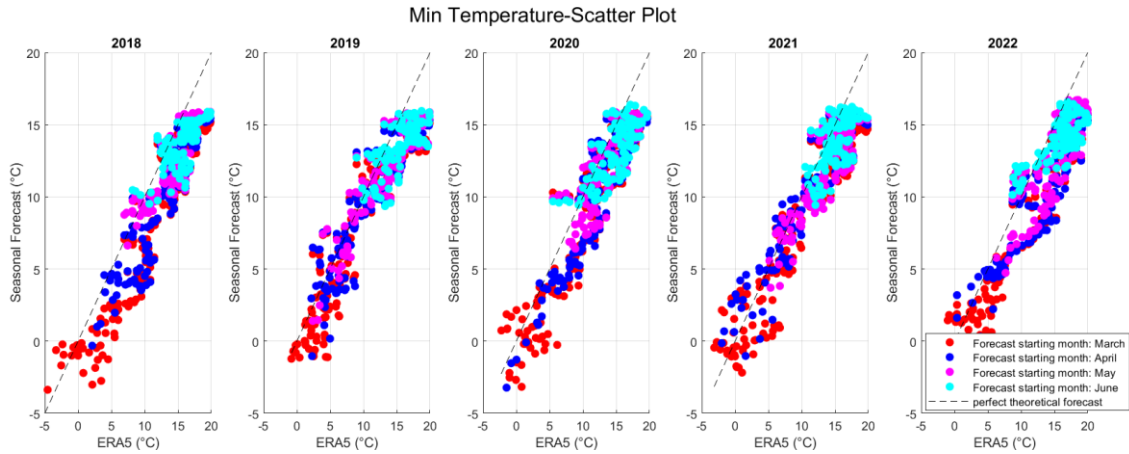
Figure 50. *Scatter plot of ensemble members average vs ERA5 data considering four months of initialisation and five years for Minimum temperature*

The dashed line represents the perfect forecast, i.e., a hypothetical perfect match between ERA5 data and the predictions. As shown in the graphs, maximum temperatures seem to approximate this behaviour more closely over the years, with a nearly symmetrical distribution along the dashed line. However, there are some temperature peaks not captured by the seasonal forecasts, such as those corresponding to the April, May, and June 2019 forecasts and most of the data for 2022.

Conversely, the forecasts for minimum temperatures generally tend to underestimate the values, as seen particularly in 2018, 2020, and 2022.

In the above graphs, all days for each forecast initialized between March and June have been plotted without distinguishing specific lead times. As a result, the amount of data plotted for March is significantly greater than that for June (7 forecast months compared to 4). Therefore, in the following figures, the errors recorded by different forecasts with fixed lead times of 0 and 2 will be compared. The residuals are plotted against the seasonal forecast, with the y-axis representing the error and the x-axis representing the seasonal forecast in °C.
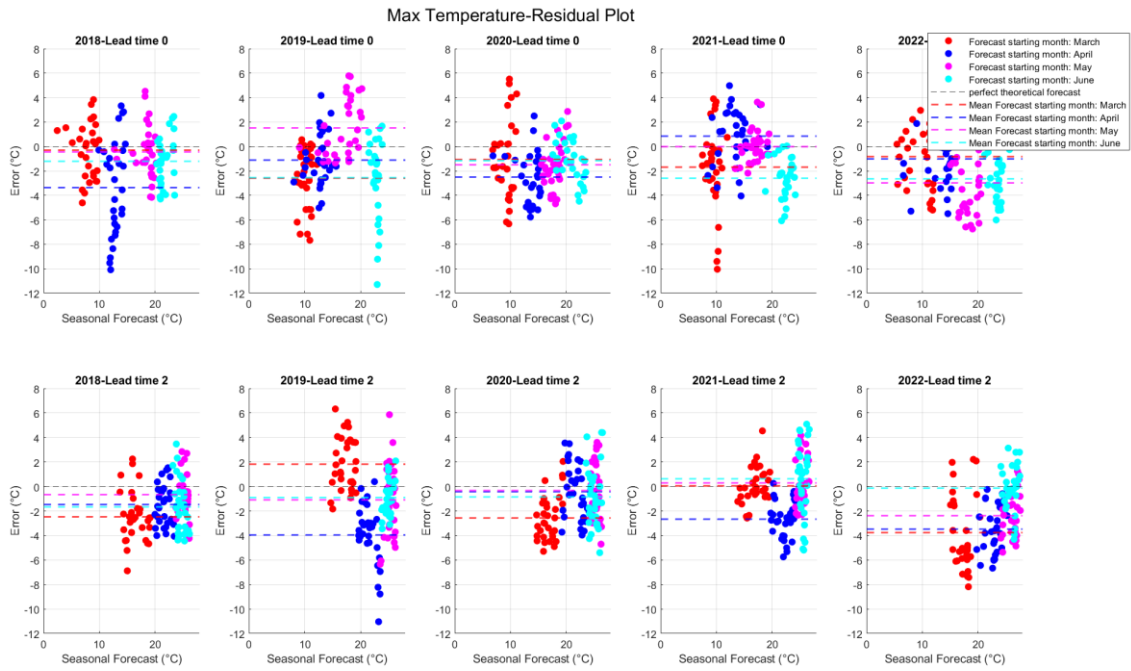
Figure 51. *Residual plot of ensemble members average considering four months of initialisation and five years for Maximum temperature*
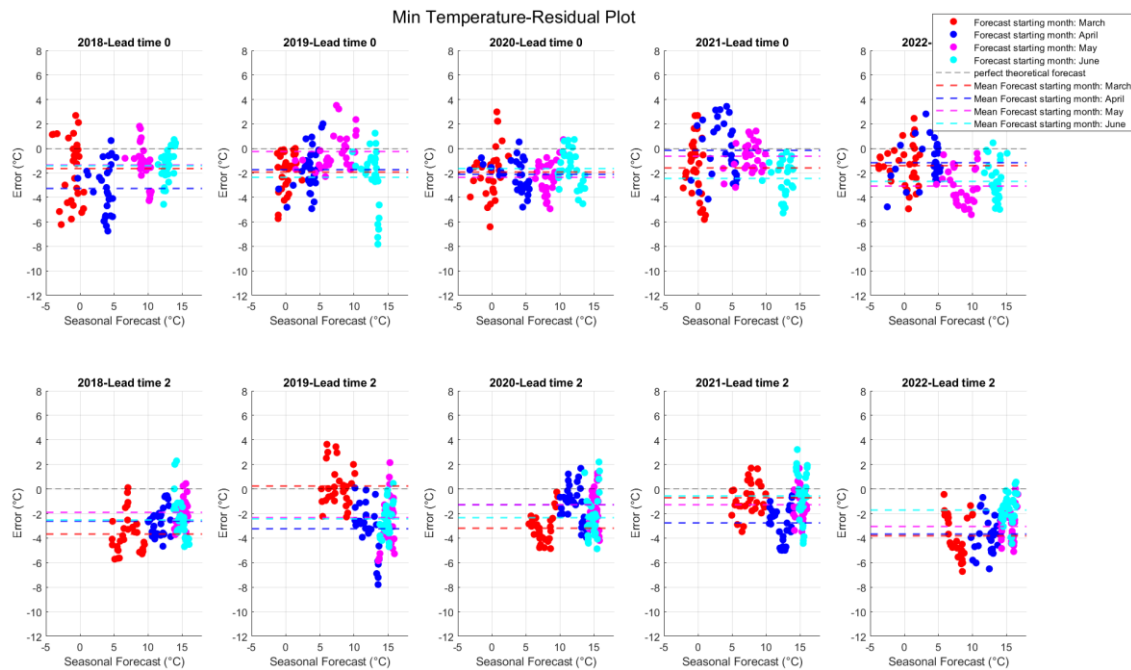

Figure 52. *Residual plot of ensemble members average considering four months of initialisation and five years for Minimum temperature*

The errors in maximum temperature forecasts appear to vary widely, with an error range generally between -8°C and +6°C. The years 2020 and 2021, particularly for lead time 2, show errors more clustered around 0, indicating better forecasts.

Outliers are present in most years, especially for higher forecast values, suggesting occasional significant overestimations or underestimations.

The errors in minimum temperature forecasts seem to be more evenly distributed around the means, although 2019 shows some deviations. Nevertheless, minimum temperatures often exhibit higher average absolute errors compared to maximum temperatures, as was already observed in the previous analysis for 2018. This behaviour is thus repeated over the years.

Comparing lead time 2, the green and pink distributions are those that show lower errors for both maximum and minimum temperatures in almost all years. This means that forecasts of May and June with lead time 2 (for July and August) are the most accurate. It is worth noting that these summer months had already shown low variability among ensemble members, suggesting potentially higher forecast accuracy. This is thus confirmed by the error analysis.

Conversely, for lead time 0, there do not seem to be any initialization months that consistently show lower errors. For example, if in 2018 the month of March shows higher maximum temperature errors compared to other months, in 2021 the same month shows no errors.

- *Bias correction*

In general, for both variables, it can be observed that in few cases the error is close to zero, but almost always it is at least greater than 1°C. For this reason, a bias correction was applied to the forecast data. For this operation, hindcast data from the ECMWF forecasting center from 1993 to 2016 were used and compared with ERA5 data from the same reference period. Through linear regression of the data, it was possible to determine the coefficients a and b and subsequently correct the real-time data from 2018 to 2022.

$$T_{ERA5\_hindcasts} = a * T_{ECMWF\_hindcasts} + b$$

to determine the coefficients a and b, where a is the y-intercept and b is the slope of the line, hindcast data were used. Subsequently, using these coefficients, the corrected real-time data were obtained through the equation:

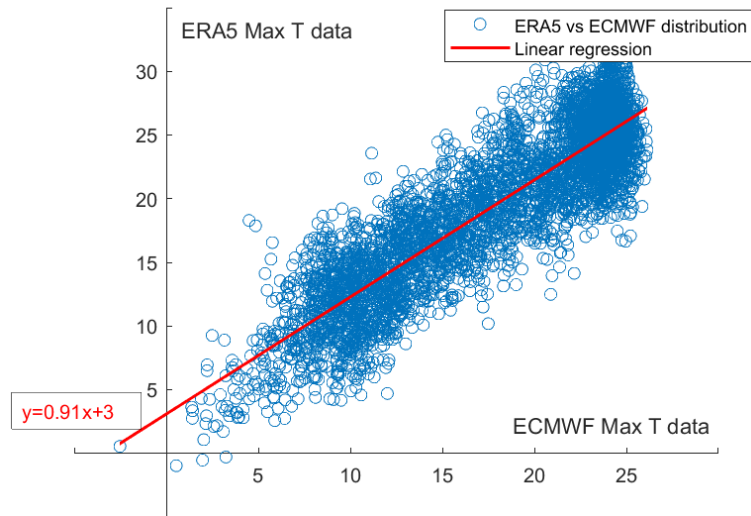$$T_{ECMWF\_realtime} = a * T_{ECMWF\_hindcasts} + b$$

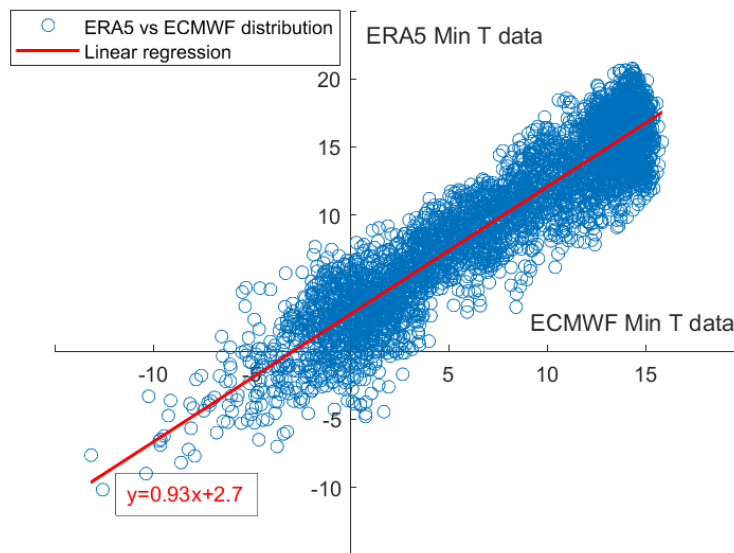***Figure 53.*** *Bias correction Maximum temperature*

Figure 54. *Bias correction Minimum temperature*

The coefficients found for maximum and minimum temperatures are respectively:

- T_{max} : a=0.91, b=3;

- T_{min} : a=0.93, b=2.7.

After applying bias correction, the errors for the two variables were re-evaluated with the following results:
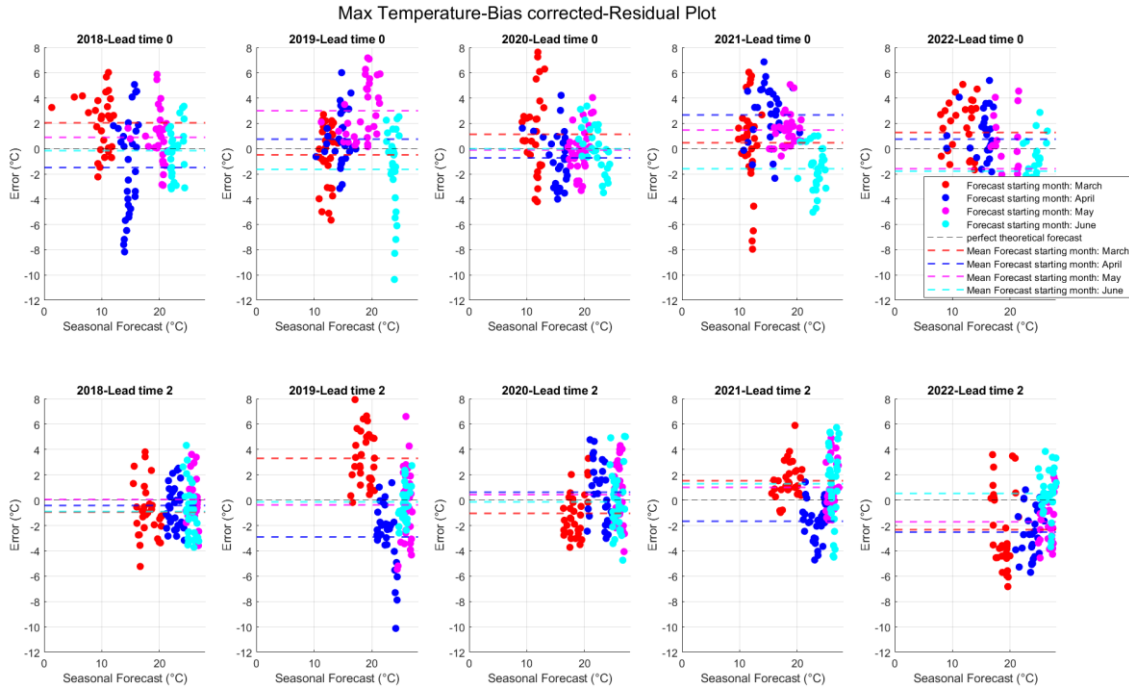
Figure 55. *Residual plot of ensemble members average considering four months of initialisation and five years for Maximum Bias-corrected temperature*
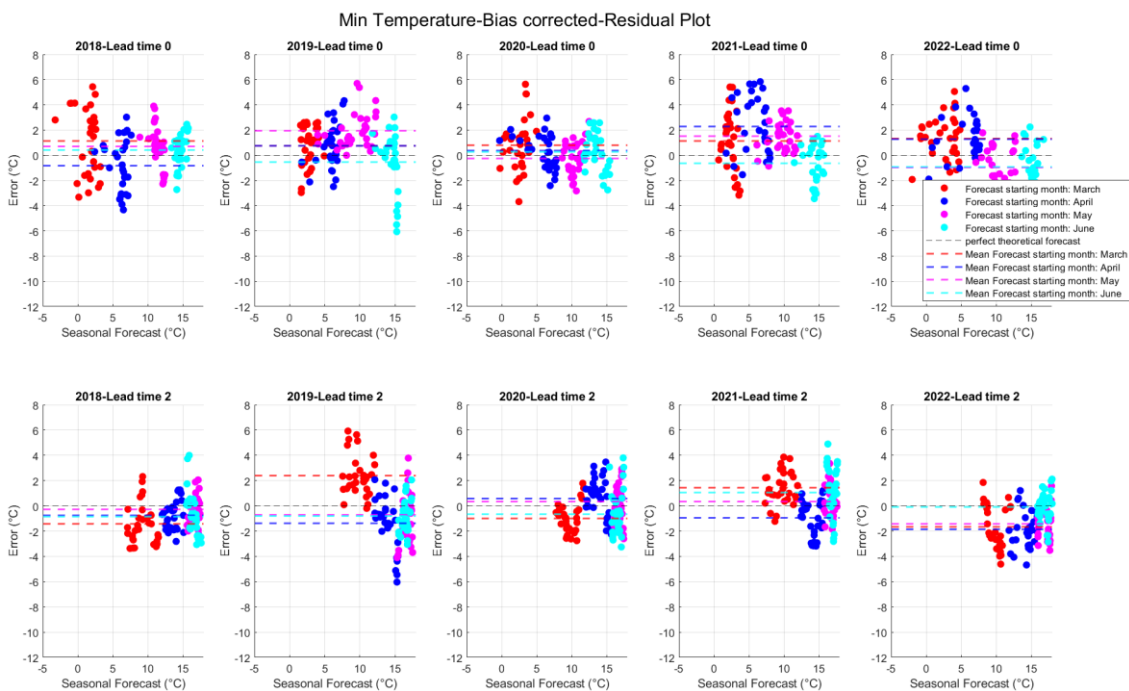

Figure 56. *Residual plot of ensemble members average considering four months of initialisation and five years for Minimum Bias-corrected temperature*

After bias correction, it is immediately evident that the errors have decreased, with maximum absolute mean errors rarely exceeding 2°C. In particular, the mean errors for

minimum temperatures are nearly zero, confirming the greater accuracy of these forecasts, as indicated by the lower variability among ensemble members.

There is also a significant improvement in the accuracy of maximum temperature forecasts. However, in some cases, there appears to be an overestimation of the forecasts, such as for the months from March to May with lead time 0.

In general, moreover, a trend in the quality of the forecasts at different lead times can be appreciated after bias correction. If we look at the graph of minimum temperatures, for example in June 2019 with lead time 0, the error recorded is about 0.25°C, whereas the forecast on June from April (lead time 2) reaches about 1.8°C. This behaviour can be seen in several cases under consideration. This means that as the lead time increases, generally but not always, there is a loss in forecast accuracy.

- *Spatial representation and comparison*

Finally, the following representations analyse the spatial distribution of errors. In this analysis, the difference between the monthly mean value of the ensemble members for each forecast cell and the corresponding value from the reanalysis data was calculated.

At this stage, only maximum temperature errors are reported, as minimum temperatures exhibit a similar spatial distribution but with generally lower and more accurate error values.
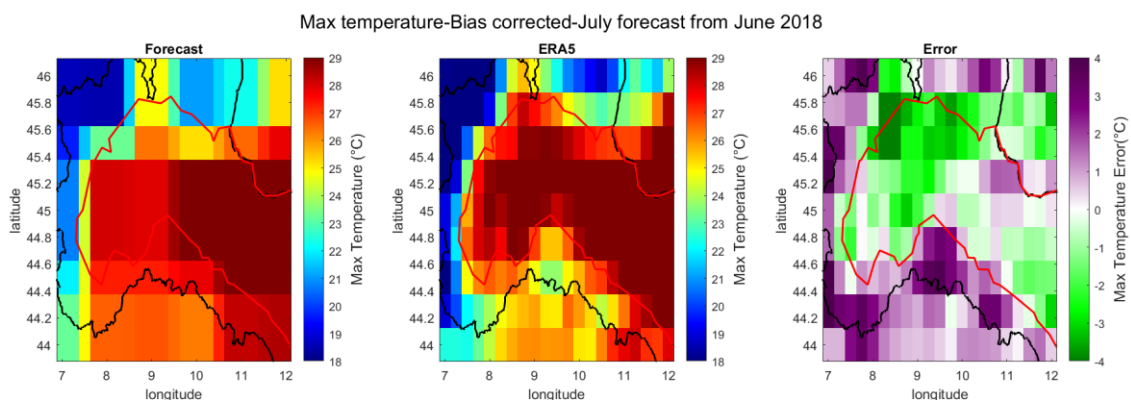

Figure 57. Spatial distribution of the monthly mean Maximum Bias-corrected temperature of ECMWF forecast, of the monthly mean Maximum Bias-corrected temperature of ERA5 and of the error with a starting month of June and a forecast month of July

From the figures above, it is observed that there is a general underestimation of forecasts compared to reanalysis data at latitudes under 45°, while further south, the errors are almost always positive. In the plain region, where agricultural areas are concentrated, forecasts generally show errors close to zero.

5.1.2 Precipitation

- *Ensemble members average and ERA5 data representation*

After a thorough analysis of temperature forecasts, we will now focus on evaluating the accuracy of precipitation forecasts. The following plots show the mean cumulative precipitation from ensemble members for ten-day intervals, compared with the ten-day cumulative precipitation from ERA5 data, for all years and initialization months.
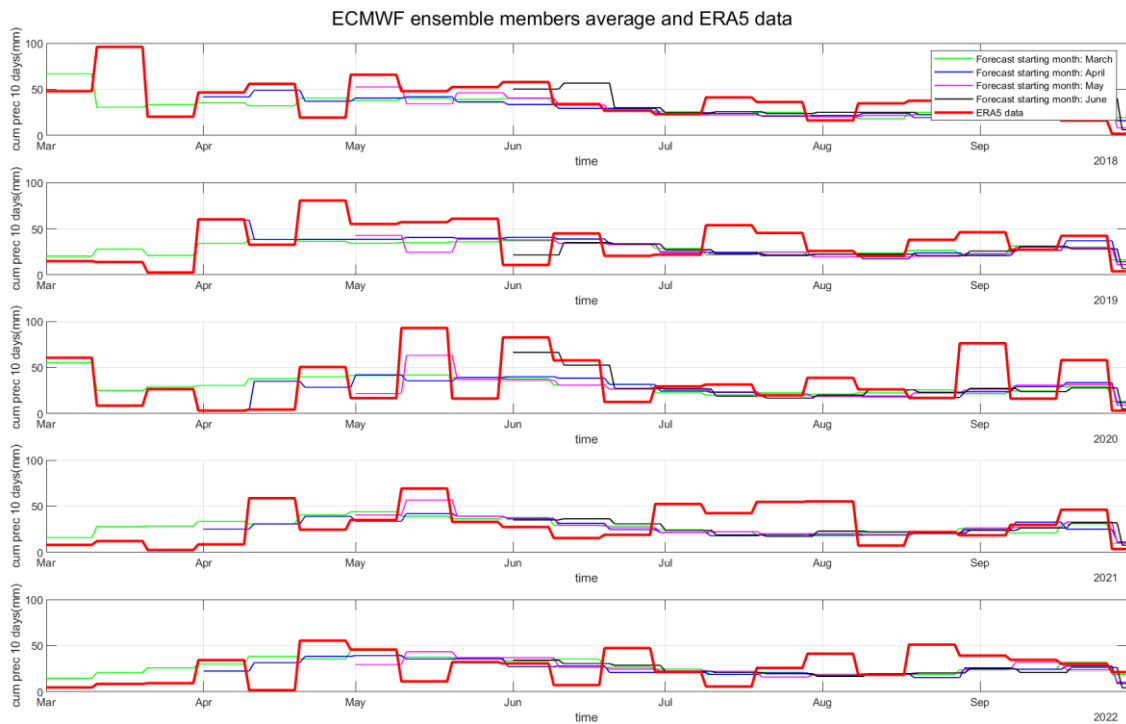

Figure 58. Representation of ensemble member averages vs ERA5 considering four months of initialisation and five years for 10 days interval cumulative precipitation

Overall, it is observed that forecasts (colored lines) often deviate from the observed data (red line). The accuracy of the forecasts varies both across years and between different model runs. Some years show better alignment between forecasts and observations, while

others exhibit more pronounced discrepancies. For example, in 2018, the forecasts appear to align more closely with the observed data, except for a few outliers, such as the second interval in March.

As with temperatures, at lead times greater than one month, forecasts tend to converge towards constant values, showing a form of stabilization. For July and August, for instance, forecasts change little between years, stabilizing around 20 mm, indicating a loss of memory of the actual system conditions and not accounting for more intense rainfall events, as shown in the July 2021 plot. However, for other years like 2018 and 2020, forecasts for July and August appear to have lower errors compared to the spring months. This may be due to greater climatic stability during the summer period.

- *Comparison with different starting months and different lead times*

After briefly highlighting the differences between the ensemble members' means and ERA5 data, this paragraph will focus on the specific behaviour of the individual ensemble members. The 90% confidence interval has also been included to better illustrate the variability among ensemble members relative to the reanalysis data. The following example presents the ten-day cumulative forecasts for the year 2018, with different initialization months ranging from March to June.
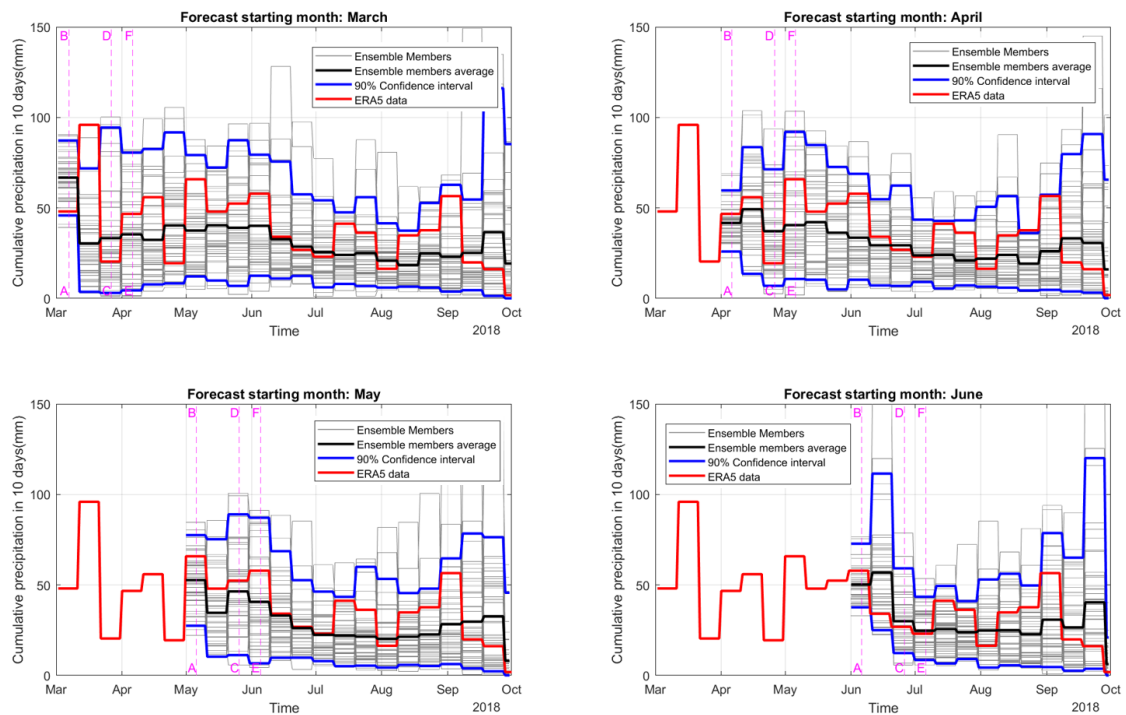

Figure 59. *10 days cumulative precipitation of ensemble members, along with the ensemble mean, including a 90% confidence interval vs ERA5 data*

In paragraph 4, we discussed the variability among ensemble members, noting that the first ten-day interval showed narrower variability compared to the second and subsequent intervals. Additionally, it was observed that spring months exhibited greater variability compared to summer months, which showed narrower confidence intervals. Theoretically, this behaviour should indicate better or worse forecast precision. However, the figure above shows that the variation in error appears to be almost random.

Specifically, when comparing forecast months April and July with March as starting month, it is observed that despite April's confidence interval being much wider than that of July, the errors recorded in April's intervals are comparable to those in July's intervals.

A closer comparison of the first and second forecast intervals relative to the starting month shows that in March, there is consistency between variability and error (greater variability corresponds to higher average error). In contrast, in April, although the second forecast interval has wider variability, the average error is similar to that of the first interval.

It can therefore be said that, unlike temperatures, precipitation does not show any significant relationship between variability and accuracy, neither as the lead time decreases (one-month forecasts by changing the month of initialisation seem to register more or less similar errors) nor when referring to summer/spring months (in some June and July intervals, the error is almost zero but in other intervals the error can be very high, not corresponding with the low variability).

Finally, the histogram of normalized errors of ensemble members is presented for three ten-day intervals with different initialization months, where the total area under all bars is normalized to 1. "Lead interval 0" refers to the AB interval from the previous figure, "lead interval 2" to the CD interval, and "lead interval 3" to the EF interval.
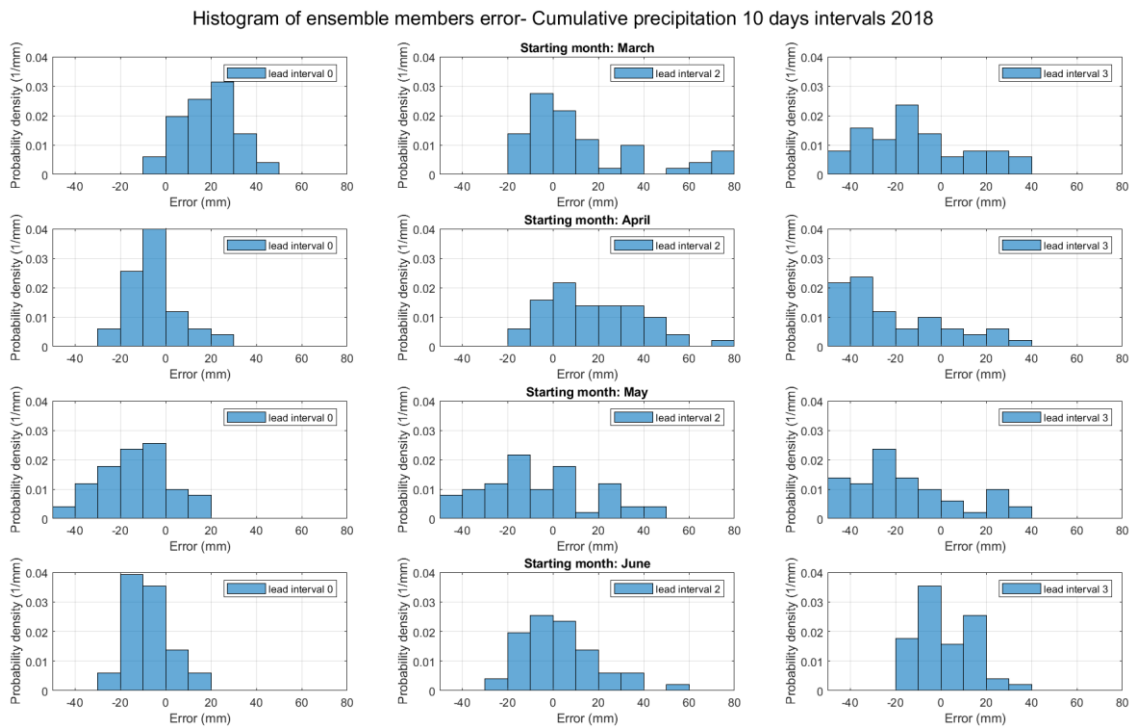

Figure 60. *Histograms of ensemble member errors, comparing four months of initialisation with lead intervals of forecast 0, 2 and 3(AB, CD, EF) of 10 days interval cumulative precipitation*

Key aspects are highlighted by the figure. Forecast intervals with "lead interval 0" show narrower error ranges compared to other intervals, with values relatively concentrated around specific values and a maximum absolute error of about 20 mm. From interval 2 onwards, however, errors appear much more dispersed, with significantly higher error values.

Finally, since the error is calculated as the difference between ERA5 values and the average value of each ensemble member, it is not possible to determine a priori whether there is a general overestimation or underestimation of forecasts for the considered intervals. To provide an initial quantitative assessment of the error for ten-day cumulative forecasts, the average errors and the absolute average errors for the three intervals from June to September are presented, comparing different starting months.
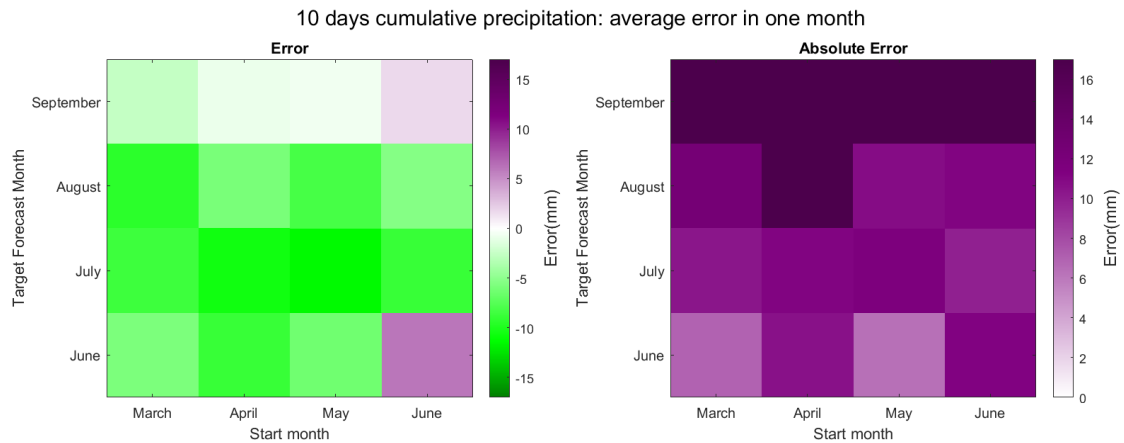

Figure 61. *Monthly average of ensemble members of three intervals in one month comparing four starting months and four forecast months*

In general, it can be concluded that, despite the high variability among ensemble members, the ten-day intervals have achieved a maximum average monthly error of 15 mm compared to the actual data. In September, very low error values are observed; however, this does not imply that every interval has an ensemble member average that closely matches the ERA5 data. As shown in Figure 61, September shows the highest absolute errors.

- *Comparison among different years*

After thoroughly examining the variability of errors among ensemble members, the next step is to compare the ensemble means (ten-day cumulative forecasts) with ERA5 data across different years. The aim is to identify any similarities and differences between the

years and quantify the errors. It should be noted that errors for all intervals corresponding to each initialization month will be depicted. Therefore, forecasts starting in March will cover 21 intervals (7 months), forecasts starting in April will cover 18 intervals (6 months), and so forth.
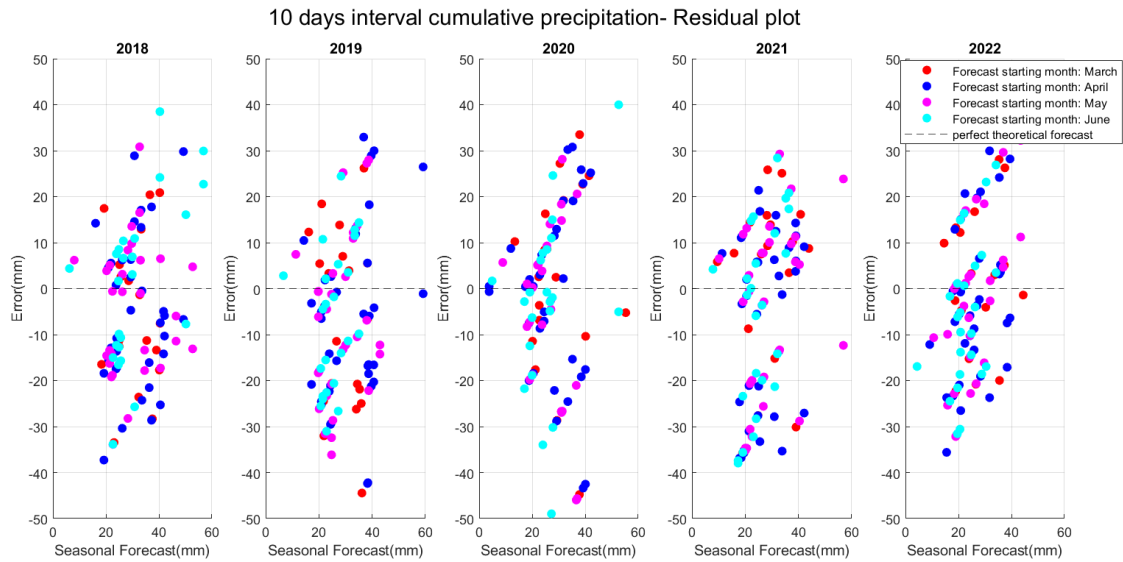

Figure 62. *Residual plot of ensemble members average considering four months of initialisation and five years for 10 days interval cumulative precipitation*

In general, it is observed that across the years analysed for all considered forecasts, there is no clear tendency towards overestimation or underestimation. For the years 2018-2020, the forecasts exhibit positive and negative errors that are evenly distributed, whereas 2021 is primarily characterized by overestimation and 2022 by underestimation of the forecasts. Additionally, when comparing different initialization months, no particular month shows significant advantages. Instead, the magnitude of the error is comparable to the forecast values themselves. Therefore, in the absence of evident systematic errors, unlike those observed for temperatures, it was deemed unnecessary to apply a bias correction.

In conclusion, it can be stated that, on a monthly average, the error of ten-day cumulative precipitation forecasts is relatively low. However, the error for individual ten-day intervals can be quite high, leading to a generally poor quality of the forecasts.

- *Spatial representation and comparison*

Finally, the spatial distribution of errors in ten-day cumulative precipitation forecasts is to be depicted. For this purpose, the error pattern is analysed for the month of July, with June 2020 as the initialization month.
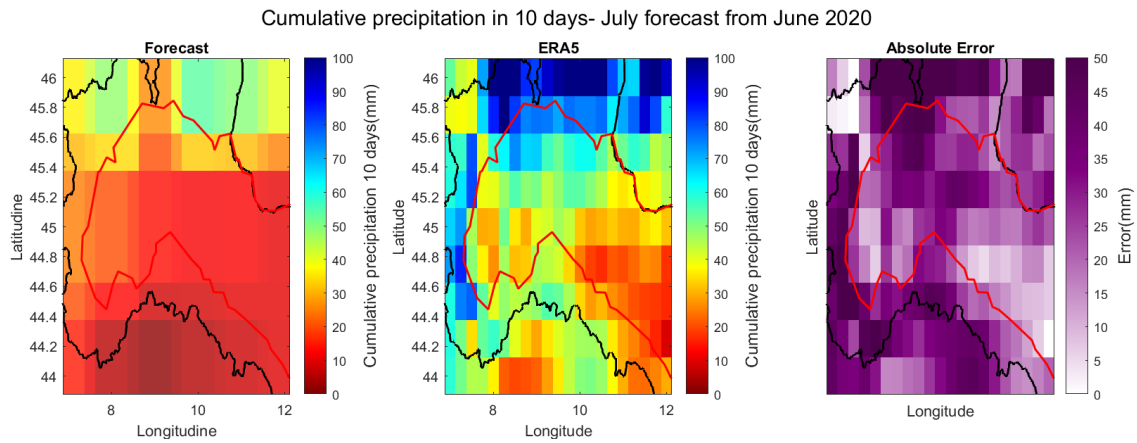

Figure 63. Spatial distribution of the monthly mean of three 10 days intervals cumulative precipitation of ECMWF forecast of the monthly mean of three 10 days intervals cumulative precipitation of ERA5 and of the error with a starting month of June and a forecast month of July

In the analysed example, the higher errors tend to be more prevalent in areas with altitudes greater than 400 meters. This pattern, combined with the spatial distribution of temperature errors, suggests a potential improvement in the accuracy of estimates for irrigation needs.

5.2 MME-MMES-ECMWF data vs ERA5 data

In this subsection, we will compare the errors produced by the single ECMWF model forecasts with those of the multi-model ensembles, MME and MMES, for the year 2022. As highlighted in Chapter 4.2, the MME multi-model showed the least variability among the ensemble members for both precipitation and temperature. Therefore, in theory, we can expect these models to have lower errors.

5.2.1 Temperature

To compare the three models, it was decided to evaluate their error at two different lead times (0 and 2) for all initialization months. The following figure shows all the daily errors for the respective forecast months, along with the average error across all days (dashed line).
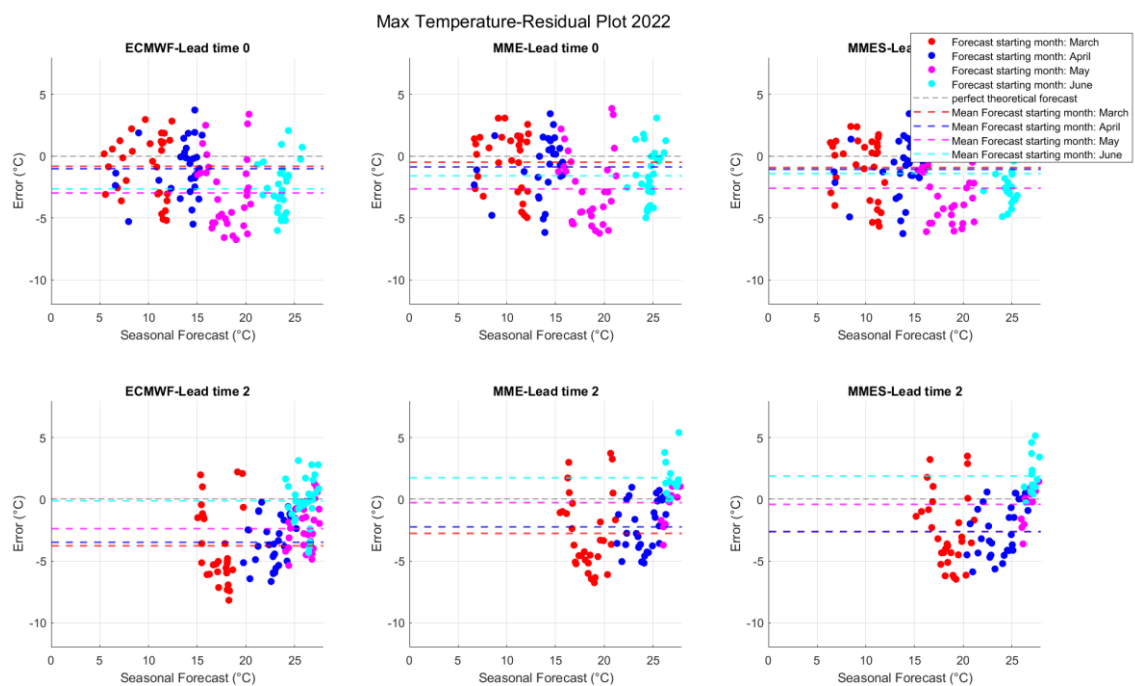
***Figure 64.** Residual plot of ensemble members average considering four months of initialisation at lead times 0 and 2 for Maximum temperature assessing ECMWF, MME and MMES behaviours*

As shown in the figure above, the three models appear to record very similar errors.

For lead time 0, the error range covered by the three models is quite similar (ranging from a minimum of -6°C to a maximum of around 5°C). However, for lead time 2, the error distribution for the MME and MMES models appears to be shifted upward compared with

the error distribution for ECMWF, suggesting that on more days, the forecasts from the two multi-models tend to be more accurate.

It is important to note, however, that the lowest average errors for both lead times are recorded by the ECMWF single-model, with differences compared to the multi-models ranging from a few tenths of a degree (as in March at lead time 0) to 2°C (as in May at lead time 2). Only in the case of June at lead time 2 are higher error values observed for the multi-models.

For a more detailed comparison among the three models, considering both the high temporal resolution (daily) and all the ensemble members, a probability plot (P-P plot) has been used. This plot represents theoretical probability values against observed values. Specifically, the process for evaluating model calibration through the construction of a P-P plot begins with calculating confidence intervals. For each day, ensemble members are sorted in ascending order, and a confidence interval is defined for each probability level using the appropriate percentiles of the ordered distribution. Subsequently, it is checked whether the daily observation falls within these confidence intervals, and this information is recorded over a period of 31 days. The observed probability is then calculated as the fraction of days on which the observation falls within the corresponding confidence interval. Finally, a P-P plot is constructed, where the x-axis represents the theoretical probability and the y-axis represents the observed probability. If the model is well-calibrated, the points on the plot should align along the diagonal, indicating a good match between theoretical and observed probabilities.

The shape of the resulting curve allows for determining whether the data follow an approximately uniform distribution (along the bisector) and for identifying the causes behind a less linear shape, as the slope of the curves increases in areas where more points accumulate.[42]
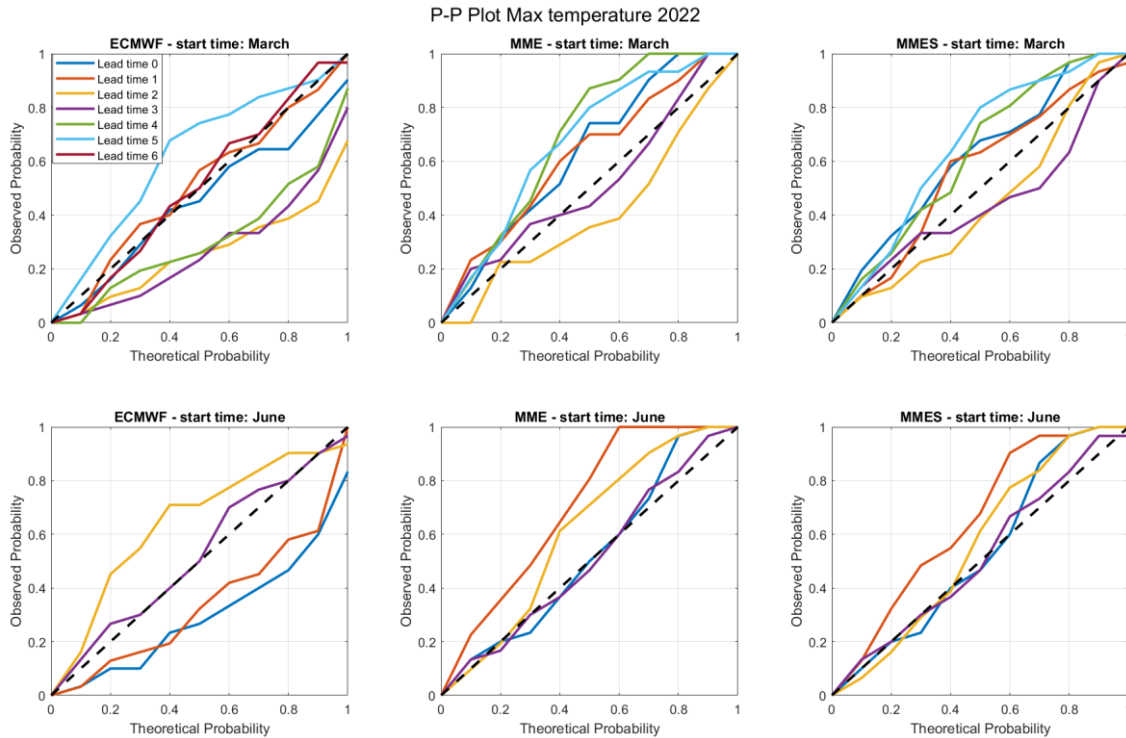
Figure 65. *Probability-Probability Plot of ensemble members considering two months of initialisation at all possible lead times for Maximum temperature assessing ECMWF, MME and MMES behaviours*

From the figure above, for forecasts with a starting month of March, it is observed that the P-P plot curves for the single ECMWF model at lead times 0 and 1 are closer to the bisector compared to those of the multi-models. Conversely, at higher lead times, such as 2 and 3, the curves for ECMWF show a greater overestimation of the proportion of days where the observation falls within the corresponding confidence interval compared to the curves of the multi-models. The multi-model curves, while still remaining below the bisector, exhibit a less pronounced deviation.

Regarding forecasts with a starting month of June, the MME multi-model appears to have a good correspondence between theoretical and observed probabilities, particularly at lead times 0 and 3. This indicates that the model's ability to predict the probability of an observation falling within the confidence intervals is accurate and reflects the actual probabilities observed in the sample.

5.2.2 Precipitation

In the previous paragraphs, the error related to precipitation was assessed in terms of 10-day cumulative totals, as a day-by-day evaluation would have resulted in excessively

large errors. In this paragraph, however, to exclusively evaluate any differences or similarities in accuracy among the models, the temporal resolution of the downloaded data, which is daily, will be utilized.
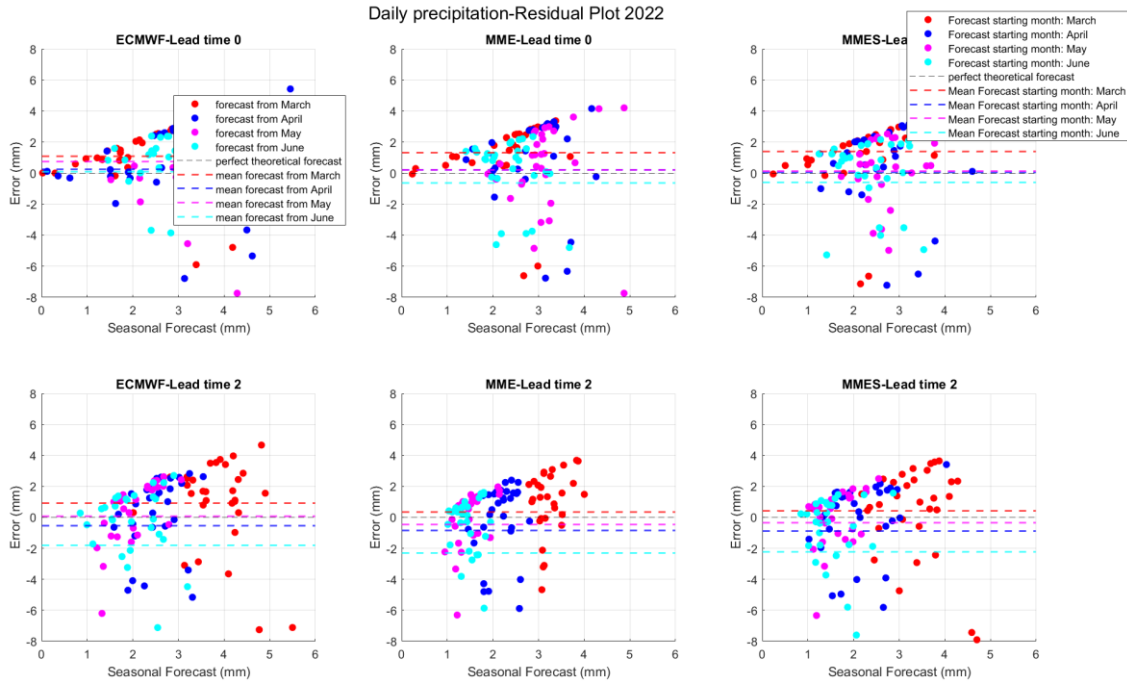

Figure 66. *Residual plot of ensemble members average considering four months of initialisation at lead times 0 and 2 for daily precipitation assessing ECMWF, MME and MMES behaviours*

As with temperatures, it is observed that the range of daily errors is quite similar across the models, except for a few outliers. Additionally, the errors seem to increase with higher daily precipitation forecasts. Specifically, a linear behaviour is noted when there is an overestimation in forecasts (positive error). Furthermore, although the error ranges vary widely, the monthly averages of these errors are comparable to the forecasts themselves.

Similarly, for precipitation, the uniformity of the data was assessed using the P-P plot.
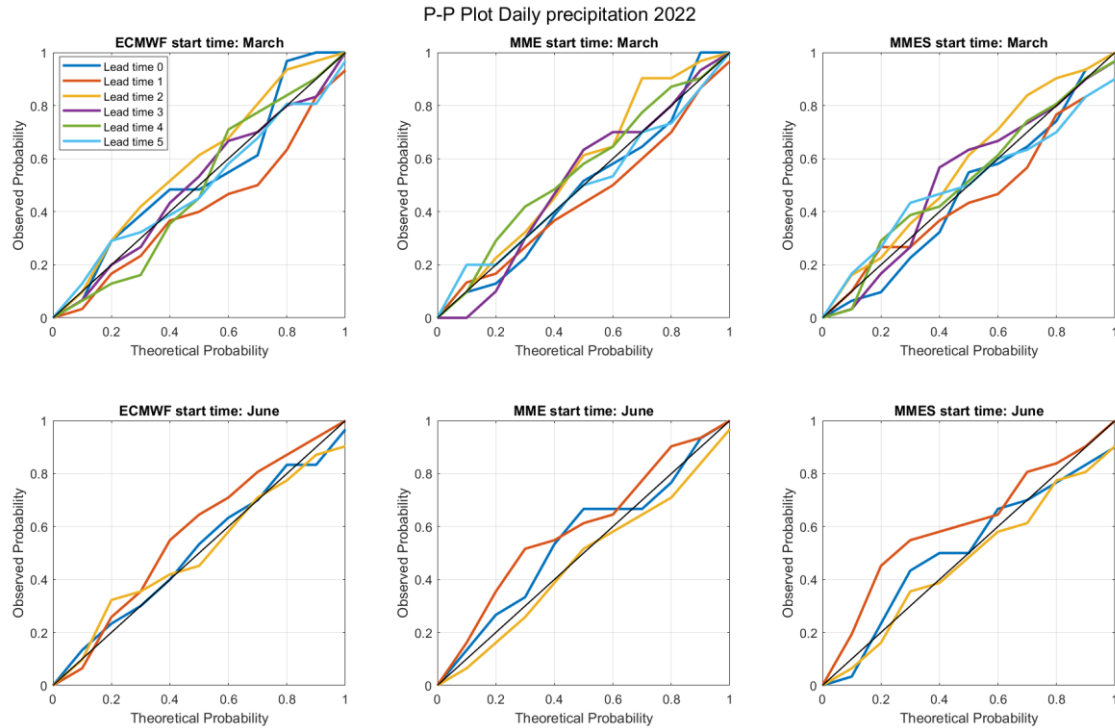
Figure 67. *Probability-Probability Plot of ensemble members for two months of initialisation at all possible lead times for daily precipitation assessing ECMWF, MME and MMES behaviours*

From the figure above, a greater overall uniformity in the precipitation data compared to temperature data is evident. However, it is important to emphasize that this does not imply that the forecasts are accurate, as demonstrated by the recorded errors.

When comparing the different models, a slightly better correspondence between the curves and the bisector is observed for the MME model, particularly for the starting month of March. Conversely, for forecasts starting in June, the predictions appear to be more uniform for the single ECMWF model.

5.3 ECMWF data vs ERA5 data vs ARPA data on two specific grid cells

Finally, a more detailed examination of the differences between two cells described in Chapter 3.3.2 will be conducted, comparing the single ECMWF model data with both ERA5 reanalysis data and observations from ARPA Piemonte meteorological stations. It is important to note that one cell is located in the "plain" area (latitude 44.75, longitude 7.75, elevation 268.59 m), while the other is in the "mountain" area (latitude 44.5,

longitude 7.25, elevation 1290.37 m). As observed in the previous sections, cells at elevations above 400 m showed larger errors for both temperature and precipitation. Therefore, to compare the quality of the forecast data for these two cells, precipitation was chosen as the representative variable. The following figure illustrates the errors of monthly cumulative precipitation, averaging the ensemble members across different lead times and initialization months.

Figure 68. *Representation of monthly cumulative absolute error respect with ERA5 and observation data comparing four different initialization months and four forecast months for precipitation at latitude=44.5 and longitude=7.25*

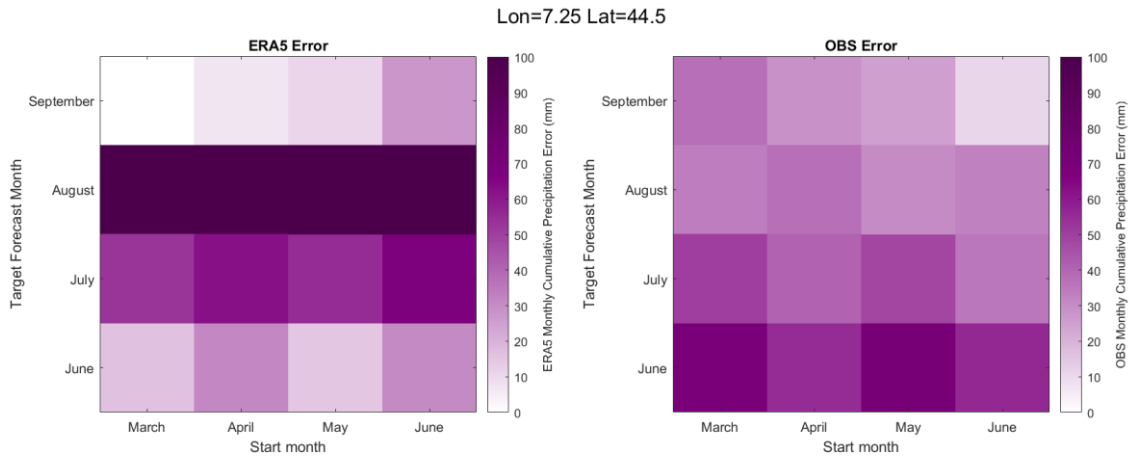***Figure 69.*** *Representation of monthly cumulative absolute error respect with ERA5 and observation data comparing four different initialization months and four forecast months for precipitation at latitude=44.75 and longitude=7.75*
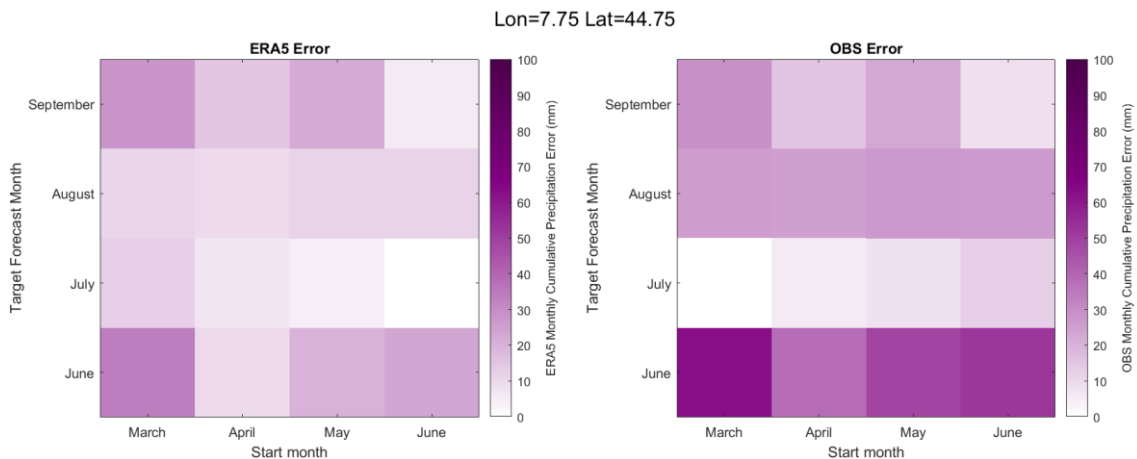
In the figure showing the spatial distribution of errors across the Po River basin, the forecast reference was taken for July based on June 2018. These images aim to

demonstrate that, regardless of the lead time and initialization month, the cell at latitude 44.5 and longitude 7.25 consistently shows much higher average errors compared to the cell at an elevation below 400 m, both when compared with ERA5 data and observational data, with errors that are nearly double.

It can thus be concluded that the larger errors observed in the previously averaged spatial data are primarily due to the cells at higher elevations. This may be attributed to the rapid altitude changes within a single grid cell in mountainous areas, which can introduce significant errors. Additionally, the high complexity of atmospheric conditions in mountainous regions poses challenges for low-resolution models to represent accurately. It is worth noting that the ECMWF data used have a resolution of 1° and were adapted to a 0.25° grid, while retaining their original resolution.

Comparing the errors with ERA5 and ARPA Piemonte data, it emerges that for both cells, the largest errors are associated with the comparison with ARPA Piemonte. For the plain cell, the difference is a few millimeters, except for forecasts in June, while in the mountain cell, the observations show more uniform errors compared to the ERA5 errors, which vary significantly depending on the forecast month.

To further understand the distribution of ensemble members relative to reanalysis and observational data, P-P plots are also presented for both cells.
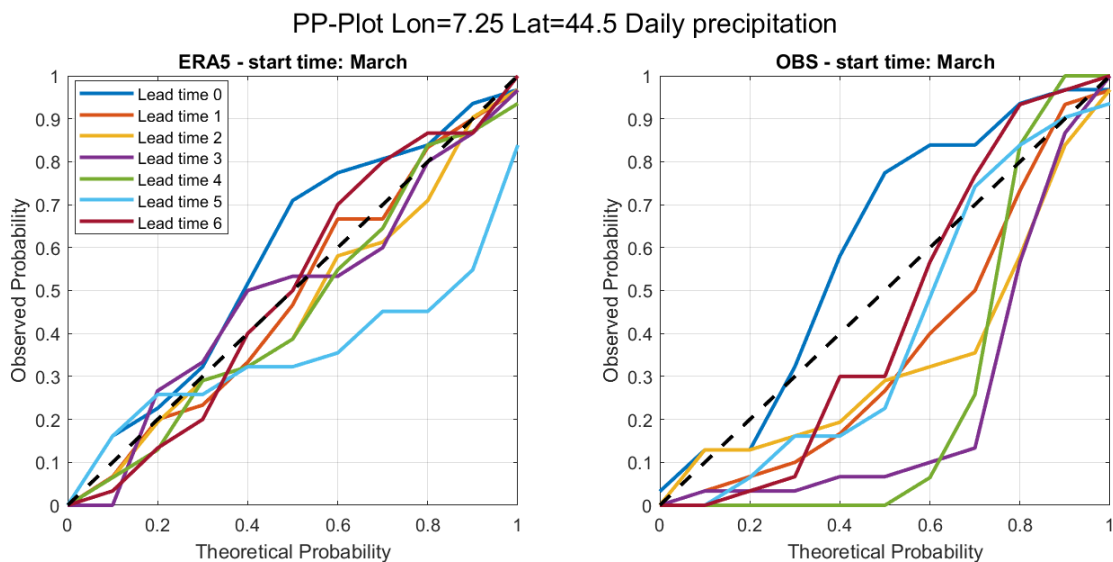

Figure 70. Probability-Probability Plot of ensemble members respect with ERA5 and observation data for one month of initialisation at all possible lead times for daily precipitation at latitude=44.5 and longitude=7.25
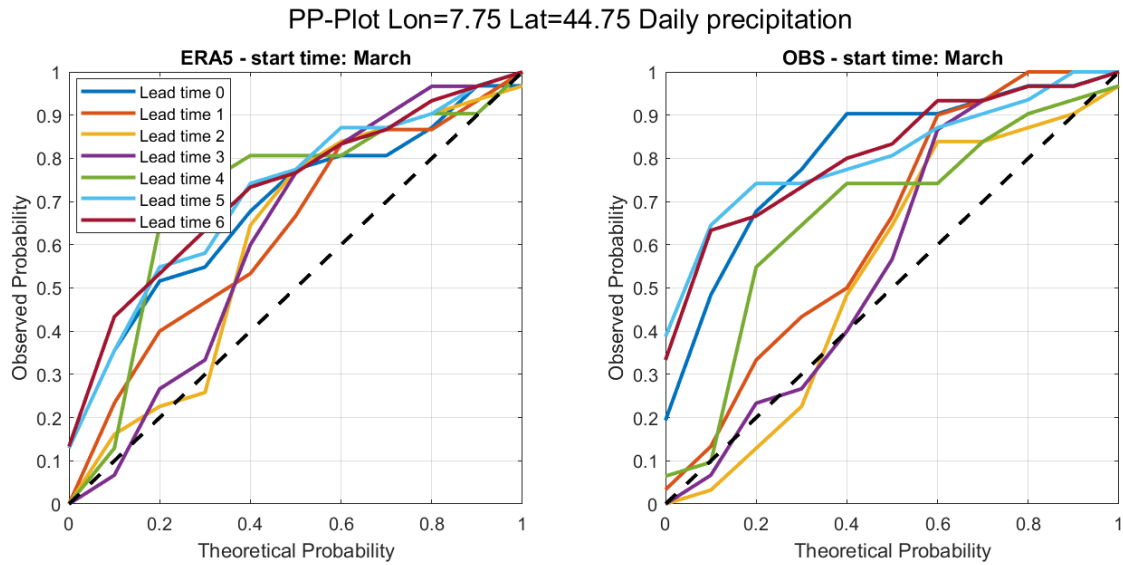
Figure 71. *Probability-Probability Plot of ensemble members respect with ERA5 and observation data for one month of initialisation at all possible lead times for daily precipitation at latitude=44.75 and longitude=7.75*

It is immediately noticeable that the curves for the starting month of March across different lead times in the plain cell suggest a more rapid variation of observed probability compared to theoretical probability in this region. This indicates that the observed probability is systematically higher than the theoretical probability.

Conversely, for the cell at an elevation greater than 400 meters, the ERA5 curves do not show consistent patterns of overestimation or underestimation of probability; rather, this varies with lead time. However, the curves for observational data consistently show observed probabilities that are systematically lower than the theoretical probabilities.

6. CROP MODEL RESULTS

The hydrological model presented in Chapter 3.4 was applied to estimate the observed and forecasted irrigation requirements for maize in the Pianura Padana. As noted in Chapter 2.2, the area of interest for this crop is primarily located at elevations below 400 meters. Therefore, in this phase of the analysis, the grid cells with a resolution of 0.25° x 0.25° within the polygon representing the Pianura Padana will be extracted.

Specifically, an evaluation of the seasonal irrigation demand forecasts for 2018 will be conducted using the ECMWF model, assuming that this year is representative of the previous five years analysed. Additionally, a comparison between the single ECMWF model and the two multi-model ensembles, MME and MMES, will be made for the year 2022. It should be noted that the growing season for maize in the area of interest extends from May 15 to September 31 (139 days), and therefore, the results obtained refer to this period.

To provide an idea of how irrigation demands are related to precipitation, Figure 72 presents the daily precipitation and irrigation requirements derived from ERA5 data.
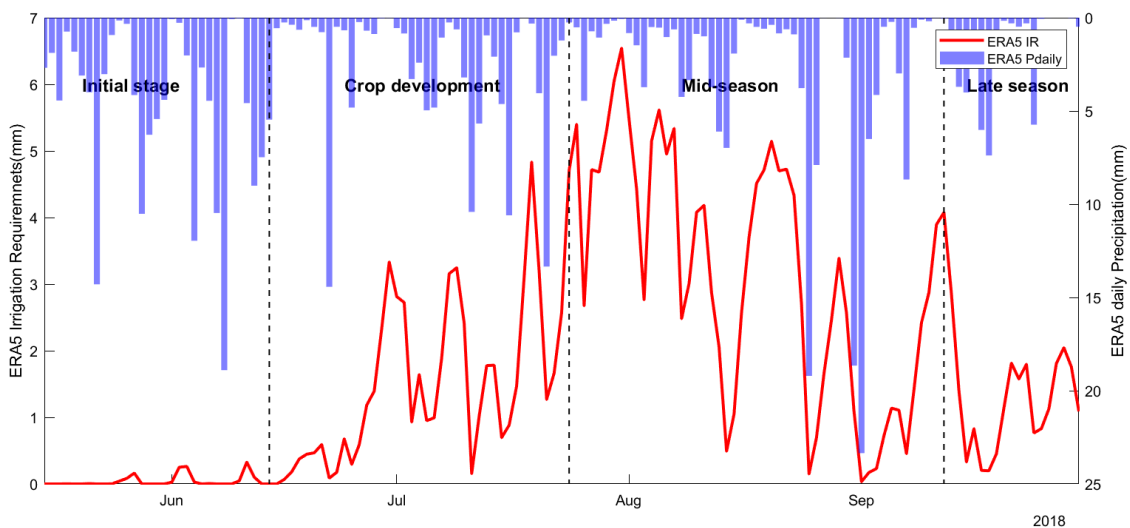
Figure 72. *Relationship between rainfall and irrigation requirements during the growing season*

In Figure 72, a clear inverse relationship between precipitation and the irrigation requirements estimated from ERA5 products emerges. Specifically, it is observed that an increase in daily precipitation is accompanied by a corresponding reduction in irrigation needs, and vice versa. This behaviour is indicative of a reactive water system, where the

availability of atmospheric water (precipitation) directly influences the amount of water required for irrigation.

6.1 ECMWF data vs ERA5 data

6.1.1 Irrigation Requirements

- *Ensemble members average and ERA5 data representation*

The results for the irrigation requirements are presented here for all ensemble members (thin grey lines), along with the respective forecast averages, differentiating between four different initialization months. The results obtained using ERA5 reanalysis data are also shown. It should be noted that, for the latter, a moving average was applied. Specifically, for each day, the average of the preceding ten days was calculated, except for the first few days where the moving average was calculated as the average of the available previous days (<10). This procedure aims to smooth out fluctuations resulting from high short-term variability, making the comparison with seasonal forecasts more meaningful.

All data are presented as the spatial average of the grid cells in the Pianura Padana.
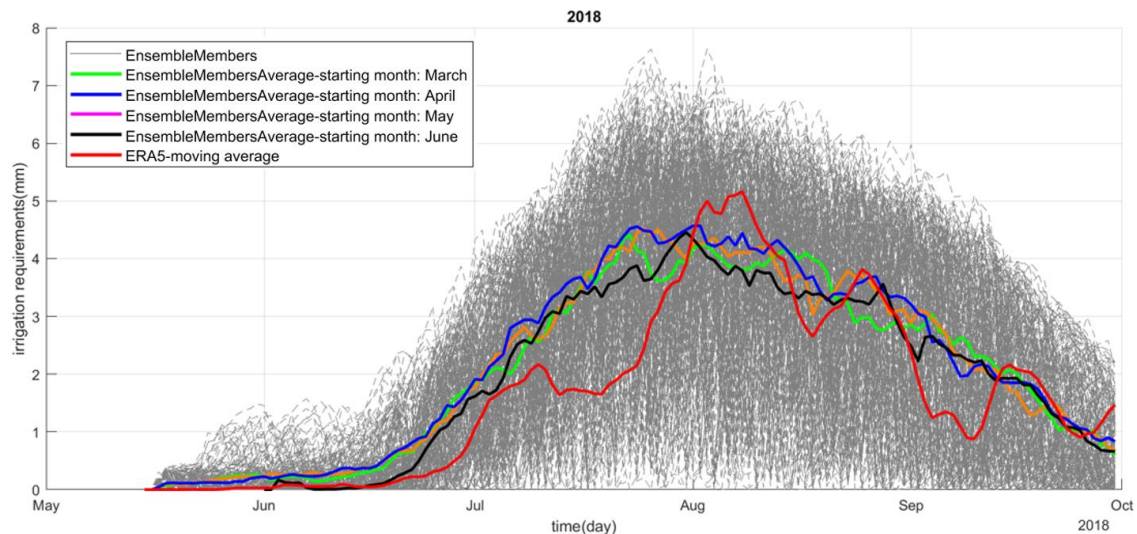
Figure 73. Representation of ensemble member vs ERA5 considering four months of initialisation and five years for Irrigation Requirements

The ECMWF products, while exhibiting a consistent pattern with an increase in irrigation requirements during the summer months of July and August, followed by a decline toward September, show a slight overestimation during the crop development phase (up to July 25). During this early phase, maize plants are still small and have a relatively shallow root

system. Consequently, the water demand from the plants is low, as the leaf area, responsible for transpiration, is still limited. Later, as the plants grow and develop more leaves, transpiration (and thus water demand) increases. The forecasts most closely aligned with reality are those initialized in June (black line), at lead time 0

During the mid-season (July 25 - September 15), the ERA5 line shows a more pronounced and temporary peak compared to the forecasts, which tend to be more smoothed. The ECMWF model outputs, therefore, reveal a steadier and more linear trend in irrigation requirements, irrespective of peaks or fluctuations in precipitation. This lack of responsiveness indicates that the model does not effectively capture the sudden changes in water needs that follow significant precipitation events. However, the forecasts tend to become more accurate as the summer period approaches.

The loss of memory observed in weather forecasts at lead times greater than zero for independent atmospheric variables is also confirmed in the analysis of irrigation requirements. In particular, no single forecast stands out in terms of consistency compared to the others when comparing different initialization months.

- *Error comparison among different starting month and lead times*

After an initial overview of the ensemble members, the focus will now shift to quantifying the bias and the mean absolute error (MAE) for each ensemble at different lead times.

Figure 74. *Bias plot of ensemble members average considering four months of initialisation and five lead times for Irrigation Requirements*
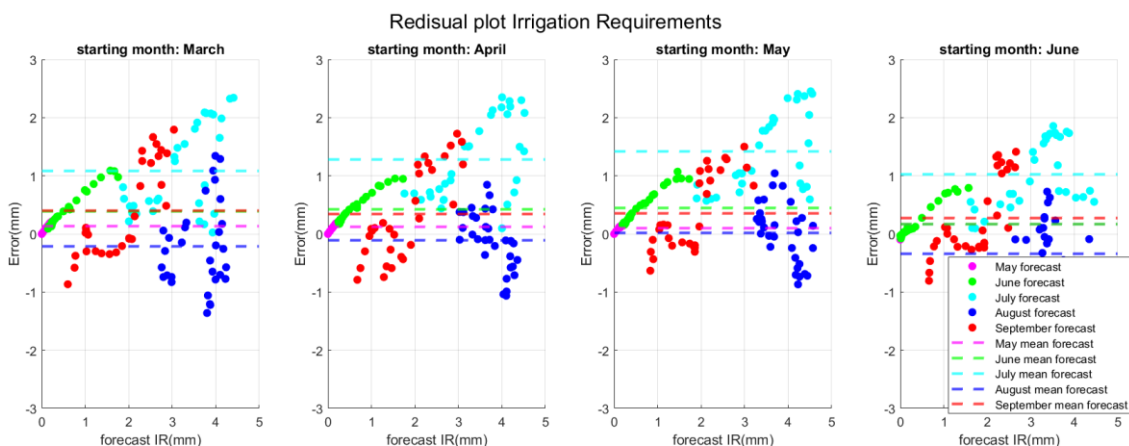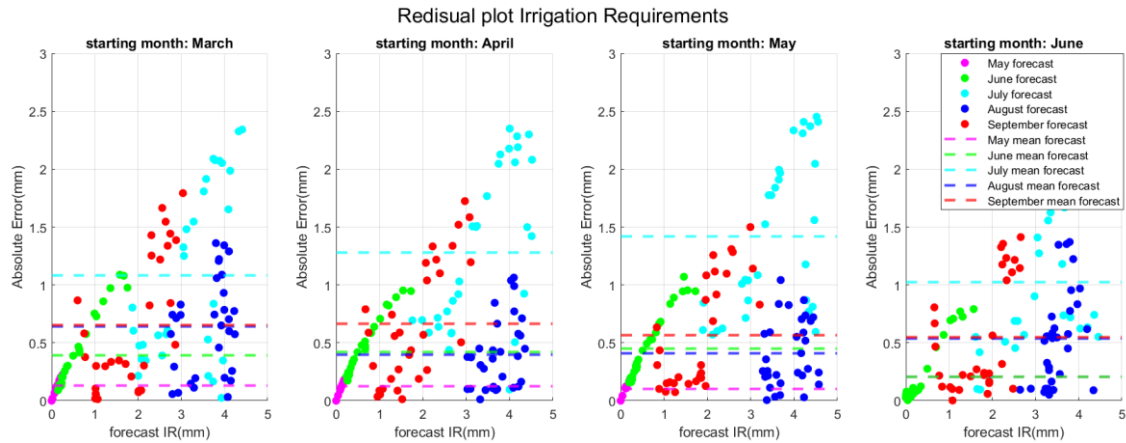
Figure 75. *MAE plot of ensemble members average considering four months of initialisation and five lead times for Irrigation Requirements*

The bias plot highlights a slight tendency for the forecasts to overestimate irrigation requirements from June to August, with a relatively symmetric distribution of points for the months corresponding to the mid-season and errors generally contained within a ±3 mm range, with a monthly average close to 0, except for July. The mean absolute error (MAE) plot reveals that the average monthly error is around 0.5 mm per day. In July, however, the forecasts show an overestimation of approximately 1.5 mm/d, which is relatively high when compared to the average forecasted value. This increase in error during the months of higher demand could be attributed to the model's difficulty in accurately capturing daily variability during periods of both positive and negative peaks in the irrigation season.

Interestingly, the June forecasts with a lead time of 0 benefit from more accurate initial conditions, as June is the only month where the forecasts start within the "growing season." This results in an improvement in forecast quality for that month and lower errors compared to other initialization months.

May also shows relatively low average monthly errors, but this is mainly due to the fact that only 17 days of forecasts are considered. The growing season starts on May 15th, and from that date, the model focuses on a time window where irrigation demand is present but still low, as it corresponds to the early growth phase.

Overall, the forecasts show a generally reasonable level of accuracy, with daily errors that, in some cases, are comparable to the predicted values.

To quantify the daily water flow required for irrigating the agricultural area dedicated to maize cultivation in Pianura Padana within the boundaries of the Po River Basin (previously estimated at approximately 438,138.00 ha), both the forecasted and observed data were multiplied by the area in question, thereby transitioning from mm/d a m^3/d.
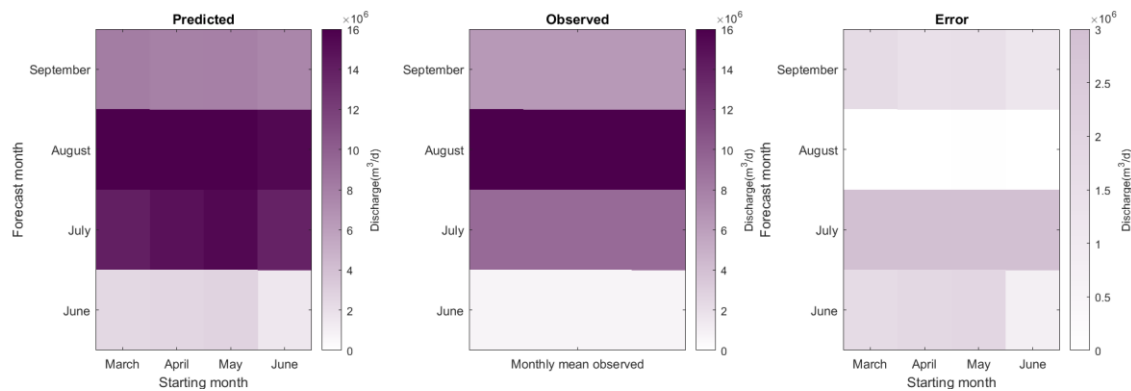

***Figure 76.** Predicted and observed water flow rate, with respective error*

In the Po Basin, the water volumes required for crop irrigation are expressed in millions of cubic meters. It is observed that the error between the forecasted irrigation requirements and the actual measurements can reach significant values, up to approximately 3 million cubic meters per day. This error can represent a considerable discrepancy when comparing the estimated irrigation needs with the actual ones, especially during periods of higher demand, such as the summer months. Such differences can have a significant impact on water resource management, requiring careful planning to avoid shortages or water wastage. In comparison, the average flow rate of the Po River is $1.8 * 10^8 m^3/d$. This analysis highlights how irrigation demand represents a significant portion of the river's available flow.

6.1.2 Irrigation Requirements vs Precipitation and Temperature

In this section, an analysis will be conducted to examine how the errors produced by the forecasts of independent variables, such as temperature and precipitation, are related to the errors arising from the estimation of irrigation demand forecasts.

- *Comparison of rainfall and temperature errors with errors in irrigation requirements*

Firstly, scatter plots were created to depict the errors resulting from precipitation against those arising from irrigation demand on the left, and the errors from average temperatures

against those from irrigation demand on the right. In this case, April was used as the starting month, with forecast months being June, July, and August. The points represented refer to all the days of the forecast month, averaging across the ensemble members. Together with the scatter plot, also the linear fitting line are shown.
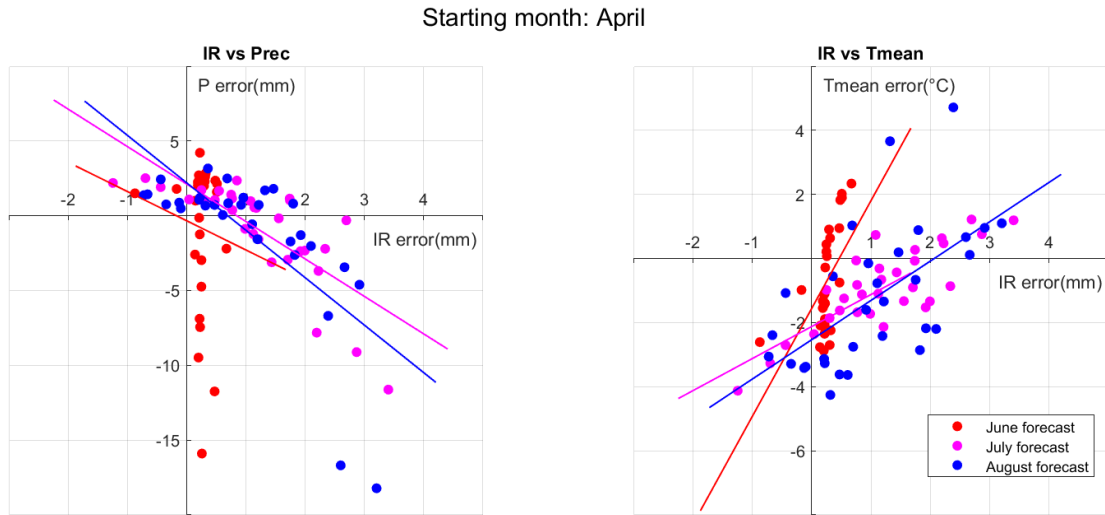
Figure 77. P error vs IR error(left) and Tmean error vs IR error(right)

To measure how well the data fit a linear regression model, the coefficient of determination R^2 was calculated. This provides an indication of the correlation between the errors of precipitation and temperature forecasts and the errors in irrigation demand. The coefficient of determination is calculated as follows:

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{y_i})^2};$$

where RSS is the Residual Sum of Squares, TSS is the Total Sum of Squares, $\widehat{y_i}$ are the errors arising from the estimated data for the irrigation demand, y_i are the errors arising from the precipitation data or temperature and $\overline{y_i}$ is the monthly average value of the y_i data.

A value of R^2 close to 1 indicates a strong correlation between the errors.

Table 7. *Coefficient of determination*

	Starting month: April	
	IR error vs P error	**IR error vs T error**
Forecast month:	R^2	R^2
June	0.01	0.31
July	0.58	0.66
August	0.49	0.36

From this example, it can be seen that on average there is a higher correlation in July, which tends to decrease in August, especially when comparing with temperatures. In the next section, the correlation will be analysed in more detail, taking into account all ensemble members and thus also the variability of the errors.

- *Correlation analysis between the errors relating to irrigation requirements and the independent variables*

In this section, we report the linear correlation coefficient (Pearson's coefficient), which was computed on MATLAB, exploiting the function corrcoef. The equation of the linear correlation coefficient is:

$$\rho(A, B) = \frac{1}{N-1} \sum_{i=1}^{n} \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right);$$

Where N is the total number of data under analysis in one month, μ_A and σ_A are the mean and standard deviation of A and μ_B and σ_B are the mean and standard deviation of B. A and B represent the values of the irrigation requirements and precipitation error in one case and of the irrigation requirements and mean temperature in the other case. Since both variables consist of a set of members, the correlation coefficient was found for each member and then the error correlation boxplot was created to visualise how the correlation between the different members varies.
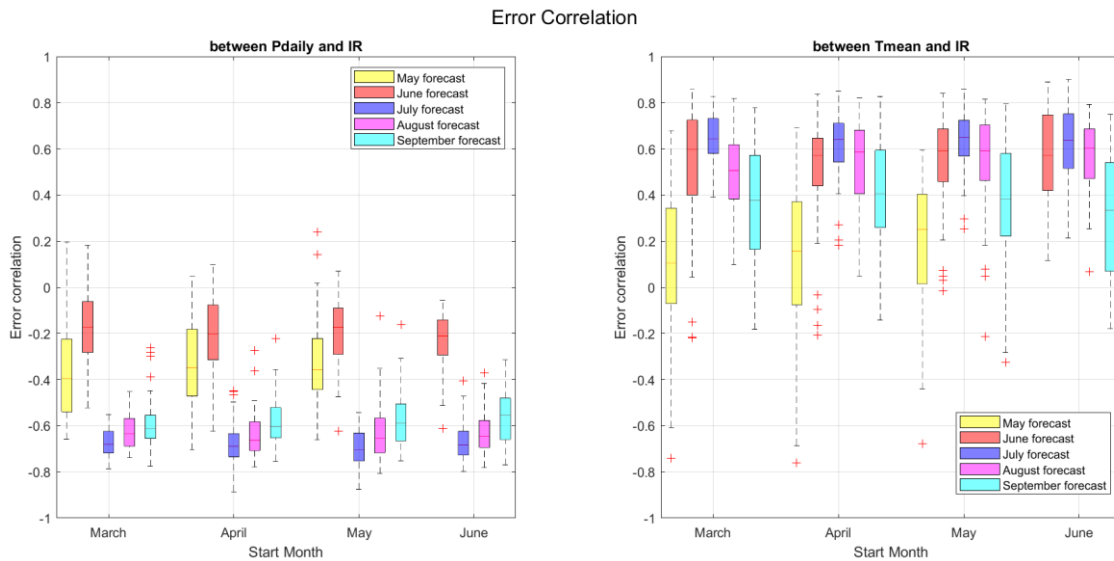
Figure 78. *Boxplot of the error correlation between irrigation requirements and precipitation and temperature*

From Figure 75, it is evident that there is a negative correlation between the errors in irrigation demand and the errors related to precipitation, along with a positive correlation with the errors related to temperature. This is linked to the fact that in the numerical model used for calculating irrigation requirements, these are proportional to temperature (as it directly affects evapotranspiration) and inversely proportional to precipitation. Consequently, a positive error in precipitation (overestimation of rainfall) leads to a lower error in irrigation demand (underestimation of the demand), hence the error in precipitation is negatively correlated with the error in irrigation demand.

In previous sections, it was noted that the lowest errors in irrigation demand were associated with May and June, while Chapter 5 revealed that the highest errors in both precipitation and temperature corresponded to the spring months. This is explained by correlation values with precipitation around -0.4 and correlation values with temperature of approximately 0.2, accompanied by strong variability. This suggests that errors in precipitation forecasts for May and June influence irrigation requirements, but in a less pronounced and more variable manner. This could be related to the fact that, although errors in temperature and precipitation forecasts are higher in spring, the overall impact of these errors on irrigation demand is more subdued because the crops are not yet in their peak growth phase or maximum water demand. Therefore, irrigation requirements are not significantly affected by spring errors.

Conversely, in July, there is a strong correlation and less variability of the coefficient with both precipitation and temperature errors, confirming what was observed through the coefficient of determination. This indicates that there is a strong negative correlation of irrigation demand errors with precipitation errors and a positive correlation with temperature errors. In August, however, the coefficient of determination was relatively low between irrigation demand errors and temperature errors. In this case, only the average of the ensemble members was evaluated; however, as seen in Figure 77, the median of the correlation coefficients remains very high but simultaneously shows significant variability.

- *Propagation among the errors*

In conclusion, the aim is to quantify how error propagation occurs when the independent variables are used to calculate irrigation requirements. To this end, the relative error of the three variables was compared, defined as:

$$RelativeError = \frac{PredictedError - ObservedError}{ObservedError}.$$
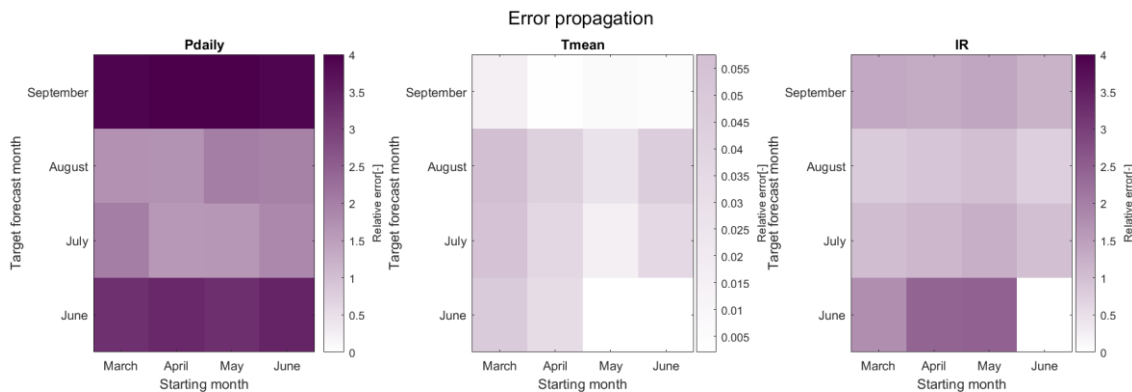
Figure 79. *Error propagation from P and T errors to IRerror*

The image illustrates a representation of the relative error associated with the propagation of errors in daily average forecasts for precipitation (Pdaily), mean temperatures (Tmean), and irrigation requirements (IR) over time, categorized by forecast initialization months and forecast months. It is immediately apparent that the predictions for mean temperature are more accurate than those for precipitation, regardless of the month in which they are made.

From the results obtained, it appears that the error in irrigation demand (IR) forecasts is closely correlated with the errors in daily precipitation forecasts (Pdaily) and, to a lesser extent, with the mean temperature forecasts (Tmean). Specifically, the error in precipitation forecasts has the most significant impact on irrigation requirements (IR). This is evident from the fact that high errors in P for earlier forecasts correspond to significant errors in IR, even if to a lesser degree.

This can be observed, for example, in the forecasts for June and September, which show very high errors for precipitation, while the errors for irrigation demand appear to have been smoothed. This smoothing may be attributed to the correlation with temperature, which, although weaker, still plays a crucial role.

6.2 MME-MMES-ECMWF data vs ERA5 data

6.2.1 Irrigation Requirements

- *Error comparison among different models*

Here, the mean absolute errors resulting from the outputs of the single ECMWF model and the two multi-model ensembles, MME and MMES, are reported, considering one initialization month and five different forecast months.
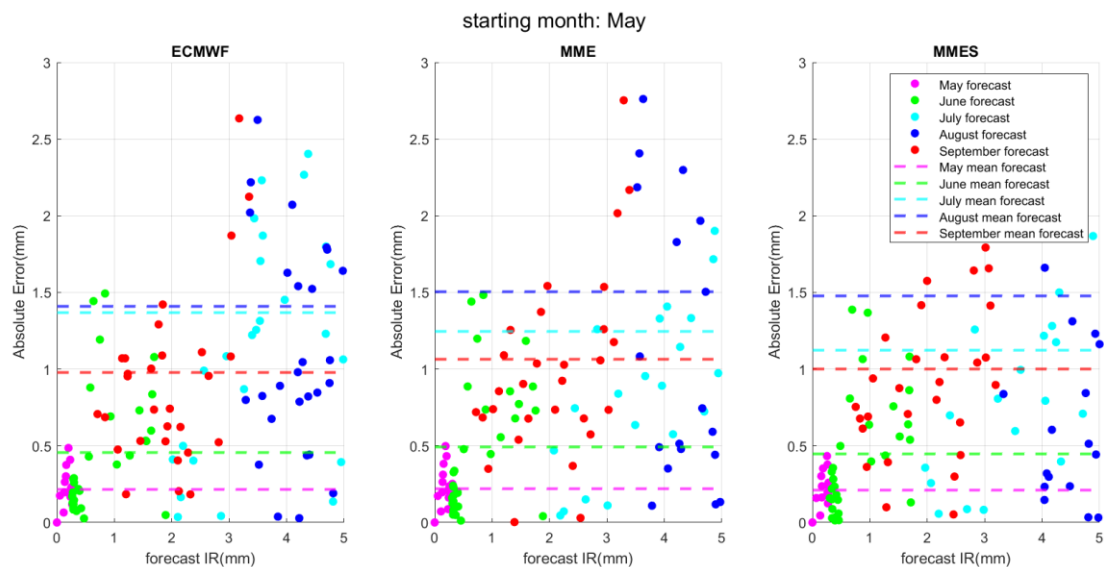

Figure 80. *MAE plot of ECMWF-MME-MMES ensemble members average considering one month of initialisation and five forecast months for Irrigation Requirements*

In Figure 80, it is evident that the three models exhibit very similar and consistent errors. Generally, MME shows slightly higher errors compared to the other two models, except for the month of July, where ECMWF demonstrates a higher error of approximately 0.3 mm.

6.2.2 Irrigation Requirements vs Precipitation and Temperature

- *Correlation analysis between the errors relating to irrigation requirements and the independent variables*
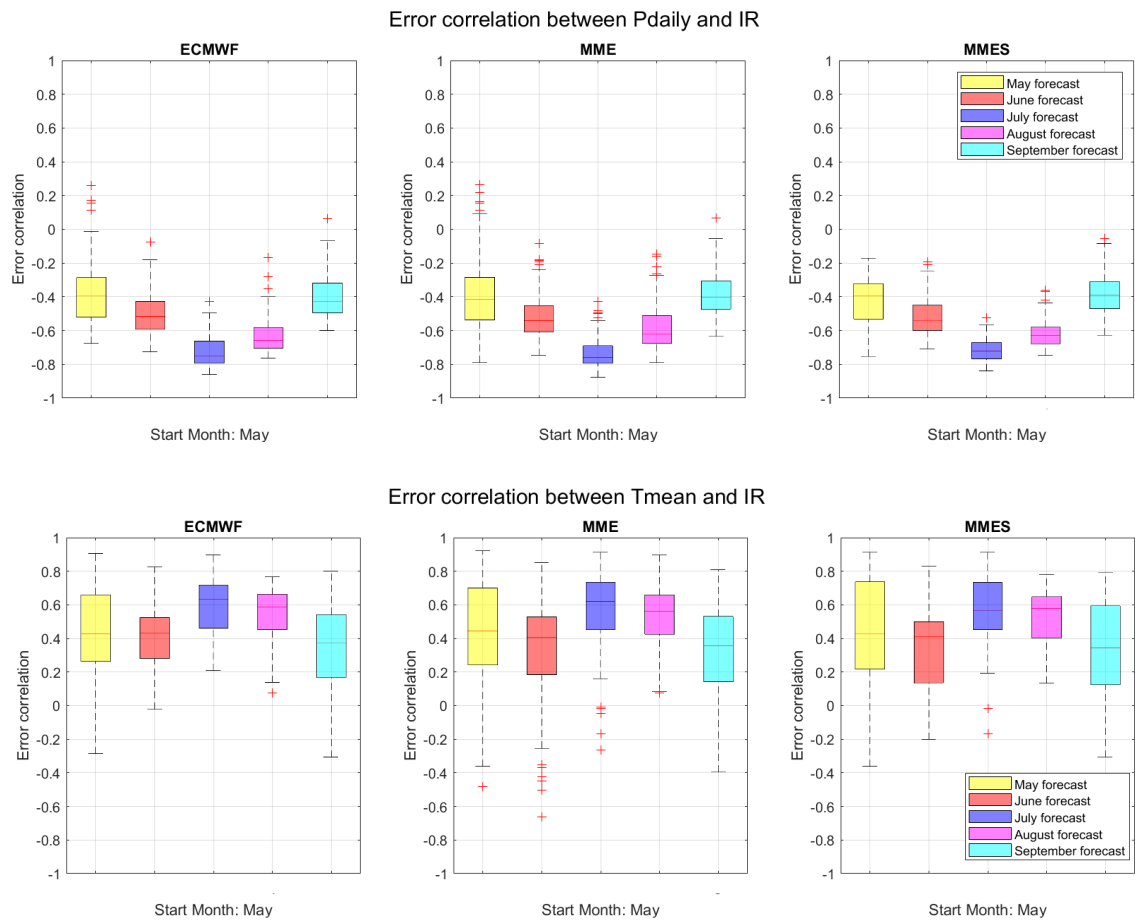
***Figure 81.** Boxplot of the error correlation between irrigation requirements and precipitation and temperature*

From Figure 81, it is observed that the absolute values of the correlation coefficients are similar across the three models, with a slightly higher correlation in the case of the single ECMWF model. Thus, there are no significant improvements when using the two multi-model ensembles; rather, there is greater variability in the correlation coefficients, with

many more outlier points observed in the case of MME. This increased variability is attributed to the larger number of ensemble members considered.

7.CONCLUSION

The study aimed to provide a detailed analysis of seasonal forecasts related to key meteorological variables, such as maximum and minimum temperatures and precipitation, with the goal of estimating maize irrigation needs in Pianura Padana. The assessments involved forecasts provided by the ECMWF SEAS5 model and two multi-model ensembles (MME and MMES) for the period 2018-2022, comparing these data with ERA-5 reanalysis and, in some case studies, with observational data from ARPA Piemonte. This comparison made it possible to verify the accuracy of the forecasts and gain a better understanding of the uncertainty associated with these estimates.

The results showed that seasonal temperature forecasts are generally more reliable than those for precipitation. In particular, maximum temperature forecasts initially proved to be more accurate, with a smaller error compared to minimum temperature forecasts. However, after bias correction based on historical data, minimum temperatures showed a significant improvement, with lower errors than maximum temperatures, reflecting greater consistency due to their lower variability. In contrast, precipitation forecasts presented greater challenges, with a significantly higher average error and less uniform spatial and temporal variability. This unpredictability in precipitation is particularly relevant for irrigation, as it can complicate water resource management during critical periods, such as summer, when water demand is at its highest. However, it was found that plain areas showed smaller errors compared to mountainous areas, with greater reliability in forecasts for these regions.

A key finding from the study is that forecast reliability is not strictly dependent on lead time. While longer lead times result in greater dispersion among ensemble members, the ensemble means tend to stabilize around similar values, regardless of the initialization month. Furthermore, there is not always a proportional correlation between increased variability and higher forecast errors. Rather, seasonality plays a crucial role in the quality of the forecasts, especially for temperature: those for the summer months, such as July and August, tend to be more reliable due to greater atmospheric stability, whereas in the spring months, errors are higher even with reduced lead times.

Another important consideration concerns the spatial resolution of the forecasts. The ECMWF SEAS5 forecasts used have a spatial resolution of 1°x1°, while the ERA-5

reanalysis data have a finer resolution of 0.25°x0.25°. For the entire Po river basin, it was necessary to apply a replication downscaling process to bridge the gap between forecasts and observations. Although downscaling improved the apparent resolution of the forecasts, it was unable to add local climatic details related to topography or microclimatic variations. Low-resolution forecasts tend to "smooth out" local variations, while, for example, local meteorological observations provided by ARPA Piemonte, although more accurate, are limited to specific points with uneven coverage. Therefore, the analysis conducted on specific grid cells in the Cuneo province confirmed that ECMWF data are more accurate when compared to ERA-5 reanalysis rather than ARPA Piemonte's direct observations. To improve the quality of the forecasts, it would be necessary to increase the spatial resolution of the models by adopting dynamic or statistical downscaling techniques, which would allow for better integration of local topographical and climatic features. Additionally, a higher density of weather stations would help bridge the gap between model forecasts and real conditions on a local scale, particularly in areas characterized by high climatic variability.

Another interesting aspect is the comparison between the single forecast model (ECMWF) and the multi-model ensembles (MME and MMES). The multi-model ensembles, which combine forecasts from multiple meteorological centres, generally proved to be more reliable for temperature forecasts, showing less variability and slightly lower forecast errors. However, for precipitation, no significant differences were found between single models and ensembles, highlighting the inherent difficulty in accurately predicting this variable, regardless of the modelling approach used.

The final goal of this thesis was to estimate the irrigation needs of maize, both observed and predicted, based on the analysed variables. The study period extended from May 15 to the end of September, dividing the plant's growth into four main stages: initial stage, crop development, mid-season, and late season. During the summer months (second and third stages), characterized by high temperatures and reduced precipitation, the irrigation demand increases significantly, leading to higher estimation errors. In contrast, in May and June, the water demand of the crop is still very low.

It is interesting to note that the correlation coefficient between the errors of irrigation needs and precipitation, as well as between the errors of irrigation needs and temperatures, approaches zero in May and June, while it is much higher in July and August.

Furthermore, it is highlighted that the strongest correlation is found with precipitation. This is further confirmed by comparing the relative errors of the three variables: precipitation seems to have a greater impact on the estimation errors of irrigation needs, while temperatures tend to reduce and mitigate such errors.

In conclusion, the thesis demonstrates that seasonal forecasts can serve as a useful tool for agriculture, particularly in planning water management in vulnerable areas like the Pianura Padana. However, there are also challenges that require further research, especially regarding the need to improve the predictive capacity of precipitation, a crucial variable for the sustainable management of water resources. Optimizing bias correction techniques, increasing spatial resolution, using multi-model ensembles, and employing probabilistic validation analyses appear to offer promising solutions to enhance the reliability of forecasts, reducing uncertainty and improving adaptive management in the face of climate change.

Moreover, extending the period of analysis would allow for a more representative sample, capable of capturing climatic variations and long-term trends that could influence forecasts. It would also be interesting to expand the study to include other crops and gather feedback from farmers regarding the use of seasonal forecasts and their experiences. This could provide additional insights into water needs and vulnerability to extreme weather events, further improving forecasts and their applications in the agricultural context.

Finally, the model for calculating irrigation needs requires various input information, which is challenging to gather at the local scale. Therefore, it would be advisable to deepen research on the most suitable parameters, for example, by optimizing the crop coefficient based on the specific characteristics of the area, moving beyond the current approach that relies solely on the type of crop and the time of year. Alternatively, using physical sensors to record actual irrigation needs could make the validation of forecasts more accurate.

References

[1]   K. Calvin *et al.*, "IPCC, 2023: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.," Jul. 2023. doi: 10.59327/IPCC/AR6-9789291691647.

[2]   W. Cramer, K. Marini, and J. Guiot, "Current situation and risks for the future First Mediterranean Assessment Report (MAR1) Edited by."

[3]   A. Toreti *et al.*, "Narrowing uncertainties in the effects of elevated CO2 on crops," *Nat Food*, vol. 1, no. 12, pp. 775–782, Dec. 2020, doi: 10.1038/s43016-020-00195-4.

[4]   M. V. , B. V. , M. S. , T. A. , A. M. , B. G. , B. F. , B. M. , C. M. V. , C. G. , E. A. , G. G. , L. T. , M. S. , M. S. Spano D., "Analisi del Rischio I cambiamenti climatici in Italia Executive summary," *Fondazione CMCC - Centro Euro-Mediterraneo sui Cambiamenti Climatici 2020*, 2020, doi: 10.25424/CMCC/ANALISI_DEL_RISCHIO.

[5]   "Piano Nazionale di Adattamento ai Cambiamenti Climatici DICEMBRE 2023."

[6]   M. Reale *et al.*, "Acidification, deoxygenation, and nutrient and biomass declines in a warming Mediterranean Sea," *Biogeosciences*, vol. 19, no. 17, pp. 4035–4065, Sep. 2022, doi: 10.5194/bg-19-4035-2022.

[7]   B. Kurnik, "Climate change adaptation in the agriculture sector in Europe," 2019. doi: 10.2800/537176.

[8]   Autorità di bacino distrettuale del fiume Po, "Siccità Pianura Padana: l'emergenza resta gravissima," 2022. [Online]. Available: https://www.adbpo.it/siccita-pianura-padana-lemergenza-resta-gravissima/

[9]   Wikipedia, "Pianura Padana." [Online]. Available: https://it.wikipedia.org/wiki/Pianura_Padana

[10]   "Italia da scoprire." [Online]. Available: https://www.italiadascoprire.net/geografia/pianura-padana.html

[11]   "Fattori che influenzano il clima italiano." [Online]. Available: https://www.schededigeografia.net/Italia/Clima/fattori_clima_italiano.htm#:~:tex t=Tre%20sono%20i%20principali%20fattori%20che%20influenzano%20il,Medi terraneo%3B%20la%20presenza%20delle%20Alpi%20e%20degli%20Appennini .

[12]   M. C. Peel, B. L. Finlayson, and T. A. McMahon, "Updated world map of the Köppen-Geiger climate classification," *Hydrol Earth Syst Sci*, vol. 11, no. 5, pp. 1633–1644, Oct. 2007, doi: 10.5194/hess-11-1633-2007.

[13]   F. and P. Reichenbach. Guzzetti, "Towards a definition of topographic divisions for Italy.," *Geomorphology 11.1*, 1994.

[14]   M. Rossi, D. Kirschbaum, D. Valigi, A. Mondini, and F. Guzzetti, "Comparison of Satellite Rainfall Estimates and Rain Gauge Measurements in Italy, and Impact on Landslide Modeling," *Climate*, vol. 5, no. 4, p. 90, Dec. 2017, doi: 10.3390/cli5040090.

[15]   Adama, "Tutto ciò che devi sapere sulla Coltivazione del Mais." [Online]. Available: https://www.adama.com/italia/it/approfondimenti-dal-blog/tutto-cio-che-devi-sapere-sulla-coltivazione-del-mais

[16]   W. J. Sacks, D. Deryng, J. A. Foley, and N. Ramankutty, "Crop planting dates: an analysis of global patterns," *Global Ecology and Biogeography*, vol. 19, no. 5, pp. 607–620, Sep. 2010, doi: 10.1111/j.1466-8238.2010.00551.x.

[17]   Istituto Nazionale di Statitistica ISTAT, "COLTIVAZIONI AGRICOL-ANNATA AGRARIA 2019-2020 E PREVISIONI 2020-2021," 2021. [Online]. Available: https://ec.europa.eu/eurostat/web/main/data/database;

[18]   ISTAT, "Censimento Agricolo 2020," 2021. [Online]. Available: https://esploradati.istat.it/databrowser/#/it/censimentoagricoltura

[19]   A. Ceglar and A. Toreti, "Seasonal climate forecast can inform the European agricultural sector well in advance of harvesting," *NPJ Clim Atmos Sci*, vol. 4, no. 1, Dec. 2021, doi: 10.1038/s41612-021-00198-3.

[20]   T. C. Portele, C. Lorenz, B. Dibrani, P. Laux, J. Bliefernicht, and H. Kunstmann, "Seasonal forecasts offer economic benefit for hydrological decision making in semi-arid regions," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-89564-y.

[21]   C. CDS, "Seasonal forecast daily and subdaily data on single levels." [Online]. Available: https://confluence.ecmwf.int/display/CKB/Seasonal+forecasts+and+the+Copernicus+Climate+Change+Service

[22]   Met Office, "Introduction to seasonal forecasting." [Online]. Available: https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/user-guide/background

[23]   Copernicus Climate Change Service (C3S), " Seasonal forecast monthly statistics on single levels," Climate Data Store(CDS). [Online]. Available: https://cds.climate.copernicus.eu/datasets/seasonal-monthly-single-levels?tab=overview

[24]   ECMWF, "Medium-range forecasts." [Online]. Available: https://www.ecmwf.int/en/forecasts

[25]   World Meteological Organization, "Global Producing Centres for Long-Range Forecasts." [Online]. Available: https://community.wmo.int/en/global-producing-centres-long-range-forecasts

[26]   F. Calì Quaglia, S. Terzago, and J. von Hardenberg, "Temperature and precipitation seasonal forecasts over the Mediterranean region: added value compared to simple forecasting methods," *Clim Dyn*, vol. 58, no. 7–8, pp. 2167–2191, Apr. 2022, doi: 10.1007/s00382-021-05895-6.

[27]   ECMWF, "C3S Seasonal Forecasts: dataset documentation." [Online]. Available: https://confluence.ecmwf.int/display/CKB/Seasonal+forecasts+and+the+Copernicus+Climate+Change+Service

[28]   C. D. S. Copernicus Climate Change Service, "Seasonal forecast daily and subdaily data on single levels." [Online]. Available: https://cds.climate.copernicus.eu/datasets/seasonal-original-single-levels?tab=overview

[29]   "SEAS5 user guide," 2021.

[30]   A. Ceglar, A. Toreti, C. Prodhomme, M. Zampieri, M. Turco, and F. J. Doblas-Reyes, "Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast," *Sci Rep*, vol. 8, no. 1, Dec. 2018, doi: 10.1038/s41598-018-19586-6.

[31]   S. J. L. F. M. B. S. B. Tim Stockdale, "ECMWF's new long-range forecasting system SEAS5," 2018. [Online]. Available: https://www.ecmwf.int/en/newsletter/154/meteorology/ecmwfs-new-long-range-forecasting-system-seas5

[32]   C. Shao, X. Yuan, and F. Ma, "Skill decreases in real-time seasonal climate prediction due to decadal variability," *Clim Dyn*, vol. 61, no. 9–10, pp. 4203–4217, Nov. 2023, doi: 10.1007/s00382-023-06800-z.

[33]   D. S. Davolio, "FISICA DELL'ATMOSFERA," Jan. 2012. [Online]. Available: https://www.isac.cnr.it/dinamica/davolio/tmp/Didattica/dispensa_struttura_composizione.pdf

[34]   LEARN STATISTICS EASILY, "Interpretazione degli intervalli di confidenza: una guida completa." [Online]. Available: https://it.statisticseasily.com/interpretazione-degli-intervalli-di-confidenza/

[35]   K. Trenberth, "Changes in precipitation with climate change," *Clim Res*, vol. 47, no. 1, pp. 123–138, Mar. 2011, doi: 10.3354/cr00953.

[36]   S. G. Kwak and J. H. Kim, "Central limit theorem: the cornerstone of modern statistics," *Korean J Anesthesiol*, vol. 70, no. 2, p. 144, 2017, doi: 10.4097/kjae.2017.70.2.144.

[37]   C. Shao, X. Yuan, and F. Ma, "Skill decreases in real-time seasonal climate prediction due to decadal variability," *Clim Dyn*, vol. 61, no. 9–10, pp. 4203–4217, Nov. 2023, doi: 10.1007/s00382-023-06800-z.

[38]   R. G. Allen and L. S. Pereira, "Crop Evapotranspiration (guidelines for computing crop water requirements)."

[39] P. Droogers and R. G. Allen, "Estimating reference evapotranspiration under inaccurate data conditions.," *Irrigation and Drainage Systems*, vol. 16, no. 1, pp. 33–45, 2002, doi: 10.1023/A:1015508322413.

[40] George H. Hargreaves and Zohrab A. Samani, "Reference Crop Evapotranspiration from Temperature," *Appl Eng Agric*, vol. 1, no. 2, pp. 96–99, 1985, doi: 10.13031/2013.26773.

[41] Food and Agriculture Organization, "Harmonized World Soils Database version 2.0." [Online]. Available: https://gaez.fao.org/pages/hwsd

[42] F. Laio and S. Tamea, "Hydrology and Earth System Sciences Verification tools for probabilistic forecasts of continuous hydrological variables," 2007. [Online]. Available: www.hydrol-earth-syst-sci.net/11/1267/2007/