POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Impact of multiple sources of uncertainty on Groundwater Recharge.

TESI DI LAUREA MAGISTRALE IN
INGEGNERIA CIVILE

Author: **Ludovico Guarneri**

Student ID: 10598567
Advisor: Monica Riva
Co-advisor: Andrea Manzoni
Academic Year: 2021-22

Abstract

No matter what steps society takes to mitigate its effects, climate change has already demonstrated its ability to affect the hydrologic cycle. Although science has made significant progress to model and forecast its consequences in the future, its effects on the future climate variables are still uncertain, and a comparison of several climate projections reveals a considerable deal of ambiguity. When performing a simulation on a model of future hydrologic conditions, the uncertainty deriving from climate projections propagates on all the outputs, including groundwater recharge. Uncertainty is likewise anchored in the specification of the hydrologic properties of the soil interested by hydrologic models and affects the models' results. Here, we consider multiple simulations of historical and future climate conditions and analyze how uncertainty in their definition propagates into the quantification of groundwater recharge. As test case for our analysis, we consider the full Lombardy Region (Italy). Four alternative combinations of hydrological inputs for historical climate and two emission scenarios from four general circulation models for future climate are used. At the Regional scale, the results of our simulations indicate only a marginal impact of climate change uncertainty on the annual recharge volumes of the Lombardy aquifer. However, at the sub-Regional scale climate change significantly impacts the groundwater recharge volumes. In particular, in the northern mountain area there is an estimated increases of annual recharge volumes, while the opposite is observed in the southern plain area of the Pianura Padana. Additionally, our analysis suggests that variations in climate change projections have a greater impact on groundwater recharge uncertainty than the impact produced by variations in soil characteristics.

Key-words: Groundwater recharge, Climate change impact, Soil Characterisation, SWB, Uncertainty.

Abstract in italiano

Indipendentemente dalle scelte che verranno adottate per mitigarne gli effetti, il cambiamento climatico ha già dimostrato di essere in grado di influenzare il ciclo idrologico. Sebbene la scienza abbia compiuto grandi progressi nel tentativo di modellarne e prevederne le sue conseguenze future, i suoi effetti sul clima futuro sono ancora incerti, e il confronto tra diverse proiezioni climatiche rivela una notevole dose di incertezza. Eseguendo una simulazione idrologica basata su condizioni climatiche future, l'incertezza derivante dalle proiezioni climatiche si propaga sugli output della simulazione stessa, incluse le stime di ricarica della falda acquifera. L'incertezza è in aggiunta legata alla determinazione e definizione delle proprietà idrologiche del suolo interessate dai modelli idrologici. In questa tesi, si considerano diverse simulazioni di condizioni climatiche storiche e future per analizzare come l'incertezza nella loro definizione si propaghi nella quantificazione della ricarica delle acque sotterranee. Come caso studio per la nostra analisi, consideriamo l'intera regione Lombardia (Italia). Sono stati utilizzati quattro combinazioni alternative di input idrologici per le serie storiche e due scenari di emissione di quattro modelli climatici per le serie future. A scala regionale, i risultati delle nostre simulazioni indicano solo un impatto marginale dell'incertezza dovuta ai cambiamenti climatici sui volumi di ricarica annuale dell'acquifero lombardo. Tuttavia, alla scala sub-regionale i cambiamenti climatici hanno un impatto significativo sui volumi di ricarica delle falde acquifere. In particolare, nell'area montana settentrionale si stima un aumento dei volumi di ricarica annuale, mentre il contrario si osserva nell'area meridionale corrispondente alla Pianura Padana. Inoltre, la nostra analisi suggerisce che le variazioni nelle proiezioni dei cambiamenti climatici hanno un impatto maggiore sull'incertezza della ricarica delle acque sotterranee rispetto all'impatto prodotto dalle variazioni delle caratteristiche del suolo.

Parole chiave: Ricarica della Falda Acquifera, Impatto del Cambiamento Climatico, Caratterizzazione dei Suoli, SWB, Incertezza.

Contents

Introduction

The increase in greenhouse gas concentration in the atmosphere is undoubtedly the cause of several changes in global climate systems. According to data produced by the Intergovernmental Panel on Climate Change (IPCC), global temperatures have increased significantly in recent years. In the last two decades (2000-2020) global surface temperatures were measured to be 0.99 °C higher on average than in the years 1850-1900 [1] and the recent trends of change have continued to increase this alteration. A similar variation is taking place for precipitations, which have likely increased since 1950 and are expected to produce more intense and more frequent extreme rainfall events [1]. Greenhouse gas concentrations have shown to be constantly increasing following a simple power law [2], but the elaboration of future projections of greenhouse gas concentrations and their impact on the Earth's climate has multiple "storylines" that consider every possible scenario of socioeconomic reactions to climate change [1].

With the IPCC Fifth Assessment Report (2014)[3], the scientific community defined a set of four scenarios, called Representative Concentration Pathways (RCP), used to model the potential impacts of different levels of greenhouse gas emissions on future climate change. The four RCPs are identified by the accumulated radiative forcing (the total quantity of energy added to the Earth system from anthropic activities) from 1750 to the year 2100: for example, the scenario RCP4.5 assumes a radiative forcing of approximately $4.5 \, W/m^2$. Each scenario represents a different greenhouse-gas emissions trajectory based on different assumptions about population growth, economic development, and energy use. The four RCPs are:

- **RCP 2.6:** This scenario assumes that global emissions peak around 2020 and then rapidly decline, leading to a global average temperature increase between 0.3 - 1.7°C above pre-industrial levels by the end of the 21st century.

- **RCP 4.5:** This scenario assumes that global emissions continue to increase until around 2040 and then start to decline, leading to a global average temperature increase of $1.1 – 2.6$ °C above pre-industrial levels by the end of the 21st century.

- **RCP 6.0:** This scenario assumes that global emissions continue to increase until around 2080 and then start to decline, leading to a global average temperature increase between $1.4 – 3.1$°C above pre-industrial levels by the end of the 21st century.

- **RCP 8.5:** This scenario assumes that global emissions continue to increase throughout the 21st century, leading to a global average temperature increase between 2.6 – 4.8 °C above pre-industrial levels by the end of the 21st century.

Of these four scenarios available, the RCP4.5 and RCP8.5 have been the ones most used for climate simulations. The RCP4.5 is known as the intermediate scenario because it assumes that some initiatives do take place to reduce greenhouse gas emissions: CO_2 concentrations should drop below current levels by 2070, while by the end of the century, concentrations should stabilize to double pre-industrial levels. On the contrary, RCP8.5 is known as the business-as-usual scenario, because it assumes that no mitigation strategy is adopted, and greenhouse-gas emissions are expected to grow at the current trend. This scenario estimates that by 2100, CO_2 atmosphere concentrations will reach values of 850 ppm, triplicating the preindustrial concentration of 280 ppm [3].

To incorporate RCPs into climate models, scientists use the RCP scenarios as inputs to the models. Climate models use physical-based mathematical equations to simulate how changes in greenhouse gas concentrations will affect the Earth's climate over time, including changes in temperature, precipitation, sea level, and other variables. The type of models available vary depending on the level of complexity of the representation of the physical processes they take into consideration and the way they represent the spatial variability of their estimates [4]. General Circulation Models (GCM) are computer-based models that take into consideration the dynamics of flow systems in the biosphere, hydrosphere, cryosphere, atmosphere, and geosphere to provide estimates of climate parameters such as precipitation, temperatures, or wind velocity [5]. These models can simulate climate over long periods (decades to centuries) both for future and past conditions. GCMs performances are constantly improved by continuously validating the output of the model (temperature, precipitation, etc.) to actual observations of past weather conditions measured from weather stations, satellites and other sources [6].

GCMs are valuable tools to understand the complex interactions that shape the Earth's climate, but they are also subject to some inevitable uncertainty and limitations related to the intrinsic uncertainty of the forecast problem. Nevertheless, the IPCC specified that quantitative estimates of future climate models provide sufficient levels of confidence, especially at continental scales and above [7]. One source of confidence comes from the fact that models are based on accepted physical laws such as the conservation of mass, energy, and momentum. A second source of confidence is to be found in the ability of models to reproduce observed features of current and past climate prediction. For example, global temperature projections of the last two decades have proved to be in overall agreement with subsequent observations over that period. Still, some flaws have been identified in the prediction capacities of these models: Shi et Al. (2018) [8]analyzed the performance of 18 GCMs in an extended area in the Yellow-Huai-Hai region (China) and proved that GCMs simulate with more accuracy

temperatures rather than rainfall and the majority of GCMs tend to underestimate temperature and overestimate precipitation. In addition, even though GCMs provide information on the direction of change for the simulated climate variables, their resolution scales (typically around 200 km) make it unfeasible to use these prediction models directly on a hydrological-hydraulic simulation, usually characterized by a much finer resolution scale. To fix this and provide enough spatial detail for use in regional climate impact assessment, GCMs are typically employed as a scaling factor between historical and future climatic variables to be used in a downscaling process. Several different methods have been developed to perform downscaling processes on GCMs: for example, dynamical downscaling techniques use the output from GCMs as initial and boundary conditions to simulate the regional climate over a smaller spatial scale with the help of a Regional Circulation Model.

Application of GCMs and their downscaled projections show that climate change will alter meteorological conditions such as precipitation and temperatures and consequently affect processes involved in the hydrologic cycle like evapotranspiration patterns, snow cover and melting, and surface runoff. For example, Eckhardt and Ulbrich (2003) [9]evaluated the effects of two climate change scenarios on a central European low mountain range from 2070 to 2099 and predicted that the rise in temperatures will naturally reduce precipitation in the form of snow; this decrease will reduce the spring-snowmelt peak and will increase the flood risk in winter. At the same time, the analysis by Chiew and MacMahon (2002) [10] on the impact of climate change on runoff quantities in several populated regions of Australia proved that climate scenarios predicted by general circulation models have the potential to increase evapotranspiration and generate important runoff modifications of up to ±25% by the year 2030.

As surface water could potentially decrease in the future, it is more and more essential to study how climate changes propagate to the subsurface water bodies and estimate the effects on aquifer recharge. There is not a universally accepted theory on the impact of climate change on groundwater replenishment. The general concern is that unless more water reservoirs are created to counteract the lack of surface water during dry periods, expected higher temperatures will decrease groundwater levels, both directly by increasing evapotranspiration patterns and indirectly by causing higher groundwater withdrawals. Whatever the effect will be, as of now groundwater recharge modeling carries great deals of uncertainty related to the precision of soil characterization, and to the correct representation of recharge processes through mathematical equations.

Uncertainty is a concept deeply rooted in every hydrological model, but especially in subsurface groundwater models where a proper calibration of the model is not always possible for subsurface flow conditions. To understand the origin of the uncertainty, it is essential to comprehend how groundwater recharge quantities can be estimated through different methods and how these methods have been applied to study the

effects of climate change on groundwaters. Depending on the complexity and the inputs used for the estimation, the methods can be split into empirical models, and physical-based models [4].

Empirical methods are the simplest techniques that aim at finding a mathematical correlation, usually a direct linear relation, between aquifer recharge and a set of climatic variables such as precipitation. This relationship tends to be established on scales of annual recharge and it is extremely dependent on the local climate condition and the hydrologic response of the territory, so it cannot be applied outside of the area where it has been generated. For example, Mistry and Suryanarayana (2021)[11] used different empirical equations to estimate groundwater recharge in the North Gujarat region (India) and calculated a percentage of recharge to total precipitation varying between 13% and 20%. Attempting to predict aquifer recharge with this approach will produce results with low accuracy, but it is a simple way to identify the possible range of variation of recharge quantities.

Physical-based recharge models can consider (i) only the water balance equations or (ii) include the momentum balance equation (i.e., Richards equation).

In water balance models, the net infiltration (groundwater recharge) is calculated as the difference between water inputs (ex. precipitation or surface runoff) and water outputs (ex. evapotranspiration and runoff). The term net infiltration or potential recharge is used to describe the amount of water recharging the aquifers, i.e. reaching the zone below the root zone depth[12]. Even though the potential recharge estimated is not necessarily equivalent to the actual recharge happening on the groundwater surface, its quantitative measurements can be considered a starting point to compare the results obtained by applying different conditions in the climatological or soil input data. In conclusion, by applying physical equations to calculate the different components of hydrological processes such as surface runoff, interception, or evapotranspiration, these methods can produce an estimation of net infiltration equivalent to potential recharge. Despite limitations related to the non-accounted surface/groundwater interaction, water balance methods have proven of being able to give reasonable annual or monthly groundwater recharge estimations that can be used to assess and compare the differences in recharge trends over the years[13]. There are many codes and software developed to estimate potential recharge with water balance models: their main differences are the modeling time step (hourly, daily, monthly…), the equations used to quantify each physical process, and the spatial resolution of the estimates. Some models provide a single estimate of average recharge in the entire watershed [14], while others can split up the spatial domain into hydrological response units (HRUs) [15]. Modeling time steps are usually set to daily, monthly, or yearly: shorter time steps of simulation are preferred to better interpret the effect of intense precipitation, but the analysis of groundwater recharge volumes is more significant at monthly, seasonal or yearly time scales.

The natural evolution of water balance models is found in groundwater numerical models. These models simulate and predict the behavior of groundwater coupling the study of water flow both on the surface, through rivers, lakes, and other surface water bodies, and on the subsurface, through aquifers and other underground water sources. This type of modeling applies Darcy's law to create a mathematical representation of the groundwater system that simulates the flow of water through the unsaturated zone and estimates the aquifer replenishment rate. The modeling process involves several initial steps including conceptualization of the system, building the numerical model, and calibrating the model to fit observed data. After these operations have been performed, the model can be used to make predictions on the behavior of the groundwater. Groundwater numerical modeling is a powerful tool to simulate and predict the behavior of water in the subsurface and provides a much more accurate estimate of groundwater recharge. However, even though calibration of the model takes place, uncertainty remains, deriving from input data definition and the abstract mathematical representation of the model dynamics.

Creating a numerical model can be a challenging and complex process due to data availability, model complexity, model calibration and validation, and computational requirements. In light of this, most of the studies analyzing the effects of climate change on groundwater concentrate on using water balance models thanks to their low cost, their simplicity of application, and the possibility of being used for large dimensional and temporal scales. The precision of water balance models has been discussed in several studies, as no information is available on the accuracy of their estimates of potential recharge. The general consensus is that even though these models produce an estimate of potential recharge that might differ from actual groundwater recharge, their feedbacks are sufficient as a starting point to compare the results of different models stressed under different condition (ex. different climatological data) [16].

Several studies have analyzed and demonstrated how climate change will most likely alter recharge dynamics and produce long-term effects on recharge rates, but consensus on the direction of change is not necessarily the same in all aquifers. Simulated recharge trends are variable spatially, depending on the type of soil, vegetation, or land use of an area, as well as temporarily, depending on GCMs and the choice of emission scenario. In many areas of the world, it is unclear whether recharge will increase or decrease under climate change conditions.

Atawneh et Al. (2021) [17] recently analyzed seventy-three papers of regional studies that quantified long-term projected changes in groundwater recharge conditions. Depending on the emission scenario and the season analyzed, the research displayed contrasting projections for Oceania, North America, Europe, and Asia but confirmed a general decrease in the estimation of future groundwater recharge rates for North Africa, Latin America, and Southern Europe. Southern Europe is expected to observe increasing winter recharge rates but decreasing annual and summer recharge rates.

The paper also pays close attention to the identification of the sources of uncertainty considered in these studies: only 20% of the studies quantified uncertainty related to the different responses of GCMs, only 10% of studies quantified uncertainty associated with the choice of greenhouse-gas emission scenarios and none of the studies quantified the internal uncertainty associated with the model physics and processes.

Despite the lack of knowledge on the influence of the sources of uncertainty that affect projections of future recharge conditions, several studies in the field assessed the difficulty in evaluating a one-directional effect of climate change in a specific geographic area. For example, Crosbie et Al. (2013) [18] developed an innovative strategy to interpret and summarize the wide range of results produced when using different GCMs and different emissions scenarios. An ensemble of three emission scenarios for 16 different GCMs, for a total of 48 climate datasets, is employed to run water balance simulations of the entire Australian continent under a 2050 climate. The range of variation of the results is quite large and for some areas, it entails both increases and decreases in recharge rates depending on the GCM. To fix this discordance, a risk analysis framework is applied by fitting the results of the simulations to a probability distribution: this facilitates the communication of the likelihood of change in recharge for a specific era and in a specific location. This statistical analysis also evidenced a greater instability of recharge predictions when increasing the emissions scenarios.

Meixner et Al. (2016) [19] also quantified uncertainty in the estimation of recharge across eight representative aquifers located across the Western United States. Uncertainty is estimated as the range of variation of projected change in the estimated total recharge: analysis shows that two northern aquifers have high uncertainty in the direction of change of future recharge. For example, in the Williston (WA) basin aquifer system, a future increase in precipitation suggests in some models an increase in recharge, but the general low permeability of the soils results in an extreme growth of evapotranspiration quantities in other models; these two competing factors produce inconsistent outlooks of the different model projections analyzed, thus high uncertainty in future diffuse recharge subsists in the area. One other innovative aspect of the research by Meixner is the partition of recharge budget components into four mechanisms: diffused recharge coming from infiltration on the land surface, focused recharge produced by superficial water bodies, irrigation, and mountain-systems recharge. The available data indicates average reductions of future potential recharge between 10% to 20% for southern aquifers, and lower increases of recharge for northern aquifers (between 5% to 10%). However, the direction of change for the same aquifer is not unique over all recharge mechanisms: for example, in the southern San Pedro valley, under future climate conditions, focused recharge is expected to increase due to forecasted high precipitation intensities, but mountain system recharge is expected to decrease due to decreasing snowpack volumes. Still, recharge in the San Pedro valley is expected to drop by 30% as mountain system recharge takes up a great

fraction of the total recharge in the area. This kind of observation is recurring in many studies. Smerdon (2017) [20] reviewed six articles published between 2011 and 2016 on the interaction between groundwater and climate change: his findings confirmed that the most sensitive regions to recharge will be mountain areas, where shifts in timing and duration of seasonal weather will have a greater impact on recharge habits.

Apaza-Coria et al. (2022) [21]also found results that highlight the weight of mountain recharge on the evaluation of total recharge in the region of the Valle Alto Basin in Bolivia, characterized by a semi-arid climate with hot and dry summers, and mild winters with occasional rainfall. Performing a soil water balance model based on monthly balance, the model estimated an annual average recharge rate of 65.68 mm/year for the 2006-2020 period, with peaks of 250 mm/year in the mountain basins. It is important to observe that higher recharge values in mountain areas are directly related to higher annual precipitation, but in addition to this, the percentages of recharge to precipitation were higher in these areas than anywhere else in the region (~16%) probably due to topographical and climatical heterogeneity that facilitated recharge processes.

Other studies overlook uncertainty or spatial distribution of estimates to focus their attention on identifying a potential trend in the development of projected recharge on longer time frames. Hughes et al. (2021) [22] used a national-scale recharge model for the British mainland to investigate the impacts of future climate on groundwater resources. Potential groundwater recharge calculations are performed using a water balance model (ZOODRM) developed by the British Geological Survey (BGS) with a coarse grid of 2 x 2 kilometers cells that estimates recharge on a daily time step from 1950 to 2099. The results are presented on graphs of simulated potential recharge for every month of the years from 1950 to 2099. Recharge estimates demonstrate good stationarity for every month between 1950 and 2009, but from 2010 to 2099 trends of increased or decreased potential recharge are observed depending on the month analyzed: winter months show increased recharge rates while summer months show the opposite trend. This behavior is directly correlated to the trends observed in monthly precipitation in the area. In addition, a reduction in potential recharge at the historical start of the recharge season in September suggests that the future recharge period might shorten, even though the overall volume of yearly potential recharge will be equal or even higher (depending on the GCM) with respect to current conditions. A similar result is produced by Raposo et al. (2013) [23]when using the Soil Water Assessment Tool (SWAT) model to simulate the impact of 56 different climate models in the district of Galicia- Costa in Spain from 2071 to 2100. All models confirmed that the principal and most certain impact of climate change would be a modification of the temporal patterns of recharge, with a concentration in the winter season and a drastic decrease in the spring and autumn seasons: the length of the dry season would increase around 30%.

Shifting the focus to the Italian peninsula, despite the current problems of drought that in the past years affected particularly northern Italy and the Pianura Padana, fewer studies have been performed to assess the potential effects of climate change on groundwater storage. The study from Behulu (2013) [24] investigates the effects of six climate change scenarios on the hydrological behavior of the Upper Tiber River Basin, in central Italy. SWAT is used to perform a physically based watershed simulation for a future time frame from 2071 to 2100. The results of the model suggest that there will be a modest reduction of mean annual precipitation and a significant reduction in mean surface runoff: this suggests a possible shift from surface water utilization to groundwater extraction. While the lower emissions scenarios show more consensus on the trend of reduction of mean groundwater recharge, high emissions scenarios disagree on the direction of change with some scenarios even predicting an increase of up to 30% with respect to the current condition. The study from Braca et al. (2019) [25], instead, provides a rough estimation of the impact of climate change on the availability of groundwater recharge in the entire Italian peninsula. Climate data from 1996 to 2015 have been projected every year on a regular 10km grid considering four climate change emission scenarios and three different periods of short/medium/long term projections. Groundwater balance estimation has been carried out yearly using a GIS-based hydrological water budget and scaling historical estimates of recharge with respect to the GCMs. The research shows that for RCP2.6 low emission scenario the variation of groundwater recharge rates in the whole Italian territory is practically null, as the percentage of reduction of groundwater recharge for 2090 to the period from 1996 to 2015 is only -7%. Instead, the same quantity for the RCP8.5 high emission scenario is -42%. Considering the landscape heterogeneity of the Italian territory, accurate research on this matter should be performed locally to predict the future of Italian aquifers.

Even though the acknowledgment of the presence of uncertainty in climate projection has positively encouraged researchers to understand its impact on recharge prediction, a lot has still to be done on this matter. All groundwater models are simplified representations of a real system that can't be perfectly reproduced, especially if many of the properties that characterize it are defined with a certain degree of uncertainty. Uncertainty is deeply rooted in all aspects of a hydrologic model, from the possibility of correctly representing the real system characteristics in all its details, to the possibility of correctly reproducing the climate condition that influences the model evaluations. The accuracy of these models is naturally influenced by factors that include the quality and availability of input data.

Meteorological input data provide information on the amount and distribution of precipitation and temperature patterns in the study area, which are key factors of a model that determine water balance and consequently the infiltration process. The quality and accuracy of meteorological input data is directly related to the quality and accuracy of the predictions of groundwater recharge made by a model. For this reason, it is always essential when performing a simulation of groundwater recharge to gather

the highest-quality and up-to-date meteorological data available for that scope. The quality of meteorological data does not depend only on their realistic representation of weather patterns but also on their spatial resolution. Climate datasets tend to have a low resolution for several reasons: the first is that collecting and then storing high-resolution data with a high temporal frequency can be technically challenging and expensive. Then, for certain applications such as long-term climate models, lower resolution data may be sufficient for large-scale modeling studies that do not require fine inputs or that are used to compare different models.

A similar observation can be done for soil input data, which are one of the most important factors affecting the accuracy of groundwater recharge models. They provide information about the physical and hydrological characteristics of an area and have a significant impact on the accuracy of the recharge model predictions: if soil estimates are inaccurate, it can result in incorrect predictions of hydrologic processes and lead to over or under-estimation of recharge. For example, if soil input data underestimate the infiltration capacity of the soil, the model will produce lower infiltration and greater runoff with respect to what could potentially be measured in the real field. It is therefore important to use high-quality soil input data that incorporate information on a large spectrum of parameters such as soil texture, structure, permeability, and other hydrological properties that influence the model output.

Accurate soil input and meteorological input data are crucial for building a reliable and accurate groundwater recharge model, as they both play a critical role in determining the accuracy of the model's predictions, but which of these data types has more influence on the final output of a model? This question becomes relevant when the time frame of application of a recharge model is not the present or a past time but a future condition. Climate change seems to have the effect of altering precipitation and evapotranspiration patterns as well as influencing land use changes: recent events of the past years provided evidence of how water resources are vulnerable and how their scarcity has strong consequences on the surrounding ecosystems. Even though the most visible impacts of climate change take place on surface waters, there is no doubt that climate change will also affect groundwater quantity and quality; as water supplies will decrease and the population increase, groundwater resources are likely to become essential to compensate effects of climate change. Therefore, developing good models that reproduce the mechanisms behind groundwater recharge has become essential to study future conditions and provide a tool to make predictions about water sources, but it is also essential to understand the uncertainty behind these models.

When a hydrological model is performed on a historical set of data, meteorological conditions can be supplied to the model employing reliable datasets of atmospheric reanalysis of the historic climate, but the simulation could still produce uncertainties in the output related to the soil input data. On the opposite, when studying the future

effects of climate change, simulations of recharge are conducted on a future series of climate variables, for example by using one of the downscaled climate scenarios derived from the General Circulation Models, which naturally carry a certain level of uncertainty. Under these considerations, this thesis aims to analyse the effects that a variety of soil and weather input data has on the output of net infiltration in a hydrologic model simulation performed with a water balance method. The ultimate question is: when trying to forecast groundwater recharge trends on future climate conditions, is it more meaningful to operate on increasing the precision of soil input data or climate input data? For example, if the results show that a variation of soil input data generates more difference in the average net infiltration than a variation in the future weather variables does, then it would mean that to improve the precision of net-infiltration estimates, it is more urgent to work on improving the precision of soil characterization. On the contrary, if the answer of the research shows that average net infiltration values have substantial quantitative differences when keeping soil input data constant and comparing historical and future climate conditions, then it would mean that it is more urgent to ensure the highest forecast accuracy of the downscaled Global Circulation Models.

The Soil Water Balance Model (SWB) is here employed to estimate spatial and temporal patterns of potential recharge from 2000 to 2019 and from 2080 to 2099 on an area covering the Lombardy region. The area of simulation is rectangular in an Eckert IV projection system (coordinates: left=699000 m, bottom=5520000 m, right=943000 m, top=5730000 m). Variations in the soil input data are produced by applying different empirical models to define these quantities, while variations on the meteorological data are provided through two emissions scenarios from four different GCMs. Calibration of the model is not performed because the SWB model is not designed for it: groundwater models are usually calibrated by comparison of simulation outputs with measured hydrological variables such as streamflow. However, even though SWB considers the influence of runoff processes in the water balance, it doesn't make a complete computation of streamflow. To fill this gap, a study by Hart et al. (2012) [26] simulating recharge conditions in Dane County (Wisconsin) compared SWB results to potential recharge calculated with the PRMS hydrologic modeling system based on water and energy balance: the two models agreed for the estimates of potential recharge. SWB has also been chosen because previous work on groundwater modeling with SWB1.0 had been performed by the PhD candidate Andrea Manzoni who also made available through a Github repository all the necessary elements to create the inputs and run a SWB simulation [27].

1 Description of the study area

1.1. Location

Lombardy is a region located in northern Italy between Veneto and Piemonte. Its borders are irregular as the area is enclosed by the Swiss alps to the north, the Po River, and the Apennines to the south. It is an area of approximately $24,000 \ km^2$. In terms of extension, the region is split almost equally between mountain areas (the alps) and plain areas (Pianura Padana), with altitudes ranging from a maximum of around four thousand meters in the alps to just a few meters in the eastern Pianura Padana. The same cannot be said for the distribution of population: the lowland area, with the presence of big cities such as Milan or Bergamo, shows high values of urbanization and industrialization that are hardly matched anywhere else in Italy, but the mountain areas are poorly urbanized. Along with urbanization, the presence of the Po' river which crosses the Pianura Padana from east to west makes this region also renowned for its agricultural and pastoral economy.

Figure 1.1: Area of study

The region is characterized by a variety of topographies, making it hard to identify a prevailing landscape.

Alps

A great portion of land in Lombardy (around 40%) is occupied by mountain areas. The northern area exhibits some of the highest peaks of the alps such as the Cevedale (3779 masl), the Adamello (3539 masl), and the Bernina (4049 masl) and hosts most of the glaciers present in Italy [28]. Because of the glaciers and other hydrological features such as artificial dams, these mountains are a precious resource in the management of the region contributing both with spring snowmelt and water reservoirs. The Southern Alps show lower elevations as they transition to hill areas and then to the Pianura Padana. Overall, the alps have an essential role in the environment and consequently in the economy of this region, as they influence agriculture and energy production.

Pianura Padana

The Lombardy lowland occupies part of the Pianura Padana, which consists of the hydrographic basin of the Po' river and finds its origin in alluvial deposits. The Po' river is one of the rivers that in the Pleistocene generated this lowland by dragging downstream great quantities of sediments. The Pianura Padana is divided into an upper and lower part that are differentiated for their soil characteristics, their hydrology, and their vegetation. The upper part, closer to the mountain range, has soils made mainly of sand and gravel which result in a low propensity to moisture-holding behaviors. The lower part, in the south and closer to the Po River, is mainly composed of finer materials such as clay and characterized by very low permeability [29].

Stream network

Being Lombardy a humid region with a heterogeneous morphology, superficial runoff is an important element of the regional water balance. The Po' river flows from east to west on the southern border of the region and it functions as a collector of the water that accumulates in the alps and in the Pianura Padana. As can be seen in Figure 1.2, the stream network of the region is very dense and spreads homogeneously over its territories. Of all the rivers, four of them are more important than the others in terms of their discharge: the Ticino River, the Adda river, the Oglio river, and the Mincio river. These rivers connect the mountain region to the Po' and therefore develop more or less in a north-south direction.
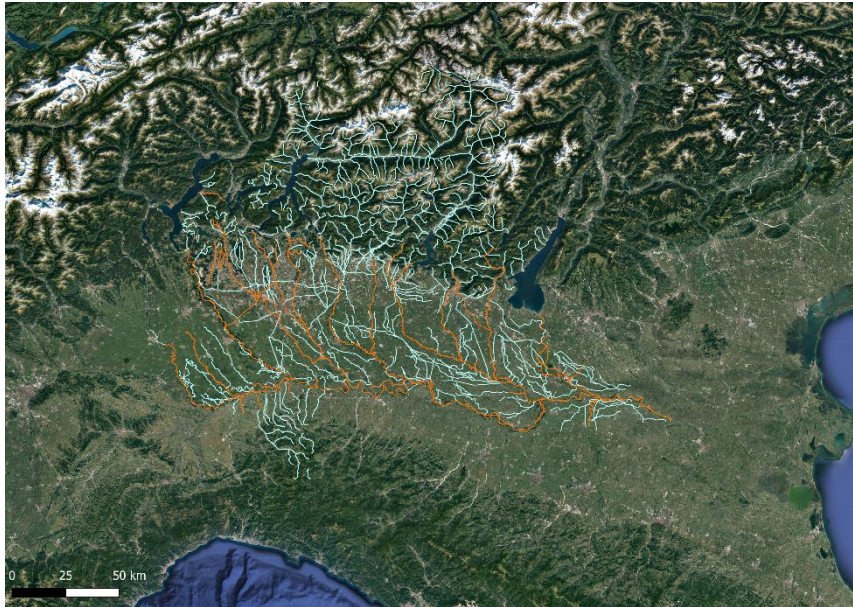
Figure 1.2: Lombardy stream network

Lake system

Lakes are widely diffused in this region, especially in the mountainous and hilly areas. Most of the lakes have a glacial origin, but some of them are artificial. For what concerns dimension, lakes range from small lakes in alpine settings to big lakes in the Prealps. All the four rivers mentioned in the previous paragraph have in common the characteristic of being effluent of one of the four big lakes of Lombardy. The Ticino River is the effluent of the Lago Maggiore, the Adda river is the effluent of the Lario lake, the Oglio is the effluent of the Lago di Iseo and the Mincio is the effluent of the Garda lake. All these lakes are characterized by high water storage that influences and mitigates the surrounding climate of that area, but more importantly, acts as an important water reservoir for the downstream areas. With these great volumes of water, it is obvious that water reservoirs in Lombardy play an important role in guaranteeing both a source of water resiliency during times of droughts and a storage element to collect water during times of high rainfall to reduce flood risk.
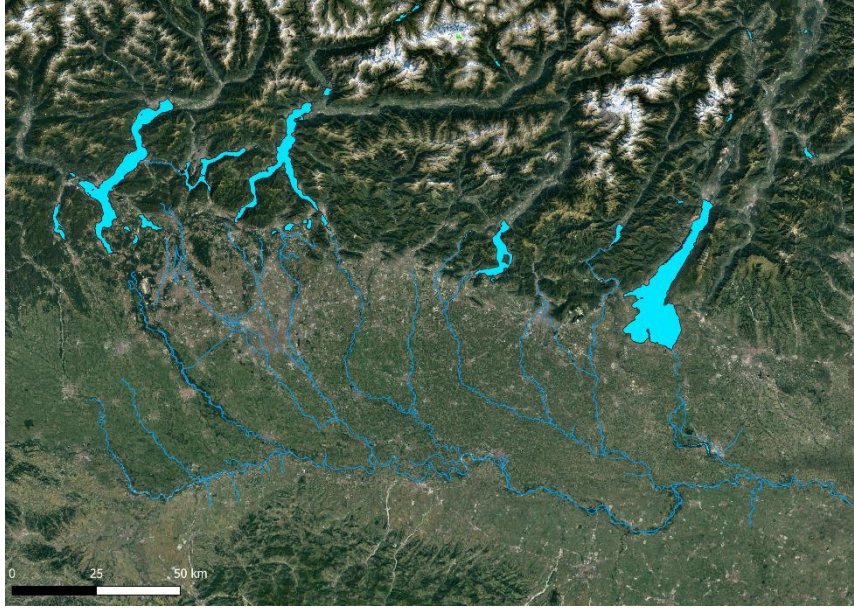
Figure 1.3: Lakes and rivers in Lombardy

1.2. Climate

The region has a temperate humid climate. Its features are heavily influenced by the location of the region inside the Pianura Padana and the enclosure in three directions by the Alpine and Apennines ranges. The mountain ranges play a crucial role in influencing the meteorological conditions of the region as they act as a barrier, diverting moist air masses from the Mediterranean sea to rise and cool as they move over the mountains. Additionally, in winter, cold air slides down from the mountains to the Pianura Padana and cause frost and fog, while in the summer, the valleys can experience intense heat waves [28].

To get a first glimpse of th climate in this area the ERA5 dataset is used. More on this dataset is specified in chapter 2. For now, it can be said that being a reanalysis dataset, it combines observations from various sources with a prediction model to produce a complete and accurate record of the state of the atmosphere. Average monthly precipitation and temperatures are extracted from the dataset over a timeframe ranging from 2002 to 2019 and presented in Figure 1.4.
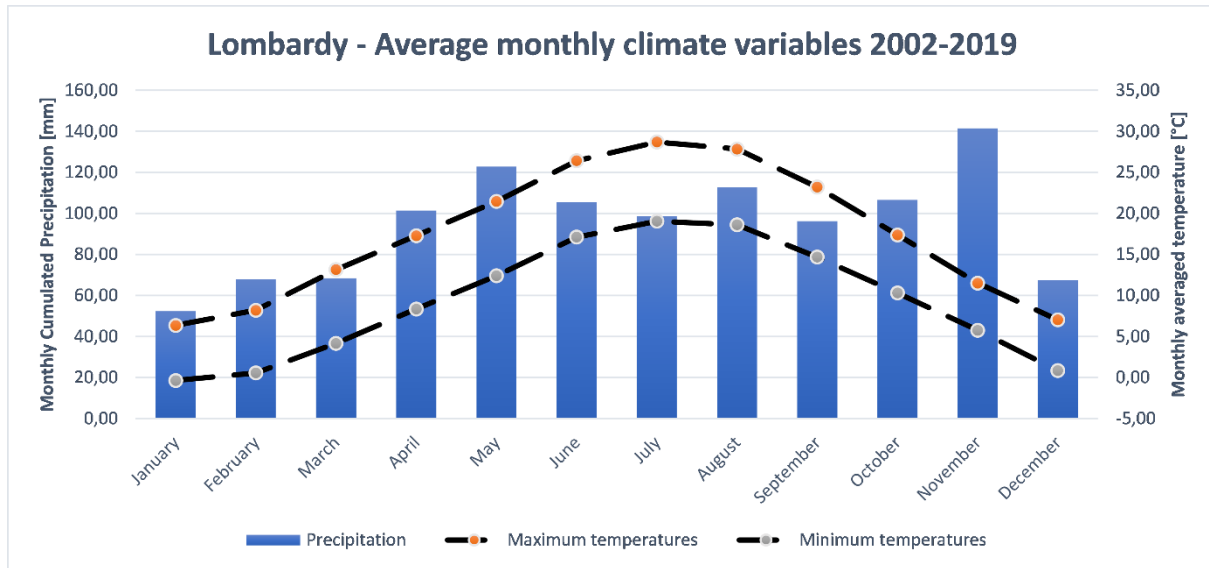
Figure 1.4: Monthly precipitation. Average cumulated annual precipitation over the region is measured to be 1140 mm.

Figure 1.4 shows the monthly distribution of precipitation and temperatures over the area. Precipitation intensities tend to be lower during the winter and summer seasons, and higher during the spring and autumn season. In some cases, the difference is so marked that the average precipitation in a dry month is less than half of the precipitation encountered in a wet month. This is the case of months such as January ($Total\ precipitation = 59\ mm$) and May ($Total\ precipitation = 122\ mm$) but also months that are closer such as November ($Total\ precipitation = 141\ mm$) and December ($Total\ precipitation = 67\ mm$).

Temperatures follow smoother trends with simply the hottest month being in July ($maximum\ temperature = 28.7\ °C, minimum\ temperature = 19.0\ °C$) and the coldest month being in January ($maximum\ temperature = 6.3\ °C, minimum\ temperature = -0.4\ °C$).

The spatial average values calculated are only indicative: one element that should not be ignored when analyzing these data is the spatial variability of the meteorological conditions depending on the landscape that constitutes the Region. The flat areas of the Pianura Padana, for example, tend to show higher temperatures and lower precipitation compared to the mountain areas. In addition, the presence of large lake basins, such as lake Como or lake Garda, contributes by mitigating the climate in the surrounding areas and creating a milder Mediterranean microclimate. The distribution of climate variables depending on the landscape will be analyzed more in-depth in the following chapters.

Droughts conditions are not infrequent in this region during the summer. For example, the year 2022 showed the worst drought in the last 70 years and a general deficit in precipitation all year round [30], and the current year of 2023 is already showing similar signs of water scarcity. In 2022 the region experienced a prolonged period of

below-average precipitation combined with high temperatures and low humidity, which led to a severe water shortage that impacted particularly the agricultural sector and consequently the economy of the region. The scarcity of water highlighted the need to prepare and research the impacts of climate change.

1.3.   Land Properties

The geologic structure of the region is again diverse considering the heterogeneity of its formation, but it can be divided into three main zones that have already been described in the previous paragraphs: upper Pianura Padana, lower Pianura Padana, and mountain ranges. Mountain ranges are a complex composition of materials mostly with metamorphic origin. From a hydrologic point of view, the hundreds of natural and artificial lakes and the thick stream network in the region ensure that the area is homogeneously supplied by water resources.

From a land cover perspective, the area is characterized by a series of different land uses, which influence the hydrological response of the area. The distribution of land use in the area is variable depending on the landscape: in the Pianura Padana, the majority of the land is occupied by agricultural areas, while in the Alps and Pre-alps the dominant land cover shifts to being mixed or coniferous forest. The distribution of land cover can be visualized in Figure 1.5, where data from the 2018 Corine Land Cover inventory are analyzed.
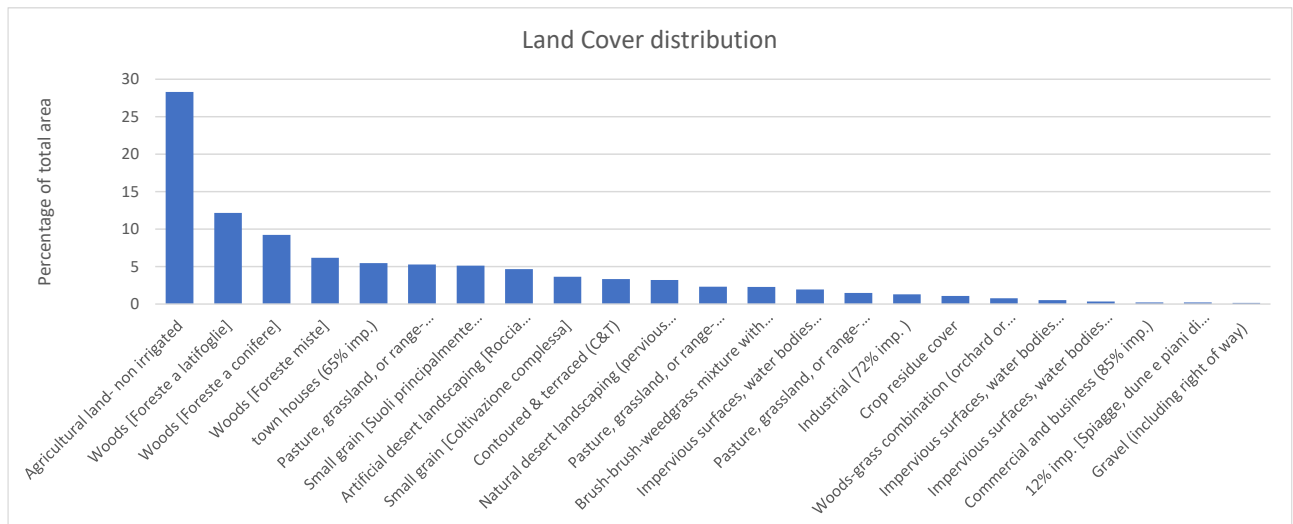
Figure 1.5: Land cover distribution in Lombardy

LC_Corine_box
Banda 1 (Gray)
- 111 - Continuous urban fabric
- 112 - Discontinuous urban fabric
- 121 - Industrial or commercial units
- 122 - Road and rail networks
- 123 - Port areas
- 124 - Airports
- 131 - Mineral extraction sites
- 132 - Dump sites
- 133 - Construction sites
- 141 - Green urban areas
- 142 - Sport and leisure facilities
- 211 - Non-irrigated arable land
- 212 - Permanently irrigated land
- 213 - Rice fields
- 221 - Vineyards
- 222 - Fruit trees and berry plantations
- 223 - Olive groves
- 231 - Pastures
- 241 - Annual crops
- 242 - Complex cultivation patterns
- 243 - Land principally occupied by agriculture
- 244 - Agro-forestry areas
- 311 - Broad-leaved forest
- 312 - Coniferous forest
- 313 - Mixed forest
- 321 - Natural grasslands
- 322 - Moors and heathland
- 323 - Sclerophyllous vegetation
- 324 - Transitional woodland-shrub
- 331 - Beaches - dunes - sands
- 332 - Bare rocks
- 333 - Sparsely vegetated areas
- 334 - Burnt areas
- 335 - Glaciers and perpetual snow
- 411 - Inland marshes
- 412 - Peat bogs
- 421 - Salt marshes
- 422 - Salines
- 423 - Intertidal flats
- 511 - Water courses
- 512 - Water bodies
- 521 - Coastal lagoons
- 522 - Estuaries
- 523 - Sea and ocean
- 999 - NODATA

Figure 1.6: Land Cover Map

1.4. Groundwater conditions

One element of the research dealt with the analysis of the groundwater settings in the region. Lombardy presents an extended network of groundwater bodies that are closely monitored due to their importance for water resources management. The levels of groundwater are influenced by a variety of factors including precipitation, evaporation, and human activities such as irrigation and groundwater pumping, but vary widely depending on the location.

Arpa Lombardy constantly checks groundwaters on the whole regional territory through a network of wells and piezometers that monitors quantitively and qualitatively the water resources in the Pianura Padana and the mountain valleys. A dataset of piezometer and wells levels was collected to study groundwater level fluctuation and compare this data with the data obtained from the water balance simulations. The first operation performed on this dataset was to filter all wells that were not part of a superficial groundwater system. The underlying hypothesis was that fluctuation levels of superficial groundwaters had to be related to the temporal distribution of the net infiltration values resulting from simulations, but it was harder to possibly find connections between deep-groundwater levels and superficial

environment variables. The filtering of the stations restricted the dataset to 237 stations, split unevenly into the 12 different districts of the Lombardy region.

Table 1.1: Distribution of observation wells of superficial groundwater systems.

DISTRICT	N° stations	DISTRICT	N° stations
BERGAMO	19	MANTOVA	30
BRESCIA	13	MILANO	36
COMO	17	MONZA BRIANZA	8
CREMONA	20	PAVIA	34
LECCO	15	SONDRIO	17
LODI	16	VARESE	12

For each station, the dataset gave information on the location of the station in WGS84 coordinates, the type of piezometer/well and its geometric features, the category of water body to which it belongs, and the type of use to which the well/piezometer is subject (industrial/potable/monitoring well/etc.). Most importantly, the dataset gave values of groundwater depth below ground level with a monthly frequency (not always respected) and within a year frame going from 2000 to 2020. From this data, it was possible to perform an analysis of the superficial groundwater in the area displayed in the following images analyzing the spatial distribution of groundwater depths.
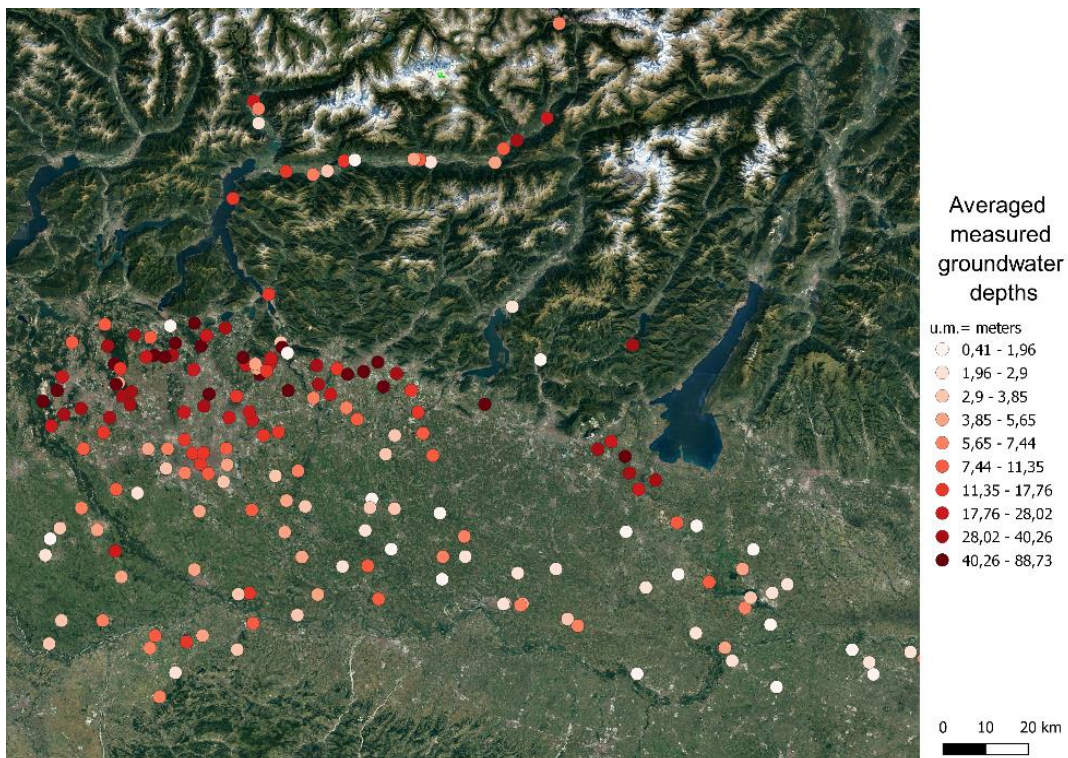
Figure 1.7: Average depth of groundwater measurements

Average depths of superficial groundwater range from less than 1 meter of depth to values of up to 80 meters in the Pre-Alps region.
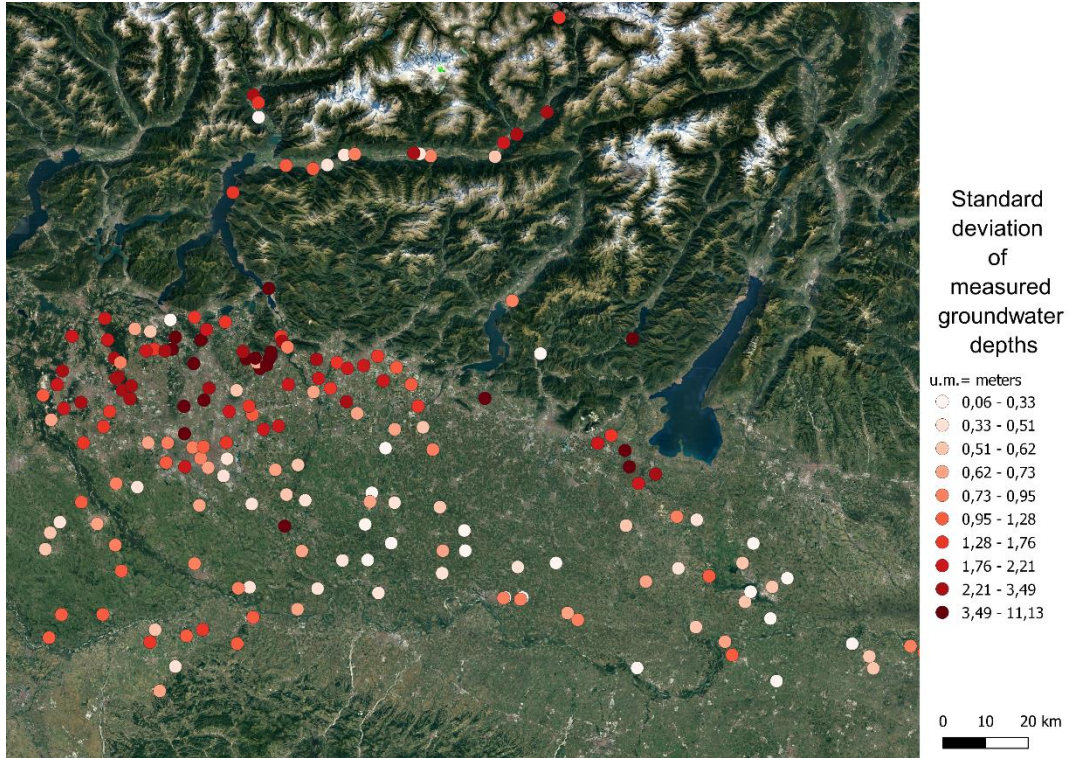
Figure 1.8: Standard deviation of groundwater depth measurements

Using data from the TINITALY digital elevation model, another image is produced to represent values of mean groundwater table elevation above mean sea level for each well in the analysis.
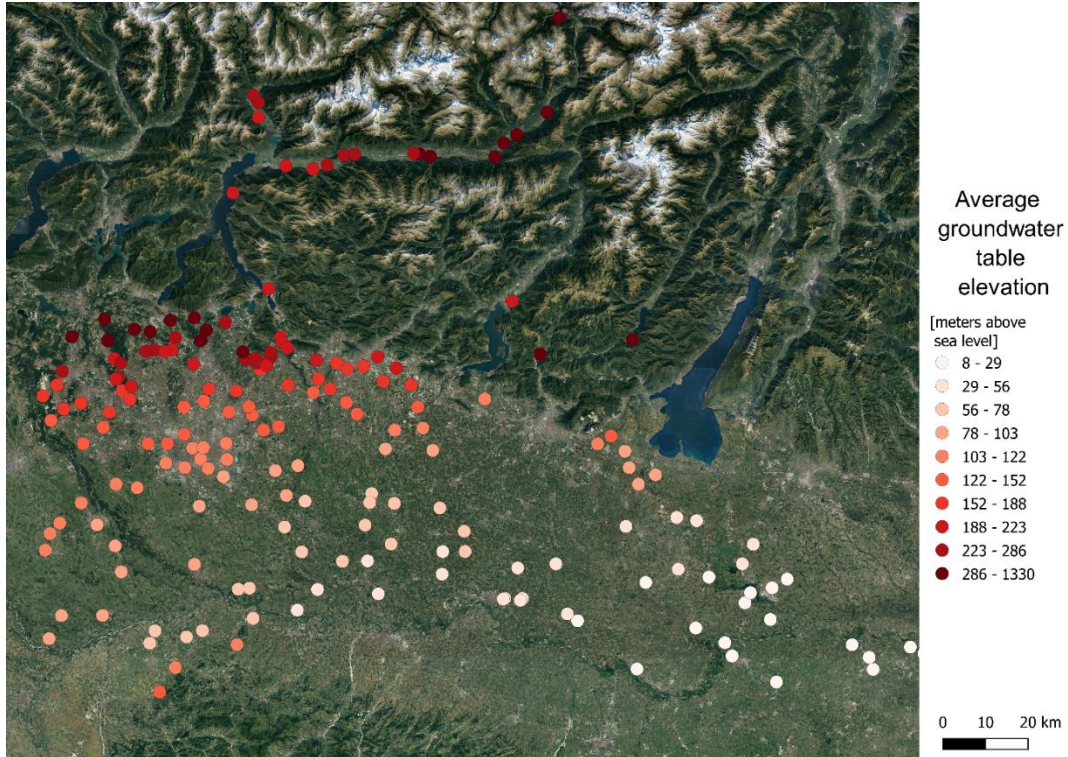
Figure 1.9: Groundwater table elevation of each well.

The last image shows a smooth trend of change of groundwater table elevation that is consistent with the orographic context of the Region. Instead, Figure 1.7 and Figure 1.8 highlight a spatial correlation between the average depths of groundwater levels and the standard deviation of these data for each station. Closer to the mountain areas, measurements of average depth are higher compared to the ones registered in the flat areas of the Pianura Padana. The same type of observation can be made for the values of standard deviation, which are higher near the mountain areas. The area of the upper Pianura Padana tends to show simultaneously higher depth of the superficial groundwater and greater oscillation of groundwater level. Correlation between the data can be estimated quantitatively using the following formula:

$$Correl(X,Y) = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} \qquad (1.1)$$

Where:

- X and Y are the two series of data of mean depth and standard deviation

- $\bar{x}$ and $\bar{y}$ are the sample means respectively of X and Y

The calculation of the correlation coefficient between the two sets of data identifies a moderate to strong correlation with a coefficient of 0,578 . Knowing that the upper Pianura Padana is an area characterized by more permeable soils, the hypothesis is that the higher permeability of the soil exposes the groundwater to be more responsive

to variations in the surface precipitation, thus it is observed a greater fluctuation of groundwater levels.

In the appendix A, the output of historical simulations on SWB is compared with the available data on groundwater depth to estimate the existing correlation between the two sets of data.

2   Soil Water Balance simulations

The Soil-Water-Balance (SWB) model was developed by Dripps and Bradbury (2007) for the US Geological Survey (USGS) as a computer program designed to simulate water movement in the atmosphere, plant, and soil system to allow reliable estimates of potential recharge in humid temperate climates [13]. The code calculates with a daily time step the water balance at the land surface using a modified version of the Thornthwaite-Mather soil moisture accounting method applied on a gridded domain. Input data such as climate, soil properties, vegetation, and land cover are used by the model to simulate the water balance and independently estimate potential recharge for each cell of the gridded domain and for each daily time step. The fundamental equation of the soil-water accounting method is the following:

$$
\begin{aligned}
Potential\ Recharge \\
= Rainfall + Runon + Snowmelt - Interception \\
- Runoff - ET - (SM_t - SM_{t-1})
\end{aligned}
\tag{2.1}
$$

Where :

- SM_{t-1} is the soil moisture in a grid cell on the previous simulation day
- SM_{t-1} is the soil moisture in a grid cell on the current simulation day
- ET is the actual evapotranspiration

To better understand how all elements in equation *(2.1)* are generated, Figure 2.1 is a conceptual diagram of the processes implemented in SWB to perform net infiltration estimates. Each of the physical processes present in the diagram is explained in the following paragraphs.
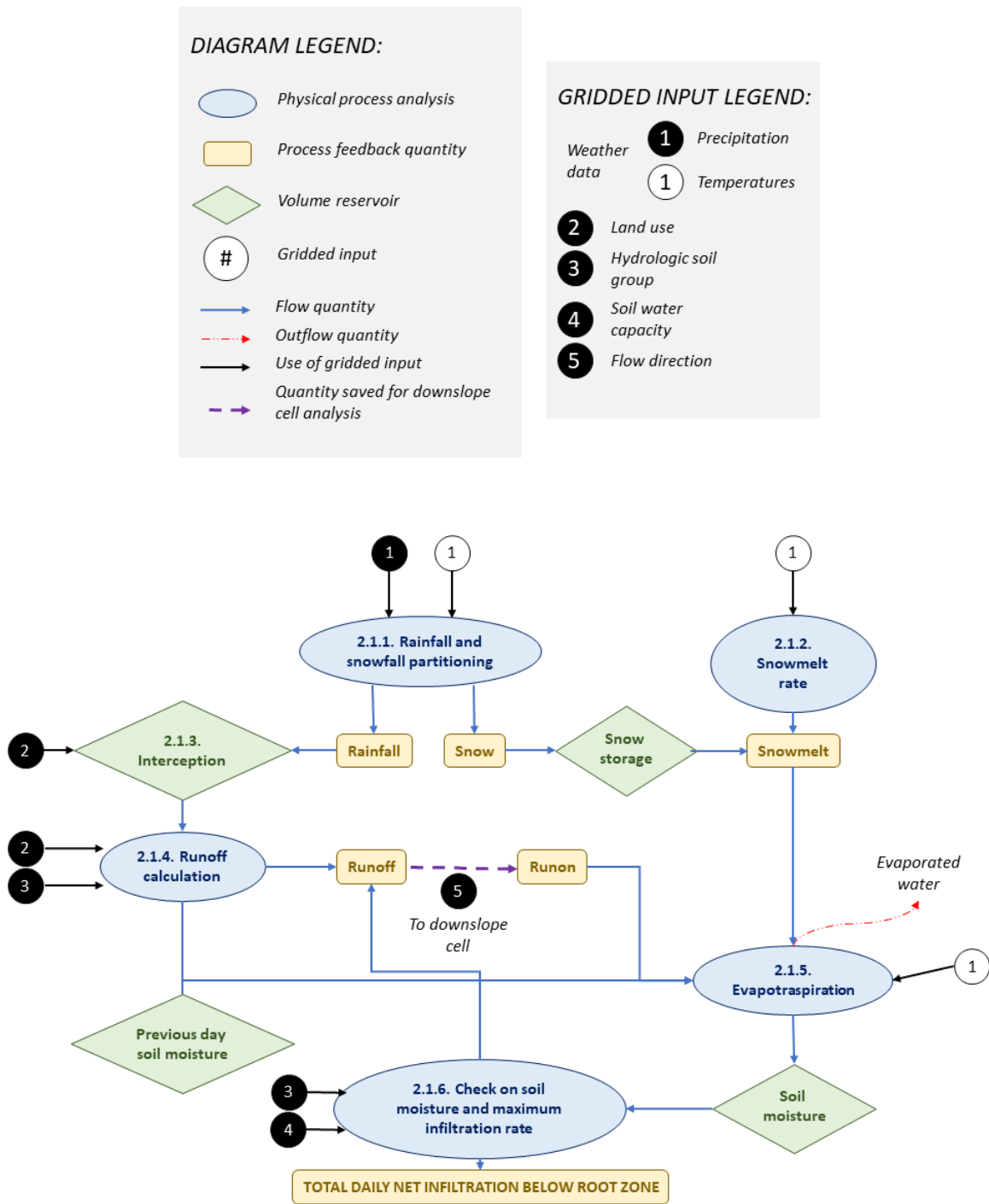
Figure 2.1: Conceptual diagram of SWB; numbers of physical processes correspond to the paragraph number analyzing it.

2.1. Quantification of physical processes

The code allows one to choose between different options to reproduce the various physical processes involved in the water balance. For example, three methods are implemented in SWB for the estimation of potential or reference evapotranspiration: the Jensen-Haise (1963) method, the Hargreaves and Samani (1985) method, or a

dataset of gridded monthly estimates of reference or potential evapotranspiration. In this case, the choice of the method used for the simulations is naturally driven to the second method because the Jensen-Haise method has been calibrated in the Southwestern United States and could result in incorrect estimates, while monthly gridded estimates of evapotranspiration are not available for the area. Similar reasoning has been applied to the remaining processes to define each equation controlling the SWB output. The following paragraphs analyze synthetically all models and equations used to reproduce the physical processes involved in the calculation of water balance.

2.1.1. Rainfall and snowfall partitioning

All precipitation is controlled by SWB to define whether it falls as snow or rain. If the daily mean temperature minus one-third of the difference between the maximum and minimum temperature is less or equal to the freezing point of water, the precipitation falls as snow[13]. From this point, water falling as rainfall or as snow takes two different paths for the analysis of their contribution to recharge. Precipitation falling as rain will be computed for the estimation of potential recharge. All the precipitation falling as snow contributes to the snow reservoir, which is allowed to expand with no limitation. The limiting factor affecting the growth of snow storage reservoir is the snowmelt contribution.

2.1.2. Snowmelt

Snowmelt contribution is automatically calculated in SWB as snow-water equivalent based on the temperature index method [13]. The method assumes that each day, 0.059 inches (1.5 mm) of snow are expected to melt for every degree °C above the freezing point of water. This implies that when snow is accumulated in that grid cell and when possible due to temperature conditions, snowmelt will take place until the snow storage reservoir is emptied. All snow that converts to snowmelt is assumed to contribute to the water balance in the same way as all other water sources.

2.1.3. Interception

The precipitation that doesn't transform into snow becomes rainfall. Interception is calculated with a simple "bucket" model approach. Before any water is assumed to reach the soil surface, daily precipitation values must exceed the specified interception amount. Interception amounts are defined for each land use and are different for growing or dormant seasons. For example, wood areas have an interception volume of 0.05 inches (1.5 mm) during the growing season and 0 inches during the dormant season. If 0.04 inches of precipitation fall in a day of the growing season, no precipitation will reach the ground, but if the same quantity of rain falls during the dormant season it will all reach the ground and contribute in some fraction to potential recharge.

2.1.4.　Runoff calculation

All daily precipitation that doesn't contribute to interception storage is directly redirected to the runoff calculation. Runoff is the quantity of water that is rejected from the soil due to poor infiltration capacity.

Runoff estimates are performed with the Curve Number (CN) method. This method estimates runoff in relation to the difference between precipitation (P) and an initial abstraction term (I_a). The CN number is nondimensional and specific for each land use and soil type: it varies from 0 for a pervious surface to 100 for an impervious surface. From the CN coefficient, maximum soil storage (S_{max}) and initial abstraction (I_a) are defined to calculate the daily runoff generated (R).

$$S_{max} = \frac{1000}{CN} - 100, \qquad \text{[mm]} \qquad (2.2)$$

$$I_a = 0.2\, S_{max}, \qquad \text{[mm]} \qquad (2.1)$$

$$R = \frac{(P - I_a)^2}{P + S_{max} - I_a} \qquad if\ P > I_a. \qquad \text{[mm]} \qquad (2.3)$$

In addition to this, the Curve Number coefficient of a cell is automatically modified by the code to consider the influence that the amount of precipitation that occurred in the previous 5 days period might have on infiltration capacity. If the rain in the previous 5 days has been higher than a certain value, the CN number is increased to account for higher runoff (less infiltration) generated when precipitation falls on saturated soil. On the opposite side, when it has rained less than a certain amount in the previous five days and soils are dry, CN coefficients are decreased to reflect a general decrease in runoff (more infiltration) compared to average conditions. The values of the Curve Number are in both cases adjusted according to constant coefficients formulas that depend on the original Curve Number coefficient.

The estimation of a runoff value becomes essential not only for the redirection of the incoming precipitation of a cell but also for the cell located downslope. The runoff quantity of a cell, which acts as a sink, becomes the run-on quantity of the next downslope cell. For this reason, Figure 2.1 displays a link between the runoff and run-on, where the dashed lines represent ideally the movement of runoff volume from the uphill to the downhill cell of the grid. The linkage between the different cells of the grid is implemented by SWB with a downhill routing method based on a D8 flow routing scheme, which transforms runoff from an upslope cell to run-on on its downslope cell. To correctly represent downhill routing, the model performs its mass

balance calculation starting always from the most upslope cells and moving downslope.

2.1.5. Evapotranspiration

Evapotranspiration (ET) is equal to the sum of all the water vapor leaving the system from snow, ice soil, or plant transpiration. Actual evapotranspiration is calculated in two steps.

First, reference evapotranspiration is estimated with the Hargreaves and Samani method (1985): SWB reads in the average latitude and longitude of the study area (to estimate solar radiation) and the values of minimum and maximum daily temperature to produce a spatially variable estimate of reference evapotranspiration:

$$ET_0 = 25.4 \cdot (0.0023 \, (T_{mean} + 17.8) \, \sqrt{T_{max} - T_{min}} \, \, R_a) \qquad [mm]$$

(2.4)

Where:

- ET_0 is the grass reference evapotranspiration [mm].

- R_a is the solar radiation that is directly calculated from the given values of latitude and longitude of the grid.

- T_{max}, T_{min} are the maximum and minimum values of temperature $[F]$ for a cell and a day analyzed.

The reference evapotranspiration differs from the actual evapotranspiration: the first one refers to the quantity of water that could potentially evapotranspirate from a well-watered grass, while the second one refers to the real quantity of water that evapotranspirates from the soil in its real soil moisture conditions. The consequence is that actual evapotranspiration will always be equal to or smaller than reference evapotranspiration: it will be equal on those days when the soil shows high-moisture values, while it will be lower when the soil is partially dry. To estimate actual evapotranspiration, the Thornthwaite-Mather (1957) method is used. The method is based on the hypothesis that the ratio between instantaneous actual and reference evapotranspiration must be equal to the ratio between current soil moisture and maximum retained soil moisture.

$$et_{actual} = et_0 \, \frac{SM}{SM_{field\ capacity}} \qquad [mm]$$

(2.5)

Where:

- et_0 is the instantaneous reference evapotranspiration [mm].

- et_{actual} is the instantaneous actual evapotranspiration [mm].

- SM and $SM_{field\ capacity}$ are respectively soil moisture in the soil and soil moisture in the soil at field capacity [mm].

Equation *(2.5)* can be integrated to evaluate evapotranspiration quantities. To do so interim soil moisture has to be evaluated.

$$SM_{interim} = SM_{t-1} + rainfall + runon + snowmelt \qquad \text{[mm]} \quad (2.6)$$
$$- runoff$$

$$ET_{actual} = SM_{interim}\left(1 - e^{-\frac{ET_0}{field\_capacity}}\right) \qquad \text{[mm]} \quad (2.7)$$

2.1.6. Check on soil moisture and maximum infiltration rates.

Now that all the sources and sinks have been estimated, it is possible to estimate the updated value of soil moisture SM_t as reported in equation (2.1).

If SM_t exceeds the maximum capacity of the soil (field capacity), SM_t is updated to the value of the field capacity and all the exceeding soil moisture is considered to generate net infiltration:

$$net\_infiltration = SM_t - field\_capacity \qquad \text{[mm]} \quad (2.8)$$

If SM_t is lower than the maximum capacity of the soil, no net infiltration occurs.

In addition, the model considers that with flow routing enabled, downslope cells could produce enormous values of infiltration that would not be compatible with their hydrological characteristics. To avoid this from happening, the model has the possibility to define a maximum recharge rate for each hydrologic group. If net infiltration in one cell is higher than its maximum recharge, the net infiltration is updated to the maximum recharge value, and the rejected net infiltration is routed to the next downslope cell in the form of runoff.

2.2. Input gridded data

All the water balance calculations are performed based on a rectangular grid, which size is defined by the user. Through means of this regular grid, there are several data requirements that the code needs to perform its simulation and return potential recharge values with a daily frequency. Inside the SWB input manual, explanations are given on how to extract and produce the necessary data. The required input gridded data are the following, reported in the same order as they are represented in the conceptual diagram of Figure 2.1:

1. Weather data: daily precipitation and daily maximum and minimum temperature

2. Land use classification

3. Hydrologic soil group

4. Flow direction

5. Soil-water capacity

The choice of the grid dimension is free of constraints for the user. The optimal length for the grid dimension of a recharge model depends on the spatial variability and the complexity of the recharge process, as well as the available computational resources. In general, a finer grid resolution can better capture spatial variations in recharge, but also requires more computational power and data to parameterize and run the model. Overall, there is no single "best" grid dimension for a recharge model, and the optimal resolution depends on the specific application and available data.

In the case of Lombardy, the data referring to the land cover characteristics are the one available with the smallest resolution (100 m). However, considering the scale of the recharge process and that a resolution of 100 m would imply long simulation periods, the decision was to set a coarser grid of dimension 250 m as derived from the resolution corresponding to the available dataset for the hydrologic soil groups.

The five input gridded datasets listed above functioned as a starting point to investigate what kind of effect a variation in the input dataset had on the output of the model, especially on the net infiltration values. By varying the GCMs used or the empirical equations that define land characteristics a set of different simulations of net infiltration is generated for the same area. An analysis of all the outputs of the different scenarios produced will answer the project scope.

2.2.1. Weather data

The first and more obvious variation in the input data was the one related to the weather conditions. A partial objective of the research is to estimate the amount of variation in the quantity of net infiltration in the face of the effects produced by climate change. A series of reliable datasets were available for download on Google Earth Engine Catalogue (https://earthengine.google.com/). This made it possible to download, clip, and reproject this data to generate an input that could be readily available for SWB simulations.

Historical series – ERA5 Dataset

The first simulations had to be performed on a historical scenario to assess current recharge conditions. Using a time window of 22 years (from 1998 to 2019) data were downloaded from the ERA5 dataset. ERA5 is a global atmospheric reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) that provides a comprehensive view of the Earth's atmospheric conditions globally

over the past four decades (from 1979 to 2020). The dataset is generated using a computer model that combines millions of measurements from a variety of sources, such as weather stations, satellites, and weather balloons and it includes a wide range of atmospheric variables such as temperature, wind, pressure, humidity, and precipitation, among others. These variables are available on a regular grid with a spatial resolution of approximately 28 kilometers and a temporal resolution of one hour. To comply with the simulation frequency of one day of SWB, ERA5 daily aggregates were used instead of the hourly data. Daily total precipitation values were calculated as daily sums, while temperature parameters were extracted as daily averages [31]. To match smoothly the SWB simulation grid, weather data have been downloaded with a square grid of 1km interpolating the original data with a bilinear method.

Future series – NASA NEX-GDDP Dataset

To satisfy all these conditions, data were downloaded from the NASA NEX-GDDP dataset. The NASA NEX-GDDP (National Aeronautics and Space Administration National Climate Assessment - Global Daily Downscaled Projections) dataset is a high-resolution climate projection dataset that provides global climate model simulations of future climate change scenarios. It was developed in a collaboration between the NASA Earth Exchange (NEX) and the University of Maryland [32].

The NEX-GDDP dataset provides daily climate projections for temperatures and precipitation at a spatial resolution of 0.3 degrees (approximately 28 kilometers) for the period between 1950 and 2100. The projections are based on the outputs of 21 global climate models and two different greenhouse gas concentration pathways (RCP4.5 and RCP8.5) produced by the Intergovernmental Panel on Climate Change (IPCC). The RCP4.5 corresponds to an intermediate scenario that assumes a stabilization of greenhouse gas emissions by the end of the century. On the other side, the RCP8.5 represents a more extreme scenario with higher levels of global warming and climate impacts compared to RCP4.5 [32].

A large set of possible data is available through Global Circulation Models to be applied as weather input data. The first decision to be made concerns the type of emission scenario to be chosen, for the NEX-GDDP dataset this choice is not only limited between RCP4.5 and RCP8.5 but is extended for each RCP to the choice of global climate model run. Each of these options will determine differences in the input weather data and consequently differences in the output. To acknowledge the uncertainty related to the choice of greenhouse gas emissions scenarios (RCP) and of global climate model run, it is recommended to take into consideration a sufficiently large ensemble of these datasets and perform a simulation for each. The choice of GCMs out of 22 could be driven in many ways, for example by previously estimating the reliability of each GCM. An example of this practice is the research conducted by Ruan et al. (2018) [33], which assessed the performance of 34 CMIP5 GCMs in the

Lower Mekong Basin between 1975 and 2004 by comparing the estimated GCM precipitation values with the real measured. This kind of analysis would contradict the scope of this project as the results obtained might not reproduce the reality of GCMs' uncertainty. The result is that the simulations performed took into consideration four random models: the ACCESS1-0 produced by the Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia, the CCSM4 produced by the National Center of Atmospheric Research (USA), the MIROC5 produced by the Atmosphere and Ocean Research Institute, the National Institute for Environmental Studies, and the Japan Agency for Marine-Earth Science and Technology (Japan) and the MRI-CGCM3 produced by the Meteorological Research Institute (Japan). For each model, the two scenarios of RCP4.5 and RCP8.5 were considered.

The following table works as a summary of the different climate projections datasets that were used in the application of the SWB recharge model.

Table 2.1: Summary of available weather data

Dataset	Representative Concentration Pathway	Model Climate Run
ERA5	-	-
NASA NEX-GDDP	RCP4.5	ACCESS1-0
		MRI-CGCM3
		CCSM4
		MIROC5
	RCP8.5	ACCESS1-0
		MRI-CGCM3
		CCSM4
		MIROC5

These datasets can be analyzed in many ways. The first analysis performed is on the temporal distribution of the yearly average precipitation for the three classes: historical dataset, RCP4.5 datasets, and RCP8.5 datasets. Figure 2.2 draws the envelope of

variation of yearly average precipitation. The main difference between the datasets is found in the range of variation: while the historical precipitation varies between a minimum of 955 $mm/year$ and a maximum of 1701 $mm/year$, the RCP8.5 scenarios vary between 866 $mm/year$ and 2734 $mm/year$. As expected, the RCP4.5 displays a general variation that sits in the middle between the historical scenario and the RCP8.5.
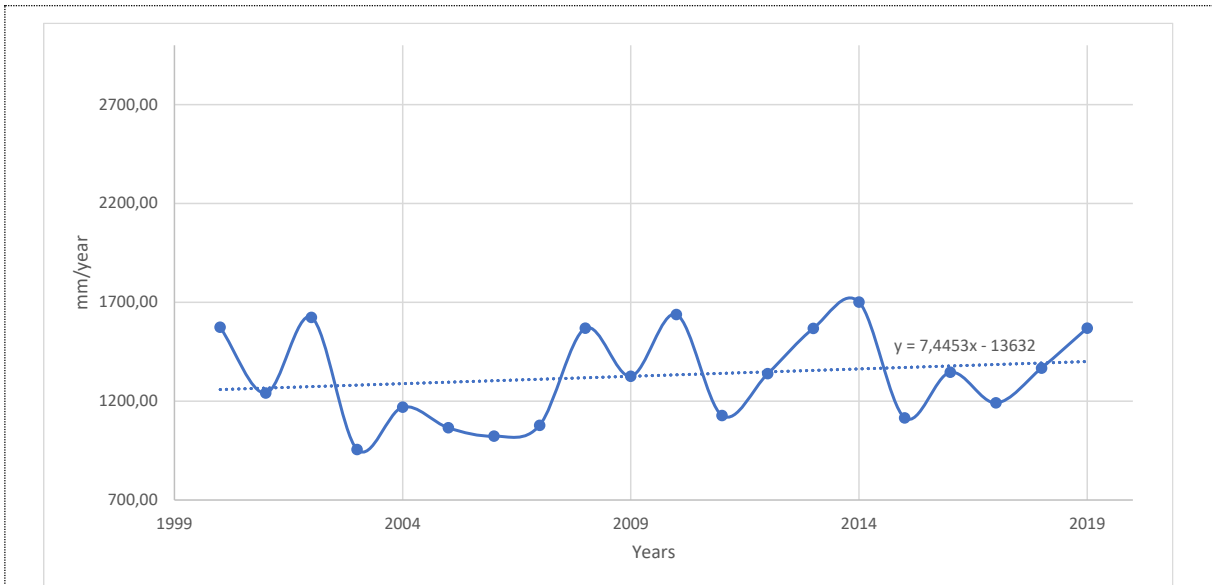
(a) Temporal distribution of yearly average precipitation for ERA5 data.
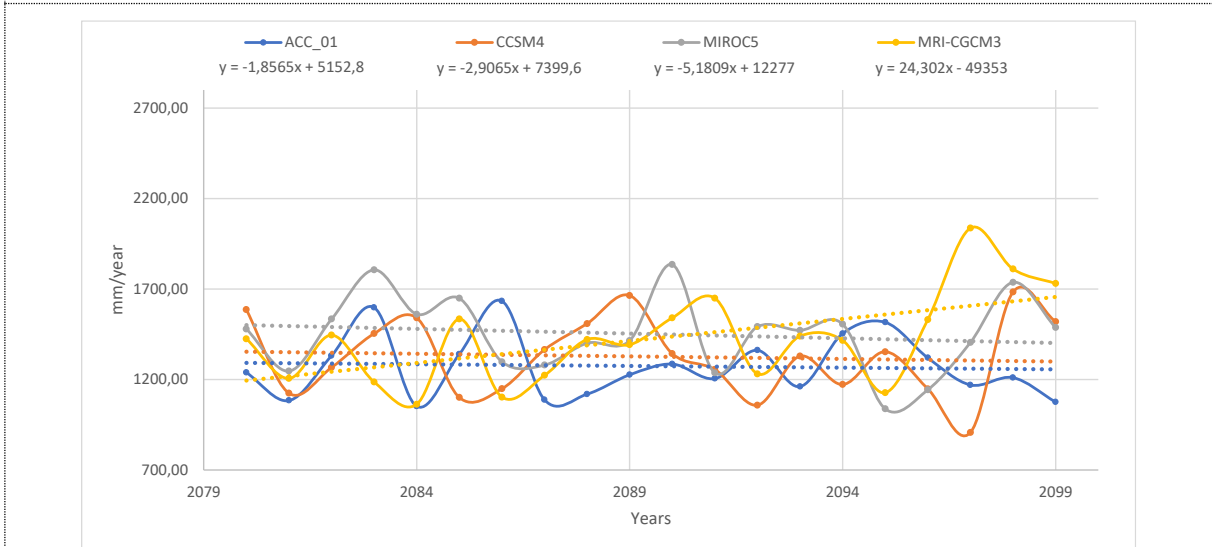
(b) Temporal distribution of yearly average precipitation for RCP4.5 data.

(a) Temporal distribution of yearly average precipitation for RCP8.5 data.
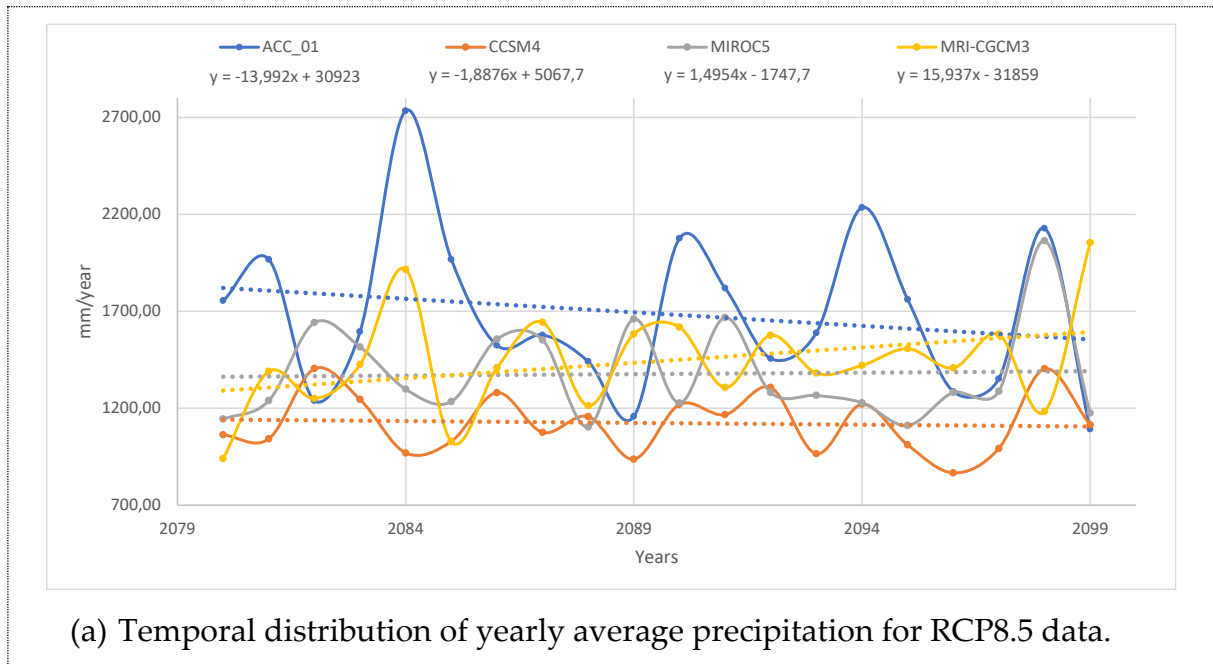
Figure 2.2: Development of average precipitation for different datasets; dotted lines represent the linear trend and their defining equations are displayed under the legend.

It is noticed that, especially for the RCP8.5 scenario, specific general circulation models tend to overestimate or underestimate precipitation with respect to the average estimate of others. For example, for the year 2084, the ACCESS-01 model forecasts an average precipitation of 2734 $mm/year$, close to 1000 mm/year more than the second highest precipitation of 1916 $mm/year$ of the MRI-CGCM3 model and almost triple the precipitation estimated for the same year by the CCSM4 of 969 mm/year. In general, the RCP 4.5 model shows smaller variations of precipitation from one model to the other. Table 2.2 shows an analysis of mean values of each dataset. The uncertainty of climate projections is highlighted by the fact that the RCP4.5 and the RCP 8.5 predict both an increase and a decrease in average precipitation depending on the model used.

Table 2.2: Statistical analysis of precipitation datasets; in the 'AVERAGE' column, the yellow color means average precipitation lower than historical reference, blue color means average precipitation higher than historical reference; in the standard deviation column the blue data bars are scaled with respect to the standard deviation values.

SCENARIO	MODEL	AVERAGE [mm/year]	% VARIATION FROM HISTORICAL
HISTORICAL	ERA5	1329,4	0,0
RCP4.5	ACCESS-01	1273,7	-4,2
	CCSM4	1326,4	-0,2
	MIROC5	1451,2	9,2
	MRI-CGCM3	1425,6	7,2
RCP8.5	ACCESS-01	1687,5	26,9
	CCSM4	1123,6	-15,5
	MIROC5	1376,9	3,6
	MRI-CGCM3	1441,6	8,4

Similarly to Figure 2.2, it is possible to estimate the average precipitation for each month of the year to analyze the eventuality of a variation in the seasonality of precipitation patterns in the future.
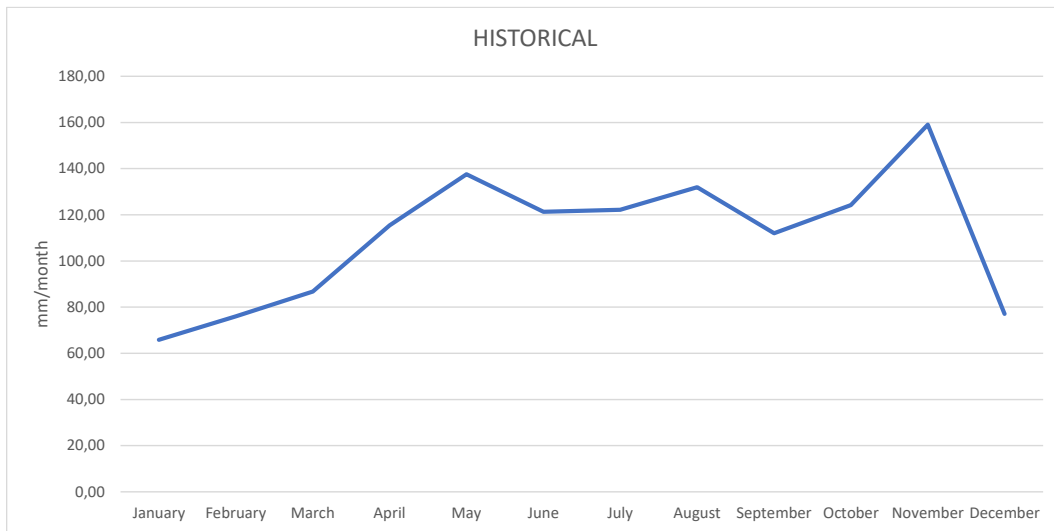
(a) Monthly distribution of precipitation values for ERA5 dataset
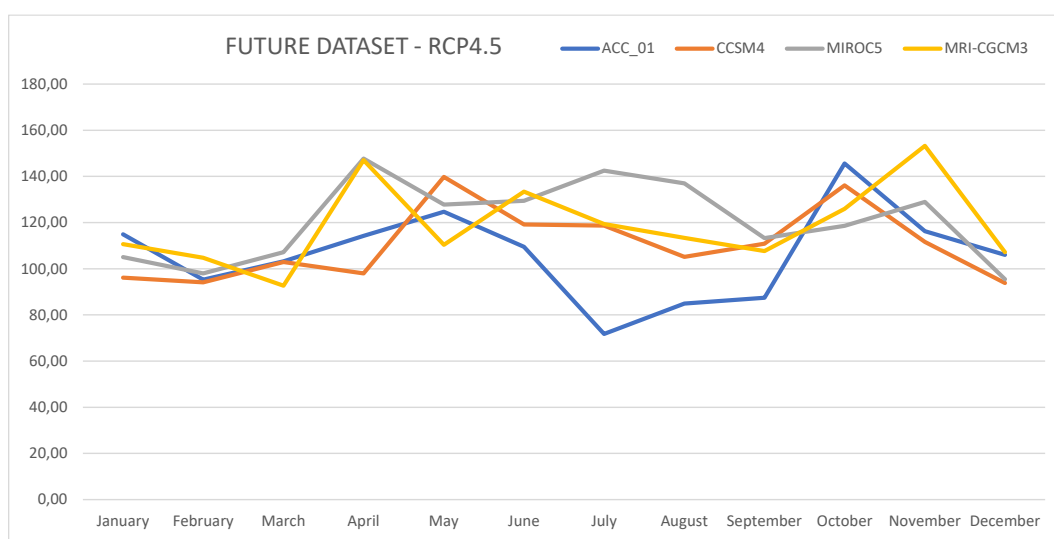
(b) Monthly distribution of precipitation values for RCP4.5 dataset
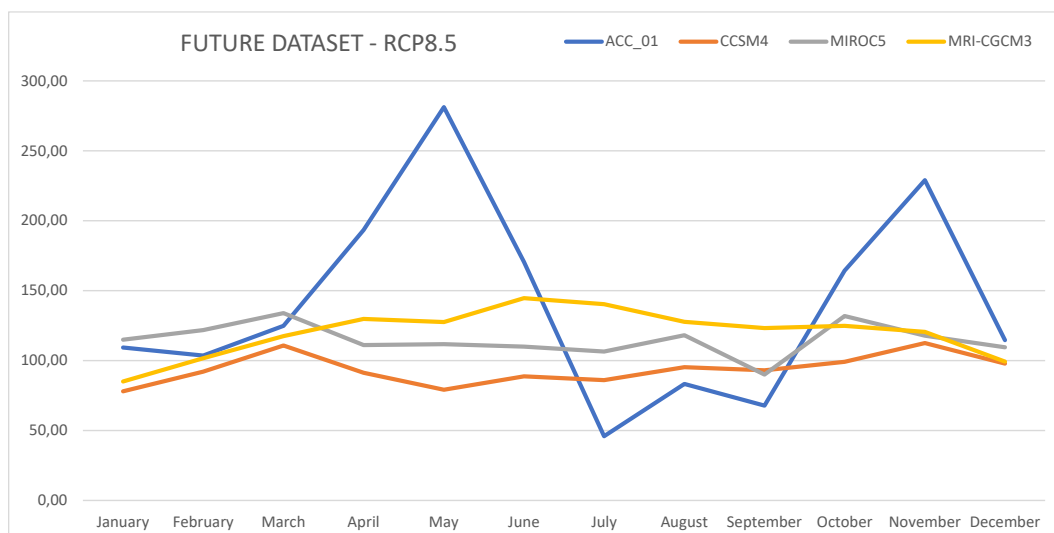
(c) Monthly distribution of precipitation values for RCP8.5 dataset

Figure 2.3: Average precipitation for each month of the year

The historical dataset confirms the information presented in chapter 1.2: the two rainiest months in the past two decades have been on average May and November. August has also shown significantly high values of cumulated precipitation, probably due to intense rainstorms, while January appears to be the driest month of the year. This trend becomes much more dispersed moving to the analysis of GCMs predictions: the rainiest months are anticipated or postponed depending on the model chosen. For both emissions scenarios, the ACCESS-01 model departs significantly from the other GCMs both in the absolute values and in the temporal distribution of the peaks and lows of precipitations. This is another proof of the uncertainty rooted in the GCMs. One other observation is that while the RCP4.5 scenarios seem to agree more or less with the seasonal distributions of precipitation in the area, the RCP8.5 scenarios all

show different patterns of precipitation across the year. Intuitively, this should also have an effect on the seasonal distribution of recharge estimates that will be analyzed in the following chapters.

The last analysis of the precipitation data considers the spatial distribution of precipitation in the entire domain modeled. While Figure 2.4 displays the mea annual precipitation for the historic dataset, Figure 2.5 analyses the relative variation of GCMs' forecasts of mean annual precipitation with respect to the historic data of figure Figure 2.4.
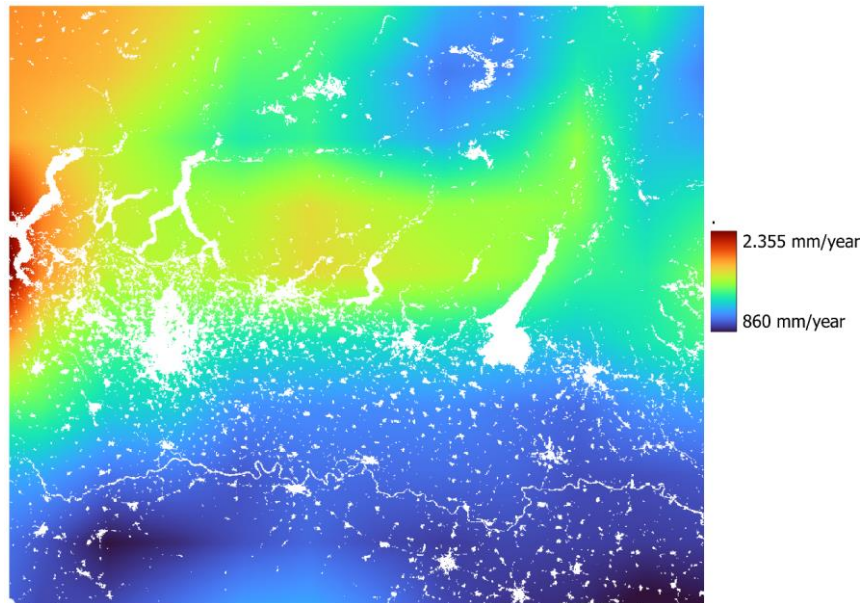
Figure 2.4: Spatial distribution of average yearly precipitation from 2000 to 2019; white areas in the map are impervious areas (urban or water bodies) where no recharge prediction are performed.

Figure 2.4 and Figure 2.5 present respectively white and dark areas partially covering the displayed raster. These areas will be present almost in all the reported images and correspond to the location of water bodies or urban areas where recharge calculations are not performed.

MODEL	RCP4.5	RCP8.5
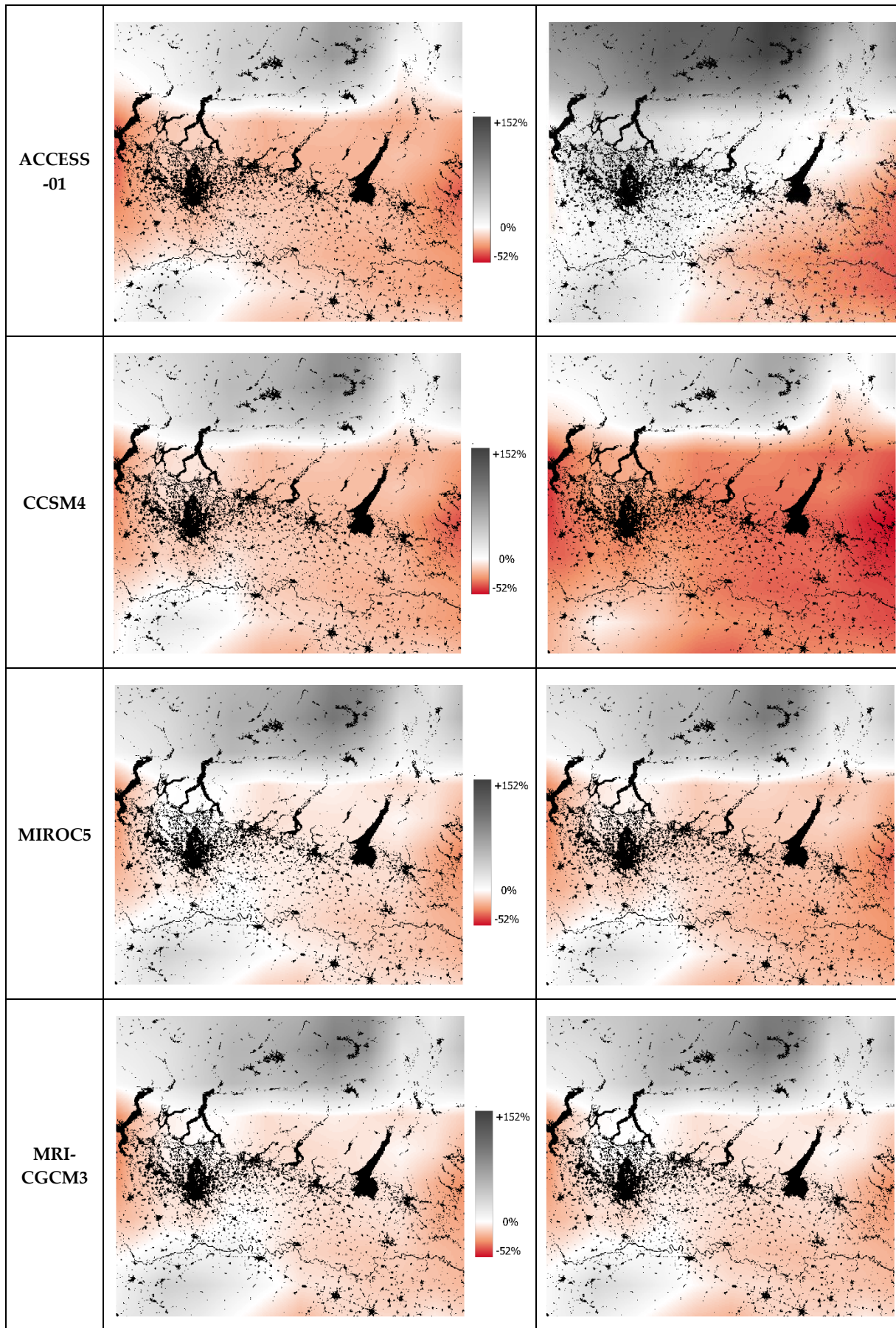
Figure 2.5: Average yearly precipitation (from 2080 to 2099) for all GCMs expressed as a percentage of variation with respect to the data from Figure 2.4; Black areas represent impervious areas where no recharge simulation are performed.

All GCMs predict some variation of precipitation with different trends depending on the landscape. On average, all models predict a small reduction ($\sim -10/-20\%$) of yearly precipitation in the flat areas of the Pianura Padana and a significant increase of precipitation in the mountain areas (~60/100%). Again, the model ACCESS-01 produces out-of-trend prediction with respect to other models: the RCP8.5 projections of the other models seem to emphasize the distribution produced by the RCP4.5 projections, but the ACCESS-01 model has an RCP8.5 projection that doesn't match the RCP4.5. For example, in the ACCESS-01 RCP8.5 projection, only the eastern part of the Pianura Padana is expected to observe a decrease in yearly precipitation. In addition, the CCSM4 RCP8.5 predicts a decrease in precipitation that is significantly lower than all the other models which will probably propagate to estimates of potential recharge.

Precipitation is one factor influencing aquifer recharge, but it is not the only one. The other meteorological parameters downloaded from the online datasets are the maximum and minimum daily temperatures. Theoretically, temperatures influence a lot of processes involved in the definition of recharge, from infiltration capacity to other phenomena such as mist and haze formation, but in practice, temperatures are used in SWB balance only to define reference evapotranspiration and manage the processes involving snow generation and snowmelt.

The Figure 2.6 represents the trend of change of maximum and minimum temperatures for different climatological conditions averaged over twenty years. The average temperatures represented serve as a reference to compare different scenarios: differences of 10/20 °C could be found in the same month for the same dataset moving from the mountains to the flat areas in the Pianura Padana. Despite this, the graph shows some interesting elements of change between the different scenarios. Temperatures estimates have a wider range of variation and uncertainty in the hottest months of the year (around July and August) than in the coldest months of the year. All RCP8.5 scenarios produce average temperatures, both in the minimum and maximum field, that are higher than their corresponding RCP4.5 scenario. The only exception to this is the MRI-CGCM3 model that for the RCP4.5 scenarios predicts temperatures that are very close to the historical dataset. However, estimates from one RCP8.5 can intersect the estimates of an RCP4.5 scenario from another model. This happens for the ACCESS-01 model: in July, minimum and maximum average temperatures predicted for the moderate RCP4.5 scenario are respectively 17,34 $°C$ and 28,93 $°C$ and are higher than the prediction made for the extreme RCP8.5 scenario by the MRI-CGCM3 model which shows minimum and maximum temperatures respectively of 16,59 $°C$ and 25,36 $°C$. This feature shows that, against expectations, a moderate emissions scenario of a GCM can estimate temperature changes that are

more intense than an extreme emissions scenario of another GCM and proves the uncertainty rooted in GCMs.
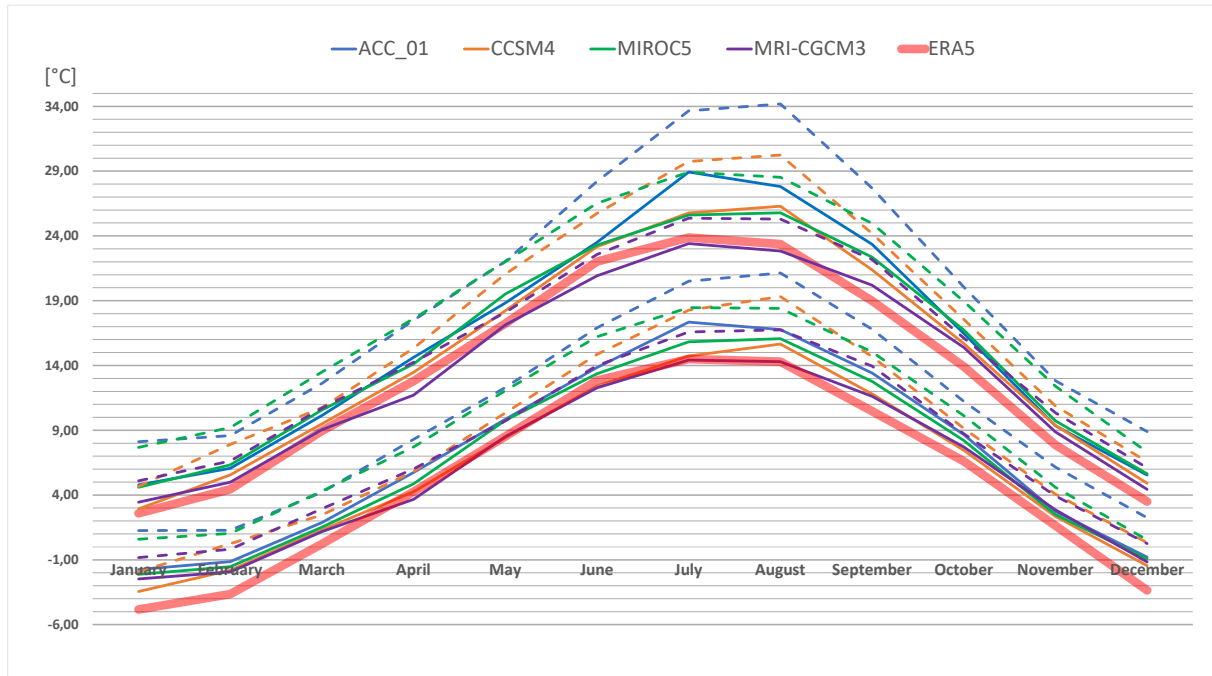
Figure 2.6: Average monthly maximum and minimum temperatures for each dataset over the twenty years available (2000-2019/2080-2099); the thick red line represents the ERA5 historical dataset; continuous thin lines represent RCP4.5 scenarios while dashed thin lines represent RCP8.5 scenarios.

Figure 2.7 presents annual mean temperatures from the ERA5 dataset, averaged over 20 years from 2000 to 2019. Results show that mean temperatures in the area range from a mean annual value of 15 °C in the low Pianura Padana up to a minimum of -1 °C in the northern parts of the alps.
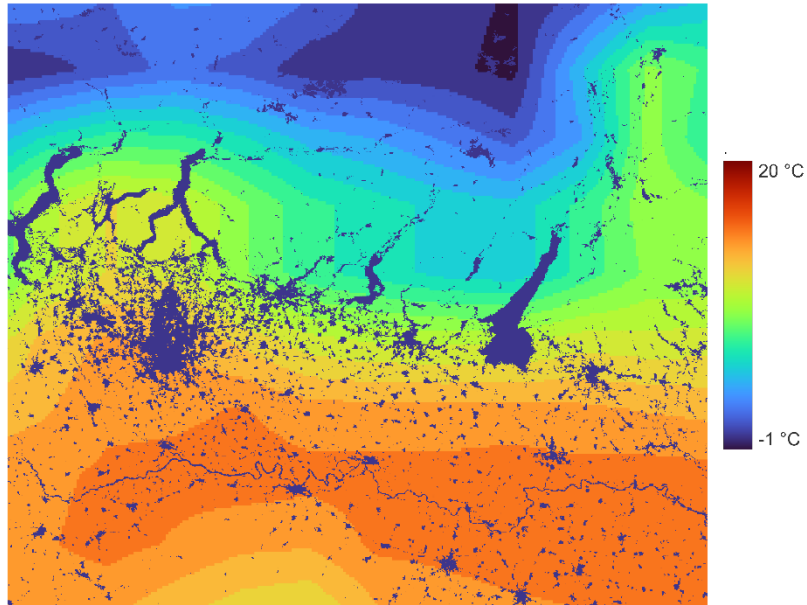
Figure 2.7: Annual mean temperatures for ERA5 dataset from 2000 to 2019

Figure 2.8 instead analyzes mean temperatures for the eight climate models investigated. Results show that all RCP8.5 models produce mean annual temperatures with 3 °C or 4 °C more than the corresponding RCP4.5 model. Mean annual temperatures tend to increase with respect to historical conditions up to +10 °C, but it's not rare to find areas of reduced mean annual temperature that decreases up to -3 °C.

MODEL	*TEMPERATURES*	*VARIATION FROM HISTORIC DATA*
ACCESS-01 **RCP4.5**		
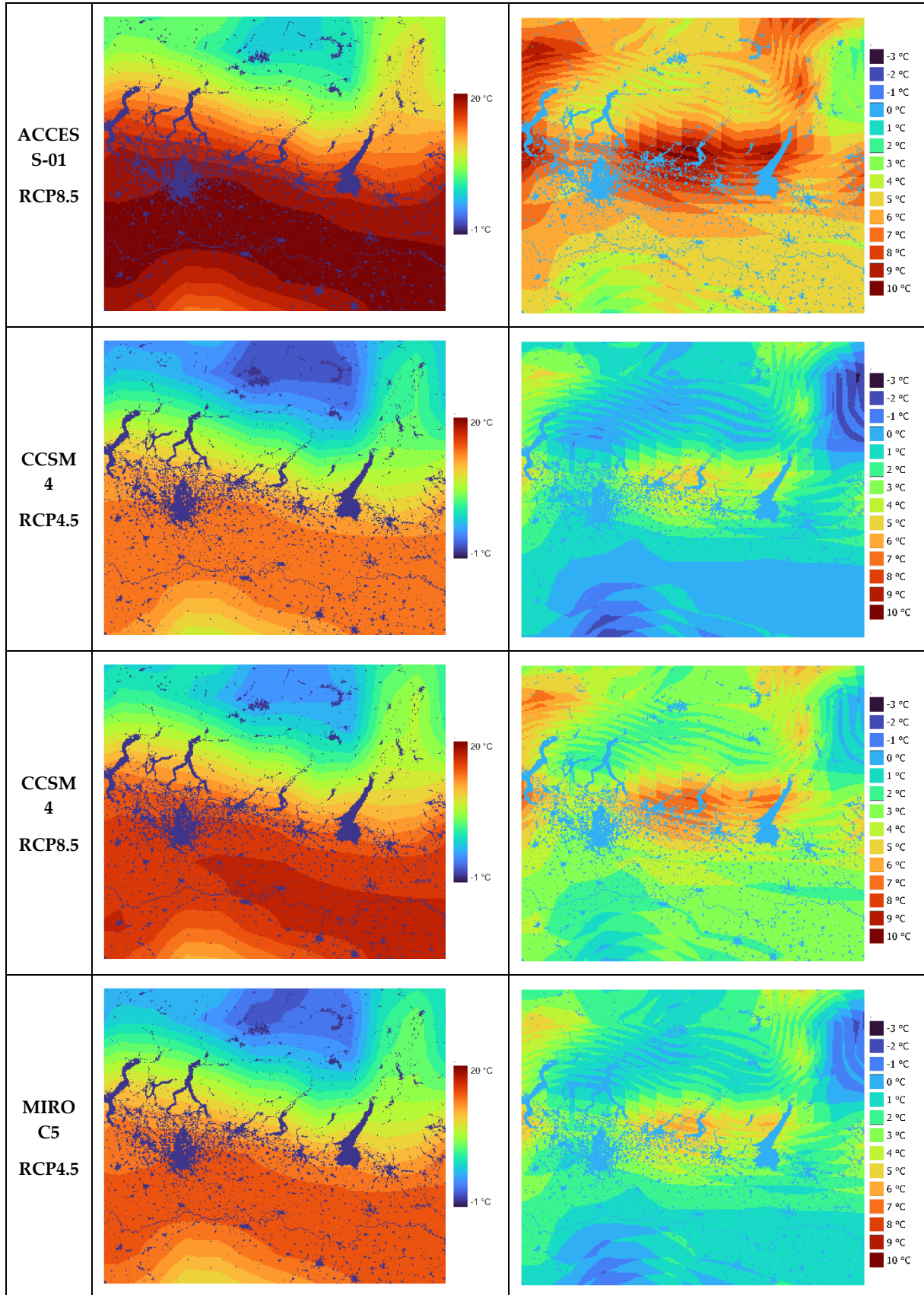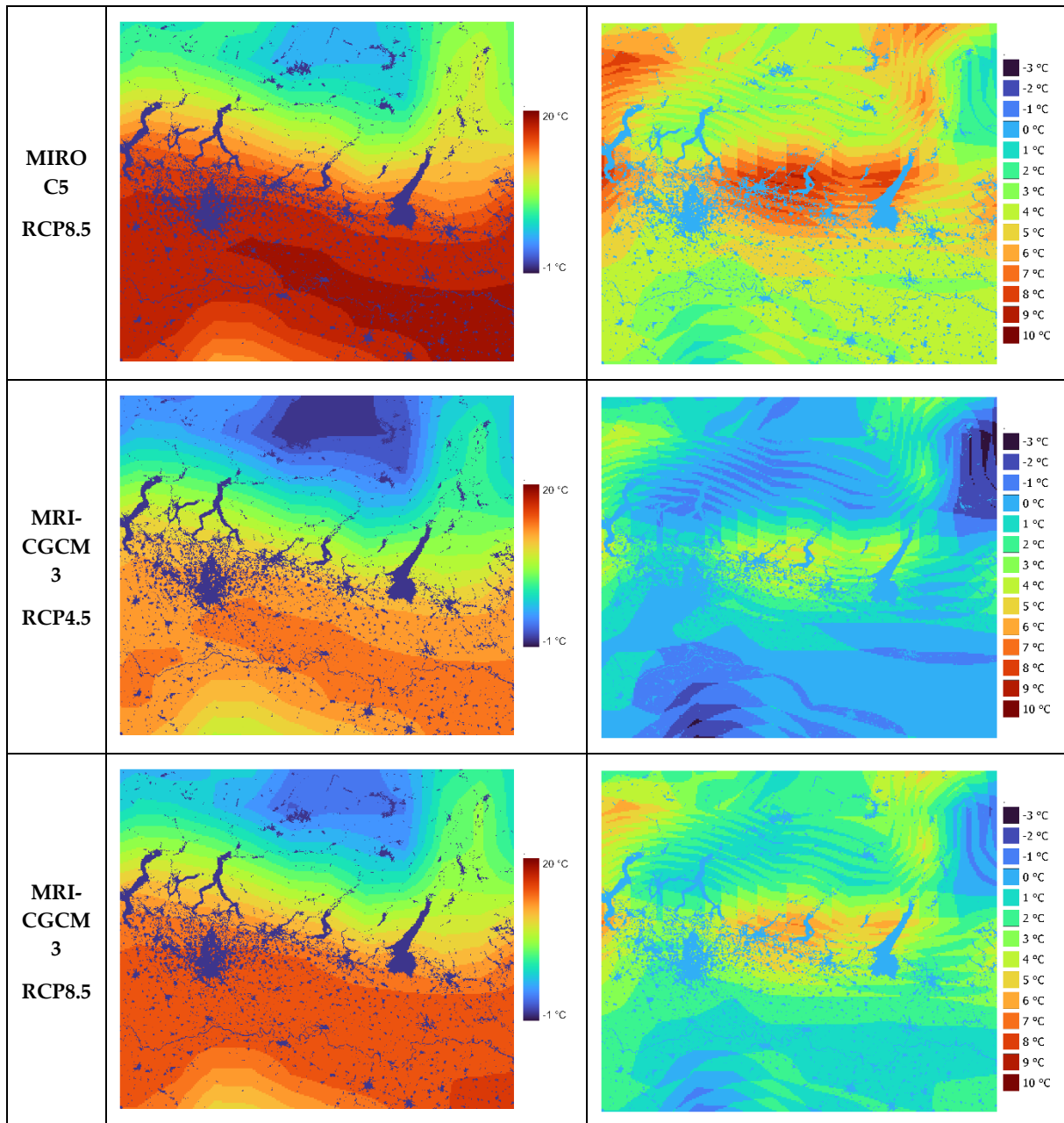
Figure 2.8: Distribution of mean annual temperature (averaged between 2080 and 2099) and of difference between future mean annual temperature and historical mean annual temperature.

2.2.2. Land cover

Land cover classification is extracted from the Corine land cover inventory. CORINE Land Cover (CLC) is a digital map of land cover in the European Union (EU) and some neighboring countries. The inventory consists of 44 land cover classes, which are defined based on the FAO Land Cover Classification System and are shown in Figure 1.6. It is produced from a combination of satellite imagery, aerial photography, and ancillary data, such as digital terrain models and climatic data. It is updated every 6 years to capture changes in land use and land cover and the most recent version,

CLC2018, was released in 2020 and covers the period from 2012 to 2018. From the EEA website, the CLC inventory is freely available for download. The dataset is distributed in the standard European Coordinate Reference System defined by the European Terrestrial Reference System 1989 (ETRS89) datum and Lambert Azimuthal Equal Area (LAEA) projection (EPSG: 3035). The raster grid size of the CORINE Land Cover (CLC) inventory depends on the specific version and product. For example, the most recent version of the CLC, CLC2018, is available in a spatial resolution of 100 meters by 100 meters size [34].

To adapt the CORINE inventory to SWB use, the data needs to be reprojected into World Eckert IV projection (ESRI:54012), clipped to the extension of the model area, and resampled to the desired grid dimensions (from the 100 meters grid dimension of CORINE to the 250 meters grid dimension of SWB). All these operations are performed on Q-GIS. In particular, the resampling of the data is performed through a nearest neighbor interpolation method to take into account the discrete nature of the data.

The process of resampling the input data to a coarser grid brought to the attention the question of what the optimal grid is to be used for the model. All simulations have been performed with a regular square grid of 250 meters. This grid is coarser than the landcover input which has a finer size of 100 meters. To study the potential effects of the use of a finer input grid on the SWB model, we considered appropriate to perform one simulation in the historical series applying a finer grid (100 m) derived from the finer land cover raster and compare its results to those of other simulations.

2.2.3. Hydrologic soil group

The hydrologic soil group is a classification system used to categorize soils based on their infiltration and runoff characteristics. This classification was developed by the Natural Resources Conservation Service (NRCS) of United States Department of Agriculture (USDA) to help estimate the runoff potential and erosion susceptibility of soils in a given area. The hydrologic soil group classification takes into account several factors that influence infiltration, including soil texture, structure, and permeability, as well as land use and cover. Soils are classified in four groups (A, B, C, and D equivalent to 1,2,3 and 4 in SWB) based on their ability to infiltrate water [13]. Group A soils have the highest infiltration rates and for the same conditions, could potentially generate more recharge compared to group D soils, which have the lowest infiltration rates. In the same way, group A soils have lower surface flow potential compared to group D soils.

Table 2.3: Hydrologic soil group infiltration rates as described in the SWB manual.

Soil Group	Infiltration Rate [mm/hour]
A	> 7.62
B	3.81 - 7.62
C	1.27 - 3.81
D	< 1.27

The HSG classification is used in a variety of applications, including stormwater management, flood control, and erosion control. In SWB, the information from the hydrologic soil group and the land use of each grid cell of the model domain are combined to define several parameters necessary for the hydrologic simulation.

The first property that depends on the hydrologic soil group is the curve number coefficient related to the CN method described in paragraph 2.1.4. The curve number is an empirical nondimensional coefficient used to predict the runoff response to rainfall excess: when a soil has a high infiltration rate, the curve number coefficient tends to lower values and consequently produces lower runoff quantities for the same precipitation. The second property that is influenced by the value of the hydrologic soil group is the root zone depth of a grid cell, which is the depth of soil that roots do effectively penetrate and extract water and nutrients from. Generally, plants growing in soils with hydrologic groups A or B with high infiltration rates have deeper root zones than if growing in soils of groups C and D [35]. The last element defined by the hydrologic group is the maximum infiltration rate as derived from Table 2.3. As described in paragraph 2.1.6, when preparing a SWB simulation, a daily maximum recharge rate may be specified for each combination of land cover and land use. In the case where daily recharge exceeds the maximum value, all recharge in excess is rejected and transforms to contribute to water runoff.

The U.S. Department of Agriculture, Natural Resources Conservation Service classified more than 14,000 soil series into the different hydrologic soil groups (A–D) based on infiltration capacity. In the case where information on the type of soil of an area is not available, models need to be used to derive the hydrologic soil group class from other properties of the soil itself.

The first step to estimate the hydrologic soil group of a soil is to estimate its saturated hydraulic conductivity. There are several methods to estimate saturated hydraulic conductivity: the most intuitive ones are field or laboratory tests, these tests can provide accurate estimates of soil permeability, but they can be expensive and time-consuming and cannot be applied on a large scale such as the one considered in the model. To overcome these problems, empirical equations are used. Several empirical equations can be employed to estimate soil permeability based on soil properties such

as texture, porosity, and bulk density. These kinds of parameters are available on Soil Grids (https://soilgrids.org/).

SoilGrids is a global soil mapping project that provides digital soil maps and related soil data for the entire world. The project is a collaboration between the International Soil Reference and Information Centre (ISRIC) and the Joint Research Centre (JRC) of the European Commission [36]. The database uses a machine learning approach to predict soil properties such as pH, organic carbon, texture, and bulk density at a high resolution. It provides soil information at different depths: 0-5 cm, 5-15 cm, 15-30 cm, 30-60 cm, 60-100 cm, and 100-200 cm. The predictions are based on a combination of remote sensing data, climate information, and soil profiles from various sources. The data is provided in raster format and for the area of Lombardy the resolution is 250 m. Through this database, it is possible to download data regarding sand, silt, and clay content and bulk density to be employed in empirical equations that estimate saturated soil permeability. Even though empirical equations provide quick estimates of soil permeability, they may not be accurate for all soil types and conditions. For this reason, we employ two methods to estimate the average saturated hydraulic conductivity to define the hydrologic soil group. From the value of saturated hydraulic conductivity, grid cells were separated according to the corresponding hydrologic soil group as in Table 2.3. The empirical methods applied are the Rawls and Brakensiel method and the Rosetta method. Both methods estimate a hydraulic conductivity value for each soil depth range available between 0 m and 2 m and then calculates the mean of these values.

Rawls and Brakensiel

The Rawls and Brakensiel method [13] is the simplest method used to estimate saturated hydraulic conductivity (K_{sat}) and is based on the assumption that soil texture and bulk density are the only factors controlling K_{sat}. The method has the advantage of being simple and requiring only basic soil information, which makes it useful for assessments of soil hydraulic properties. To implement it, thousands of soil horizons from the United States have been analyzed to produce two sets of coefficients which are summarized in the tables shown in the figures below. The first set of tables (Figure 2.9.a.) is used to identify the bulk density class of a soil (high, medium, or low), while the second set of tables (Figure 2.9.b) identifies the saturated hydraulic conductivity class.

To perform the analysis sand percentage and clay percentages are reported on Fig. a1-a3: the bulk density class is identified by checking in which of the three graphs the bulk density of the soil analyzed is coherent with the bulk density displayed in the isolines on the graphs. The bulk density class is then used to choose the corresponding graph from Fig. b.1-b.3: K_{sat} class is found intersecting clay and sand percentage in this graph. For example, if a soil has 35% of sand, 17% of clay and a bulk density of 1.43 g/cm^3. The steps to define its K_{sat} class are the followings:

(d) Intersecting clay and sand percentages in Fig. a.1, a.2 and a.3, the result is that Fig. a.1 is the only one where the point of intersection falls in a range of bulk density isolines (1.46 g/cm^3-1.59 g/cm^3) compatible with the bulk density of the soil (1.43 g/cm^3). Thus, the soil belongs to the high bulk density class.

(e) The high bulk density class corresponds to Fig. b.1. K_{sat} class is found intersecting the two values of sand and clay percentage and looking at the legend present in Fig. b.3: in this example, K_{sat} class is "moderately low".

(a.1) High bulk density group.

(b.1) High bulk density.

(a.2) Medium bulk density group.

(b.2) Medium bulk density.

(a.3) Low bulk density group.   (b.3) Low bulk density.
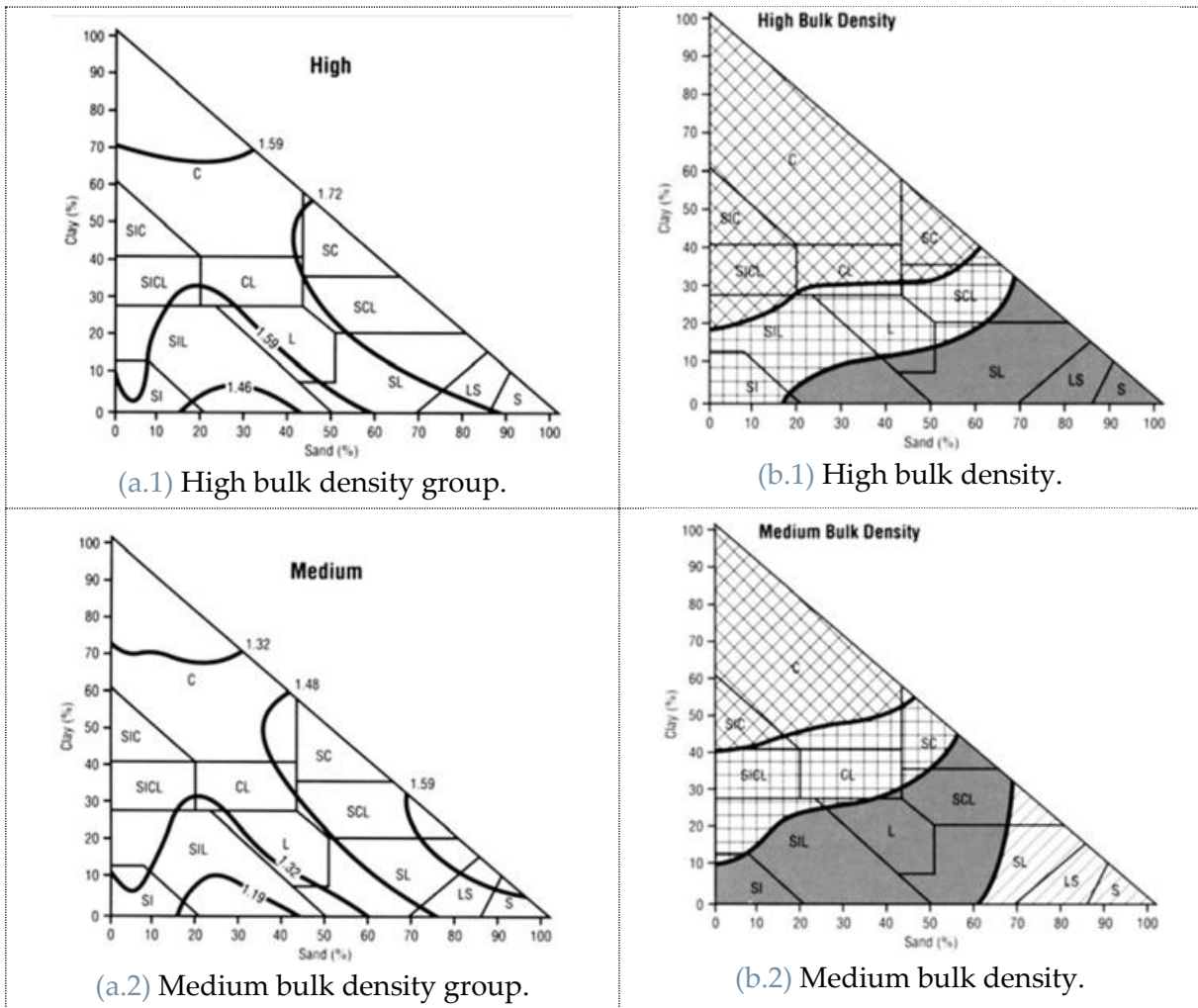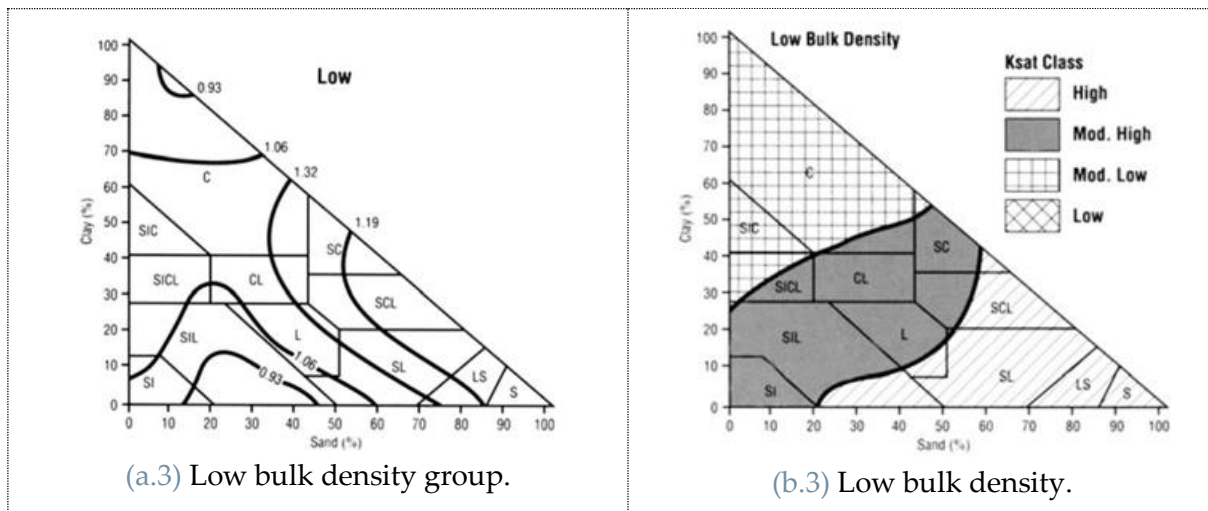
Figure 2.9: Figures representing the coefficients used to apply the Rawls and Brakensiel method [27].

Rosetta

Rosetta is a code to estimate soil hydraulic properties, including water retention and hydraulic conductivity, using pedotransfer functions (PTFs). PTFs are hierarchical, meaning they can estimate soil properties based on limited or more extensive input data, and are based on neural network analyses combined with the bootstrap method to provide uncertainty estimates of the predicted hydraulic parameters. With the application of the Rosetta code, bulk density and soil texture are used as input to estimate a value for K_{sat} [37]. Then from the estimated value of K_{sat}, Table 2.3 is used to set a hydrologic group class to each SWB grid cell. For example, a soil with $K_{sat} = 5\ mm/h$ will fall into hydrologic group B.

Correction techniques

In addition to this, the output of both methods undergoes a few correction steps as suggested by the Soil Survey Manual contained in the USDA Handbook. First, the values of K_{sat} produced by the empirical models for each soil depth (0-5 cm, 5-15 cm,…) are analyzed: a minimum value of K_{sat} occurring for depths between 0.6 and 1 m is considered "a severe limitation as it causes slow water movement" [38] and requires to downgrade the hydrological soil group to a lower infiltration rate class (ex. from HG=A to HG=B).

The output is also influenced by the superficial groundwater level, estimated on the basis of the European Soil Data Centre (ESDAC) and the digital elevation model. If the GW depth is smaller than 1 m, areas are downgraded of one class with respect to the original hydrologic soil group identified in the first place to keep track of the reduced infiltration capacity produced by the shallow water table.

Model comparison

Before commenting, in the next chapter, on the effect that these two empirical methods have on the actual output of the SWB code, it is interesting to study the similarities and

the differences between these two different rasters of hydrologic soil groups produced with the two methods. Both will show fixed values ranging from 1 to 4 to display the hydrologic soil group. As can be seen in Figure 2.11 and Figure 2.10, the models generate significant differences between the outputs. From a quantitative point of view, the histogram of Figure 2.10 shows that on average, the Rawls and Brakensiel method tends to generate an output with values that are more evenly spread on the four hydrologic soil groups, with a peak frequency of around 45% for the hydrologic soil group 2 (which translates to B as 1 to 4 is equivalent to group A to D). On the contrary, the Rosetta method tends to generate values that are less distributed over the four hydrologic groups: its peak frequency also occurs at hydrologic soil group 2 but with an occurrence of 60%. The effect of the peak frequency in terms of spatial distribution is explained by looking at Figure 2.10: for both models, the hydrologic soil group in mountain areas tends to be lower and prevalently equal to a value of 2, while in flat areas hydrologic soil group appears to be more variant and the output of the two models becomes more different.

(a) Histogram of hydrologic soil group in Lombardy.
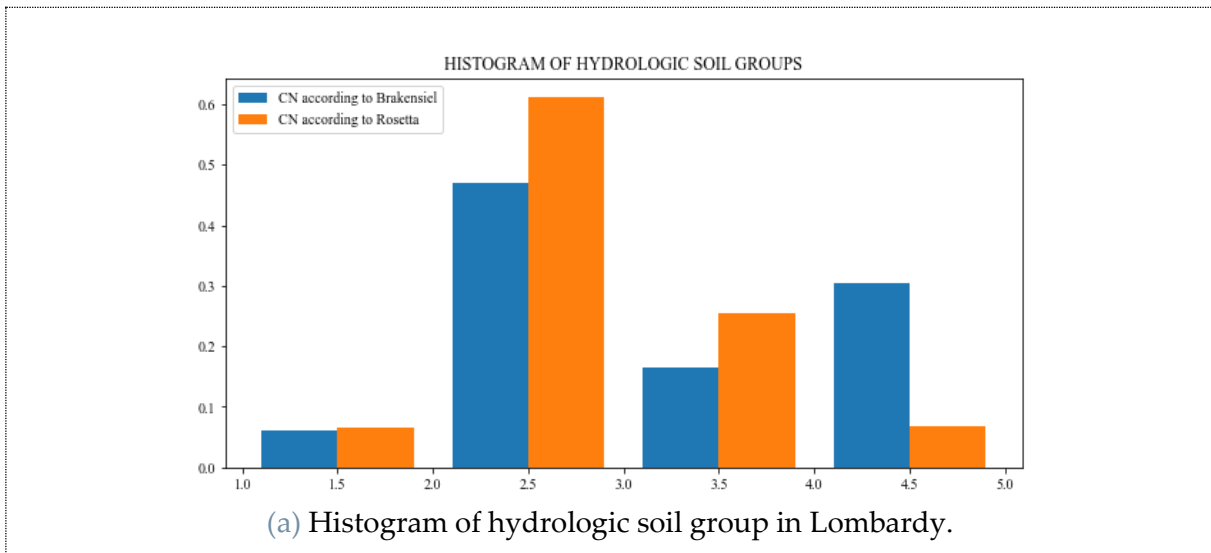
Figure 2.10: Histogram of hydrologic soil groups produced by two different empirical methods.

(a) Rawls and Brakensiel.

(b) Rosetta.
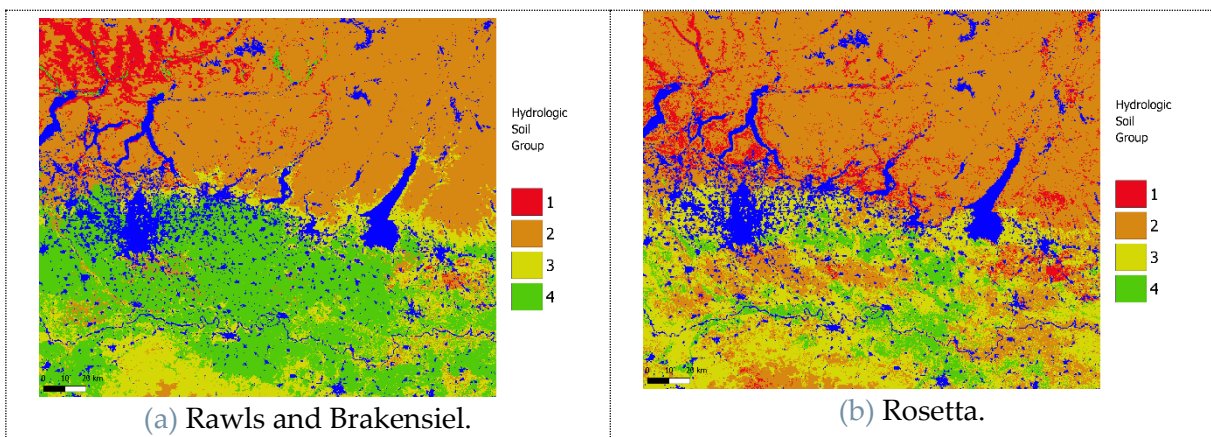
Figure 2.11: Hydrologic soil group maps according to the two methods; the blue color represents impervious zones as urban areas or water bodies.

A careful analysis of the spatial distribution of hydrologic soil groups leads to generating two different histograms: using a digital elevation model, the geographic domain is split between mountain areas above 600 m a.m.s.l. and plain areas below 600 m a.m.s.l., two different histograms are produced and displayed in Figure 2.12 for mountain and flat-land areas.

(b) Histogram of hydrologic soil group in mountain areas (above 600 m.s.l.m) in Lombardy.

(c) Histogram of hydrologic soil group in flat areas (below 600 m.s.l.m) in Lombardy.
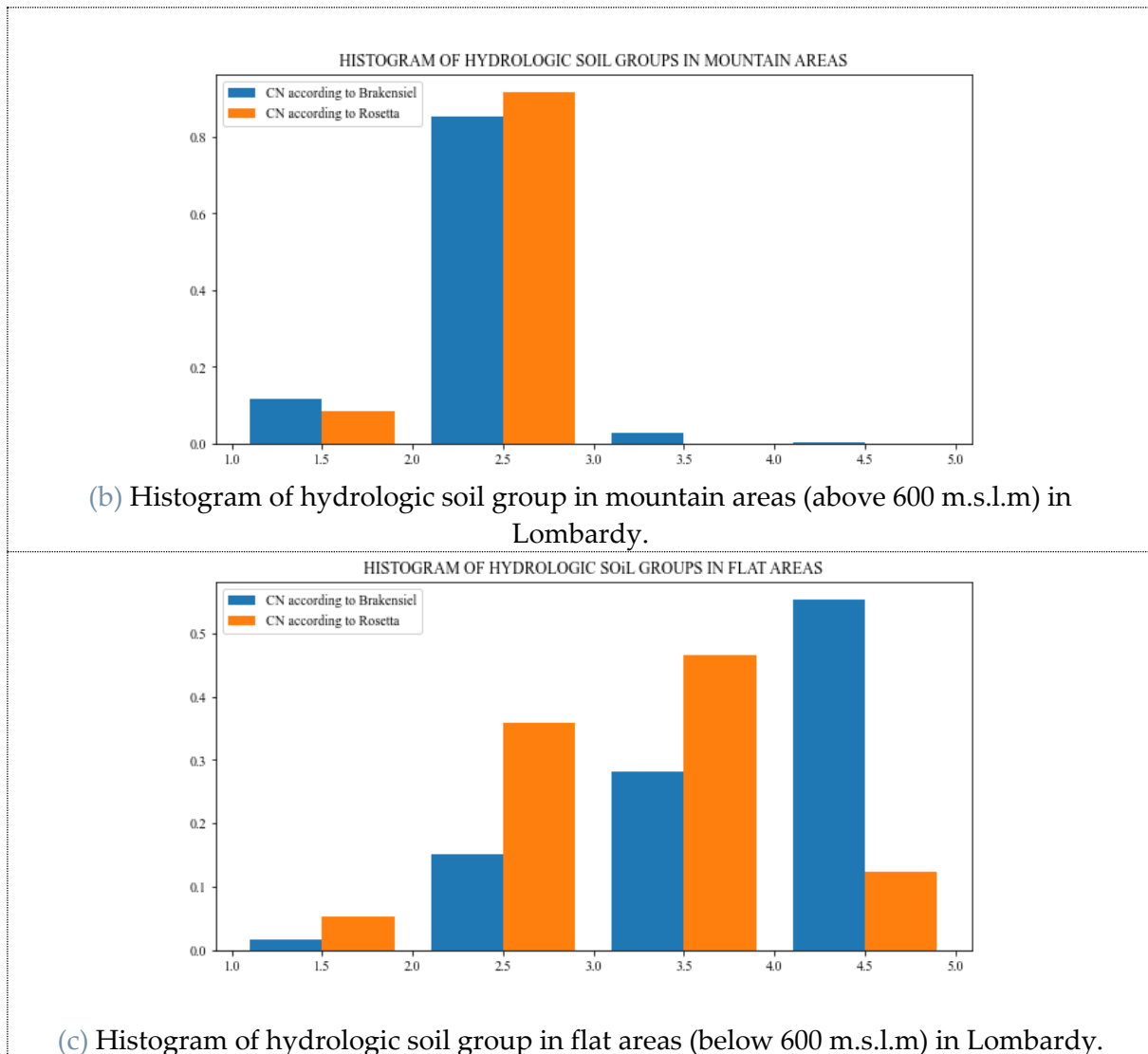
Figure 2.12: Histogram of hydrologic soil groups split between mountain and flat areas.

For what concerns mountain areas: the two methods generate an almost equal and homogeneous output, with around 85% of the land belonging to hydrologic soil group 2. The convergence of the output of both two methods on hydrologic group 2 could be attributed to the low bulk density area (high infiltration capacity) that is observed all across the mountain region.

In the flat area of the Pianura Padana, values are distributed over all four hydrologic groups. Rawls and Brakensiel method displays a peak frequency of around 55% of

cells having hydrologic soil group 4 (equivalent to the lowest soil permeability class) while Rosetta method displays a peak frequency of 45% cells having hydrologic soil group 3.

Summarizing this analysis, it is possible to conclude that the Rosetta method tends to generate a raster with lower values (higher infiltration capacity) of hydrologic soil groups with respect to Rawls and Brakensiel.

2.2.4. Available water capacity

Available Water Capacity (AWC) refers to the amount of water that can be stored in the soil and be available to plants for uptake [13]. AWC represents the difference between the maximum quantity of water that can be accumulated and kept in the soil known as field capacity, and the minimum quantity of water in the soil that guarantees the survival of the plants known as wilting point:

$$AWC = \theta_{FC} \text{-} \theta_{WP} \qquad \left[\frac{cm \ of \ water}{m \ of \ soil \ thickness}\right] \qquad (2.9)$$

Where θ_{FC} represents field capacity [cm/m] and θ_{WP} is the wilting point [cm/m]. AWC represents the total centimeters of water holding capacity per meter of soil thickness. In SWB, when the AWC capacity value of a grid cell is multiplied by the value of the root-zone depth of that specific cell, the product obtained is the maximum water storage available in that cell. In general soils with a higher AWC can support a greater volume of water and therefore plant growth, while soils with a lower AWC may require more frequent irrigation or rainfall to support plant growth. Soil texture, compaction, and organic matter content affect AWC, and it can vary widely between different types of soils.

As previously described for the definition of hydrologic soil group, the definition of AWC can take place either through laboratory samples or through empirical methods. In the case of this research, the possibility of using laboratory samplings must be discarded and AWC needs to be defined only through empirical methods. Two models have been identified in the literature to estimate the value of AWC: Thornthwaite and Mather's procedure and Saxton and Rawls's equations.

Thornthwaite and Mather procedure

This method is the only one suggested in the SWB manual as it is based on the original Thornthwaite and Mather (1957) [35] instructions for computing water balance. Soils are classified into different soil texture classes using a soil texture triangle, as shown in Figure 2.13, similar to the one used in the Rawls and Brakensiel method to estimate the hydrologic group. As in the Rawls and Brakensiel method, a set of sampled soils from the United States was used to associate an estimated value of AWC to each soil texture class. However, in this method, soils are classified solely based on their soil texture and not on other variables such as bulk density.
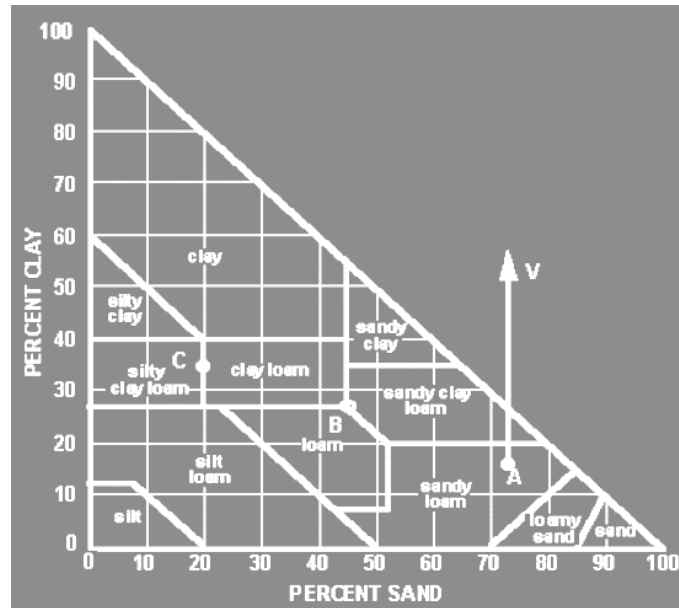
Figure 2.13: Table proposed by Thornthwaite
and Mather method to estimate soil class [27].

AWC is directly defined by a table that assigns to each soil class a value of AWC, as shown below:

Table 2.4: AWC for different soil textures [13].

SOIL TEXTURE	AWC [cm/m]
Sand	10,0
Loamy sand	11,7
Sandy loam	13,3
Fine sandy loam	15,0
Very fine sandy loam	16,7
Loam	18,3
Silt loam	20,0
Silt	21,3
Sandy clay loam	22,5
Silty clay loam	23,8
Clay loam	25,0
Sandy clay	26,7
Silty clay	28,3
Clay	30,0

Saxton and Rawls

The Saxton & Rawls (2006) [39]method defined a set of equations that directly link soil texture and organic matter content of the soil to its wilting point and field capacity. As for the Rawls and Brakensiel method, this method has been calibrated through a statistical average on a broad range of soils provided by the United States Department of Agriculture (USDA) soil database.

Its equations calculate the volumetric soil water content at field capacity and wilting point using the soil texture and the organic matter content. It estimates the AWC as the difference between the volumetric soil water content at field capacity and wilting point.

The following equations are used to estimate wilting point θ_{WP} and field capacity θ_{FC}:

$$\theta_{WP} = \theta_{1500t} + (0.14\,\theta_{1500t} - 0.02) \qquad [\%V] \qquad (2.10)$$

$$\theta_{1500t} = -0.024\,S + 0.487\,C + 0.006\,OM + 0.005(S \cdot OM) \qquad [\%V] \qquad (2.11)$$
$$- 0.013\,(C \cdot OM) + 0.068\,(S \cdot C) + 0.031$$

$$\theta_{FC} = \theta_{33t} + (1.283\,\theta_{33t}^2 - 0.374\,\theta_{33t} - 0.015) \qquad [\%V] \qquad (2.12)$$

$$\theta_{33t} = -0.251\,S + 0.195\,C + 0.011\,OM + 0.006\,(S \cdot OM) - (\qquad [\%V] \qquad (2.13)$$
$$0.452\,(S \cdot C) + 0.299$$

Where:

- θ_{1500t} [% of volume] is the lower limit of water availability for plants that takes place for a suction head pressure of 1500 kPa

- θ_{33t} [% of volume] is the upper limit of water availability for plants that takes place for a suction head pressure of 33 kPa

- S represents the sand content of the soil [% of mass]

- C represents the clay content of the soil [% of mass]

- OM represents the organic matter content of the soil [% of mass]

Sand content and clay content are known from SoilGrids. No information is available through SoilGrid for what concerns the organic matter content of the soil, but data on organic carbon content can be downloaded.

Being soil organic matter (OM) primarily made of organic content (OC), a linear relationship exists between these two variables as [39]:

$$organic\ matter = constant \cdot organic\ carbon \qquad (2.14)$$

The van Bemmelen conversion factor coefficient assumes that 58% of OM is made up of OC [40], so that a constant factor of 1.724 is employed in Eq. (2.14). The accuracy of this coefficient has been extensively analyzed. For example, the research of Heaton et al. [40] on 278 samples of loamy sand soils in the UK determined experimentally the distribution of the conversion coefficient between OM and OC: conversion factors ranged between 0.36 and 0.98 with a mean value of 0.66 and a STD=0.55. Therefore, in the absence of further data, it is acceptable to employ the van Bemmelen theory: considering an average content of 58% of OC/OM and using its reciprocal 1.724 in equation (2.14), it is therefore possible to estimate AWC for each grid cell in the domain.

Model comparison

The two models show fundamental differences both in the values and in the spatial distribution of AWC values.

(a) Thornthwaite and Mather. (b) Saxton and Rawls.
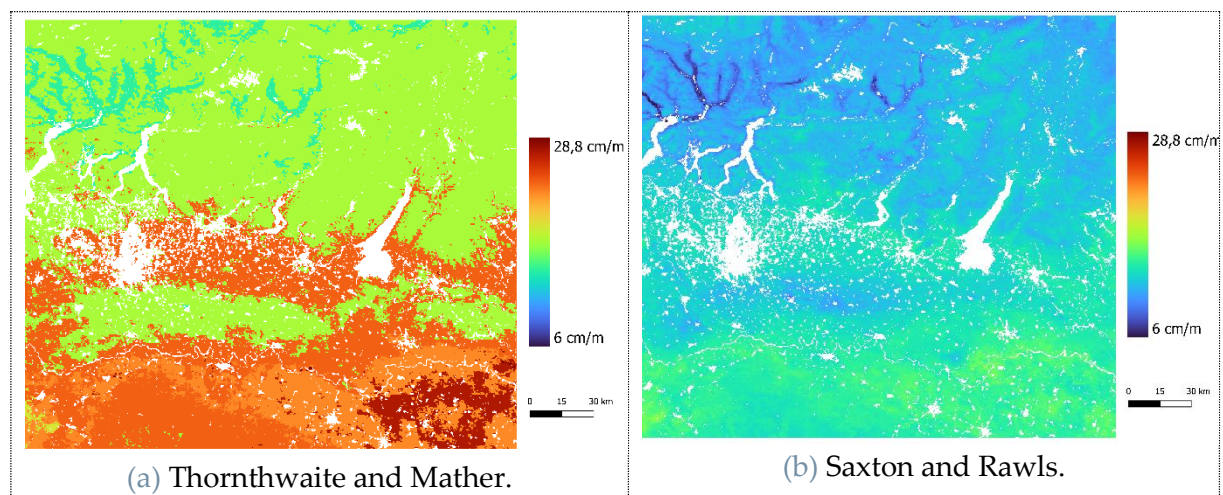
Figure 2.14: Analysis of AWC outputs; white zones of the images are urban areas or water bodies where no information are available.

The Saxton and Rawls method produces values of AWC that are significantly lower with respect to the Thornthwaite and Mather method. Lower values of AWC in the Saxton and Rawls generated input will have the effect of reducing the capacity of water accumulation in the soil. This naturally lowers the threshold of precipitation needed to produce infiltration in the soils, thus has the effect of increasing potential recharge quantities.

2.2.5. Flow direction

SWB uses flow direction as a hierarchy to define a flow routing process able to transport the calculated precipitation runoff from a cell into precipitation run-on of the

next downstream cell. If the flow routing process option is activated in SWB, as in the case of these simulations, all runoff from a cell is allowed to infiltrate in downslope cells on the same day in which it originated as rainfall or snowmelt. At the end of each day, all runoff quantities are either streamed outside of the model domain or eliminated. No stream calculations are performed from one day to the other. The flow direction grid is generated from a Digital Elevation Model, applying a D-8 algorithm. Since this data derives from elevation data that respect a high level of accuracy, it was not considered necessary to perform any modification on this type of input.

3 Results

The following images describe the development of the estimates of mean annual potential recharge for historical and future climate over the 20 years of simulation. As stated in the caption, historical simulations, RCP4.5 simulations and RCP8.5 simulations are examined both in terms of mean annual recharge as well as in terms of mean annual recharge to precipitation ratios. A summary of acronyms used to indicate the different simulations in the legend of the following graphs and table can be found in the List of symbols at the end of this thesis.

(b)
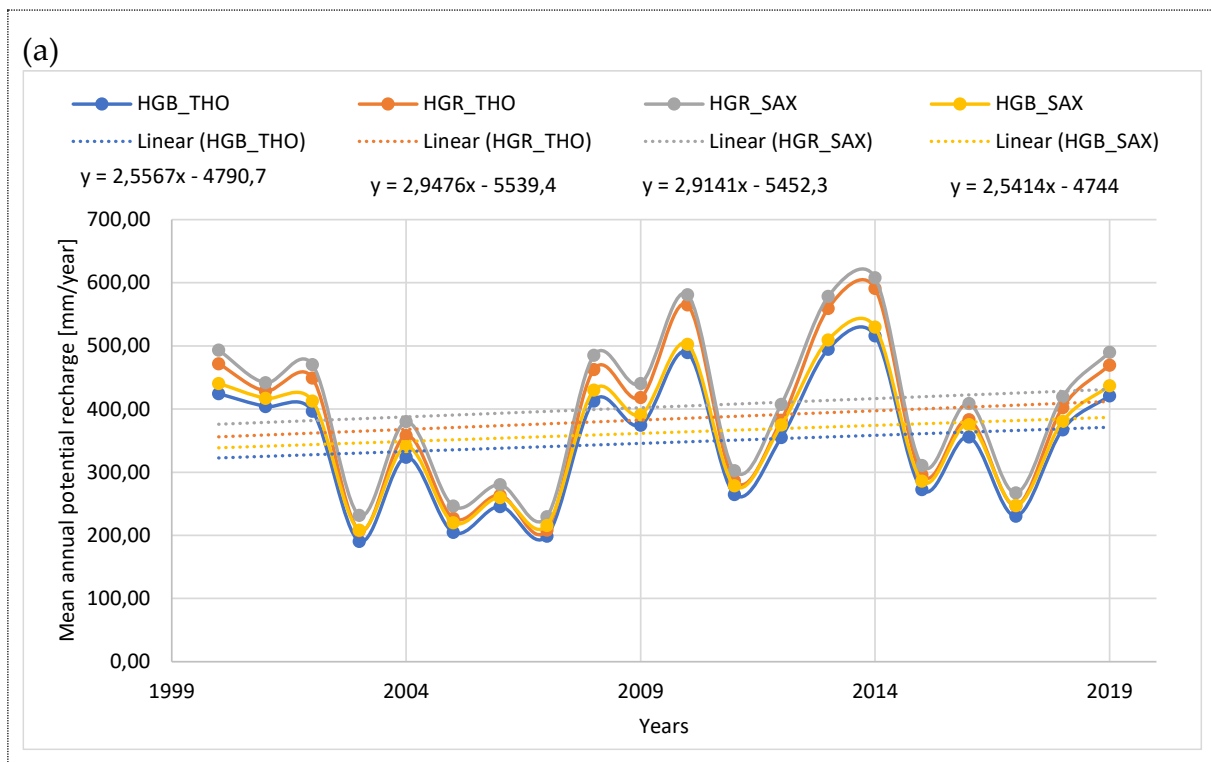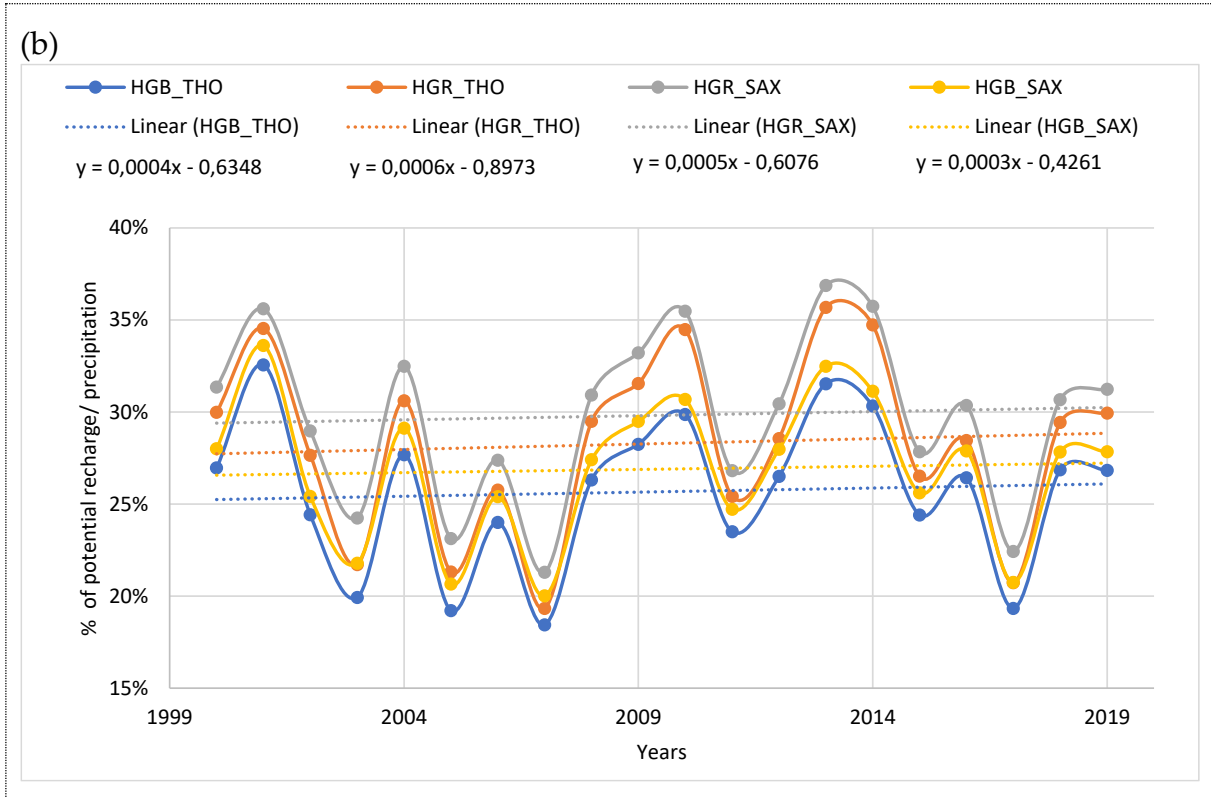
Figure 3.1: Development of mean annual potential recharge estimate of historical models. Fig. (a) represents the development of mean annual potential recharge estimates, Fig. (b) represents the mean annual fraction of potential recharge to total precipitation; Equations in the legend represent the linear interpolations of data from each model.

(a)
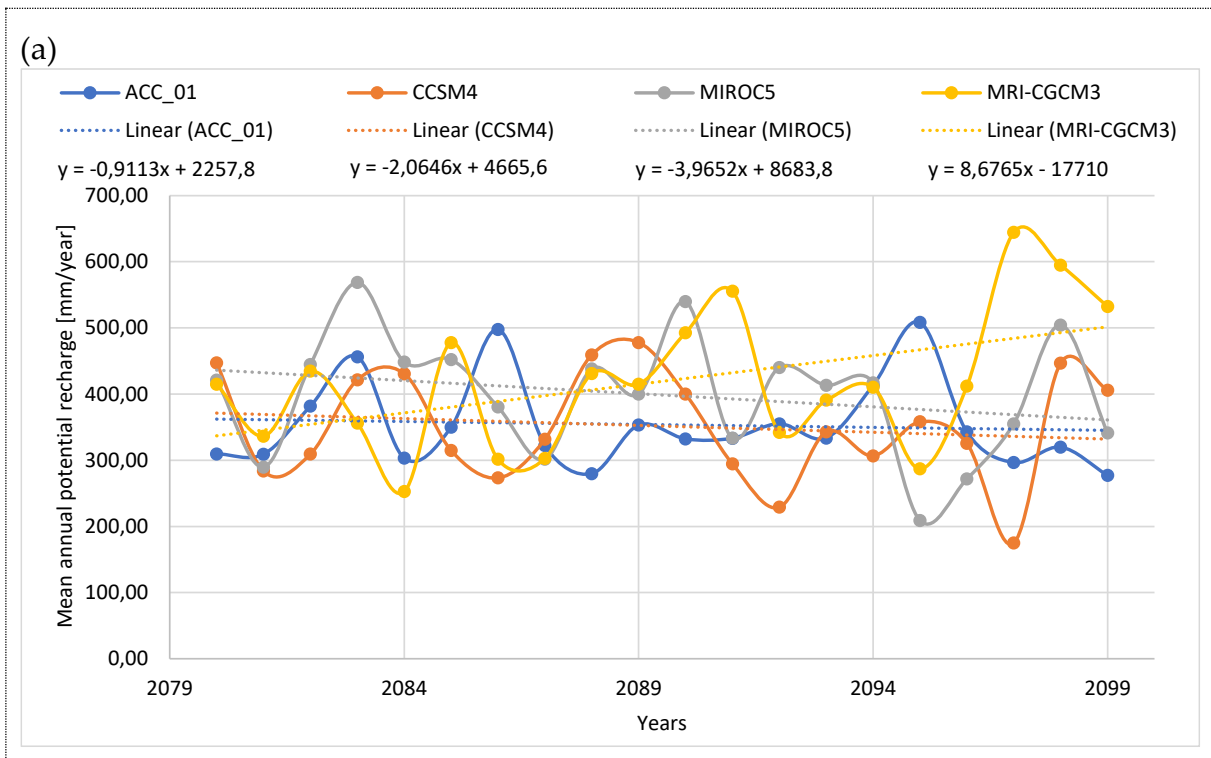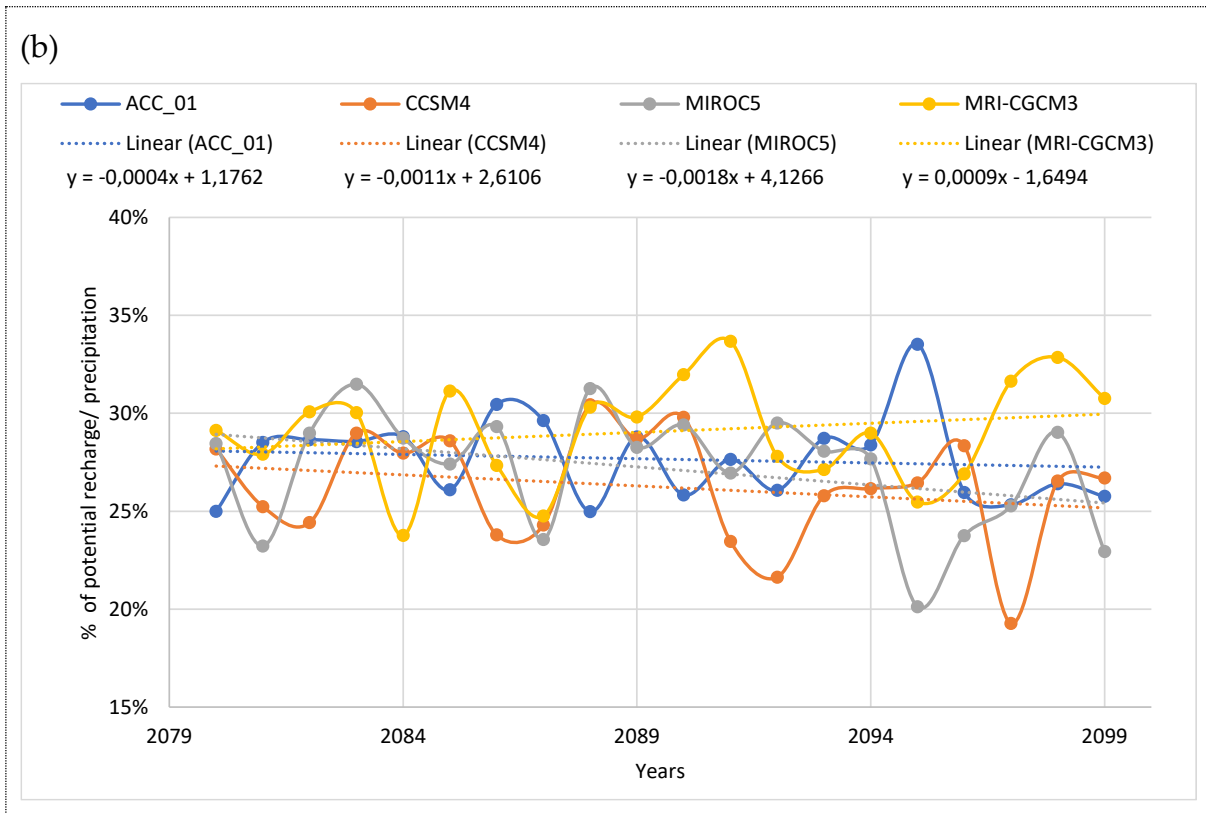
Figure 3.2: Development of mean annual potential recharge estimate of RCP4.5 climate models. Fig. (a) represents the development of mean annual potential recharge estimates, Fig. (b) represents the mean annual fraction of potential recharge to total precipitation; Equations in the legend represent the linear interpolations of data from each model.
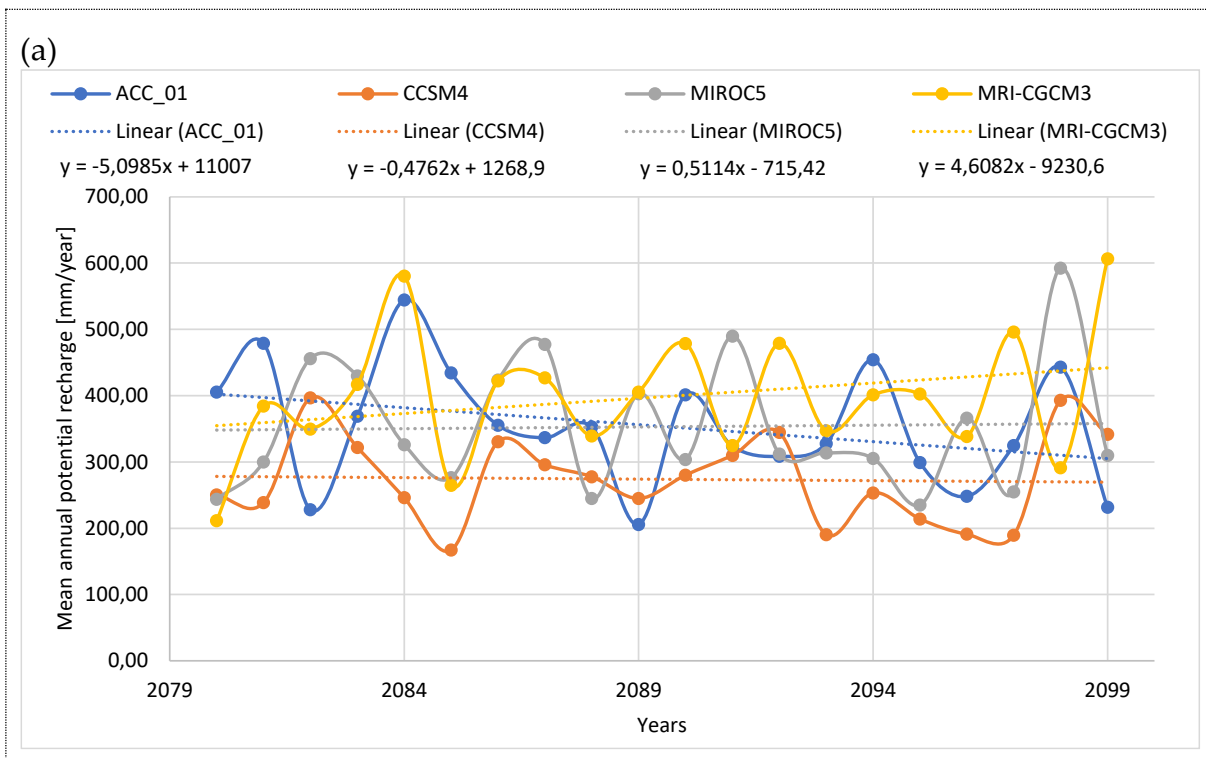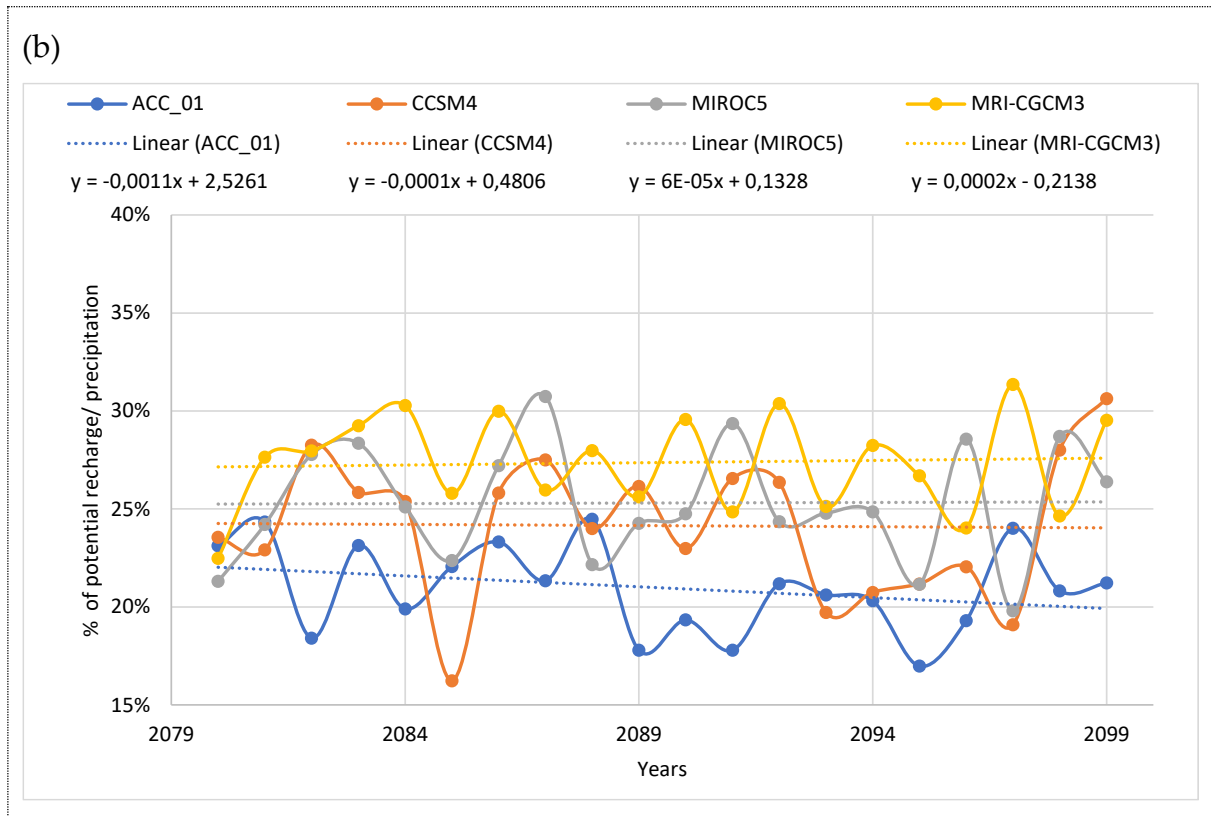
Figure 3.3: Development of mean annual potential recharge estimate of RCP8.5 climate models. Fig. (a) represents the development of mean annual potential recharge estimates, Fig. (b) represents the mean annual fraction of potential recharge to total precipitation; Equations in the legend represent the linear interpolations of data from each model.

Figure 3.1, Figure 3.2, and Figure 3.3 show that the ranges of variation of the estimates of mean annual potential recharge are similar for historical simulations (between 190 mm and 608 mm), RCP4.5 simulations (between 175 mm and 644 mm), and RCP8.5 simulations (between 167 mm and 606 mm). Ratios between mean annual potential recharge and mean annual precipitation (%R) vary between 15% and 35%: RCP8.5 simulations produce low values of %R (between 0,16% and 0,31%), the historical simulations produce high values of %R (between 0,18% and 0,37%) and the RCP4.5 simulations results fit in the middle of this trend with values of %R ranging between 0,17% and 0,34%. From these findings, it can be inferred that the recharge mechanism's efficiency in relation to input precipitation will, on an annual average, decline as the intensity of climate change effects on the environment increases.

Additionally, the development of the estimates throughout the duration of the 20 years of simulation shows that %R estimations are far more stationary, as can be seen from the slope of the interpolating lines, whereas estimates of mean annual recharge exhibit tendencies of positive or negative increase during the 20 years of simulation. For example, Figure 3.1 shows an average trend of potential recharge increasing at a rate of 2.5 mm/year.

Development of mean annual potential recharge gives an overview of the variation of recharge estimates from year to year. After identifying the years with the highest and lowest mean annual recharge, the images below display the corresponding estimates of potential recharge to assess the impact of the variation of potential recharge from years to years. For the historical series, the minimum and maximum mean annual potential recharge are produced respectively by the 2003 HGB_THO simulation (190 mm/year) and by the 2014 HGR_SAX (608 mm/year) simulation, displayed in Figure 3.4.
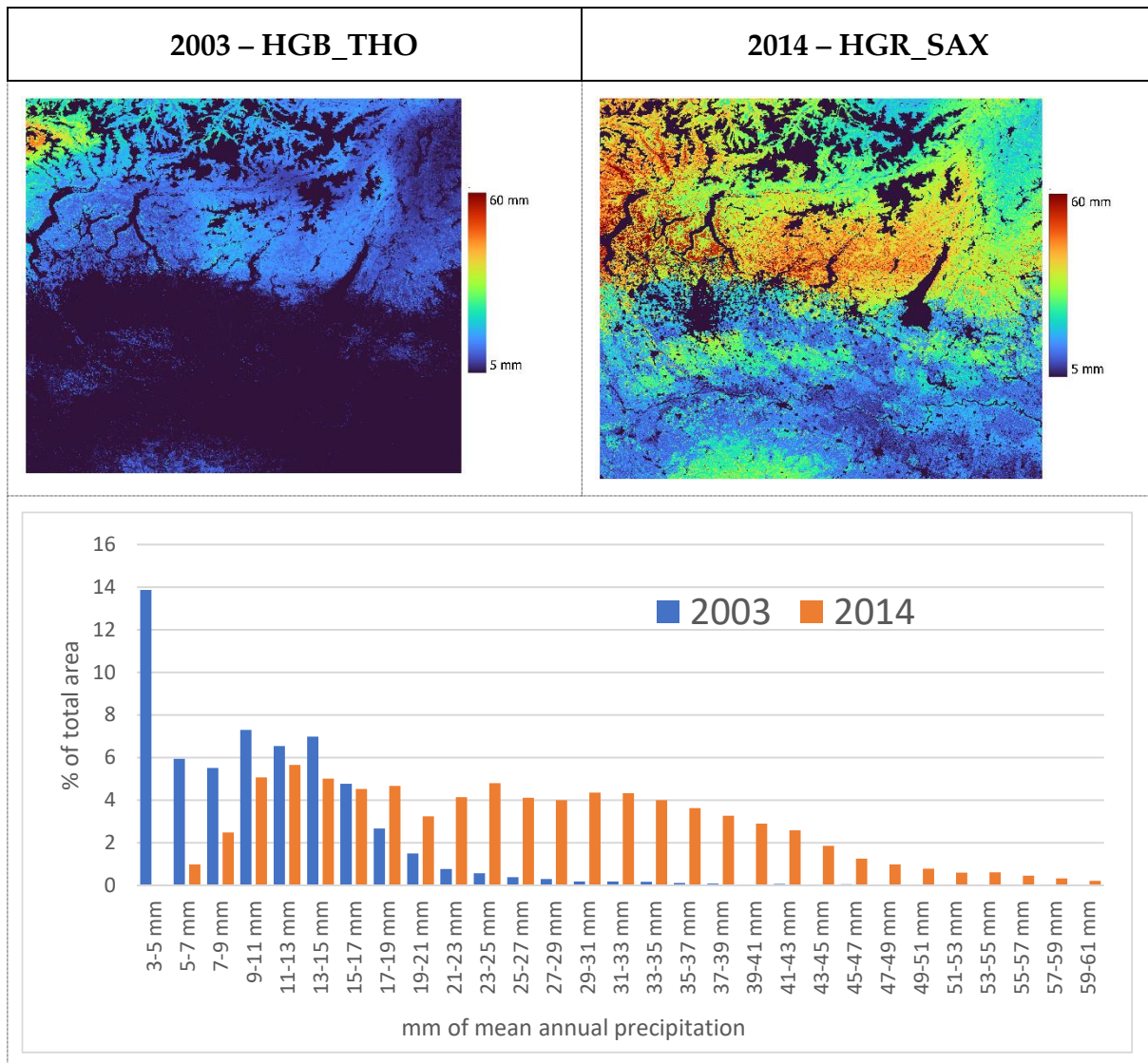
Figure 3.4: Estimates of years of maximum and minimum mean annual potential recharge for historical simulations.

Figure 3.4 shows that recharge conditions can change significantly from one year to the other. Changes in recharge conditions are heavily influenced by the changes in precipitation trends: 2003 has the lowest mean annual potential recharge estimate because it is also the driest year in terms of precipitation, as shown in the analysis of

Figure 2.2. The opposite applies for 2014, being the wettest year in terms of precipitation.

Conducting a similar analysis for the RCP4.5 and RCP8.5 scenario we selected a year of analysis with the criteria of having a high variation in terms of mean annual potential recharge between the different climate models applied. For RCP4.5 the year selected is 2097, where CCSM4 produces the lowest mean annual recharge estimate of the twenty years analyzed (145 mm/year), while the MRI-CGCM3 produces the highest mean annual recharge estimate of the twenty years analyzed (644 mm/year). Figure 3.5 displays all the outputs of annual potential recharge produced by the different climate models for the year 2097.
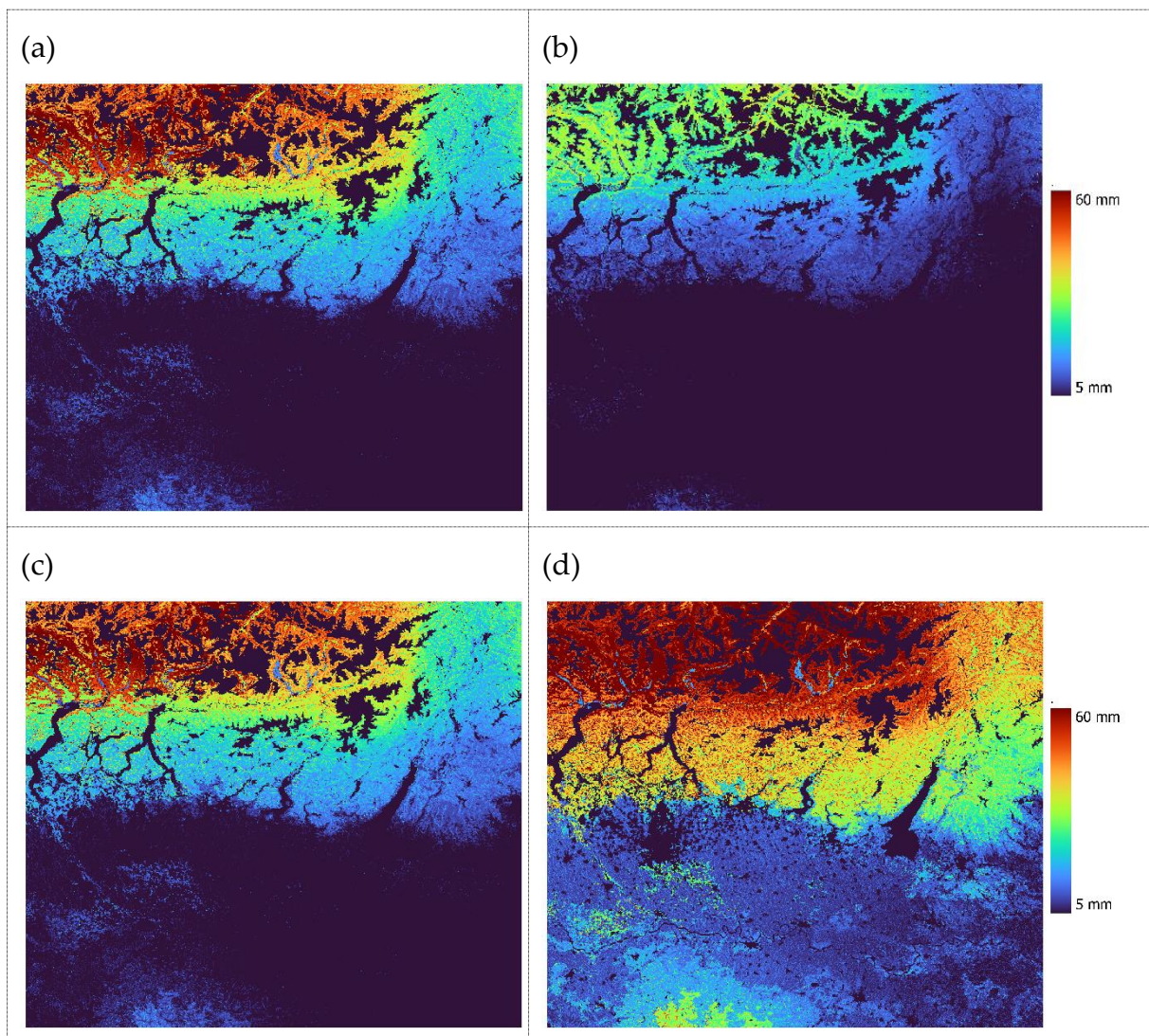
Figure 3.5: Estimates of mean annual recharge for the year 2097 by the four RCP4.5 climate models: (a) ACCESS1-0, (b) CCSM4, (c) MIROC5, (d) MRI-CGCM3.

For RCP8.5 the year selected is 2094, where CCSM4 is the GCM producing the lowest potential recharge estimate (253 mm/year), while ACCESS1-0 is the GCM producing the highest mean potential recharge (454 mm/year).
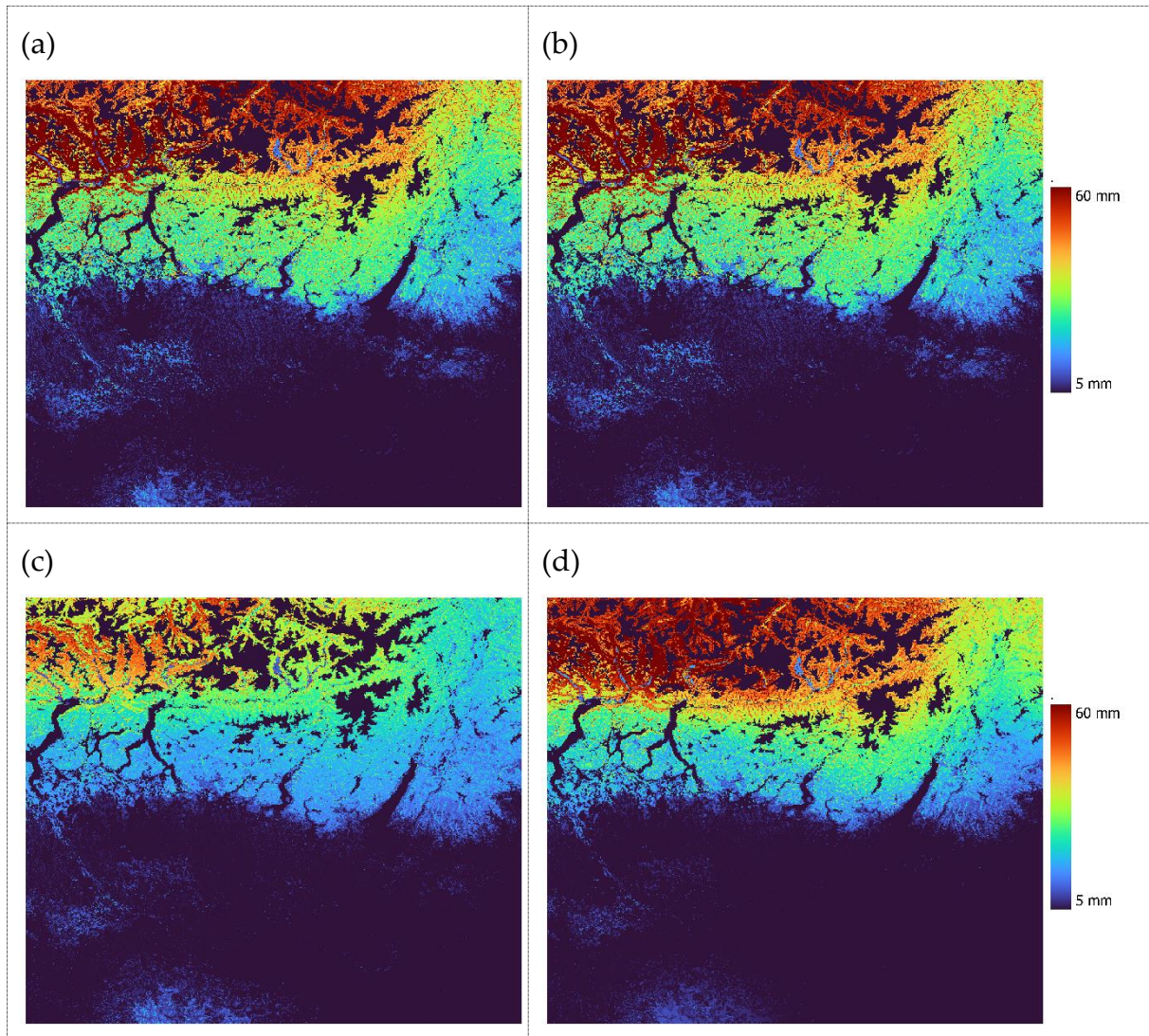
Figure 3.6: Estimates of mean annual recharge for the year 2094 by the four RCP8.5 climate models: (a) ACCESS1-0, (b) CCSM4, (c) MIROC5, (d) MRI-CGCM3.

Figure 3.5 and Figure 3.6 show some of the difference in the estimates of mean annual potential recharge produced by different climate models for a single year of simulation. To average the effects of the variation of climate from one year to the other, and to examine and compare the long-term trends of climate variables, we analyzed the effects of climate change by calculating the average of each variable over the twenty years of simulation. Table 3.1 summarizes the results of mean yearly net infiltration averaged over the 20 years of each simulation.

Table 3.1: Analysis of yearly-average cumulated net infiltration results. Cells are colored based on the deviation from the reference simulation (yellow representing a decrease and blue an increase).

SCENARIO	MODEL	AVERAGE [mm/year]	PERCENTAGE OF VARIATION OF NET INFILTRATION ESTIMATE FROM REFERENCE SCENARIO (HGB_THO)	PERCENTAGE OF VARIATION OF MEAN ANNUAL PRECIPITATION FROM HISTORICAL SERIES	% RECHARGE/ PRECIPITATION
HISTORICAL	HGB_THO	347,1	0,0	0,0	26,1
	HGR_THO	383,9	10,6	0,0	28,9
	HGB_SAX	362,9	4,6	0,0	27,3
	HGR_SAX	403,6	16,3	0,0	30,4
RCP4.5	ACCESS-01	353,7	1,9	-4,2	27,8
	CCSM4	351,7	1,3	-0,2	26,5
	MIROC5	398,5	14,8	9,2	27,5
	MRI-CGCM3	419,2	20,8	7,2	29,4
RCP8.5	ACCESS-01	353,6	1,9	26,9	21,0
	CCSM4	273,9	-21,1	-15,5	24,4
	MIROC5	353,1	1,8	3,6	25,6
	MRI-CGCM3	398,3	14,8	8,4	27,6

The historical simulation performed with input generation as described by SWB is the HGB_THO. HGB_THO finds that the average net infiltration rate is 347,1 mm per year, resulting in 26% of rainfall. These values are coherent with other similar studies conducted in humid regions that registered ratios of recharge to precipitation between 21% and 34% [41].
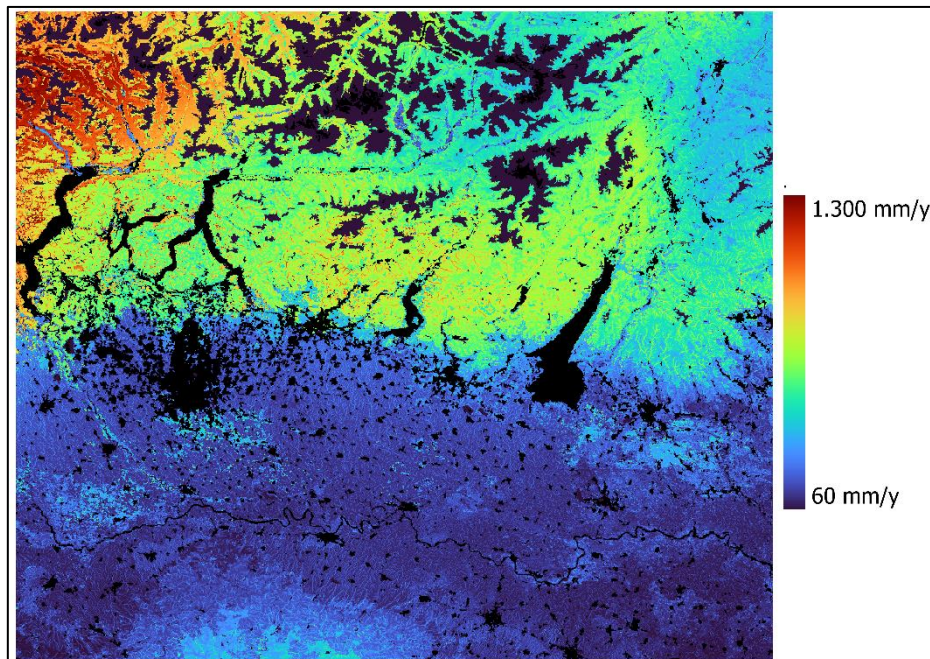
Figure 3.7: Average yearly recharge from 2000 to 2019 for reference HGB_THO simulation; white areas correspond to water bodies or urban areas where SWB calculations are not performed.
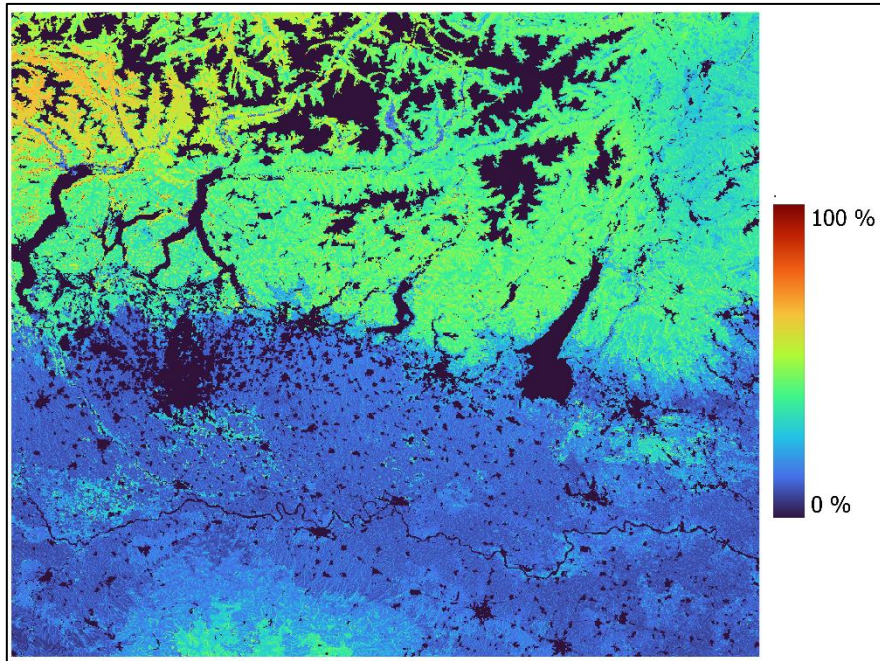
Figure 3.8: Percentage of recharge over total precipitation averaged between 2000 and 2019.

Figure 3.7 offers a point of reference to examine the trends in the spatial distribution of recharge estimates performed with SWB. The distribution of values of potential recharge estimates can ideally split the domain horizontally into two parts. In the southern part of the region, coinciding with the low Pianura Padana, estimates of potential recharge are low and tend to be around 100-200 mm/year. In the northern part, corresponding with the mountain areas, estimates of potential recharge are more variable in space and undoubtedly higher with values ranging from 600 mm/year to 1500 mm/year. The nature of this difference is to be found in multiple elements that are influencing the estimates: the first being the underlying definition of the hydrologic soil group of each cell in the model. If Figure 3.7 is compared with Figure 2.11, it is possible to see the strong spatial correlation that exists between the hydrologic soil group and the calculated net infiltration. Another element is the meteorological inputs that favor higher recharge estimates in mountain areas: there is a spatial correlation between areas of high precipitation in Figure 2.4 to areas of high potential recharge estimates. In addition, lower temperatures in the mountain areas are also contributing to a reduction of evapotranspiration in those areas, thus increasing net infiltration.

To have a finer understanding of the distribution of recharge estimates over the area, the histogram in Figure 3.9 analyzes the percentage of the total area that fits within the defined classes of net infiltration. The modal value is between 100 mm/year and 150 mm/year and the median is very close to 250 mm/year, but the histogram shows another peak value for recharge estimates between 550-600 mm/year. The two peaks are explained by looking at how values are split between mountain area estimates (elevation higher than 600 m a.m.s.l) and areas along the Po' valley (flat areas). The flat

areas confirm a modal value of recharge estimates of 100-150 mm/year, while the mountain areas histogram shows a modal value of 550-600 mm/year. The mountain areas are characterized by recharge values more smoothly distributed around the peak compared to the flat area, coherently with the greater heterogeneity of precipitation and temperature data in this zones that promotes a wider variety of recharge patterns. At the same time, mountain areas show the greatest frequency of zero recharge estimates: this effect is due to snow coverage and rocky surfaces at high altitudes as can be seen from Figure 3.7.
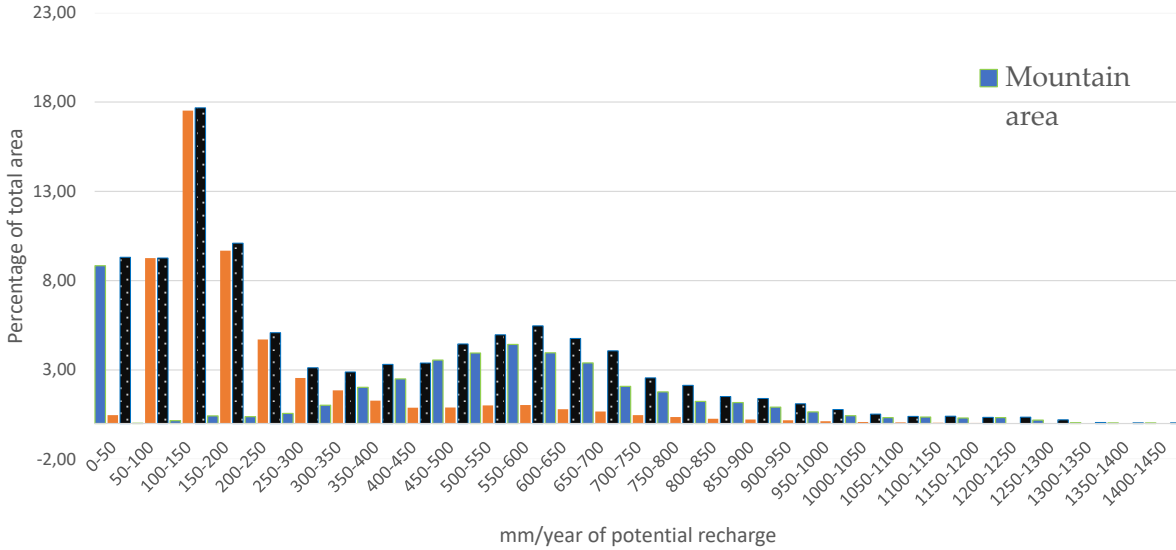
Figure 3.9: Histogram of distribution of potential recharge estimates.

The output described from the HGB_THO and analyzed in the previous paragraphs becomes essential to analyze and compare in the following sections the outputs of all other simulations summarized in Table 3.1. In several analysis, the output of an *i*-model from Table 3.1 will be compared to the output of the reference model by calculating the percentual relative variation as illustrated below:

$$\% \, relative \, variation = \frac{Z_{i-model} - Z_{i-ref}}{Z_{i-ref}} * 100 \tag{3.1}$$

Where:

- $Z_{i-model}$ is the quantity estimated for a target output by the model analyzed in each grid cell *i* of the domain.
- Z_{i-ref} is the quantity estimated for a target output by the reference model HGB_THO in each grid cell *i* of the domain.

From the analysis of model results, the effect of a variation of the discretization of the land cover input (from a grid size of 250 m to a grid size of 100 m), as described in paragraph 2.2.2, had been ignored as it was producing undetectable signs of variation in the distributions of potential recharge estimates with respect to the reference model.

The remaining four simulations on historical series (including the reference HGB_THO) are all based on the same input meteorological data, so their variation from the reference model only depends on the gridded input variation described in the previous chapter and relative to the definition of the hydrologic group and the available water capacity.

All models produce similar estimates of mean annual potential recharge. For the historical models, the reference simulation HGB_THO produces the smallest mean annual recharge out of all simulations and equal to 347.1 mm/year. The historical simulation HGR_SAX instead produces the greatest mean annual recharge value out of all the historical simulations and equal to 403.6 mm/year (16% more than the reference simulations HGB_THO). The remaining two historical models produce smaller increases in mean annual recharge but still show a significant variation from the reference HGB_THO as can be seen from Table 3.1. Between the two historical models producing the maximum and the minimum potential recharge (HGR_SAX, HGB_THO), only 4,3% of the input precipitation (equal for both models) that is infiltrated by the HGR_SAX is lost by the HGB_THO model. Paragraph 2.2.3 presented a comparison between the two different gridded inputs for hydrologic soil groups (Brakensiel and Rosetta) and showed that since the Rosetta method overestimated the infiltration capacity of the soil with respect to Brakensiel, it could have been expected to observe higher net infiltration values. A similar reasoning is conducted for the AWC definition: because the Saxton and Rawls method predicts on average lower values of AWC with respect to the Thornthwaite method, the Saxton and Rawls method offers on average less soil-moisture-reservoir and thus produces more water infiltration during rain events. The outputs of SWB's simulations confirm this hypothesis. In addition, varying only one of the two input data shows that a variation in the method to define the hydrologic soil group has a greater effect on the average net infiltration (10.6% variation) than a variation in the method to define AWC does (4.6%).

The simulations on GCMs have all been performed on gridded inputs from the HGB_THO model. The moderate RCP4.5 scenario all predict an increase in the average yearly net infiltration rates but show two different behaviors concerning the deviation of net infiltration estimates from the reference historical value. The first two models (ACCESS-01 and CCSM4) forecast a very small increase in the relative difference (respectively 1.9%, and 1.3%) despite an opposite trend of decreasing total yearly precipitation (respectively -4.2%, and -0.2%) shown in Table 3.1. The other two models (MIROC5 and MRI-CGCM3) show significantly higher deviation from the reference net infiltration estimate (respectively 15%, and 20%) that can be directly related to the increase in the estimated precipitation. In synthesis, the variations in mean net infiltration estimates produced by the RCP4.5 simulations with respect to the reference HGB_THO is comparable to the variations produced by the historical model simulations.

Considering the extreme RCP8.5 scenario, the CCSM4 model is the only one projecting a decrease (−21.1%) of yearly potential recharge with respect to the reference simulation. This reduction can be directly related to the reduction observed in the forecasted yearly precipitation of the same model (−15% of the historical series). Hypothetically, potential recharge quantities are always proportional to total precipitation quantities. However, the ACCESS-01 model contradicts this concept: its forecasted average precipitation is 26,9% higher than the reference simulation but the resulting estimates of potential recharge are only 1.9% higher than the reference simulation. Results from the two remaining models (CMIRO5 and MRI-CGCM3) are less conflicting in this sense but it's worth noting that due to precipitation quantities in the different emissions scenarios, both models produce recharge estimates that for the RCP8.5 scenario are closer to the reference historical series than for the RCP4.5 scenario.

The images in the following pages analyze the distribution of the recharge estimates and the relative percentual variation with respect to the reference model.

HGB_SAX

(1.a)

(1.b)

(1.c)

(1.d)

HGR_THO

(2.a)

(2.b)

(2.c)

(2.d)

HGR_SAX

(3.a)

(3.b)
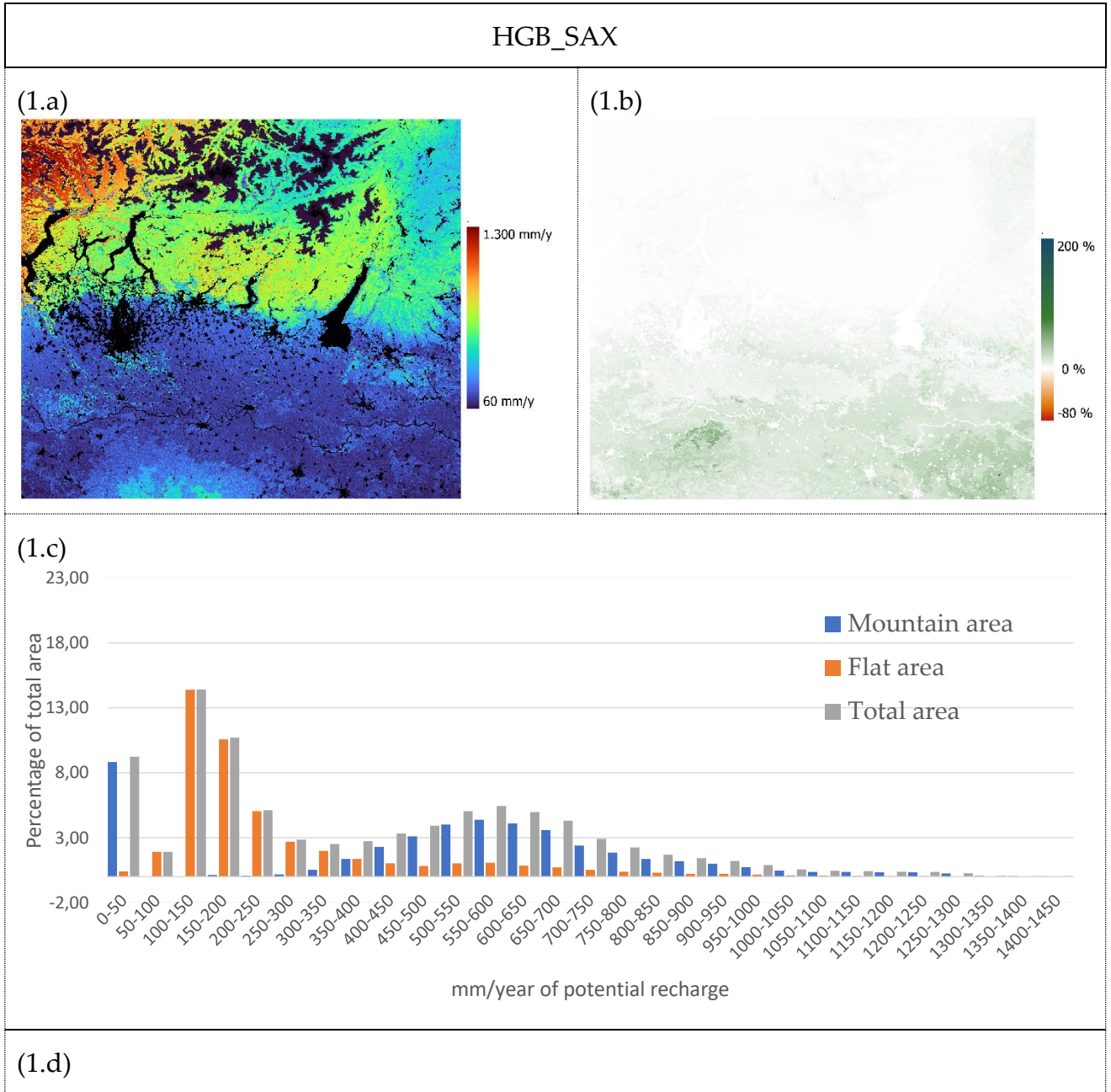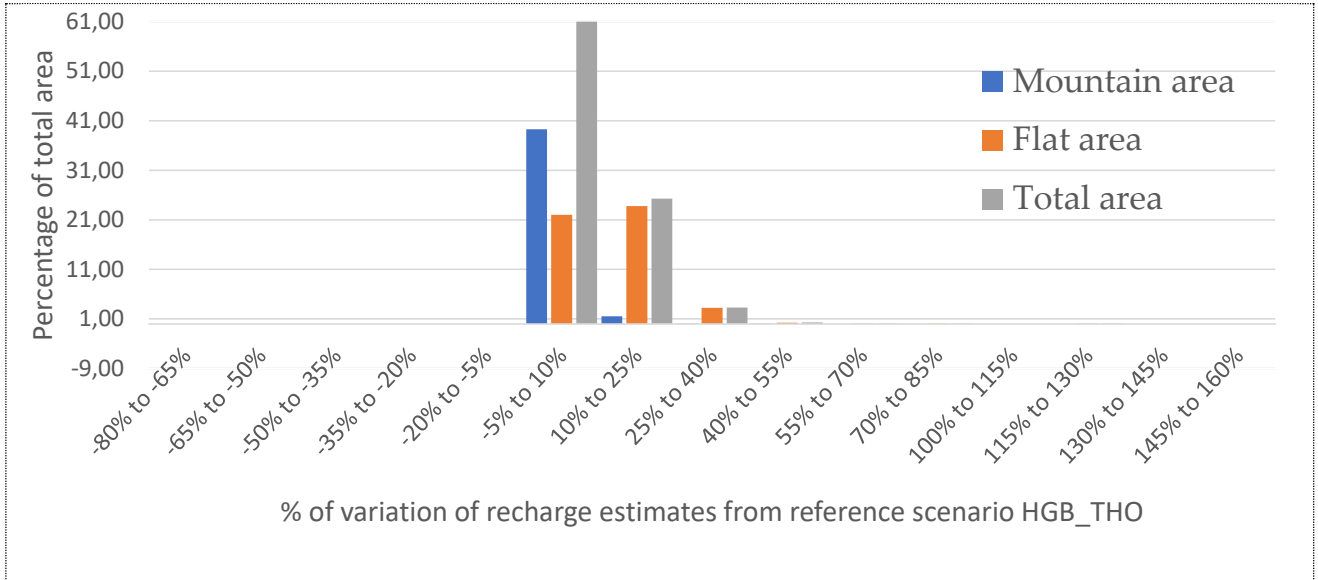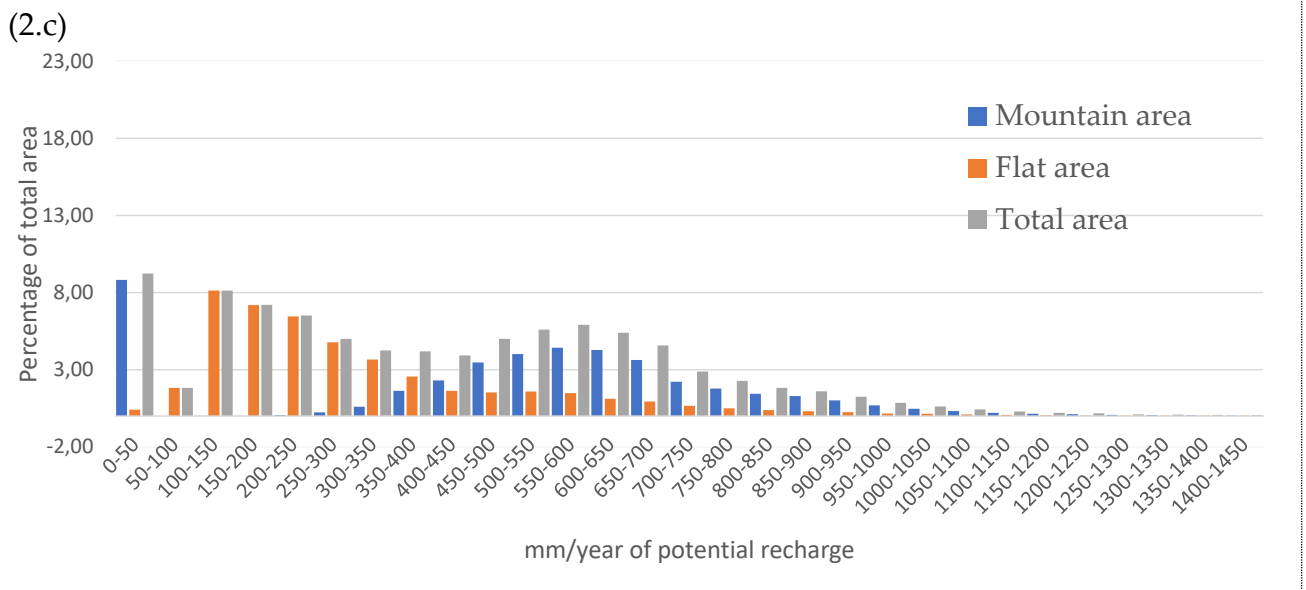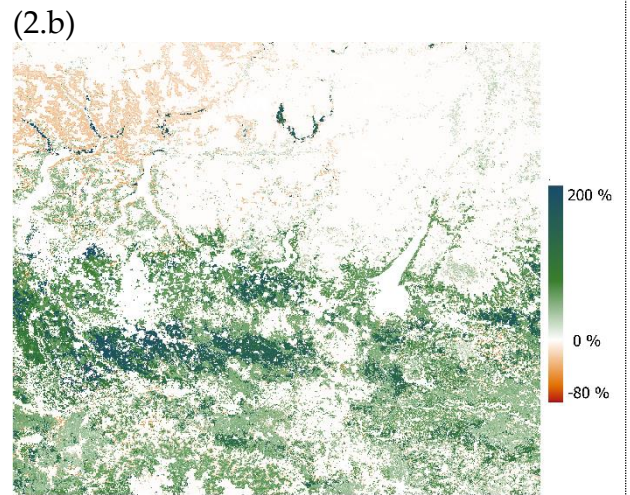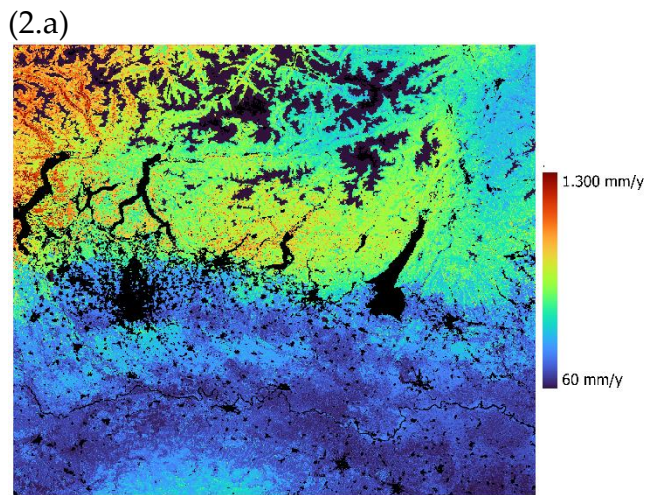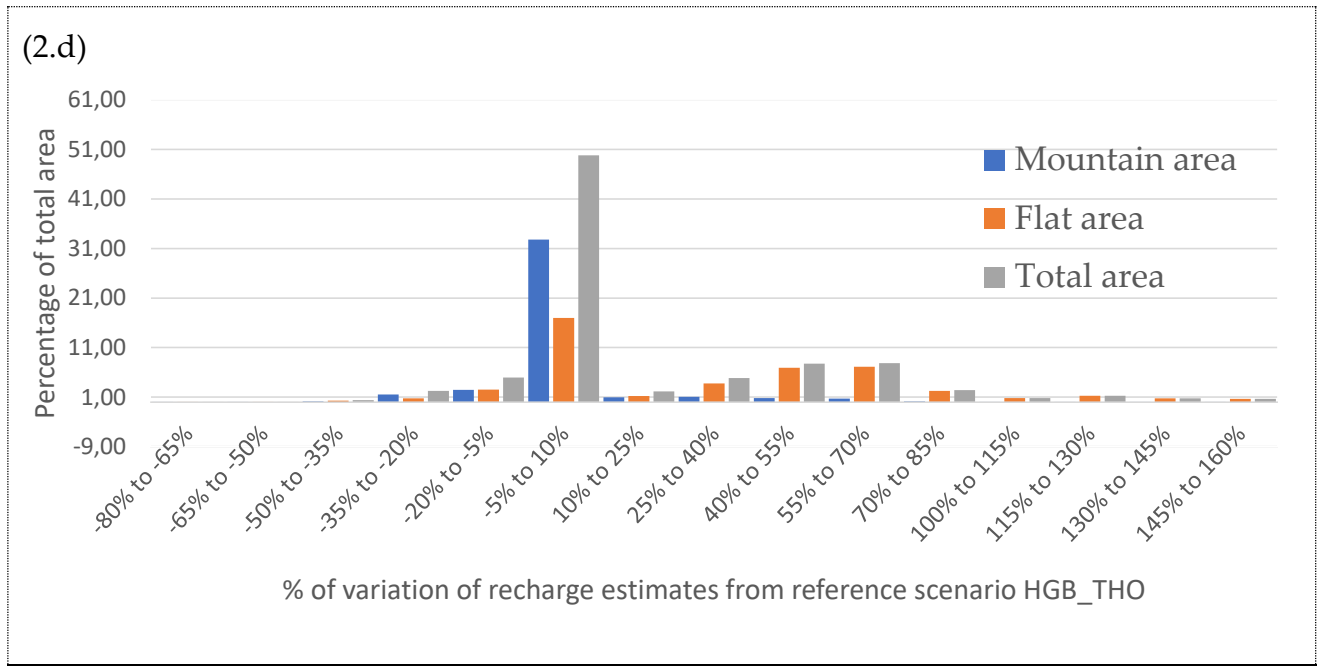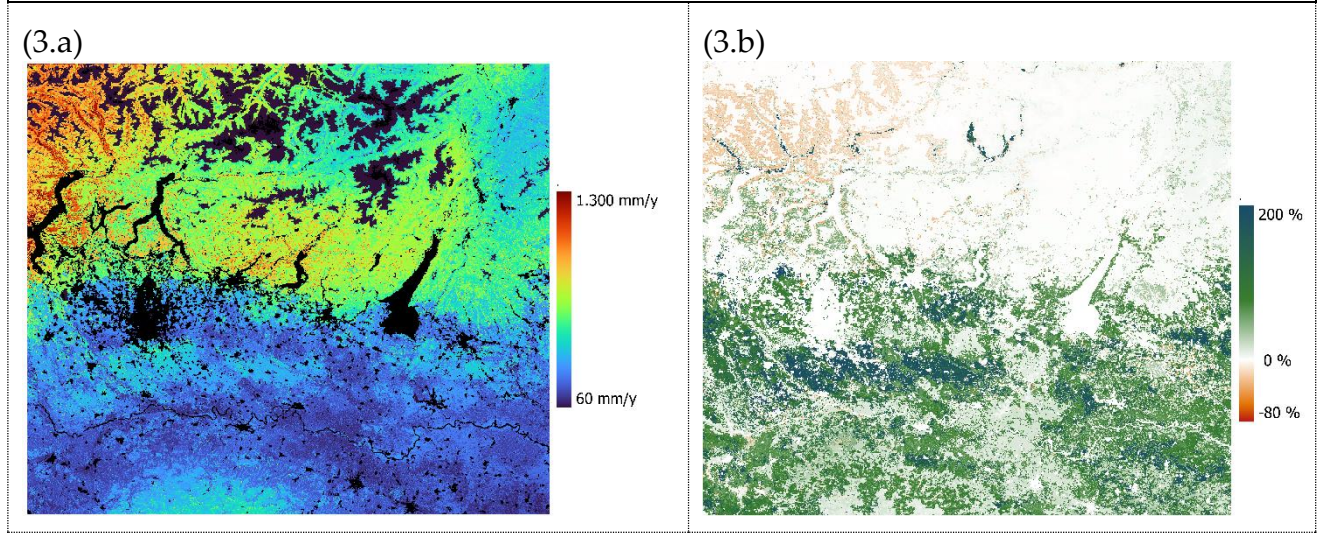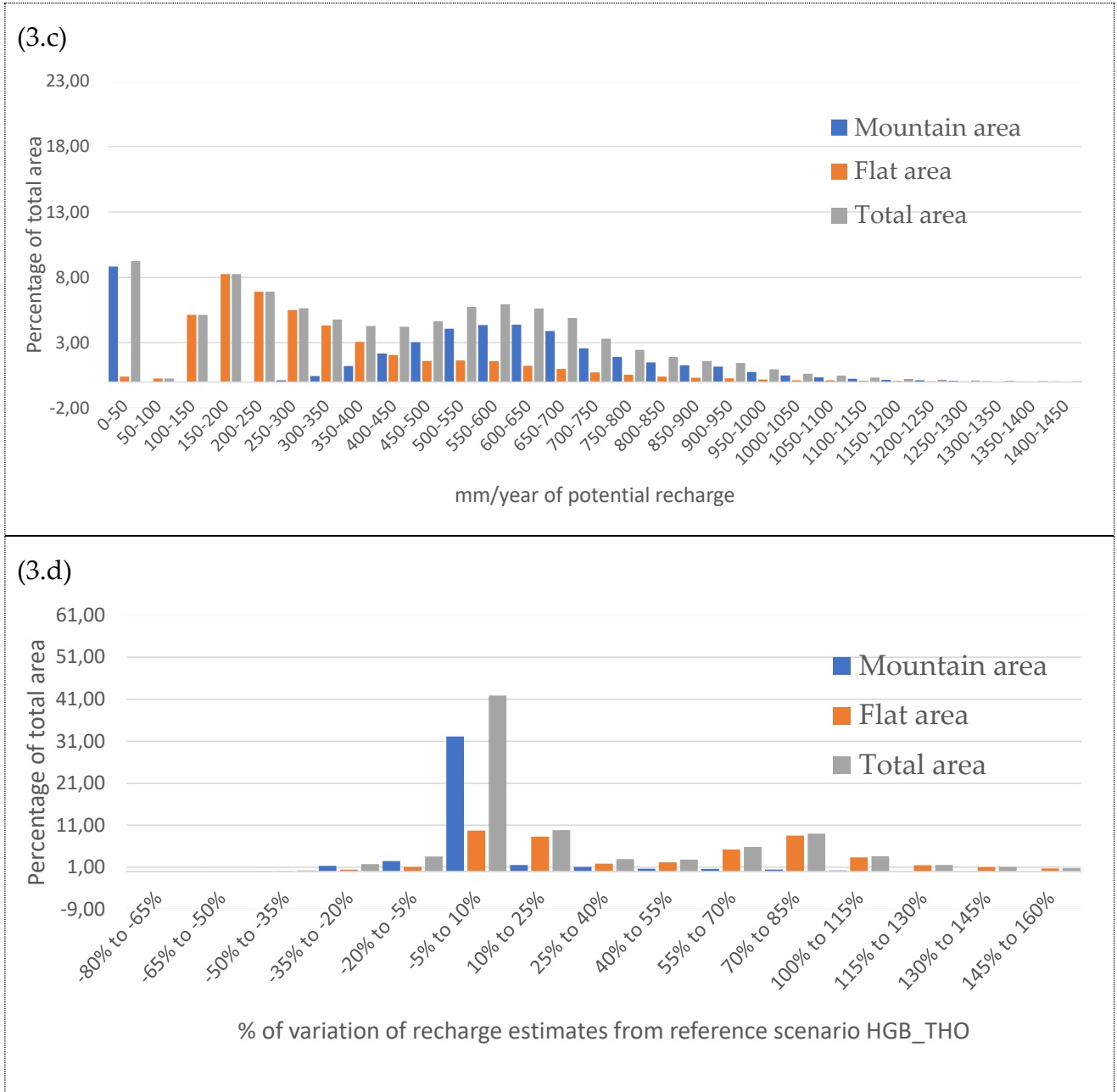
Figure 3.10: Summary of average potential recharge of the different historical simulations; figure (a) represents the potential recharge estimates, figure (b) represents the relative variation of potential recharge from the reference simulation, figure (c) represents the histogram of mountain and flatland recharge, figure (d) represents the histogram of variation of potential recharge from the reference simulation.

Figure 3.10 shows that even though all the other historical models predict a mean increase in potential recharge, the variation of potential recharge spatially can be both positive and negative. The HGB_SAX simulation shows in Figure 3.10.1.b that a modification of AWC through the Saxton and Rawls empirical equations produces everywhere in the domain only a slight increase in potential recharge. The increase of recharge is almost negligible in the mountain areas (smaller than 10%), and it increases in the plain area between 10% to 20%. Figure 3.10.1.c confirms this observation

considering HGB_THO: the mountain recharge histogram seems to be unchanged with respect to the reference simulations, but the flat areas histogram has lost most of the low recharge values. While the HGB_THO in Figure 3.9 had 9,26% of the area having potential recharge fitting in the 50-100 mm histogram range, the HGB_SAX had only 1,91% of the area with that range of potential recharge estimates. The HGR_THO shows a higher variation in potential recharge with respect to the reference scenario. The variation takes place both in the mountain area and in the flat area: there is a significant increase of recharge in the flat areas between 20% to 200% and a decrease is observed only in the north-western alps. The flat areas histogram shows a similar increase in recharge estimates to what is observed for the HGB_SAX model. In addition, flat areas recharge estimates based on the Rosetta method tend to be more extended and prolonged over higher recharge values. For example, the HGB_THO reference model distribution drops below a 2% frequency for potential recharge between 300-350 mm, while the HGR_SAX drops below 2% for potential recharge between 450-500 mm.

The following images (Figure 3.11 and Figure 3.12) analyze the spatial distribution of net infiltration estimates for future climate simulations applying the same methodology used for historical simulations. As anticipated from the analysis of previous studies in the introduction of this thesis, and as presented in Figure 3.11 and Figure 3.12, RCP8.5 simulations produce variations of net infiltration estimates that are more intense than what is observed for RCP4.5 scenarios both in the mountain areas and in the lower elevation areas.

ACCESS-01: RCP4.5

(1.a)

(1.b)

(1.c)

(2.d)

% of variation of recharge estimates from reference scenario HGB_THO

CCSM4: RCP4.5

(2.a)

(2.b)

(2.c)

mm/year of potential recharge

(2.d)

% of variation of recharge estimates from reference scenario HGB_THO
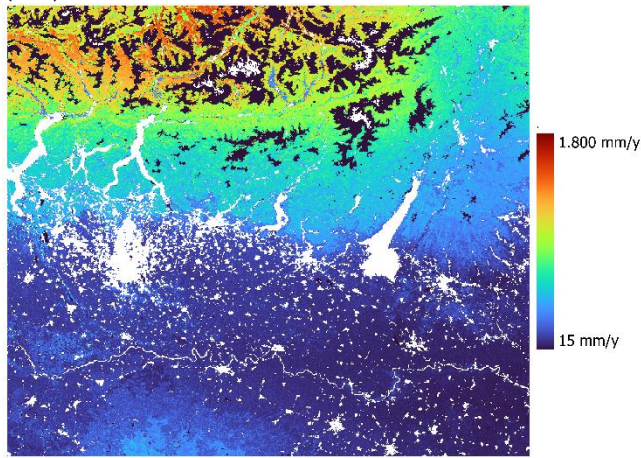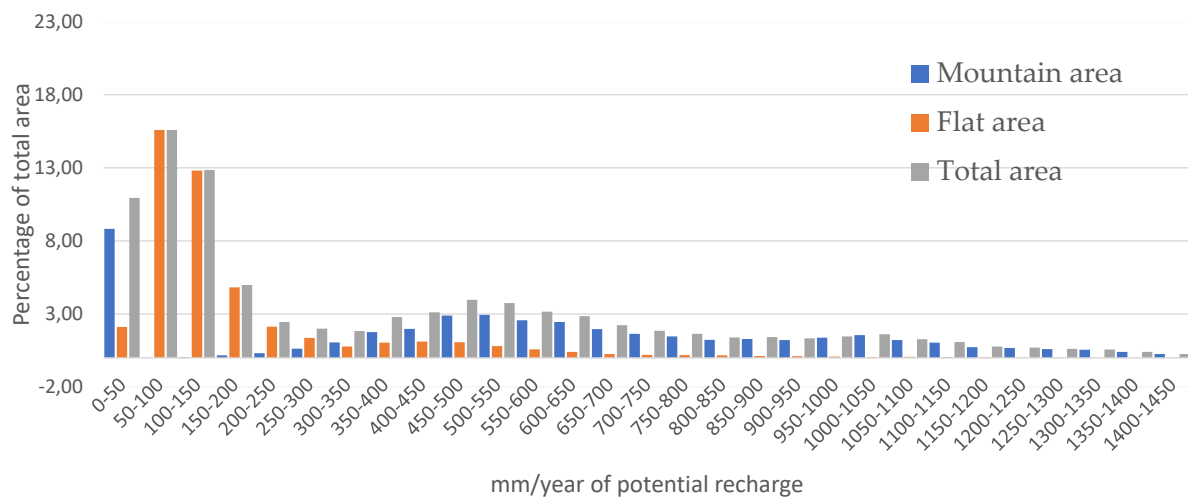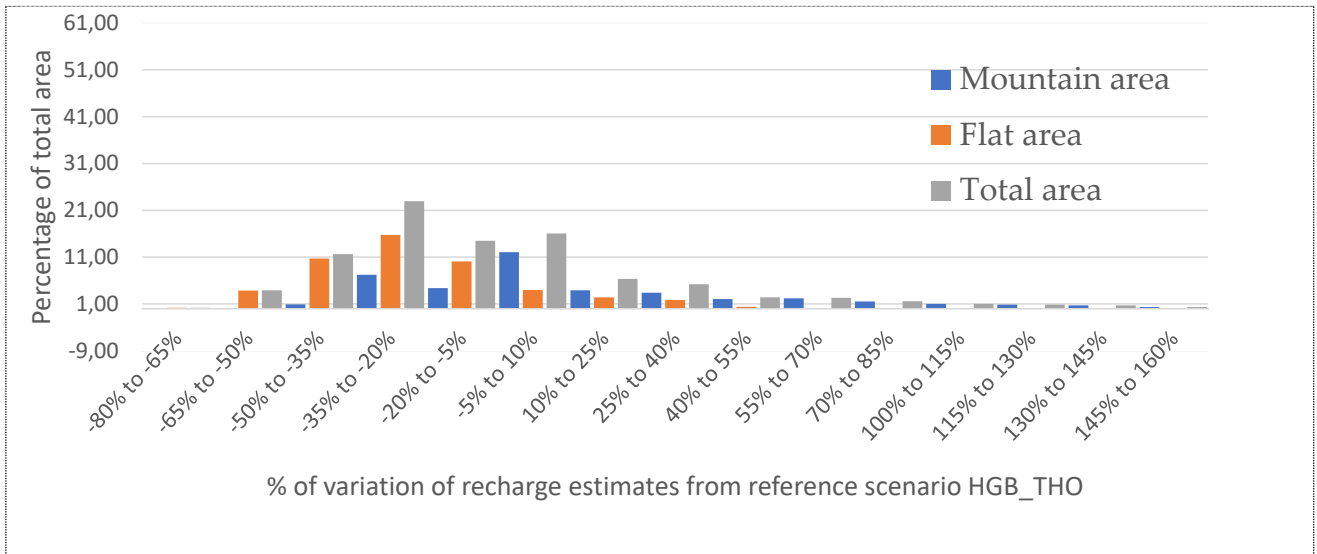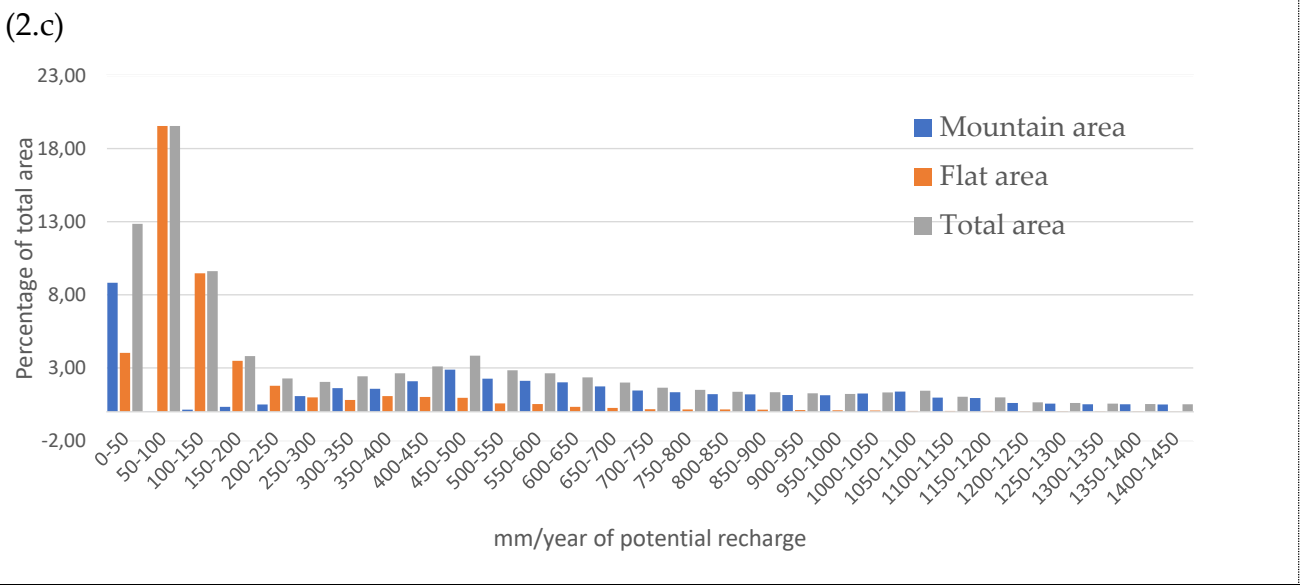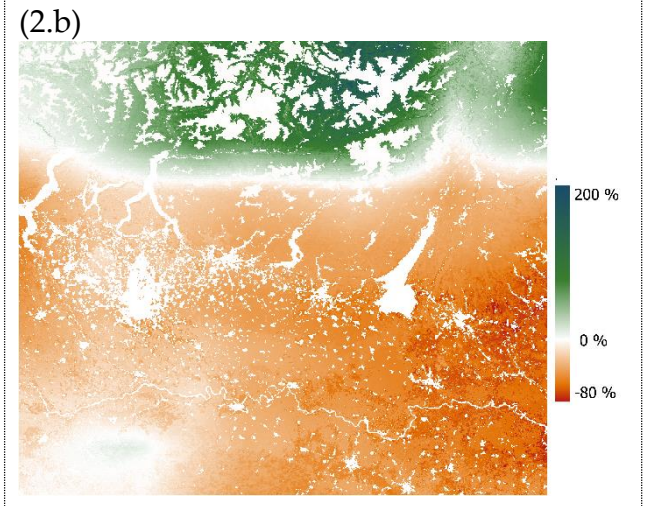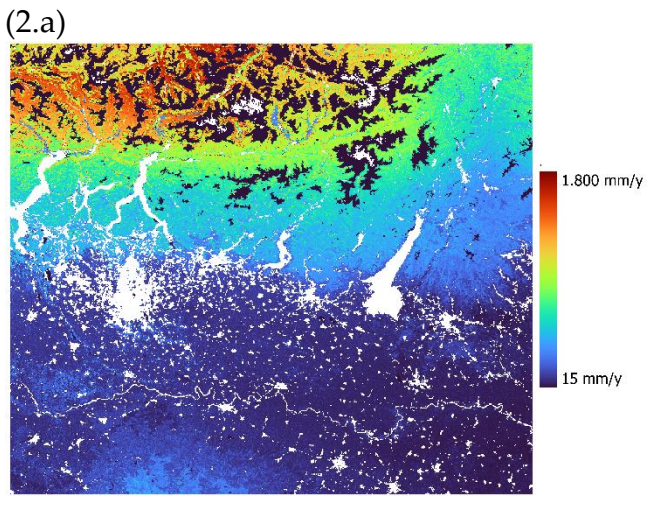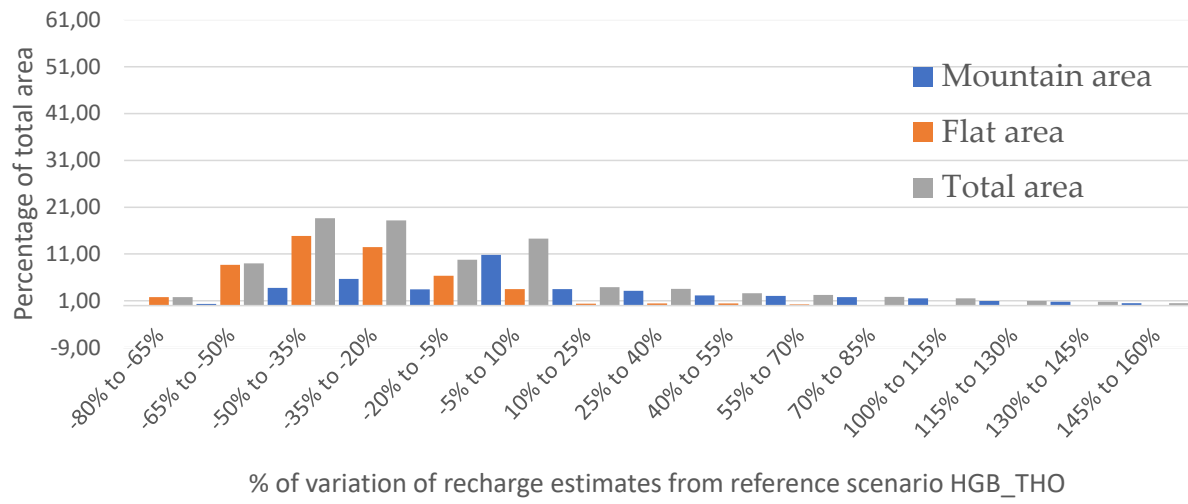
MIROC5: RCP4.5

(3.a)

(3.b)
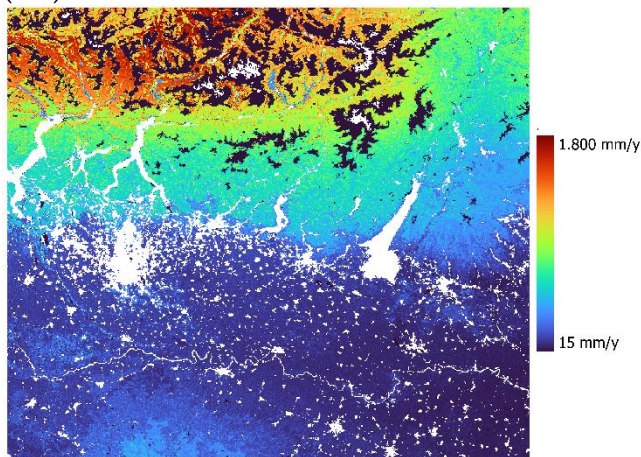
(3.c)
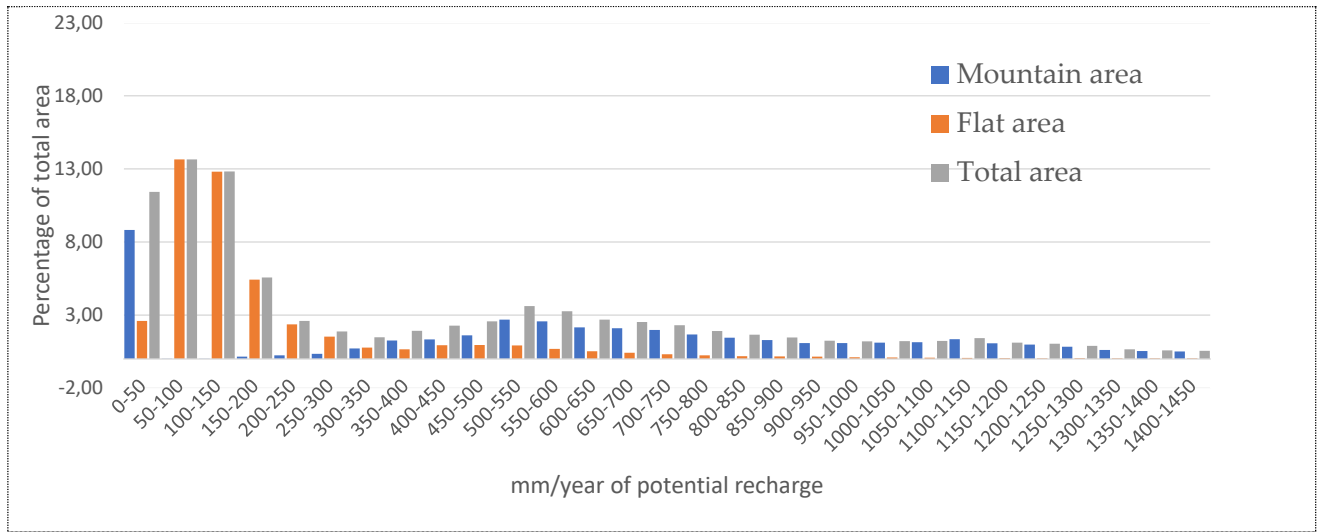
(3.d)
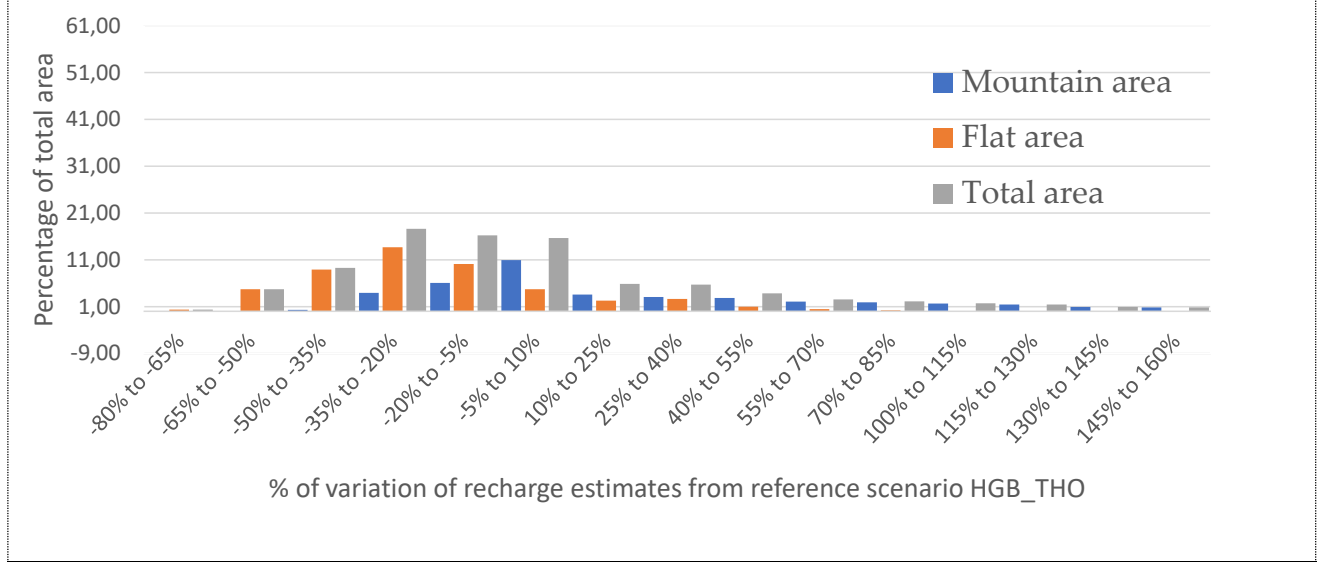
MRI-CGCM3: RCP4.5

(4.a)
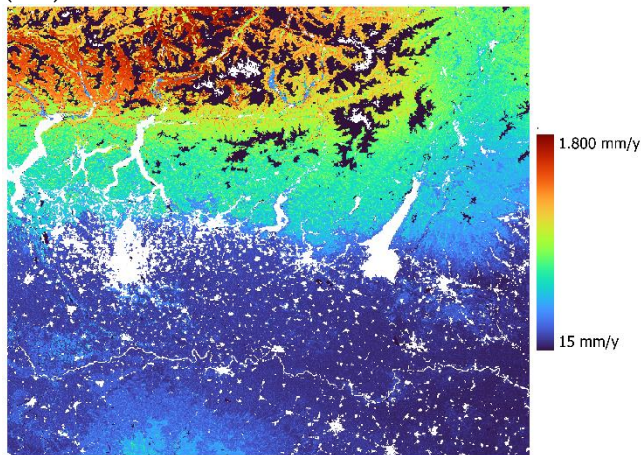
(4.b)
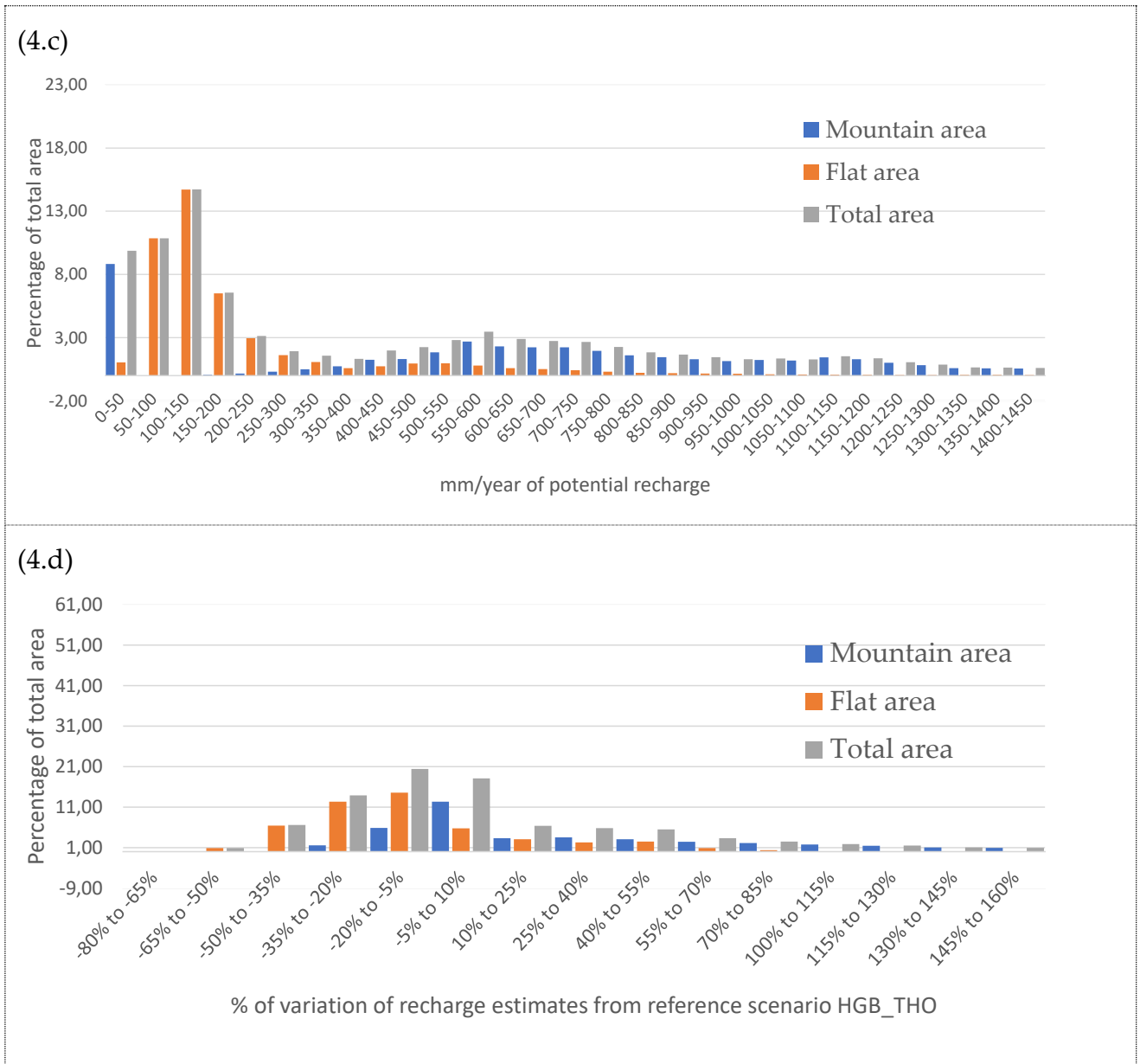
Figure 3.11: Summary of average potential recharge of the different RCP4.5 simulations; figure (a) represents the potential recharge estimates, figure (b) represents the relative variation of potential recharge from the reference simulation, figure (c) represents the histogram of mountain and flatland recharge, figure (d) represents the histogram of variation of potential recharge from the reference simulation.

CCSM4: RCP8.5

(2.a)

(2.b)

(2.c)

(2.d)

% of variation of recharge estimates from reference scenario HGB_THO
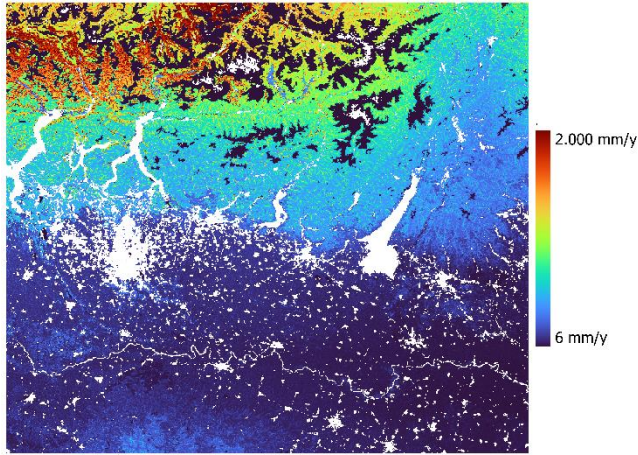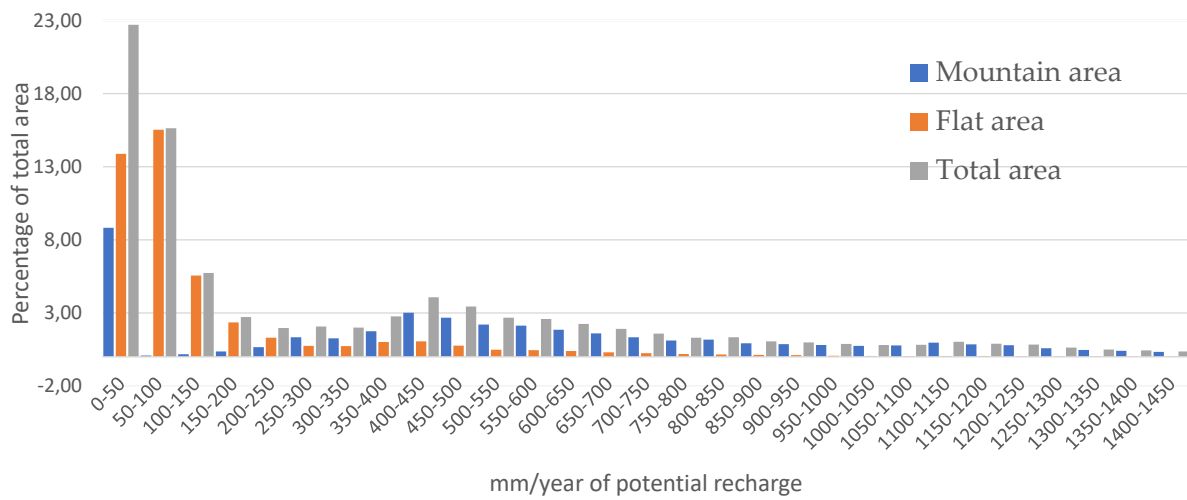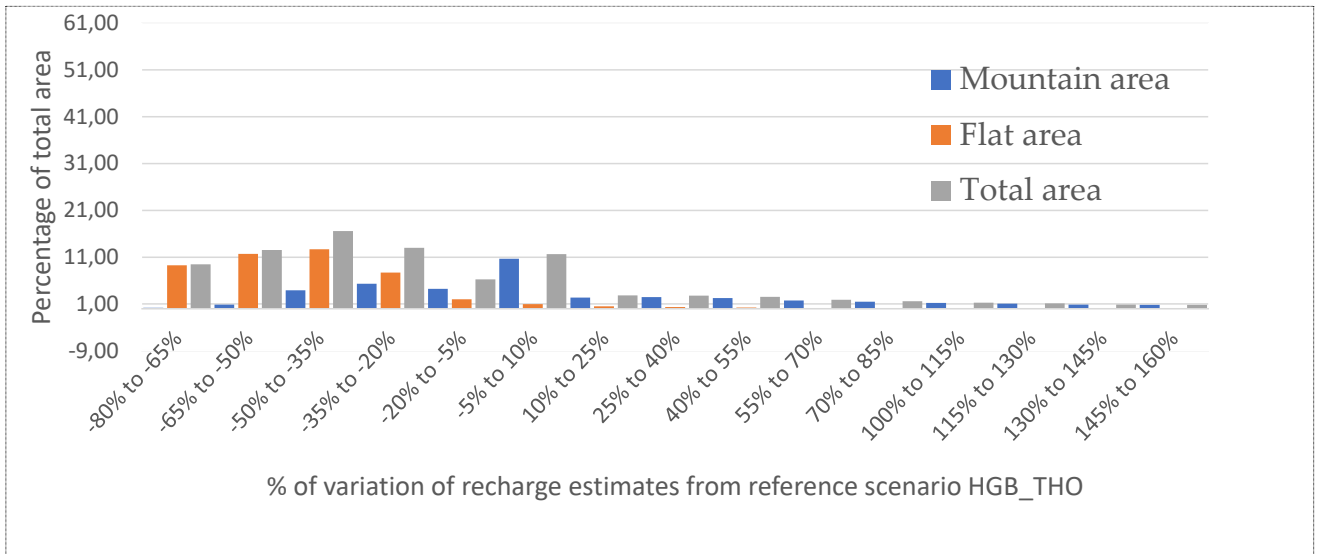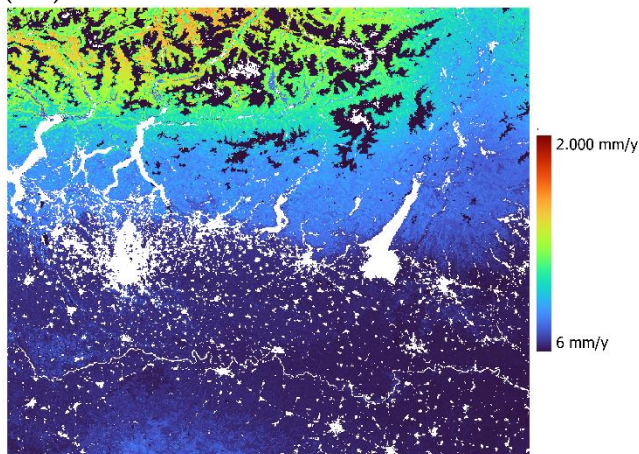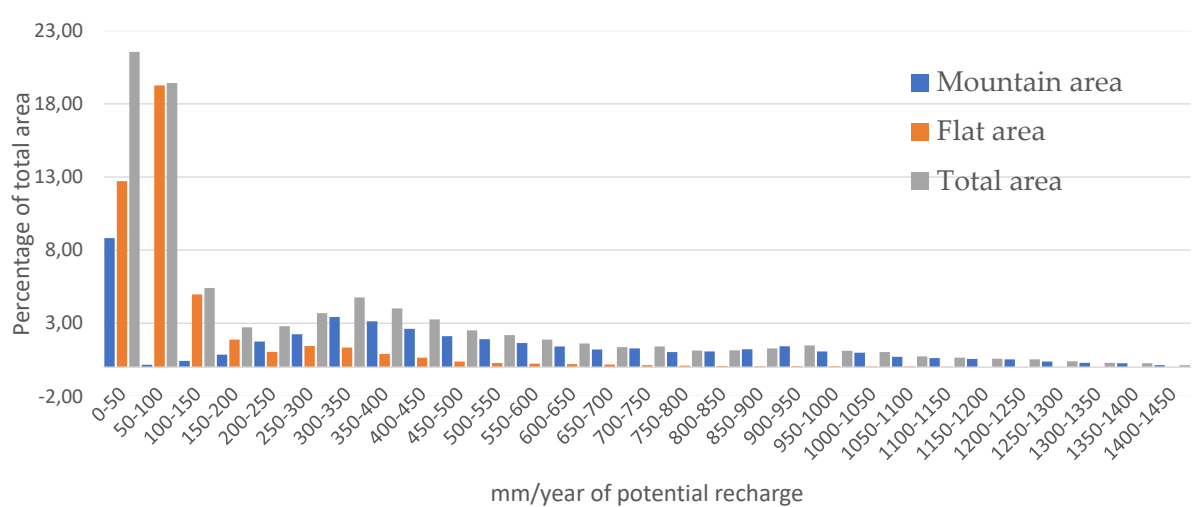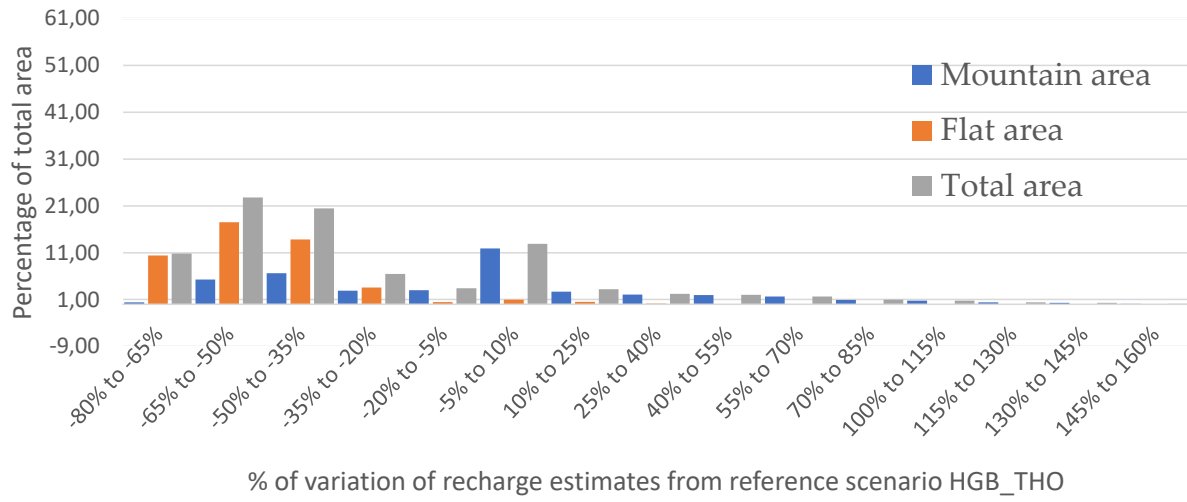
MIROC5: RCP8.5

(3.a)

(3.b)
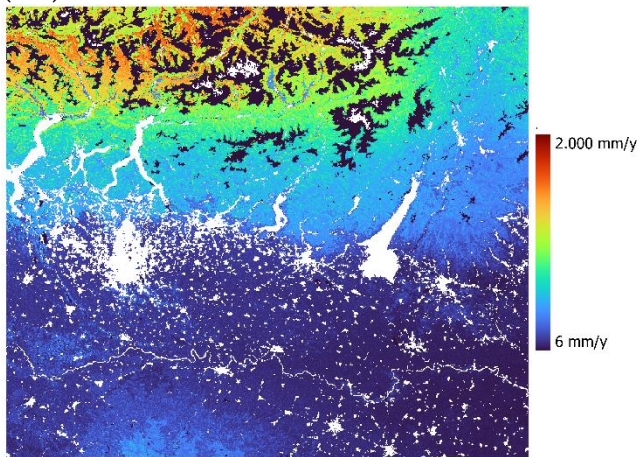
(3.c)
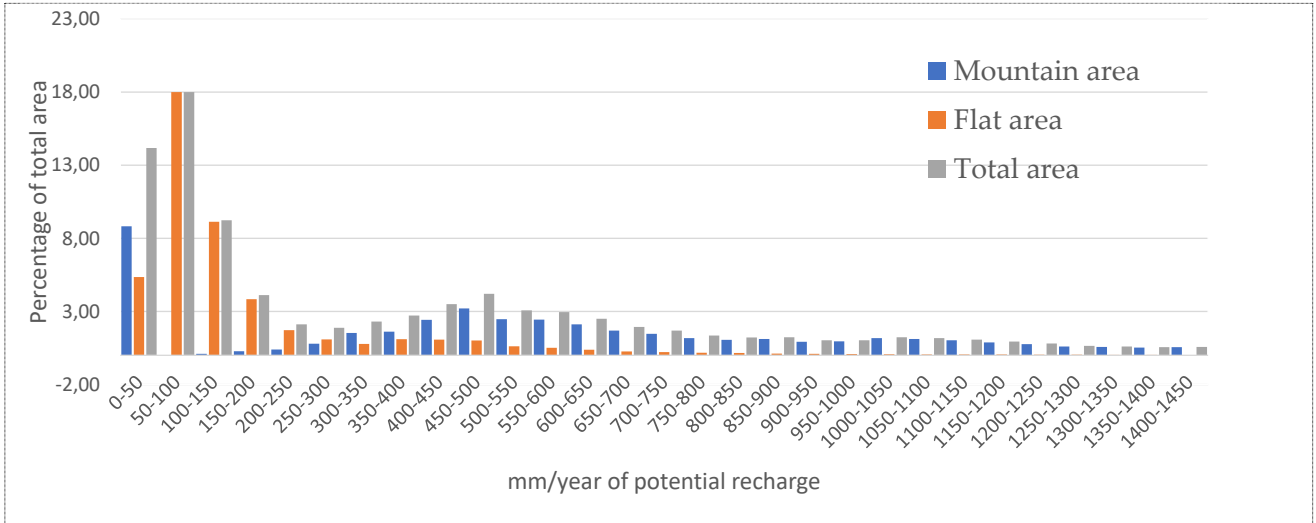
(3.d)
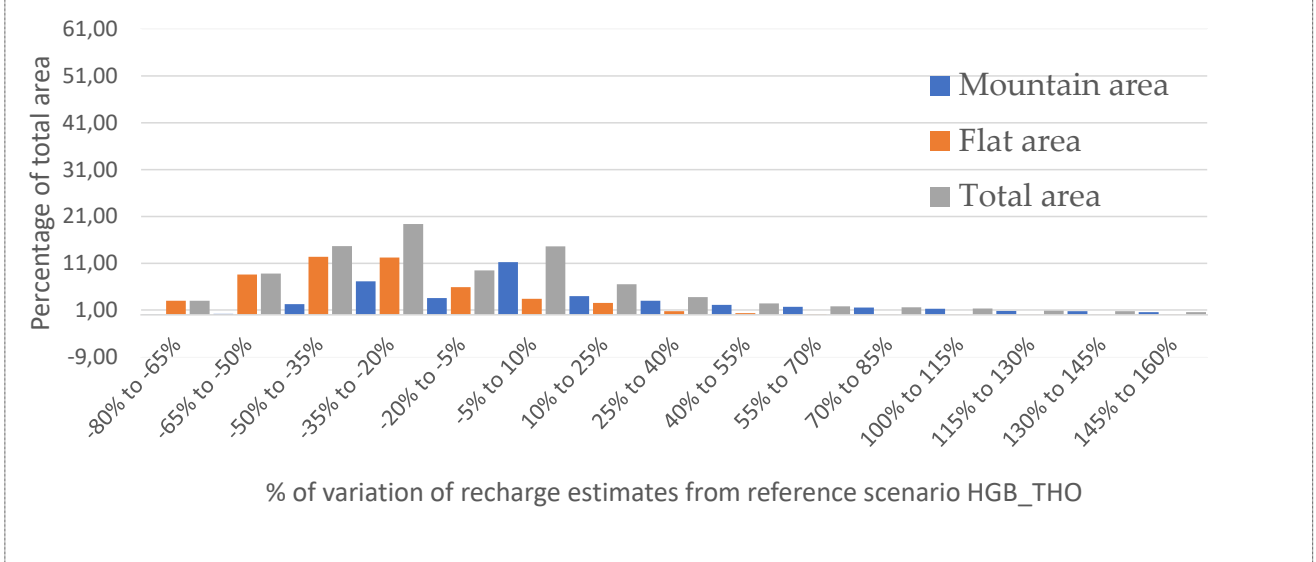
MRI-CGCM3: RCP8.5

(4.a)
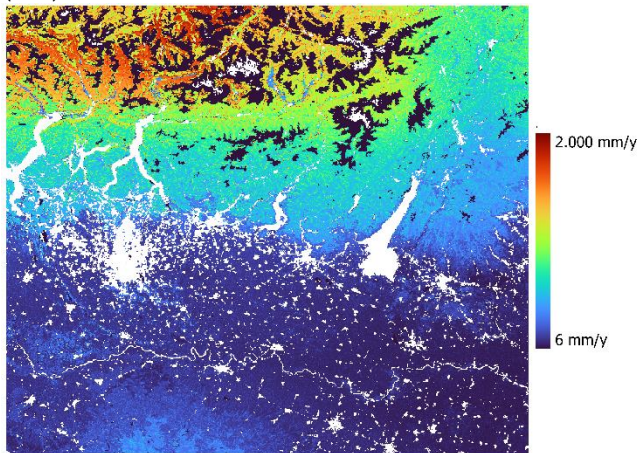
(4.b)
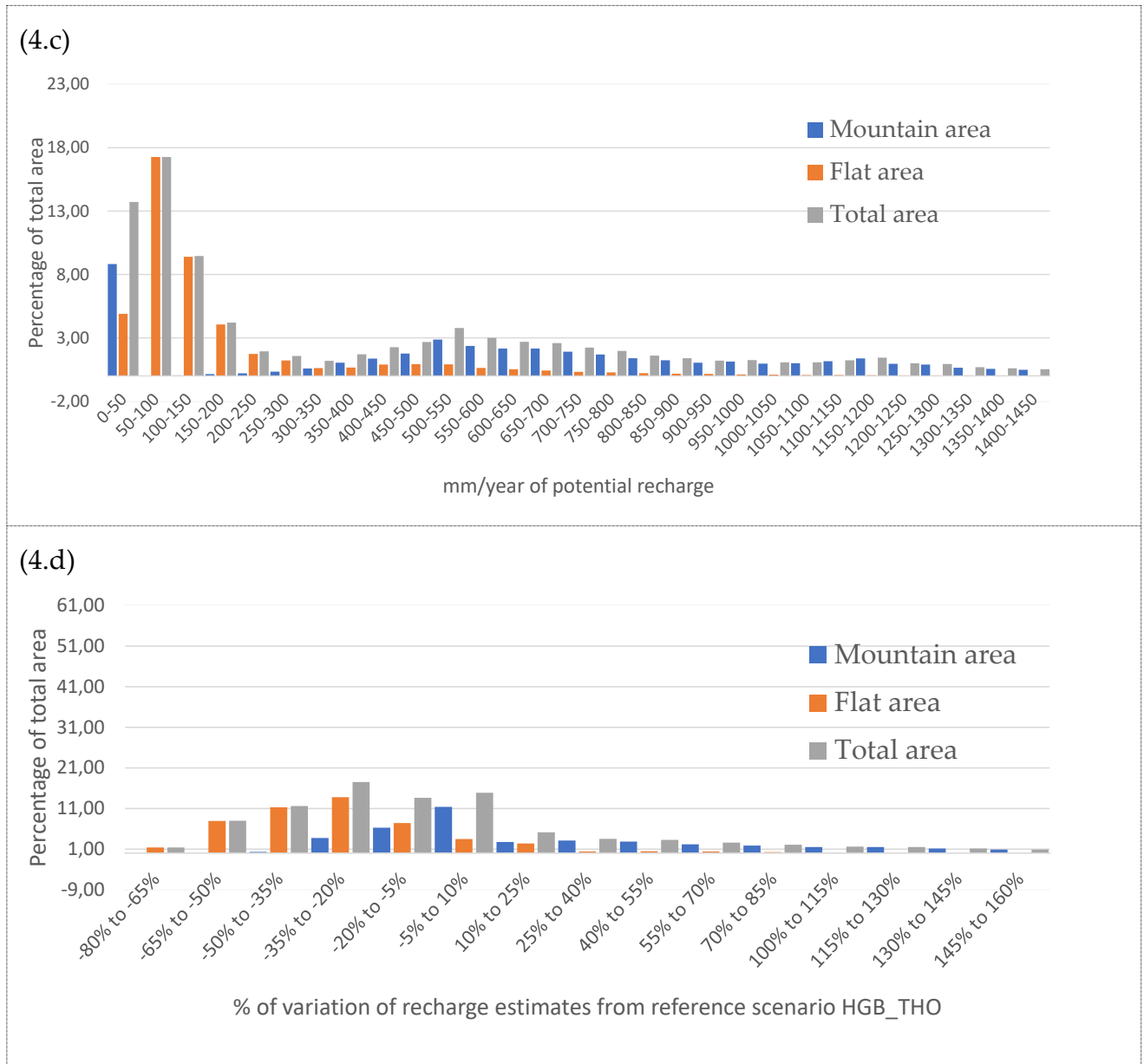
Figure 3.12: Summary of average potential recharge of the different RCP8.5 simulations; figure (a) represents the potential recharge estimates, figure (b) represents the relative variation of potential recharge from the reference simulation, figure (c) represents the histogram of mountain and flatland recharge, figure (d) represents the histogram of variation of potential recharge from the reference simulation.

The results of mean yearly potential recharge for future simulations produces variations in terms of percentage of deviation from reference scenario that are comparable to those produced by the historical simulations. All the models tend to present a reduction of potential recharge in the plain areas (the opposite occurring in the mountain areas) that is greater in the eastern half of the Pianura Padana flatland areas.

In general, future simulations forecast an increase of recharge in the mountain areas with a relative difference of up to 200% with respect to the historical reference model, and a decrease of up to -80% is predicted for the flat areas.

Regardless of the emissions scenarios, the histograms of future climate simulations show some similarities between them that emphasize the differences from historical climate simulations. Most of the flat-land area histograms of future simulations (except for the MRI-CGCM3 RCP4.5) present a modal value (50-100 mm/year) lower than the historical one (100-150 mm/year), in agreement with the prediction of lower recharge quantities for these areas. The histogram of mountain areas presents roughly the same modal value of historical simulations, but its distribution is much more spread over net-infiltration cluster classes than what is observed for historical simulations. This could potentially suggest a higher variability in future recharge conditions in mountain areas that will include both areas of decrease and increase of potential recharge.

An additional analysis on the effects of GCM on the estimates of potential recharge has showed that mountain areas are also subject to a greater uncertainty in terms of definition of potential recharge volumes. In fact, considering a single emission scenario and a single year of simulation, four different outputs of mean annual potential recharge are available to be used to estimate the spatial distribution of the standard deviation of recharge estimates. Figure 3.13 shows the spatial distribution of the standard deviation for the years 2080, 2090 and the two years picked for the analysis of recharge estimate in Figure 3.5 and Figure 3.6, respectively 2097 and 2094.

2080	2090
RCP 4.5		
(1.b)	(2.b)
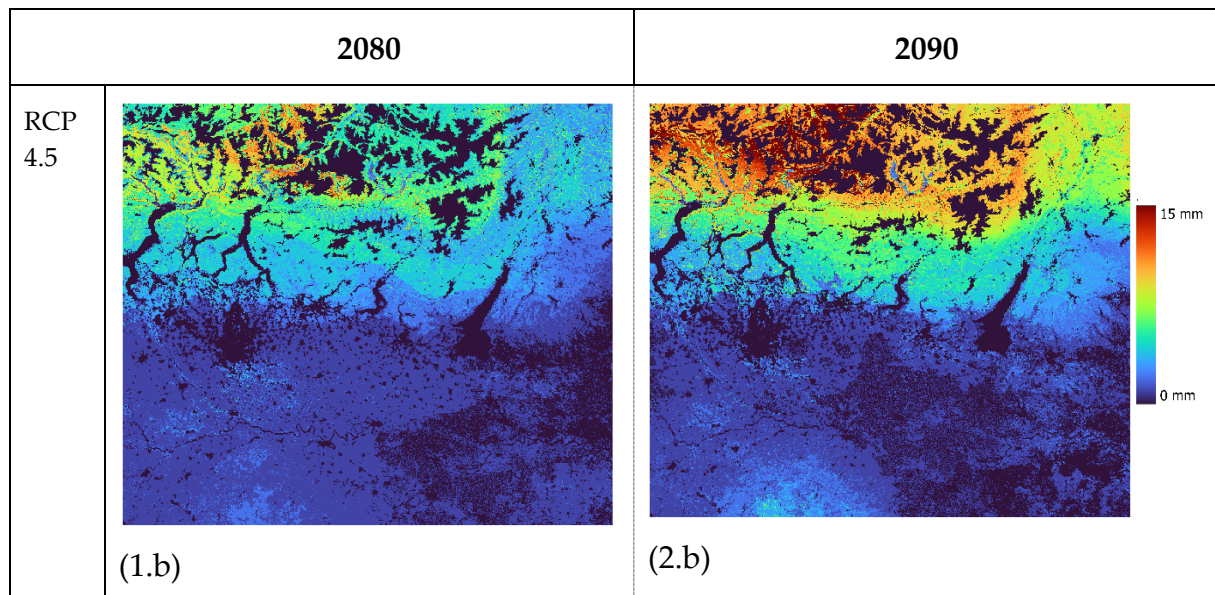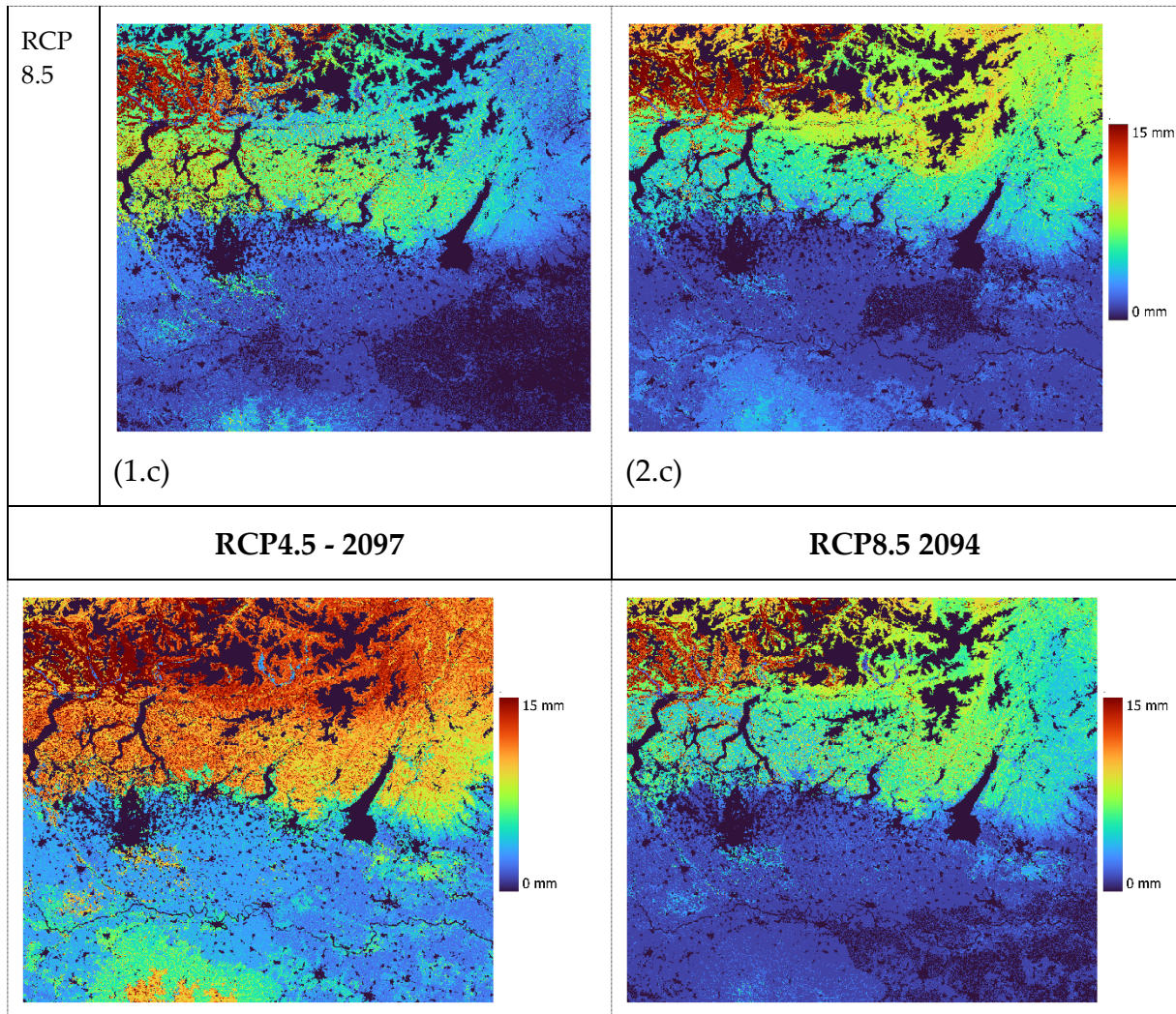
Figure 3.13: Spatial distribution of standard deviation of potential recharge estimates for different years of simulation.

The analysis of the standard deviation of potential recharge estimates shows that not only the effect of climate change will be more intense in the mountain regions of Lombardy in terms of increased volumes of potential recharge, but also the uncertainty of the magnitude of change in these areas will be greater.

Despite the variations in recharge patterns between historical and future climate simulations, Table 3.1 shows that ratios of potential recharge to total precipitation (%R) tend to remain constant. The images below represent spatially the distribution of %R averaged over the 20 years of simulations.

HGB_THO	HGB_SAX
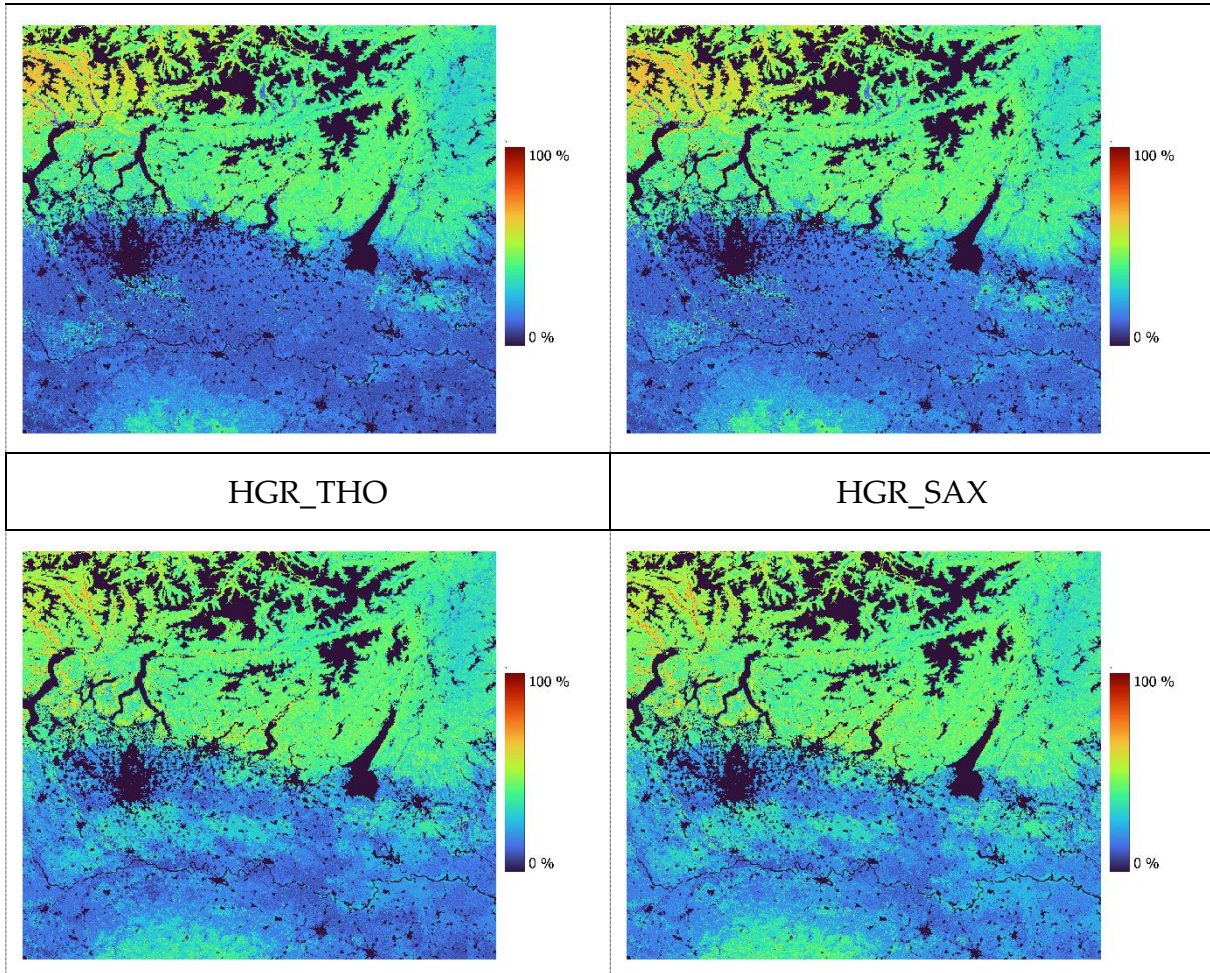
Figure 3.14: Percentages of yearly cumulated potential recharge to precipitation for historical simulations

The results of the historical simulations produce outputs coherent with the previous comments made. Considering an input total precipitation that is identical for all four models, models executed with the Saxton & Rawls method tend to show values of %R that are higher than the corresponding model executed with Thornthwaite Mather input. However, the difference is more evident when comparing models' outputs obtained with the Rosetta method or the Brakensiel method. In general, histogram analysis of these models shows that flat areas have a %R modal value of around 10%-15% while mountain areas have a %R modal value of 40%-45%. If all four historical models tend to agree on the spatial distribution of %R, this is not true for future simulations.

RCP4.5 – ACCESS-01	RCP8.5 – ACCESS-01

RCP4.5 – CCSM4	RCP8.5 – CCSM4
RCP4.5 – MIROC5	RCP8.5 – MIROC5
RCP4.5 – MRI_CGCM3	RCP8.5 – MRI_CGCM3
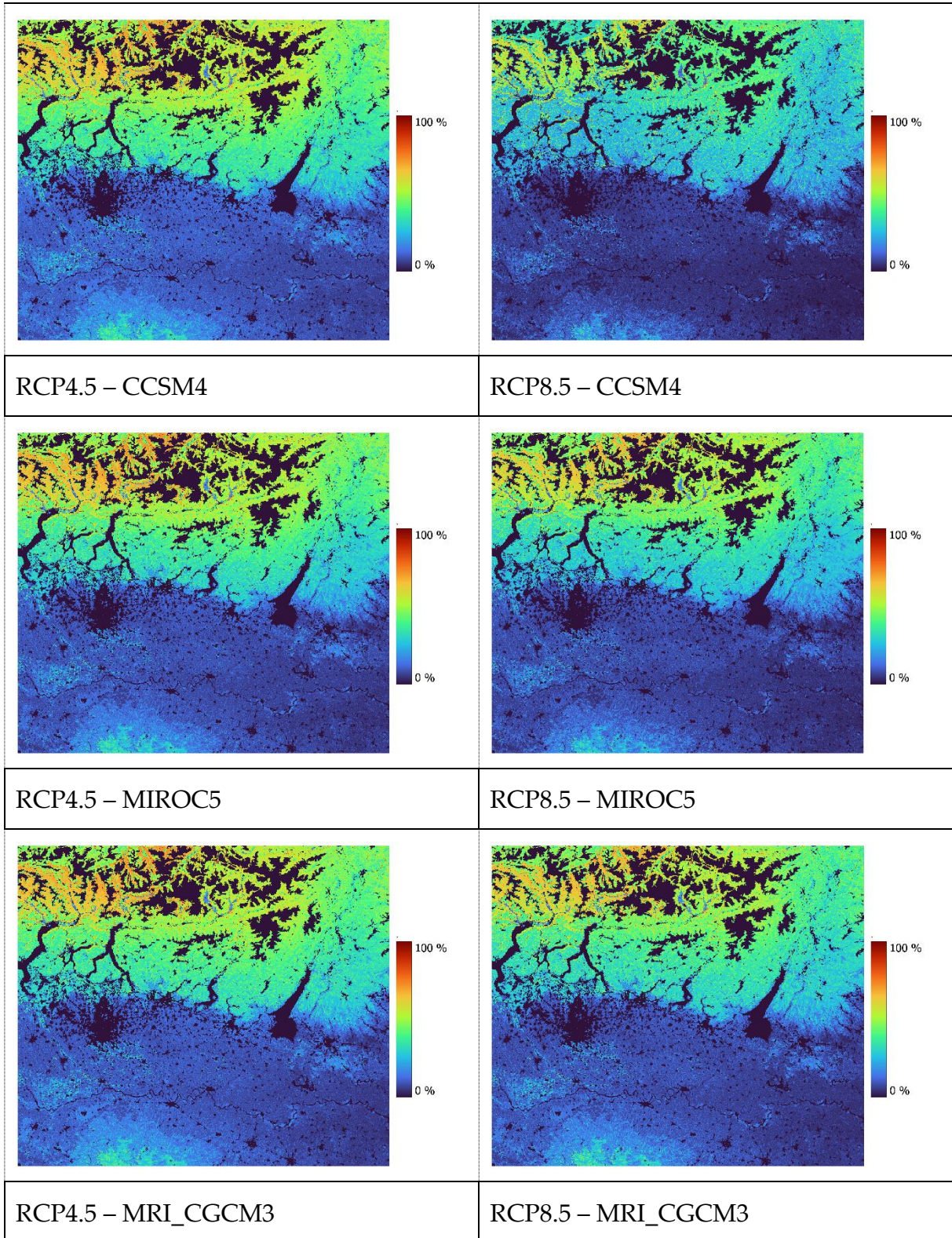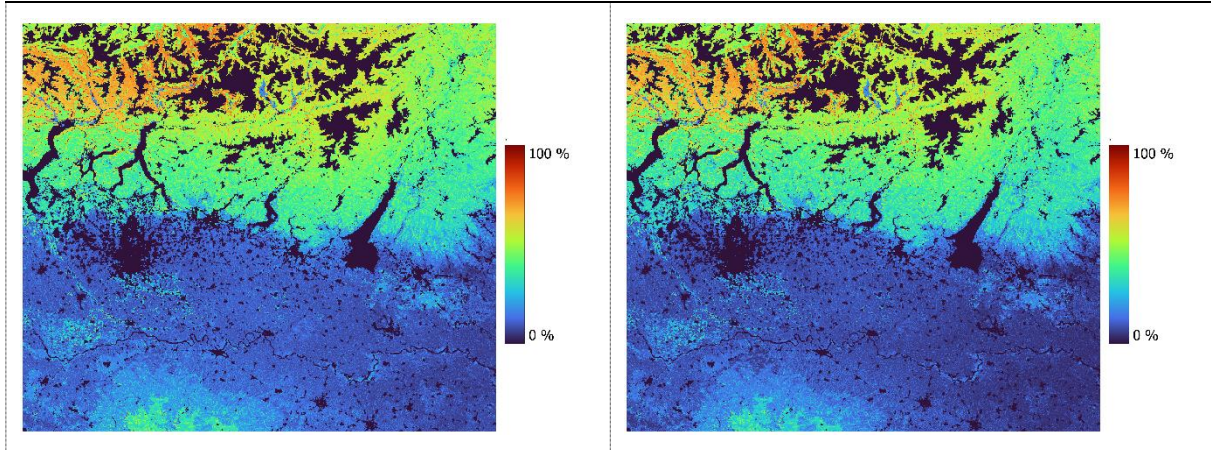
Figure 3.15: Percentages of yearly cumulated potential recharge to precipitation for future simulations

The %R produced by future simulations is similar for the different models but differs from the one produced by historical simulations. In comparison to historical simulations, %R decreases in flatland regions and increases in a significant portion of the mountain regions. Flat areas %R modal values are around 5%-8% and are lower than what is observed for historical simulations (10%-15%). Likewise historical simulations, mountain areas have %R modal values around 40% for future scenarios simulations, but similarly to what is displayed in Figure 3.11 and Figure 3.12, their values are much more distributed on higher and lower classes and range from 28% to 50% of %R.

Figure 3.16 analyzes monthly mean values of cumulated recharge for all the 12 different simulations. In all climate change scenarios, future potential recharge is expected to exceed current conditions in the first months of the solar year until the beginning of the spring season. Instead, in the months following the summer season all models predict a drop of potential recharge below current average conditions.
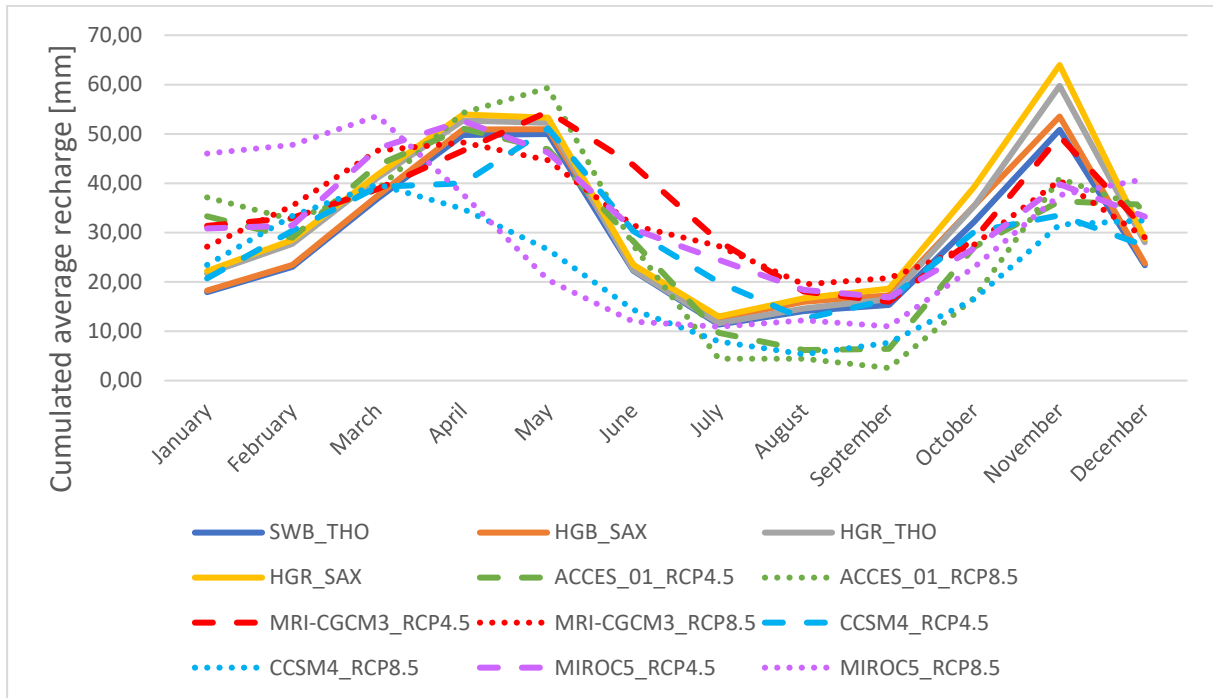
Figure 3.16: Average monthly recharge values for different simulations.
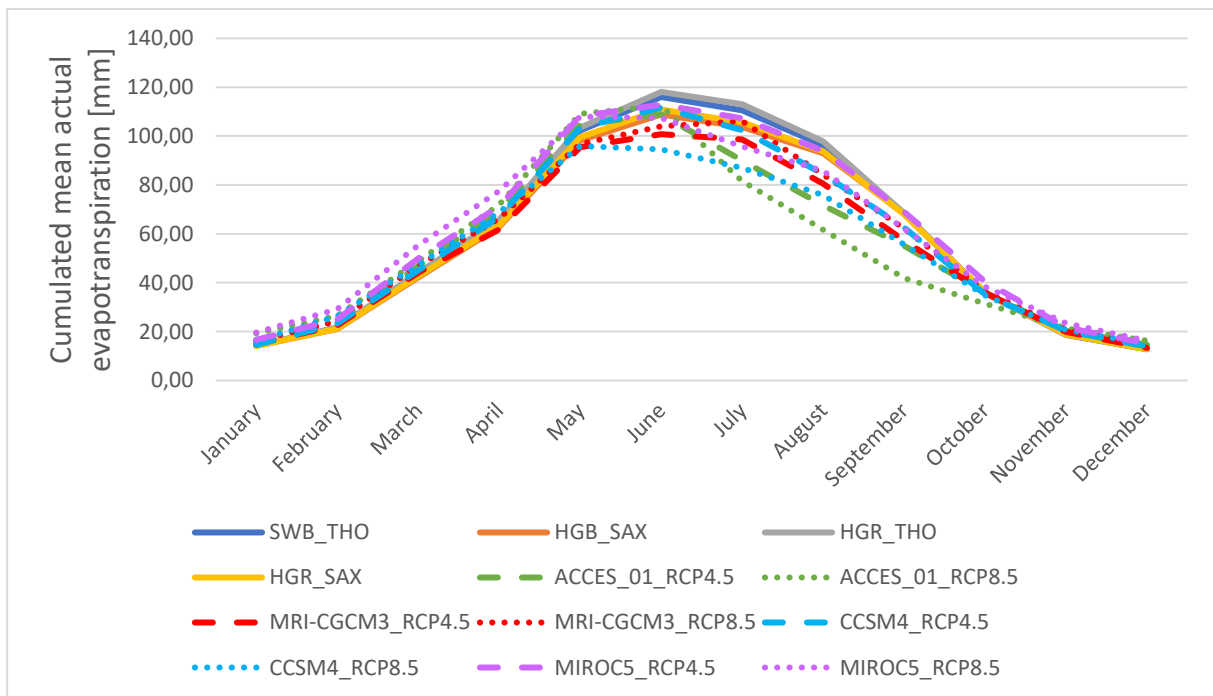
Figure 3.17: Average monthly evapotranspiration values for different simulations.
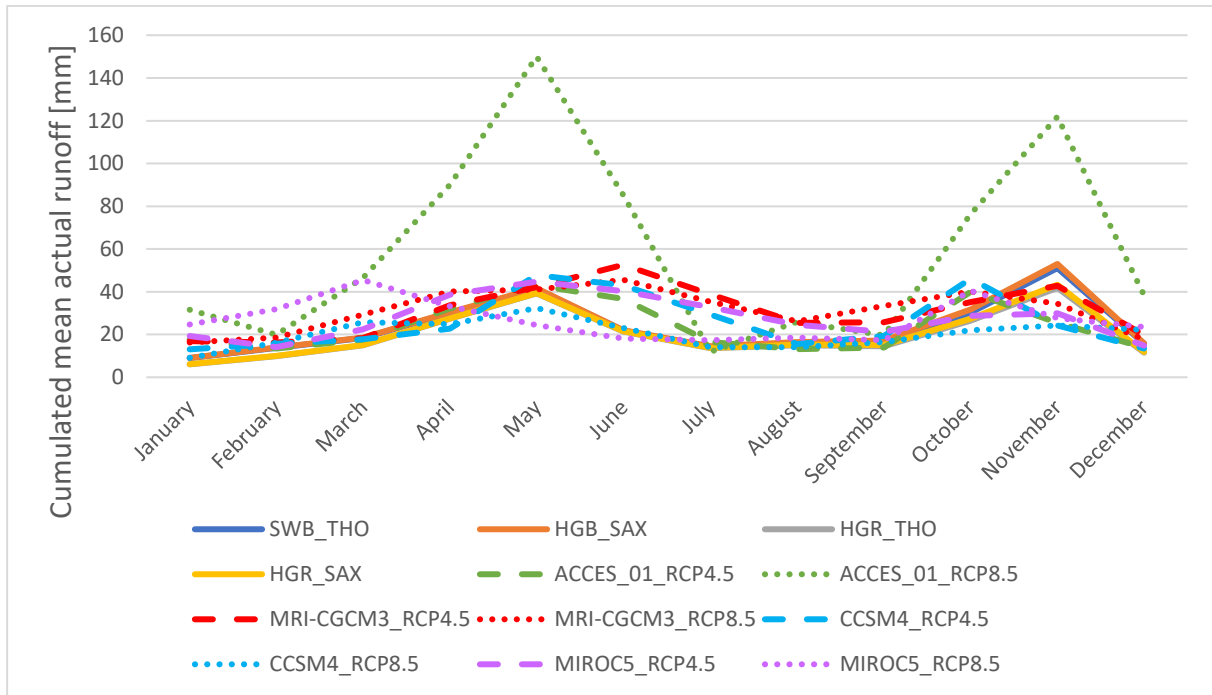
Figure 3.18: Average monthly runoff quantities for different simulations estimated by the CN method.

Figures above analyze monthly mean values of cumulated ET and runoff. If evapotranspiration patterns simulated with GCMs are more or less similar with respect to historical data, future runoff quantities are estimated to be greater the first two thirds of the solar year and then drop below current conditions only for the last months of the year. In the runoff graph, the ACCESS-01 RCP8.5 scenario produces estimates of runoff that are significantly high in the spring and autumn seasons: since a similar trend was observed for the precipitation quantities of the same model in Figure 2.3, it can be inferred that the surplus of surface runoff in the ACCESS-01 RCP8.5 simulation compensates the excess of precipitation for the same year. Additionally, it may be deduced that a surplus of precipitation amounts does not always translate into a surplus of net infiltration estimates.

Comparing and analyzing the output of simulations based on variables cumulated annually and averaged over 20 years of simulation is a good strategy to identify general pattern changes in the recharge mechanism. However, analyzing data only on yearly time frames can obscure other important patterns and phenomena happening at smaller time scales. The following graphs present the analysis of potential recharge estimates for all the twelve different models on a seasonal frequency, instead of annually, by averaging the results of the twenty years of simulations separately for each season. The seasonal investigation of model output is frequently chosen in hydrologic analysis to correctly match the variations observed in water fluxes for one solar year. The following graphs aim at analyzing how potential recharge estimates deviate positively or negatively from the reference model HGB_THO. This type of analysis is performed following two methods and produces two graphs for each

season. The first graph presents the mean percentage of relative variation of potential recharge estimates from the reference model HGB_THO for the two areas of positive and negative deviation of results from the reference scenario. The second graph is similar but instead of presenting the mean variation in percentage, returns the sum of positive and negative variations (in percentage) with respect to the reference model. While the first graph returns a value that defines the mean change observed for each season, the second graph gives information on the magnitude of this change by taking into account how positive or negative trends of variation are widely spread on the domain. For example, the positive variation of recharge for the winter season of a *i-*model is calculated generating a raster of relative variation from the reference model of net infiltration during the winter season (averaged over all winters of twenty years of simulation) similarly to the yearly analysis performed in the figures *b* of Figure 3.10, Figure 3.11, and Figure 3.12. Then, all elements of this raster that present a positive variation are either averaged, to generate the first graph, or summed, to generate the second graph. The following formula summarizes the calculations used to produce the following images:

$$\textit{Mean positive variation} \tag{3.2}$$
$$= \frac{\sum_{i=0}^{N-positive} \% \; \textit{relative variation if variation is positive}}{N - positive}$$

$$\textit{Sum of positive variation} \tag{3.3}$$
$$= \sum_{i=0}^{N-positive} \% \; \textit{relative variation if variation is positive}$$

Where:

- N-positive are all the cells analyzed in the raster of variation of winter recharge that present a $\% \; relative \; variation > 0$
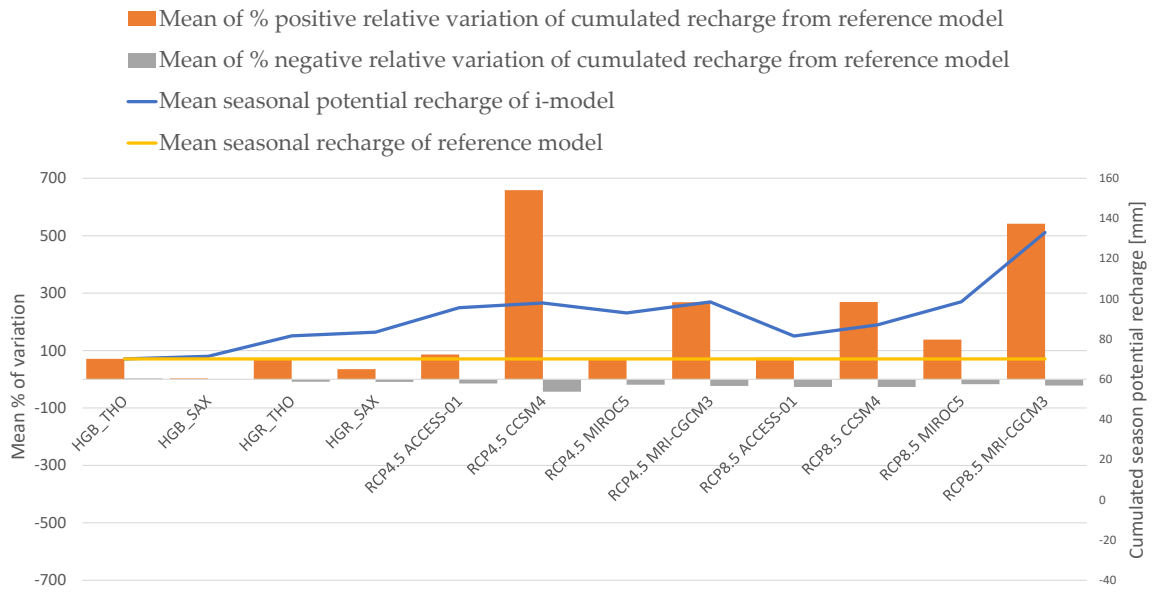
Figure 3.19: Mean % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for winter season averaged over the twenty years of simulation.
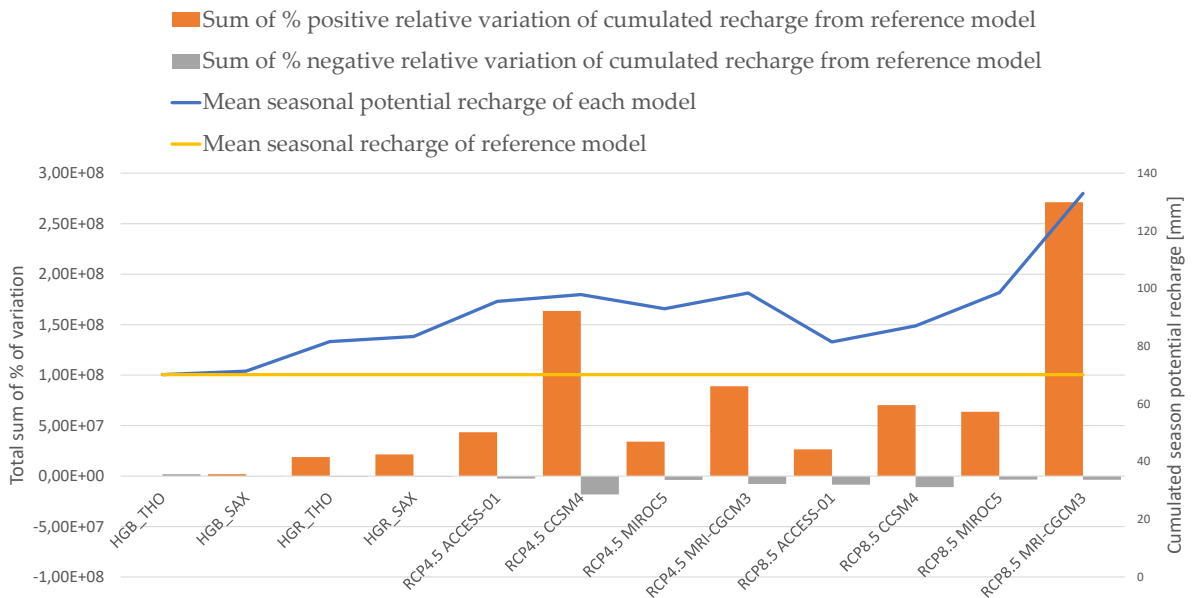
Figure 3.20: Sum of % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for winter season averaged over the twenty years of simulation.
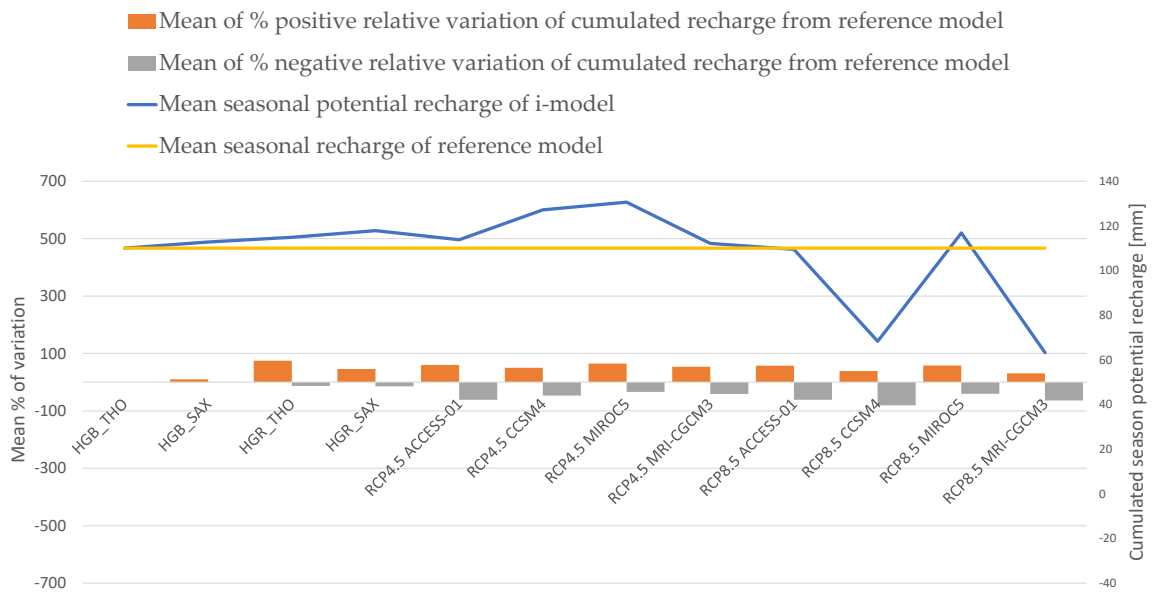
Figure 3.21: Mean % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for spring season averaged over the twenty years of simulation.
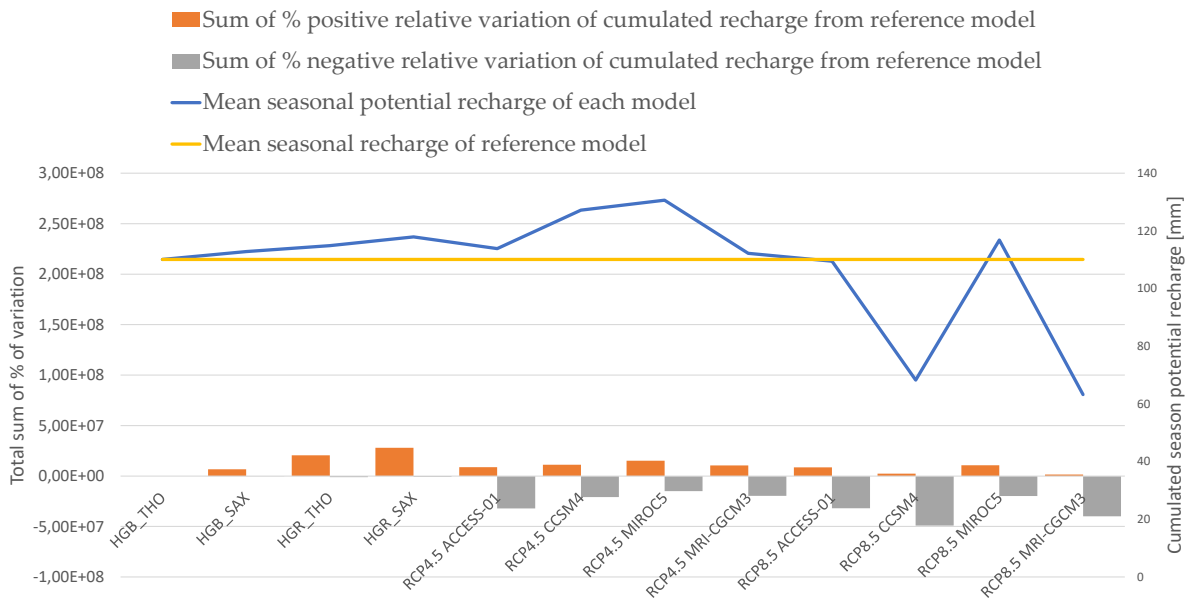
Figure 3.22: Sum of % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for spring season averaged over the twenty years of simulation.
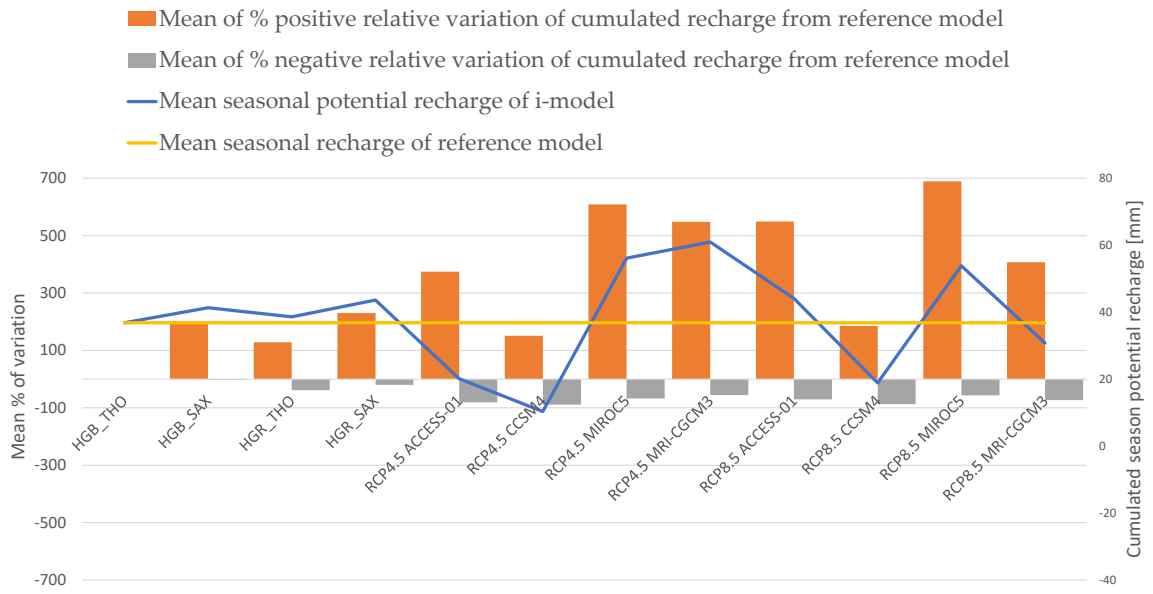
Figure 3.23: Mean % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for summer season averaged over the twenty years of simulation.
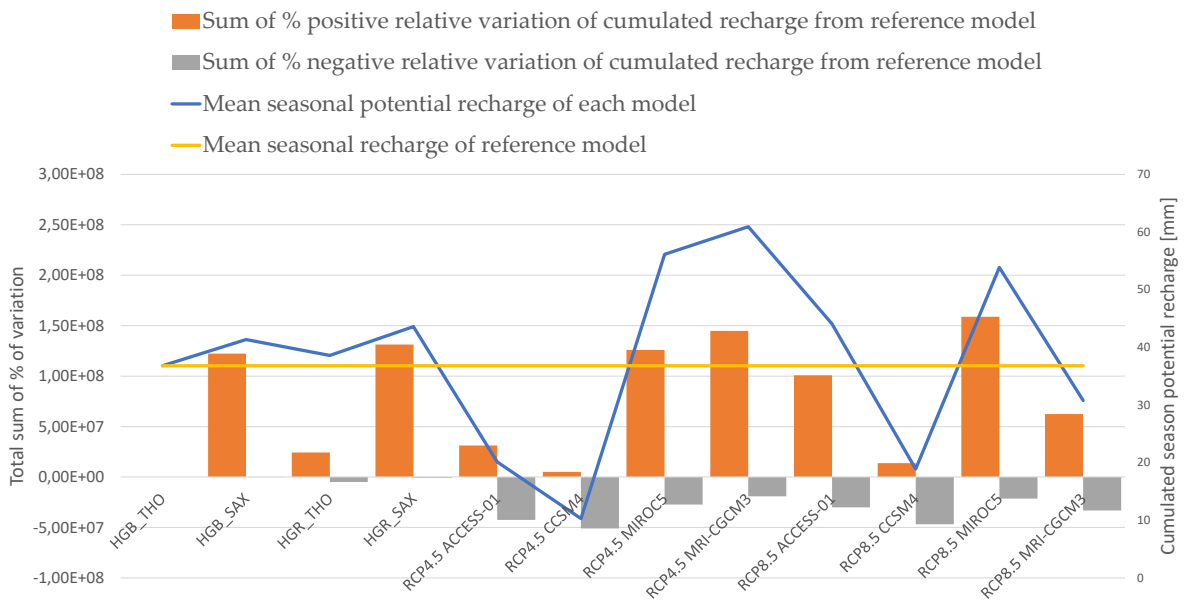
Figure 3.24: Sum of % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for summer season averaged over the twenty years of simulation.
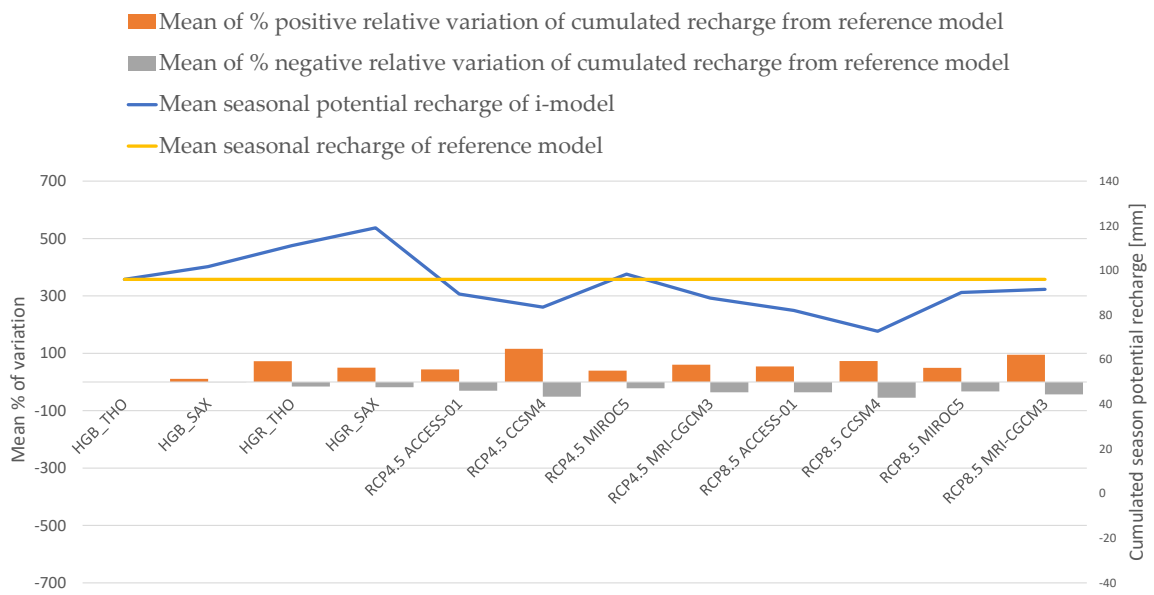
Figure 3.25: Mean % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for autumn season averaged over the twenty years of simulation.
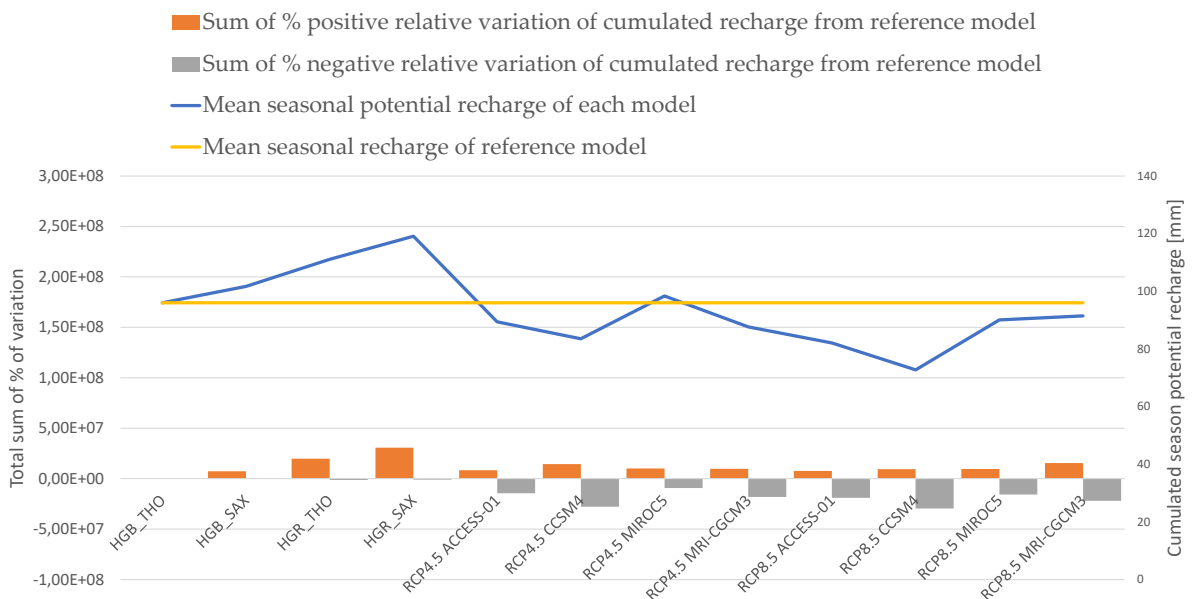
Figure 3.26: Sum of % positive and negative variations of estimates of potential recharge (with respect to the reference historical model HGB_THO) for autumn season averaged over the twenty years of simulation.

The seasonally cumulated potential recharge demonstrates great uncertainty in the model predictions, especially for GCMs forecasts that do not always agree on the direction of change. The winter season is the only one where all future models agree on an increase in mean recharge quantities with respect to historical conditions.

Summer season shows the largest variability and uncertainty in model results, while spring and autumn show some smaller variations in estimated potential recharge.

In general, winter and summer are the seasons more affected by the influence of future climate in terms of mean variation of potential recharge. Winter recharge analysis shows that all future scenarios will produce higher recharge quantities. This is seen both in the blue line representing seasonal cumulated potential recharge and in the orange bars representing the positive relative variation with respect to the reference model. In terms of mean increase, future climate scenario produce mean increases between 100% and 700%, while historical scenarios have smaller mean increases (the highest is 93% for HGR_SAX). In terms of sum of % increases, future climate models produce results that are higher than those of historical models: some models like the RCP8.5 MRI-CGCM estimate a sum of positive relative variation up to $2.71 \cdot 10^8$ % for the winter season (averaged over the 20 years of simulation); some other models as the RCP8.5 ACCESS-01 estimate a sum of positive relative variation one order of magnitude smaller ($2.65 \cdot 10^7$ %) that is comparable with the sum of positive relative variation observed for the historical models such as the HGR_SAX. Similar results are obtained for the analysis of summer potential recharge estimates: mean variation of potential recharge is on average higher for future climate models, but the sum of percentage of variation shows smaller signs of deviation of future models from the historical ones. In the summer season, models disagree more on the quantities of average cumulated recharge: as shown in Figure 3.25. The blue line of seasonally cumulated potential recharge shows that both RCP4.5 and RCP8.5 outputs present a lot of uncertainty and variability in the estimates of mean summer recharge. The increase of positive variations and the reduction of negative variations in winter is attributed to the general increase of temperatures analyzed in the paragraph 2.2.1 that might reduce winter snowpack and consequently increase winter runoff (as shown in Figure 3.11) and increase availability of water for potential recharge. In fact, the positive variation in the winter season is compensated by an increased negative relative variation in the spring season, where most of the future models produce significant reductions of recharge contrarily from the historical models. The autumn season shows increases and decreases of the same entity in the mean % variation of potential recharge. Despite this, almost all future models predict mean potential recharge that is lower than the historical models. The three historical models HGB_SAX, HGR_THO, HGR_SAX always produce positive % variations that are naturally higher for models executed with the Rosetta method.

Having identified the ambiguity of climate models in the definition of the direction of change of future potential recharge for spring and summer season, the focus is on winter and autumn season. In comparison to historical conditions, these two seasons show a respective increase and decrease in potential recharge. Therefore, the graphs below analyze the temporal evolution of mean seasonal potential recharge for the

winter and autumn season similarly to the analysis performed on mean annual recharge at the beginning of this chapter.
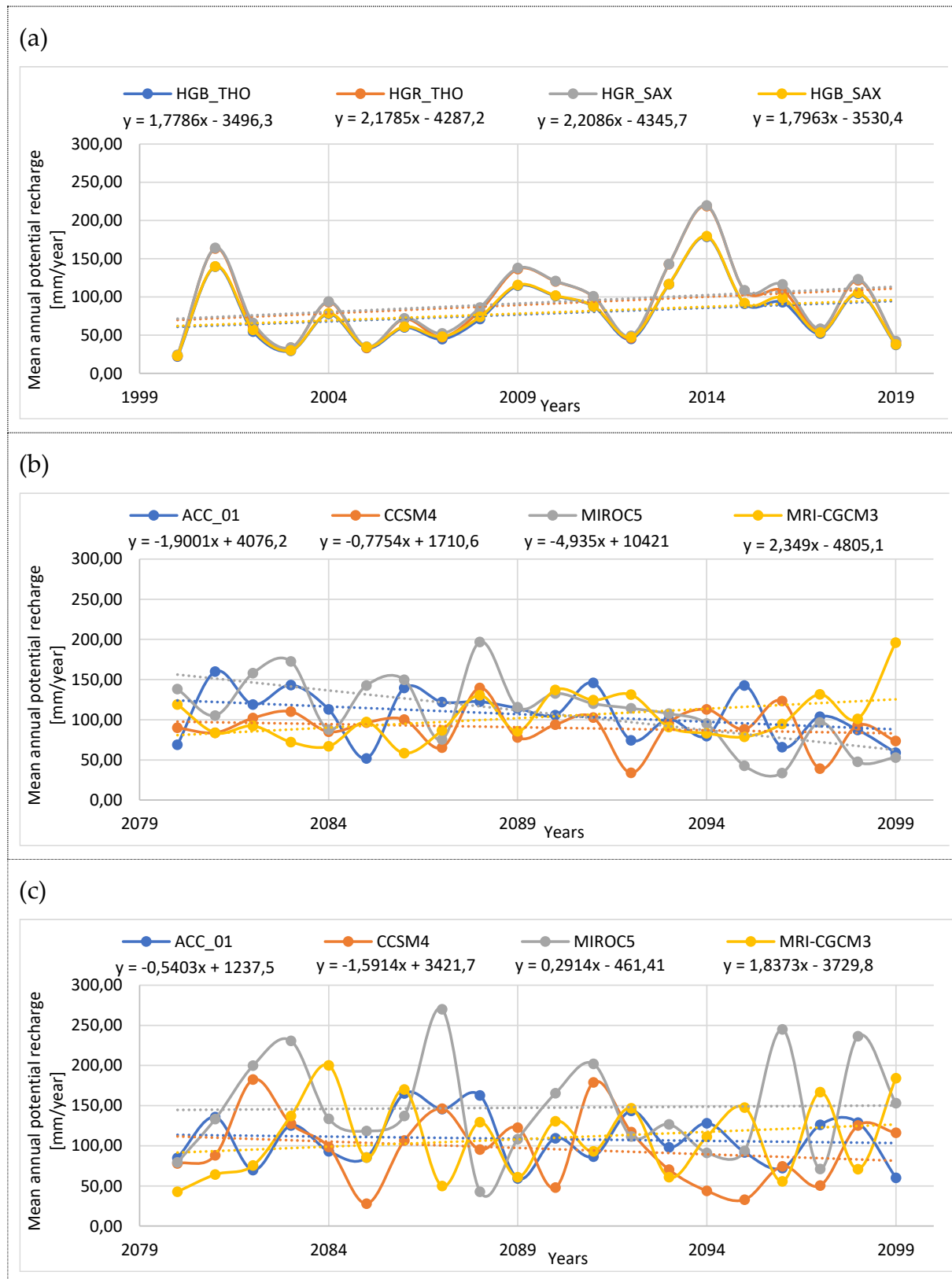
Figure 3.27: Evolution of mean winter potential recharge for (a) historical simulations, (b) RCP4.5 simulations, (c) RCP8.5 simulations.

Winter recharge is increasing during the period 2000-2019 as shown by the trendline in figure (a), while it shows little signs of consistent trends of long-term variation for the RCP4.5 and RCP8.5 climate models in figure (b) and figure (c). The increase of winter recharge as identified by linear interpolation has a magnitude between 1,77 mm/year and 2,20 mm/year, which is significantly high considering that the annual increase examined in Figure 3.1 was around 2.5-2.9 mm/year.

(a)
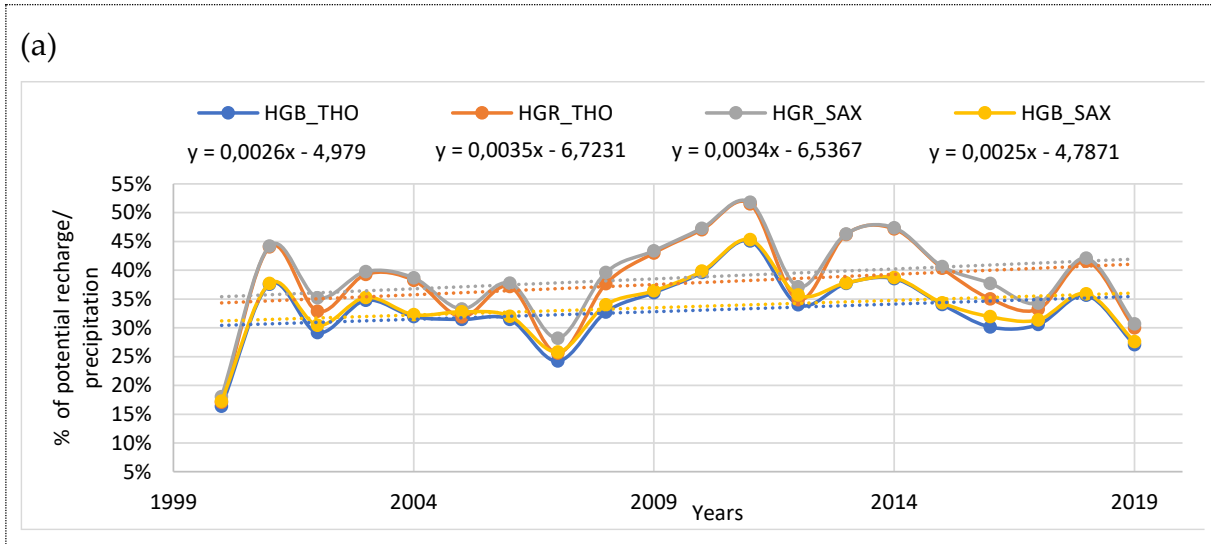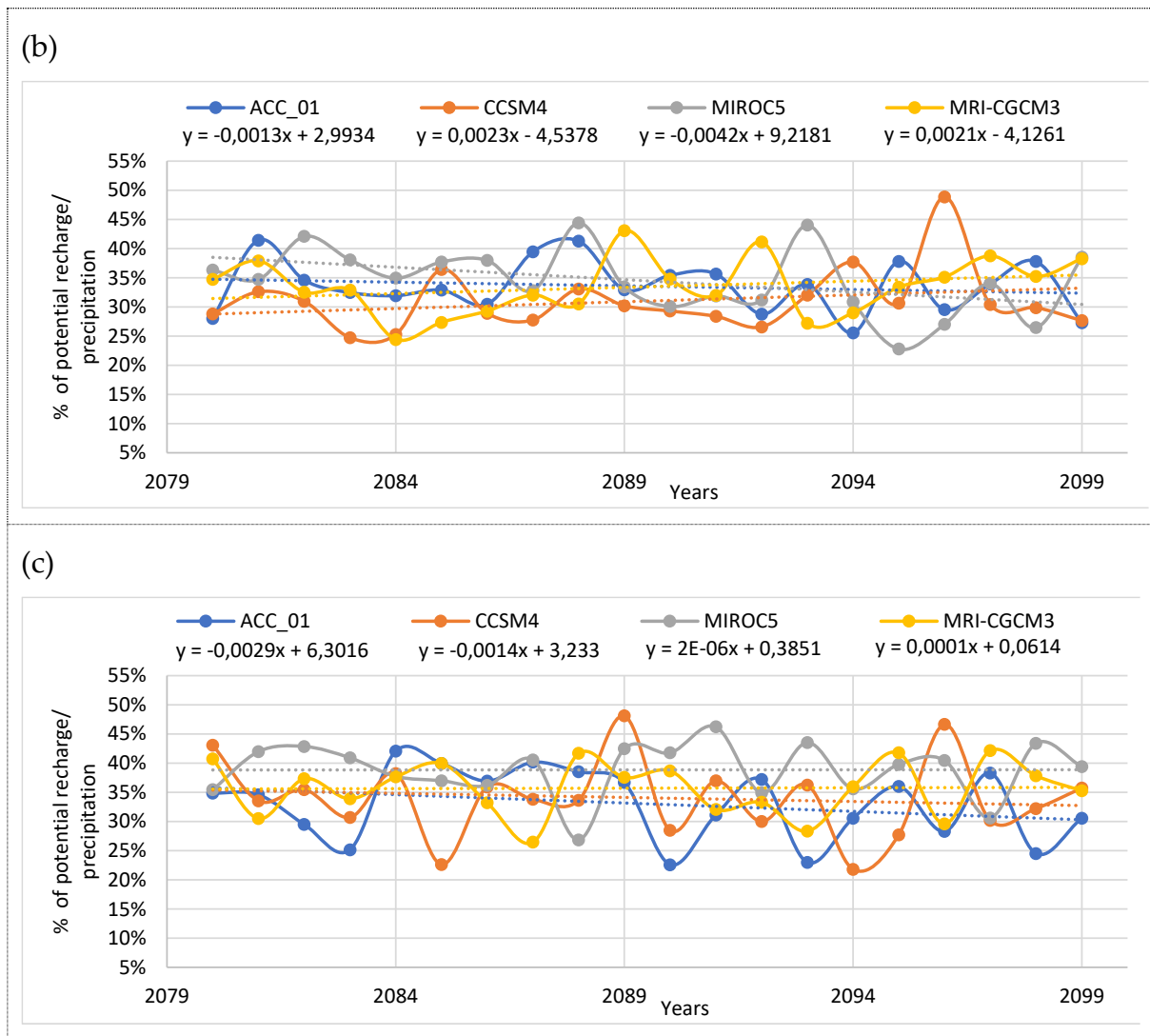
Figure 3.28: Evolution of ratio of winter potential recharge to winter precipitation for (a) historical simulations, (b) RCP4.5 simulations, (c) RCP8.5 simulations.

Ratios of potential recharge to total precipitation for the winter season are similar moving from historical climate simulations to future climate simulations: historical simulations present a higher range of variation of the ratios (min=16% to max=52%) with respect to the future climate models (min=22%, max=49%).

(a)
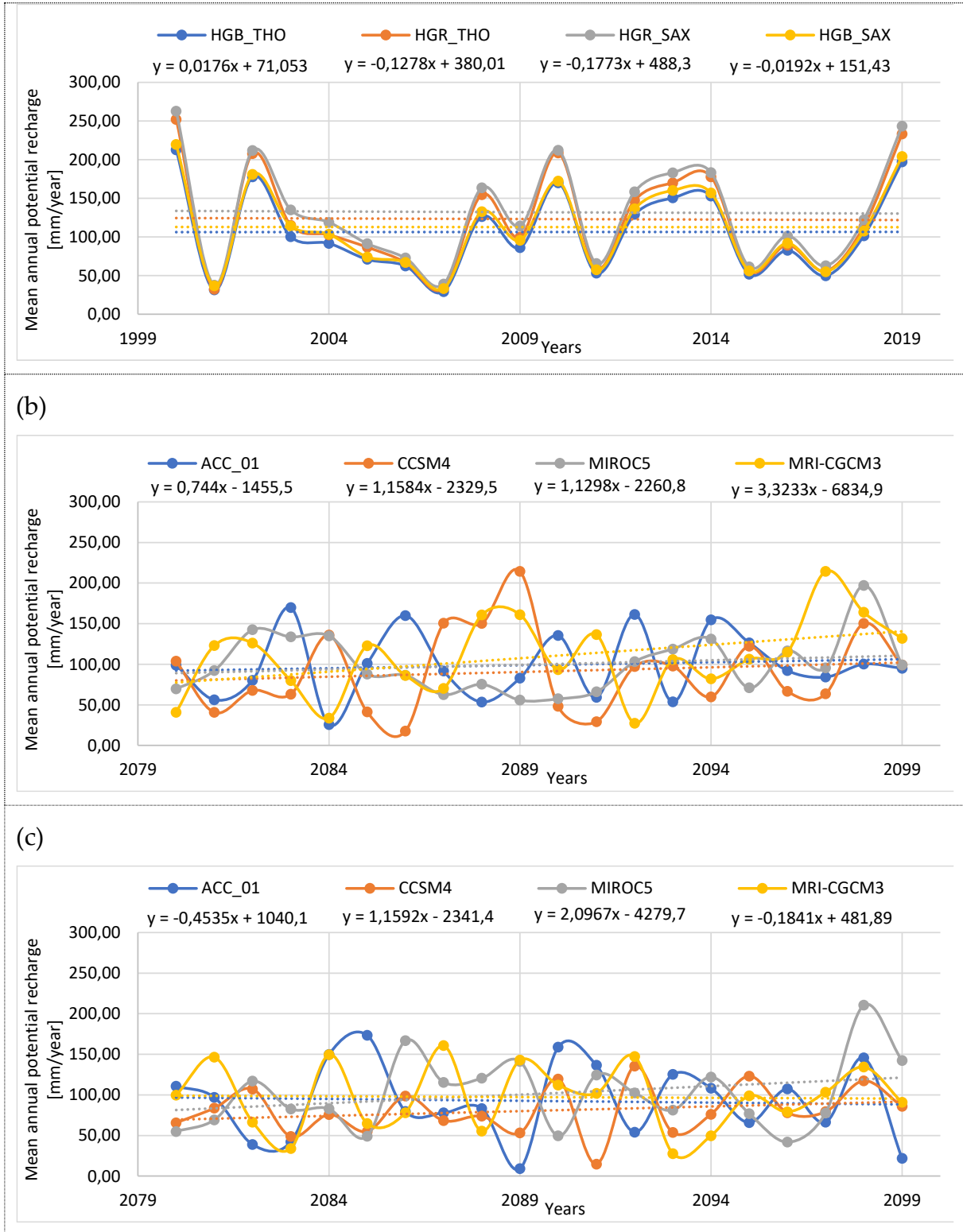
Figure 3.29: Evolution of mean autumn potential recharge for (a) historical simulations, (b) RCP4.5 simulations, (c) RCP8.5 simulations.
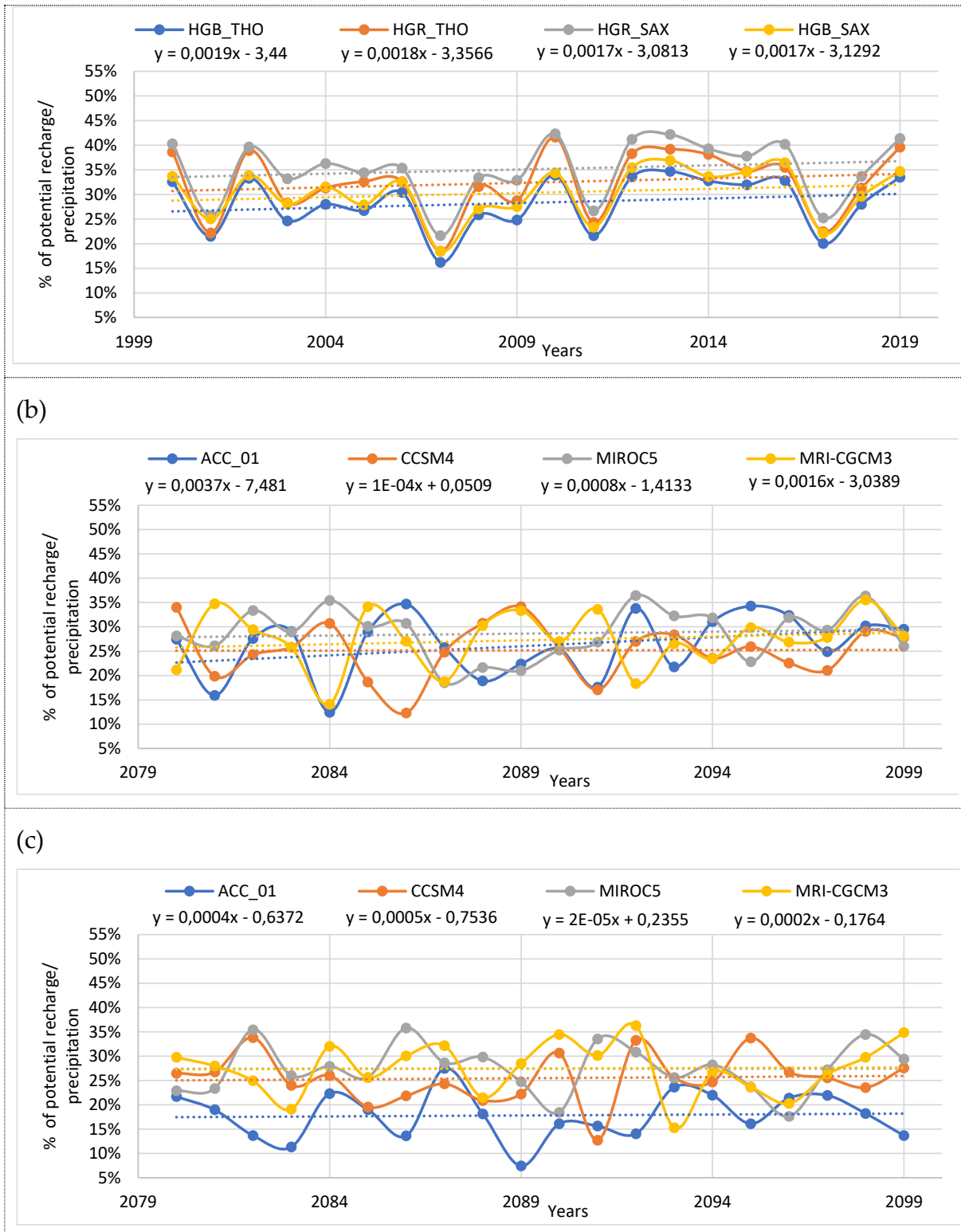
Figure 3.30: Evolution of ratio of winter potential recharge to winter precipitation for (a) historical simulations, (b) RCP4.5 simulations, (c) RCP8.5 simulations.

Differently from winter recharge, autumn potential recharge estimates show stationarity in the time range of historical simulations as there are no evident trends of variation.

Finally, we compare the examined estimates referring to the entire rectangular domain used for the simulations, to the estimate produced in the sole region of Lombardia. This is done by masking the simulation results with a raster defining the administrative borders of the region. The results are analyzed in Table 3.2.

Table 3.2: Comparison of results of mean annual recharge for rectangular simulation domain and for the administrative region of Lombardy.

SCENARIO	MODEL	MODEL DOMAIN AVERAGE POTENTIAL RECHARGE [mm/year]	PERCENTAGE OF VARIATION OF NET INFILTRATION ESTIMATE FROM REFERENCE SCENARIO (HGB_THO)	LOMBARDIA AVERAGE POTENTIAL RECHARGE [mm/year]	PERCENTAGE OF VARIATION OF NET INFILTRATION ESTIMATE FROM REFERENCE SCENARIO (HGB_THO)
HISTORICAL	HGB_THO	347,1	0,0	334,4	0,0
	HGR_THO	383,9	10,6	389,2	16,4
	HGB_SAX	362,9	4,6	349,6	4,6
	HGR_SAX	403,6	16,3	408,6	22,2
RCP4.5	ACCESS-01	353,7	1,9	321,5	-3,8
	CCSM4	351,7	1,3	315,3	-5,7
	MIROC5	398,5	14,8	359,2	7,4
	MRI-CGCM3	419,2	20,8	379,4	13,5
RCP8.5	ACCESS-01	353,6	1,9	307,2	-8,1
	CCSM4	273,9	-21,1	237,5	-29,0
	MIROC5	353,1	1,8	315,8	-5,6
	MRI-CGCM3	398,3	14,8	352,5	5,4

The table shows that shifting the analysis from the rectangular domain used for simulations to the Lombardy domain, future recharge estimates tend to reduce. In fact, it becomes more common to see for both RCP4.5 and RCP8.5 scenarios estimates of mean annual recharge that are lower than those produced by the reference HGB_THO. This type of result is expected considering the examined negative contribution of the plain area to the future estimates of recharge: compared to the rectangular area of simulation, the Lombardy Region has a greater fraction of plain areas (contributing negatively to the increase in recharge) than mountain area (contributing positively to the increase in recharge).

4 Conclusions

We investigated how soil characteristics and climate change projections might affect the potential groundwater recharge in the Lombardy region. Based on these estimates, analysis of the outputs provided relevant data to compare the influence of the sets of input on the output variance.

For all scenarios, the cumulated average annual recharge is not significantly affected by climate change: the historical mean value (evaluated in the interval 2000-2019) being 347.1 mm/year, and the mean value (evaluated in the interval 2080-2099) being 380.7 mm/year and 344.7 mm/year respectively for RCP4.5 and RCP8.5 emission scenarios. Varying the characteristics of the soil, annual potential recharge estimates increase from a minimum of 4.3% to a maximum of 16.6%. All investigated climate scenarios predict an increase of annual potential recharge with respect to the reference model in the interval 1.3% - 20.8% for the rectangular area analyzed by the simulations. The only exception to this is the RCP8.5 CCSM4 model which estimates a reduction of 21.1% of annual potential recharge with respect to the reference model. All investigated climate scenarios present a clearly visible spatial trend of increased recharge in the northern mountain areas of the region and decreased recharge in the southern parts, located in the Pianura Padana. This type of distribution has the effect of mitigating the increase of mean annual potential recharge (with respect to the historical trend) for future climate scenarios on the Regional scale but has a substantial effect on the variation of mean annual potential recharge on the sub-regional scale. For example, after clipping the output to only take into account the Region Lombardy's political boundaries, the number of future climate models predicting a decline in the mean annual potential recharge increases. This result can be attributed to the fact that Lombardy has a lower proportion of mountainous terrain (contributing positively to increased recharge) than the model rectangular domain. Furthermore, every future model predicts that the potential recharge during the winter will rise (in comparison to the historical trend), while recharge volumes during autumn will drop.

All models based on GCMs show an increase in mean seasonal recharge (averaged over 20 years of simulation 2080-2099) for the winter season and a decrease for autumns, but are inconsistent on the direction of change produced during other seasons of the year depending on the GCM employed. This confirmed the element of uncertainty rooted in climate prediction. In most cases (the only exception being the CCSM4 model) the mean annual volumetric increase of recharge in the mountain areas was able to exceed the decrease of recharge in the Pianura Padana and generate a total

trend of increased potential recharge with respect to the historical reference scenario. The variation of future estimates of potential recharge from historical conditions was more intense in the mountain areas (up to +650% in the winter) than in the lower plain areas: this suggests that, as already mentioned in the introduction, mountain areas will be the regions more sensitive to recharge under future climate conditions. A standard deviation study of model results reveals that mountainous regions are also the ones most impacted by uncertainty in the description of the extent of change to which they are subject in terms of potential recharge estimates.

It is important to recognize that due to the nature of the model and the nature of the inputs, the estimates obtained for potential groundwater recharge are approximate and the methodology is affected by some limitations. The SWB method requires several variables (AWC, HG, etc.) that are affected by uncertainty. Furthermore, no data are generally available on these variables for the area of interest, therefore they could not be calibrated, and we relied on information available in the literature. A source of additional error is also engrained in the physical model adopted to reproduce several processes. Reference evapotranspiration modelling has been conducted with a simple model (chapter 2.1.5) that only requires gridded climate data of maximum and minimum air temperatures. However, the most accurate reference evapotraspiration method is thought to be the FAO-56 Penman Monteith method: a combined temperature and energy-based approach that is not implemented in SWB as it requires gridded datasets on solar radiation, wind speed, and relative humidity [42]. Surface runoff has been evaluated adopting the widely used Curve Number method, i.e., without solving the coupled surface-subsurface flow problem.

However, using a discretely long study period of 20 years and a fine spatial resolution of 250 m, the results obtained can serve as valuable information for water management and future research.

Finally, the evidence from this study suggests that in order to improve long-time predictions on groundwater recharge it is necessary to focus on decreasing the inherent uncertainty of GCMs predictions, the impact of soil characteristics being minor. However, for the area studied the choice of geographical domain generated an exceptional singularity of dual increase and decrease of recharge that restrains the difference in the total average net infiltration volumes between future and historical simulations. Independently from this, the estimated future drop of potential recharge in the Pianura Padana and the increase in the northern mountains would provoke a change in the patterns of the hydrologic cycle that would have an impact on the ecosystem of the area.

Bibliography

[1]   Intergovernmental panel on climate change. Working group 1., V. (1971-. . . .). Masson-Delmotte, P. (19. .-. . . .). Zhai, and A. (19. .-. . . .). Pirani, *Summary for policymakers : working group I contribution to the sixth Assessment report of the Intergovernmental Panel on Climate Change*. 2021.

[2]   T. R. Green *et al.*, "Beneath the surface of global change: Impacts of climate change on groundwater," *Journal of Hydrology*, vol. 405, no. 3–4. pp. 532–560, Aug. 05, 2011. doi: 10.1016/j.jhydrol.2011.05.002.

[3]   Intergovernmental Panel on Climate Change, *Summary for policymakers: climate change 2014*. 2014.

[4]   A. Jakemann·, O. Randall, J. Huntt·, and J.-D. Andrewwross, "Integrated Groundwater Management Concepts, Approaches and Challenges," 2016.

[5]   I. P. Holman, D. M. Allen, M. O. Cuthbert, and P. Goderniaux, "Towards best practice for assessing the impacts of climate change on groundwater," *Hydrogeol J*, vol. 20, no. 1, pp. 1–4, 2012, doi: 10.1007/s10040-011-0805-3.

[6]   R. K. P. and L. A. M. (eds.) Core Writing Team, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland, 2014.

[7]   D. A. Randall *et al.*, "Climate Models and Their Evaluation," 2007.

[8]   F. Shi, Z. Wang, L. Qi, and R. Chen, "An assessment of GCM performance at a regional scale using a score-based method," *Advances in Meteorology*, vol. 2018, 2018, doi: 10.1155/2018/7641019.

[9]   Eckhardt K. and U. Ulbrich, "Potential Impacts of Climate Change on Groundwater Recharge and Streamflow in a Central European Low Mountain Range.," *J Hydrol (Amst)*, vol. 284, pp. 244–252, 2003.

[10]   F. H. S. Chiew and T. A. McMahon, "Modelling the impacts of climate change on Australian Streamflow.," *Hydrol. Processes*, vol. 16, pp. 1235–1245, 2002.

[11]   Mistry P.B. and Suryanarayana T. V. M., "Study of Groundwater Recharge Estimation Using Empirical Equations in North Gujarat, India," 2021. [Online]. Available: https://ssrn.com/abstract=3822573

[12]   D. Whittemore, "Best Estimates of Aquifer Recharge: Magnitude and Spatial Distribution," 2002. [Online]. Available: www.kgs.ukans.edu

[13] U.S. Department of Interior and U.S. Geological Survey, "SWB-A Modified Thornthwaite-Mather Soil-Water-Balance Code for Estimating Groundwater Recharge," Reston, Virginia, 2010.

[14] B. R. Scanlon, R. W. Healy, and P. G. Cook, "Choosing appropriate techniques for quantifying groundwater recharge," *Hydrogeol J*, vol. 10, no. 1, pp. 18–39, Feb. 2002, doi: 10.1007/s10040-001-0176-2.

[15] G. Pulighe, F. Lupia, H. Chen, and H. Yin, "Modeling climate change impacts on water balance of a mediterranean watershed using swat+," *Hydrology*, vol. 8, no. 4, Dec. 2021, doi: 10.3390/hydrology8040157.

[16] T. R. Green *et al.*, "Beneath the surface of global change: Impacts of climate change on groundwater," *Journal of Hydrology*, vol. 405, no. 3–4. pp. 532–560, Aug. 05, 2011. doi: 10.1016/j.jhydrol.2011.05.002.

[17] D. Al Atawneh, N. Cartwright, and E. Bertone, "Climate change and its impact on the projected values of groundwater recharge: A review," *Journal of Hydrology*, vol. 601. Elsevier B.V., Oct. 01, 2021. doi: 10.1016/j.jhydrol.2021.126602.

[18] R. S. Crosbie, T. Pickett, F. S. Mpelasoka, G. Hodgson, S. P. Charles, and O. V. Barron, "An assessment of the climate change impacts on groundwater recharge at a continental scale using a probabilistic approach with an ensemble of GCMs," *Clim Change*, vol. 117, no. 1–2, pp. 41–53, Mar. 2013, doi: 10.1007/s10584-012-0558-6.

[19] T. Meixner *et al.*, "Implications of projected climate change for groundwater recharge in the western United States," *Journal of Hydrology*, vol. 534. Elsevier B.V., pp. 124–138, Mar. 01, 2016. doi: 10.1016/j.jhydrol.2015.12.027.

[20] B. D. Smerdon, "A synopsis of climate change effects on groundwater recharge," *Journal of Hydrology*, vol. 555. Elsevier B.V., pp. 125–128, Dec. 01, 2017. doi: 10.1016/j.jhydrol.2017.09.047.

[21] C. J. Apaza-Coria, I. E. Rodriguez-Levy, M. Delfín Soruco, and M. Huysmans, "Determination of the natural potential groundwater recharge in the Valle Alto basin, Bolivia, through a soil water balance," *Hydrogeol J*, vol. 30, no. 8, pp. 2341–2357, Dec. 2022, doi: 10.1007/s10040-022-02553-x.

[22] A. Hughes *et al.*, "The impact of climate change on groundwater recharge: National-scale assessment for the British mainland," *J Hydrol (Amst)*, vol. 598, Jul. 2021, doi: 10.1016/j.jhydrol.2021.126336.

[23] J. R. Raposo, J. Dafonte, and J. Molinero, "Assessing the impact of future climate change on groundwater recharge in Galicia-Costa, Spain," *Hydrogeol J*, vol. 21, no. 2, pp. 459–479, Mar. 2013, doi: 10.1007/s10040-012-0922-7.

[24] F. Behulu and M. Roma, "Modelling the Hydrologic Behavior of the Upper Tiber River Basin under Climate Change Scenarios," 2013.

[25] G. Braca, M. Bussettini, D. Ducci, B. Lastoria, and S. Mariani, "Evaluation of national and regional groundwater resources under climate change scenarios using a GIS-based water budget procedure," *Rendiconti Lincei*, vol. 30, no. 1, pp. 109–123, Mar. 2019, doi: 10.1007/s12210-018-00757-6.

[26] D. J. Hart, P.R. Schoephoester, and K.R. Bradbury, "Groundwater recharge in Dane County, Wisconsin: Estimating recharge using a GIS-based water-balance model," *Wisconsin Geological and Natural History Survey Bulletin*, vol. 107, 2012.

[27] Andrea Manzoni, "Large basin scale recarge rate," *Github Repository https://github.com/manzoniandrea/Large_Basin_Scale_Recarge_Rate.git*, 2023.

[28] Globalgeografia, "Lombardia," *globalgeografia.com*, 2023.

[29] Hyperfvg, "Alta e bassa pianura padana," *http://www.hyperfvg.org/fvg/alta_bassa_pianura.html*, 2023.

[30] Nova.news, "Drought: Lombardy closes 2022 with -30 percent of water resources," 2022. Accessed: Mar. 10, 2023. [Online]. Available: https://www.agenzianova.com/en/news/siccita-la-lombardia-chiude-il-2022-con-un-30-per-cento-delle-risorse-idriche

[31] Copernicus climate change service, "ERA5 data from 1950 to present," 2022.

[32] B. Thrasher, E. P. Maurer, C. McKellar, and P. B. Duffy, "Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping," *Hydrol Earth Syst Sci*, vol. 16, no. 9, pp. 3309–3314, Sep. 2012, doi: 10.5194/hess-16-3309-2012.

[33] Y. Ruan, Z. Yao, R. Wang, and Z. Liu, "Ranking of CMIP5 GCM Skills in Simulating Observed Precipitation over the Lower Mekong Basin, Using an Improved Score-Based Method," *Water (Basel)*, vol. 10, no. 12, p. 1868, Dec. 2018, doi: 10.3390/w10121868.

[34] Copernicus Land Monitoring Service, "Corine Land Cover Inventory 2022," *European Environment Agency*.

[35] Thornthwaite and Mather, "Instructions tables for computing potential evapotranspiration water balance," 1957.

[36] T. Hengl *et al.*, "SoilGrids250m: Global gridded soil information based on machine learning," *PLoS One*, vol. 12, no. 2, p. e0169748, Feb. 2017, doi: 10.1371/journal.pone.0169748.

[37] M. G. Schaap, F. J. Leij, M. T. Van Genuchten, and G. E. Brown, "Rosetta: a computer program for estimating soil hydraulic parameters with hierarchical

pedotransfer functions," 2001. [Online]. Available: http://www.ussl.ars.usda.gov/.

[38] U. Nrcs, "Soil Survey Manual Soil Science Division Staff Agriculture Handbook No. 18."

[39] C.T.B. SungJ. Iba, "Accuracy of the Saxton-Rawls method for estimating the soil water characteristics for mineral soils of Malaysia".

[40] L. Heaton, M. A. Fullen, and R. Bhattacharyya, "Critical Analysis of the van Bemmelen Conversion Factor used to Convert Soil Organic Matter Data to Soil Organic Carbon Data: Comparative Analyses in a UK Loamy Sand Soil," *Espaço Aberto*, vol. 6, no. 1, pp. 35–44, Jun. 2016, doi: 10.36403/espacoaberto.2016.5244.

[41] B. R. Scanlon, R. C. Reedy, D. A. Stonestrom, D. E. Prudic, and K. F. Dennehy, "Impact of land use and land cover change on groundwater recharge and quality in the southwestern US," *Glob Chang Biol*, vol. 11, no. 10, pp. 1577–1593, Oct. 2005, doi: 10.1111/j.1365-2486.2005.01026.x.

[42] Stephen M. Westenbroek, John A. Engott, Victor A. Kelson, and Randall J. Hunt, "SWB Version 2.0-A Soil-Water-Balance Code for Estimating Net Infiltration and Other Water-Budget Components," 2018.

| Appendix A: Comparison of
groundwater depth data and simulated
recharge estimates.

107

A   Appendix A: Comparison of groundwater depth data and simulated recharge estimates.

The results of the reference SWB simulation HGB_THO have been compared with the actual measurements of groundwater level fluctuation. In particular, SWB net infiltration estimates have been reorganized in monthly cumulated estimates being that groundwater data were available with a monthly frequency. While some wells showed groundwater level fluctuations that couldn't be in any way related to the SWB estimates, some other produced interesting results. The well identified by the code PO0161770U0001, in the village of Pumenengo (BG) located at the coordinates 567748 east and 5036468 north, showed a moderate correlation between Groundwater depth and simulated net infiltration. Extracted net infiltration estimates from SWB output between 2002 and 2020 in the model grid cell containing the well are plotted in Figure A.1.
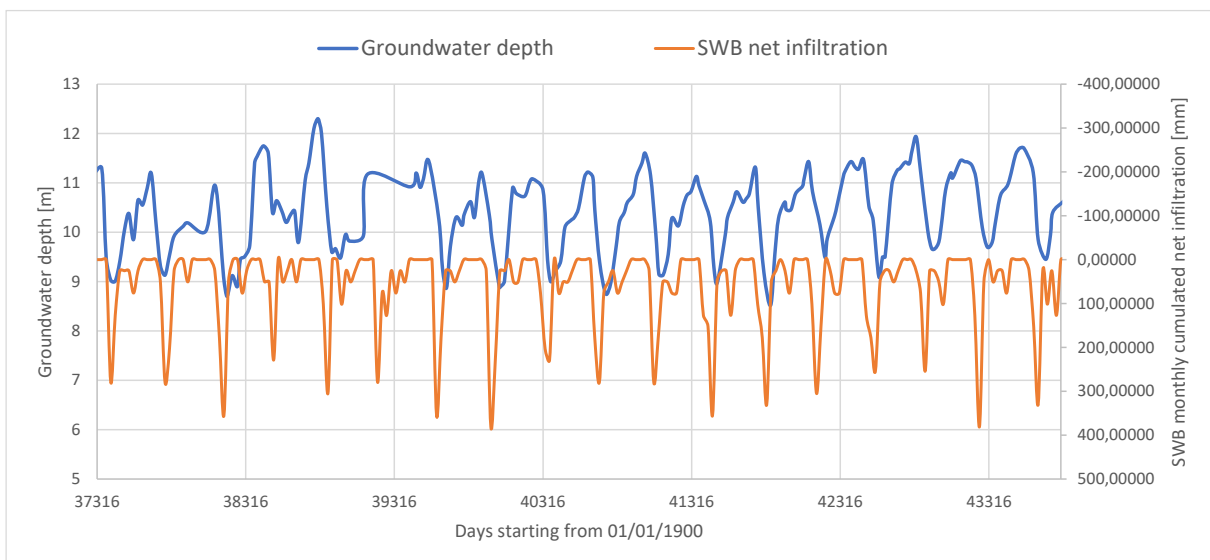
Figure A.1: Comparison of measure groundwater depth and SWB simulated net infiltration.

| Appendix A: Comparison of
groundwater depth data and simulated
recharge estimates.

108

The two quantities are expected to have a negative correlation. This trend is fairly respected in the graph: good performance is observed especially for the temporary distribution between peaks and lows of the two quantities. Finding an approximated value of the lag between the point at which net infiltration is recorded and the aquifer response, represented by the change in groundwater level, is a challenging task. Ideally, this could be performed with an infiltration process analysis, but the data on this matter are too scarce for the location analyzed. Therefore, an easier statistical approach is performed maximizing the absolute value of the correlation coefficient calculated with equation (1.1) while varying the response delay of groundwater depth variation to net infiltration estimates. In other words, the series of data related to groundwater depth are referenced to a specific day that remains unchanged, but the series of data related to SWB infiltration is shifted forward in time to fill the unknown response delay time attributable to the period of infiltration of water through the soil before reaching the groundwater table.

The graph in Figure A.2 reports the analysis of the variation of correlation coefficient on a range of response delay values ranging from 0 to 81 days. Results show that the comparison of depth to net infiltration estimates produces a minimum correlation coefficient of −0.565 for a response delay of 43 days.
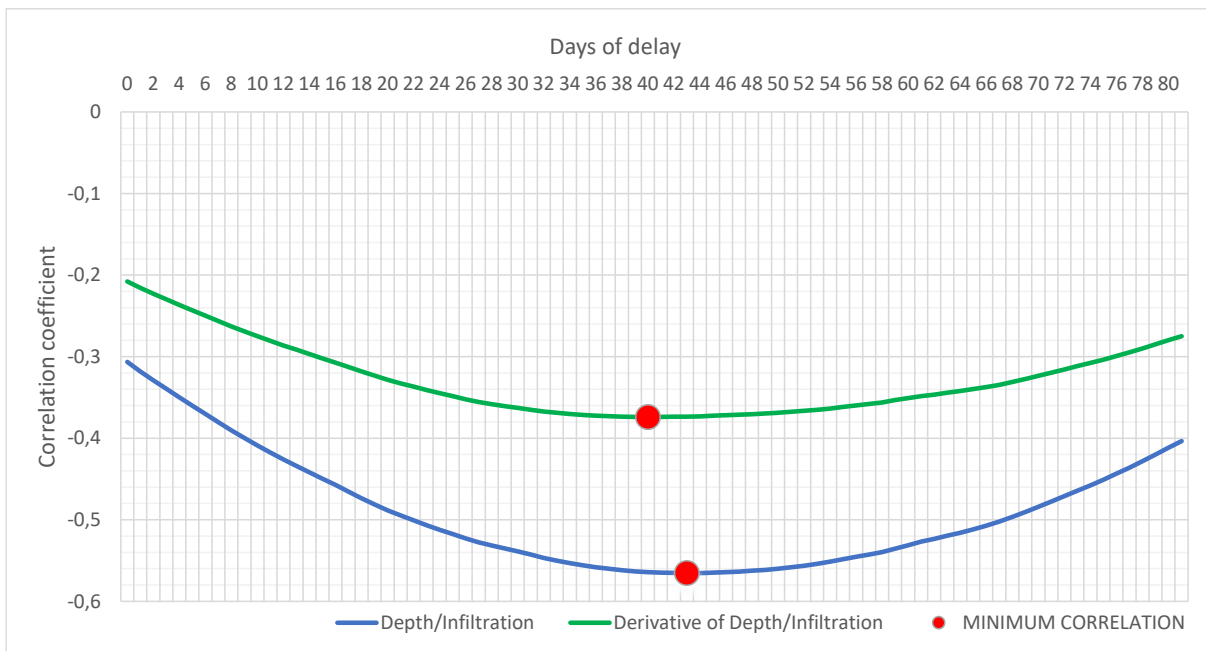
Figure A.2: Variation of correlation coefficient depending on response delay.

A promising result of this type of approach is that the graph shows that by varying the response delay from 0 days to 43 days, the correlation coefficient of the two series is almost doubled. Similar results are produced by comparing with the same method the

| Appendix A: Comparison of groundwater depth data and simulated recharge estimates.

109

derivatives of the two sets of data: a minimum correlation is found for a delay of 40 days with a minimum correlation coefficient of $-0,374$. The effect of the two delay times of 0 and 43 days is presented on two scatter plots of groundwater depths to SWB net infiltration estimates in the following graphs.
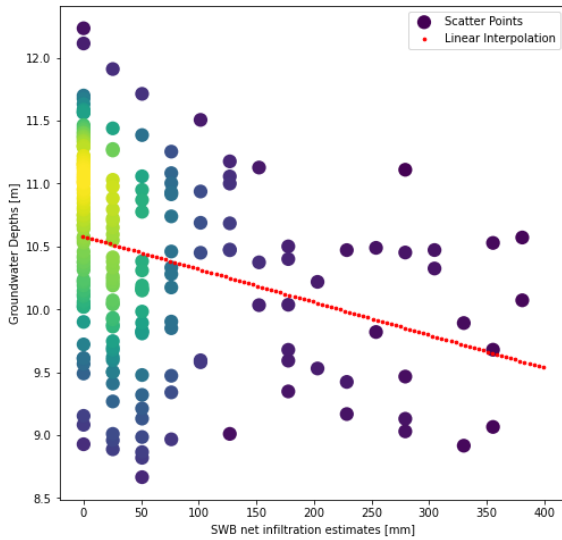
(a) Response delay of 0 days;
Interpolating equation $y = -0.0026\,x + 10.582$; $R^2 = 0.094$
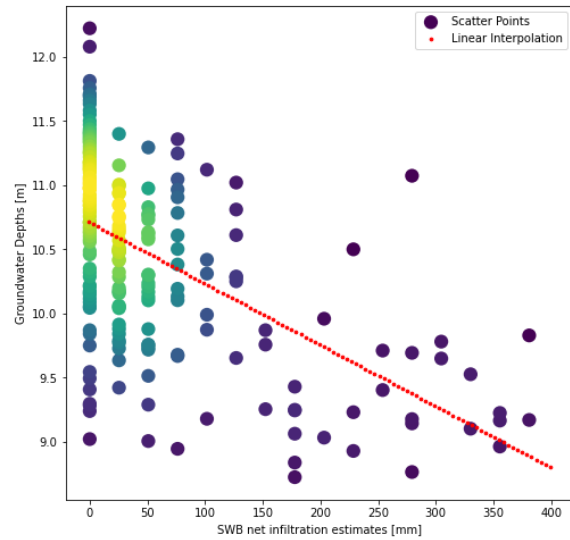
(b) Response delay of 43 days;
Interpolating equation $y = -0.0048\,x + 10.71$; $R^2 = 0.3196$

Figure A.3: Scatter plots of Groundwater depth compared to SWB net infiltration estimates.

The two estimations of delay time (40 and 43 days) are quite long considering that the groundwater depths of the well are around 8-10 meters below surface level. The conclusion is that an analysis of this type performed on monthly values might be too approximate and to provide consistent conclusions would require finer input data on the temporal distribution of water depths.

List of Figures

List of Tables

List of symbols

Variable	Description	SI unit
AWC	Available water capacity	cm/m
HG	Hydrologic Group	-
GCM	General Circulation Model	
SM	Soil moisture	mm/m
ET	Evapotranspiration	mm
CN	Curve number	-
R	Runoff	mm
P	Precipitation	mm
S_{max}	Maximum soil storage	mm
I_a	Initial abstraction	mm
ET_0	Reference evapotranspiration	mm
T_{max}	Maximum temperature	°C
T_{min}	Minimum temperature	°C
Ra	Solar radiation	W/m2
K_{sat}	Saturated hydraulic conductivity	mm/h
Θ_{FC}	Field capacity	cm/m
Θ_{WP}	Wilting point	cm/m
Θ_{33t}	Soil water availability for head pressure of 33 kPa	-
Θ_{1500t}	Soil water availability for head pressure of 1500 kPa	-
S	Sand content of the soil	mass %
C	Clay content of the soil	mass %
OM	Organic matter content of the soil	mass %
OC	Organica carbon content of the soil	mass %

SIMULATION ACRONYM		EMPIRICAL MODEL FOR HYDROLOGIC GROUP EVALUATION		EMPIRICAL MODEL FOR AVAILABLE WATER CAPACITY EVALUATION		CLIMATE MODEL
		Rawls and Brakensiel	Rosetta	Thornthwaite and Mather	Saxton and Rawls	
HGB_THO		✕		✕		ERA5
HGR_THO			✕	✕		ERA5
HGB_SAX		✕			✕	ERA5
HGR_SAX			✕		✕	ERA5
ACC-01	RCP4.5	✕		✕		ACCESS1-0 RCP4.5
	RCP8.5	✕		✕		ACCESS1-0 RCP8.5
CCSM4	RCP4.5	✕		✕		CCSM4 RCP4.5
	RCP8.5	✕		✕		CCSM4 RCP8.5
MIROC5	RCP4.5	✕		✕		MIROC5 RCP4.5
	RCP8.5	✕		✕		MIROC5 RCP4.5
MRI-CGCM3	RCP4.5	✕		✕		MRI-CGCM3 RCP4.5
	RCP8.5	✕		✕		MRI-CGCM3 RCP8.5