Royal Netherlands
Meteorological Institute
Ministry of Infrastructure
and Water Management

Improving the GRADE weather generator by using synthetic datasets from RACMO and SEAS5

L. van Voorst, H. van den Brink

De Bilt, 2022 | Technical report; TR-398

Improving the GRADE weather generator by using synthetic datasets from RACMO and SEAS5

Leon van Voorst
Henk van den Brink

March 2022

Abstract

The weather generator (WG) as currently used in GRADE[1] shows several methodical imperfections, potentially leading to either under- or overestimation of large precipitation extremes. This is caused by the relatively short lengths of the observational records used as source data for the weather generator. A potential solution is to replace the presently used observational dataset with much longer datasets, generated by weather and/or climate models like RACMO and SEAS5. The goal of this study is to remove the bias from the RACMO and SEAS5 datasets and compare them with the current source dataset and the timeseries calculated by the WG either using the current source dataset as input or RACMO and SEAS5 datasets respectively,, with special attention on their performance to describe meteorological (high) extremes. The dataset comparison is executed for the Rhine, Meuse and Vecht basins, which differ in size and consequently in hydrological response time.

A qualitative comparison of the datasets based on several climate variables shows good similarities for all basins. Quantitative differences have been corrected by means of quantile mapping, leading to comparable results in all catchments and enabling use of RACMO and SEAS5 data in quantitative analysis as well.

Comparison of extreme multi-day precipitation conditions between the datasets shows excellent climatological agreement up to return periods of 65 years (i.e. within the length of the available observations). An important finding is that, for larger return periods, multi-day precipitation extremes are higher in the RACMO and SEAS5 dataset compared to the WG associated to the observational dataset (except the 10-day precipitation extremes in the Rhine basin). This strongly hints on underestimation by WG based on observational records. The underestimation decreases for sums over more days and increases for larger return periods. The effect is visible in all basins, but is more pronounced in smaller basins (like the Vecht). If this underestimation is passed on to the hydrological results, the current GRADE procedure may lead to a false sense of security regarding flood hazard.

Comparison of summer and winter extremes between the dataset shows considerably larger summer precipitation extremes in RACMO and SEAS5 compared to the WG associated to the observational dataset for all basins, particularly in the Vecht basin. Annual extremes in RACMO and SEAS5 consist of a larger fraction of summer events than the WG associated to the observational dataset indicates. This implies that the assumption that extreme precipitation events mainly occur during winter has to be reconsidered.

Spatial and temporal comparison of the datasets shows that RACMO is able to generate intense precipitation events in summer. Methodically the WG is not capable of simulating the extreme daily peaks that characterise these events, as the concept limits the extreme 1-day extremes to the observational events. This implies that the effect of extreme summer events (which seems to be essential in smaller basins) on hydrology can better be analysed using the RACMO dataset than the WG. Additionally, the RACMO (and SEAS5) dataset allow for spatial analysis of extreme events, which is impossible using the WG dataset.

[1]Generator of Rainfall And Discharge Extremes.

A preliminary analysis of the hydrological discharges of RACMO and SEAS5 in the Meuse basin shows that the WG results associated to the observational dataset also underestimate hydrological high extremes. A first visual inspection shows over an order in magnitude difference in return period between RACMO and the WG result slightly smaller difference between SEAS5 and the WG. Based on the meteorological results, the differences are expected to be even larger in the Vecht basin. This shows the importance of further investigation of the hydrological results of RACMO and SEAS5 in all basins.

Regarding the results of this study, extra attention is required for the magnitude of summer events, both in hydrological and meteorological perspective.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The geographic location of the Netherlands stresses the importance of the nation's everlasting flood protection challenge. Being located in the delta of multiple rivers such as the Rhine, Meuse and Vecht and containing multiple largely populated areas of high economic value, prevention against flooding is one of the countries key tasks. The Dutch Ministry of Infrastructure and Water Management and its executive body Rijkswaterstaat are responsible for protection of the Dutch hinterland. For assessment of the existing protection system, probabilities of exceedance have been defined in the Waterwet. These maximum acceptable flood probabilities are directly related to discharge extremes, which for this report and for convenience are called design discharges. Correct coupling of design discharges to flood probability (or corresponding return periods) is therefore essential for the Dutch flood protection.

1.1 State of the art

Historically, estimation of design discharges was performed by fitting different types of probability distributions to historical data and performing statistical extrapolation (Hegnauer et al., 2014). These historical records originate from the start of the 20^{th} century and are therefore limited to a length of order 100 years for the main rivers, and less than 40 years for the Dutch-German Vecht river. As a result, extrapolations based on different probability distributions vary, whereas the choice of such a distribution is ambiguous. Furthermore, this statistical method is not capable of describing physical behaviour such as the effects of upstream flooding on a flood wave, changes of river geometry over time, changes in the river basin (e.g. land use) over time and climate change.

In order to generate a more reliable and physically-based approach for estimating design discharges, a consortium consisting of Deltares and the Royal Netherlands Meteorological Institute (KNMI), commissioned by Rijkswaterstaat, designed the GRADE (Generator of Rainfall And Discharge Extremes) instrument (Hegnauer et al., 2014). GRADE consists of three components, which are presented in Figure 1.1. Initially, a weather generator (WG) is used to generate 50,000 years of daily precipitation, temperature and potential evaporation by nearest-neighbour resampling from historic weather observations (see section 2.2.1 for more details). The result of the weather generator is

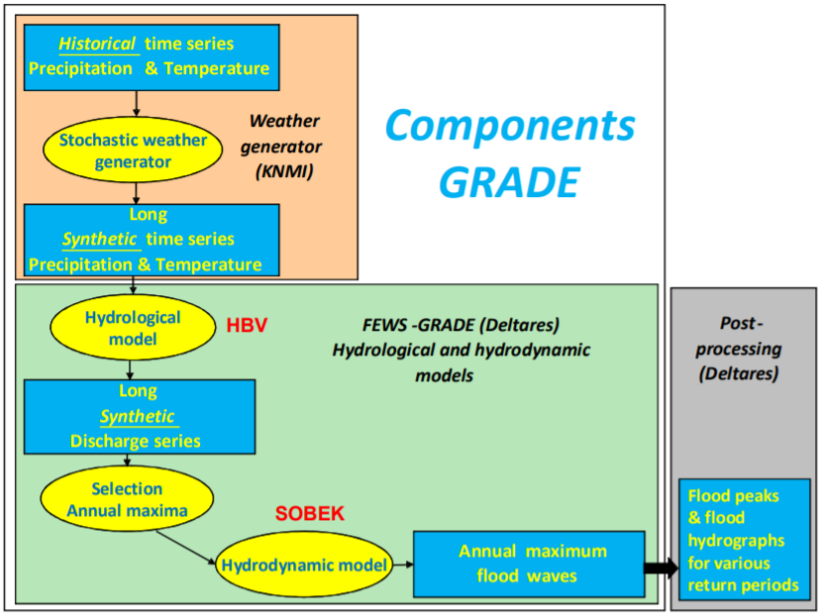

Figure 1.1: Schematisation of the GRADE components (Hegnauer et al., 2014)

then used as input for a rainfall runoff model (e.g. HBV[1]) in order to create a timeseries of 50,000 years of discharges. Annual extremes from this dataset are subsequently routed through the river stretches by using the SOBEK[2] hydrodynamic model, resulting in a time series of yearly flood peaks and associated hydrographs with their corresponding return periods.

Due to the largely increased length of the resulting discharge timeseries, statistical extrapolation is no longer necessary. Besides, multiple physical effects are accounted for. The rainfall-runoff model incorporates catchment characteristic effects such as land-use, elevation etc. (e.g. Mirus and Loague, 2013; Desta et al., 2019). Furthermore, use of a hydrodynamic model provides opportunity for including the physical characteristics of (planned) hydraulic infrastructure and the flooding of the area behind the dikes along the main river.

The WG is an essential element of GRADE, because it stochastically creates a synthetic timeseries that is much longer than the historical weather datasets that exist of approximately 70 years (e.g. Cornes et al., 2018b). Therefore, without the WG, generation of physically-based design discharges in GRADE would not be possible. A realistic description of extreme weather events up to return periods of 50,0000 years is thus fundamental for generating reliable discharge extremes.

As depicted in Figure 1.2, the WG is not capable of providing different daily rainfall amounts than those that are available in the input dataset. However, due to the creation of new temporal patterns by resampling, multi-day precipitation amounts in the synthetic dataset can and will exceed those in historic observations. According to e.g. Buishand (2007) and Beersma and Buishand (2007), this resampling procedure leads to a more accurate estimation of multi-day precipitation extremes and cumulative rainfall deficits.

[1] https://en.wikipedia.org/wiki/HBV_hydrology_model
[2] https://www.deltares.nl/en/software-solutions/sobek-and-delft3d/

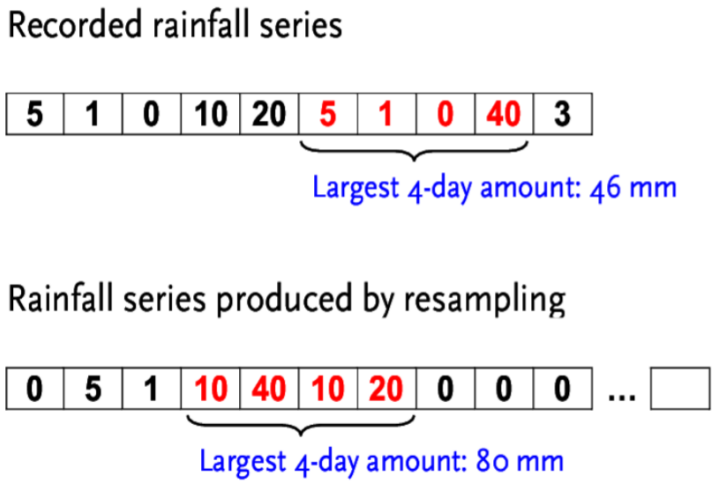

Figure 1.2: Explanation of the resampling methodology in the weather generator (Hegnauer et al., 2014)

A more detailed description of the WG, including its incorporation of the spatial patterns and seasonal variation, is provided in Hegnauer et al. (2014).

1.2 Problem statement

The major strength of the GRADE instrument is its ability to generate a synthetic daily discharge timeseries of 50,000 years[3], which is long enough for analysis of extremes. As the WG provides the input for the hydrologic and hydrodynamic models, it has a crucial position in the generation of reliable results. However, the current operational setup has several limitations. In sections 1.2.1, 1.2.2 and 1.2.3 these limitations will be explained in detail.

1.2.1 Problem 1: Underestimation of precipitation extremes

The first limitation of the GRADE WG is the risk of underestimation of extreme values due to the probable absence of high extremes in the relatively short historical observation timeseries. This is illustrated in Figure 1.3, where yearly extremes and their corresponding return periods are shown for different multi-day precipitation sums (ascending from 1 to 10 days). Note that both the historical and observation-based so-called observation-based passive dataset (see Section 2.2 for more information) and its WG equivalent are presented for the Rhine and Vecht basin[4].

All lines considering a multi-day sum are rising approximately linearly for return periods up to approximately 100 years. However, for larger return periods, the graphs of the short-period multi-day precipitation sums flatten out and will eventually even cap off. The flattening effect is most pronounced for 1- and 2-day sums.

[3]Note that this timeseries is stationary, and resembles the climate of the source data

[4]In this figure, the extremes are represented as a so-called Gumbel plot, which will be explained later.

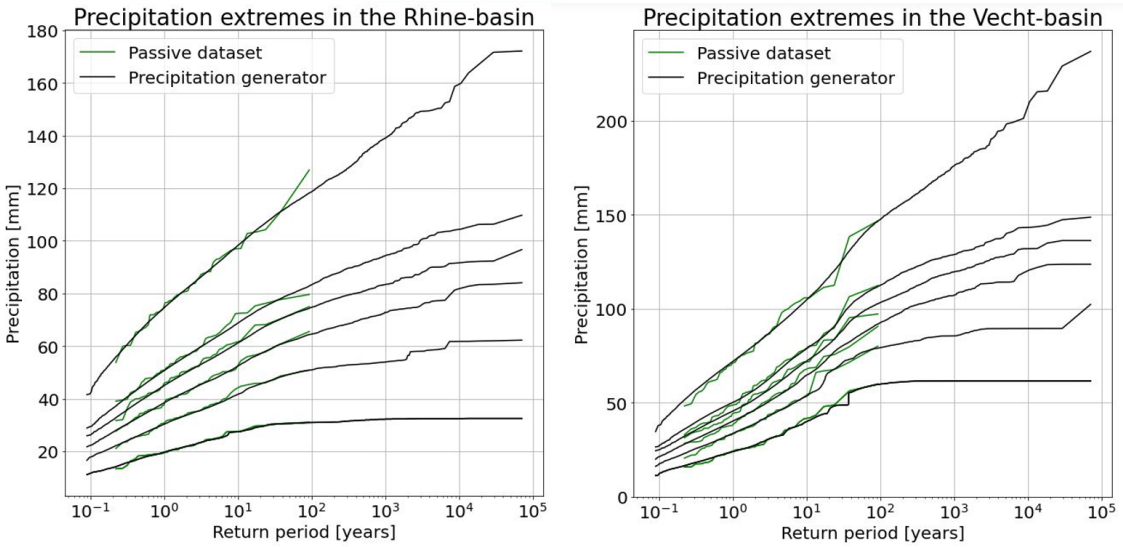

Figure 1.3: Gumbel plot for Rhine and Vecht, for the observation-based passive dataset (see Section 2.2.1) and the WG, considering a 1,2,3,4,5 and 10-day precipitation sum (in ascending order). Note that the horizontal axis has a logarithmic scale.

The effect displayed in Figure 1.3 is caused by the fact that the multi-day sum of the WG can never exceed the corresponding multitude of the maximum event in the original observation-based passive dataset. For example, the maximum 2-day sum generated with the recorded rainfall series in Figure 1.3 will never exceed 80 mm (twice the maximum observed event of 40 mm). In general, the n-day sum of the WG can never exceed n times the maximum daily precipitation in the historical record. As such, for large return periods and for short-period multi-day sums, the WG is unable to generate larger events, resulting in flattening of the curves in the Figure. Precipitation events linked to high return periods are therefore expected to be underestimated for the WG, particularly for short-period multi-day sums. This means that the synthetic dataset created with the WG is expected to consist of underestimated extremes, which could potentially influence the hydrological response of the GRADE instrument.

The main cause of the problem is the size of the historical observation datasets. A large observational dataset is more likely to contain high extremes, which enables the WG to resample higher extremes compared to a smaller observational dataset. With increasing size of the observational datset, the WG will therefore perform better. Unfortunately, accurate precipitation observations have only been performed for approximately 70 years with an adequate measurement network density, making this problem difficult to solve.

Note that for longer-period multi-day precipitation sums, the underestimation effect diminishes. This makes sense, because a longer-period multi-day sum has more possible resampling combinations and a higher maximum value. Besides, for a long-period multi-day sum, individual days do not require to be very extreme to still lead to a multi-day extreme, which means there is a larger possible variation in the resampling process. However, for what multi-day sum the underestimation is negligible is uncertain: for very large return periods the effect may still be present.

It is important to note that the response time of a river on extreme precipitation

depends on the size of the basin. The discharge of the Rhine at Lobith will therefore correlate to 6-10 day precipitation sums, whereas the smaller Meuse basin will respond to 2-4 day precipitation sums, as visualised in Figure 1.4. The effect of the WG problem described in this section on discharge output will therefore also differ per basin: it is more likely to happen for the Vecht than for the Rhine basin (which is 50 times larger).

1.2.2 Problem 2: Overestimation of precipitation extremes

The second limitation of the WG is the risk of overestimation of extremes due to the occasional occurrence of a statistical outlier in the recorded precipitation timeseries. This is graphically depicted in Figure 1.5, where the annual maximum 4-day precipitation over the Meuse basin (Calendar-year, abbreviated as CY) is plotted versus its corresponding return period.

The figure shows a very large difference between two versions of the WG, respectively excluding and including an extreme event that occurred in July 2021 in the Meuse river basin. Inclusion of the July 2021 event leads to a doubling of the 4-day sums for large return periods, which illustrates the large sensitivity of the WG to statistical outliers.

Due to the relatively short observational series, the extreme event of July 2021 is plotted at a return period of approximately 100 years, as determined by the length of the dataset. However, it is more likely to actually have a return period of approximately 1000 years or more. Arguably, due to the absence of other extremes of the same magnitude, the WG struggles to generate different events, which leads to overpresence of this event in the resulting timeseries. The result is unrealistic behaviour of the WG and large overestimation, as well as an irregular frequency distribution, as clearly depicted in Figure 1.5. Consequently, the precipitation extremes in the synthetic timeseries created by the WG are also largely overestimated, which will likely be passed on in the runoff output of the discharge models.

Although this figure shows an extreme case, it is clear that also a less extreme outlier will lead to overestimation.

1.2.3 Problem 3: Taking climate change into account

Thirdly, the operational version of the WG uses historical observations. Consequently, in the resampling process only historical events are considered, which limits the application of the WG to the current climate. The influence of climate change on precipitation extremes is therefore a relevant process to be considered in flood protection, but is not captured by the current WG.

To summarise, it is clear that the relatively short length of observational series may cause the WG to provide an unrealistic projection of rainfall extremes, which may - paradoxically - be either an over- or an underestimation. Generally, the shortness of the input dataset may cause underestimation of extreme multi-day sums, whereas outliers lead to overestimation. As these contradicting effects depend on the extremity of the outlier, on the length of the input record, on the number of consecutive days that is considered, etc., it is impossible to know the bias of the WG in advance - not only its amplitude but even its sign (except for exceptional cases like the July 2021 event).

Figure 1.4: Indication of the runtimes of the Meuse and the Rhine at the Dutch border to upstream rainfall events. The response times will be larger than the runtimes.

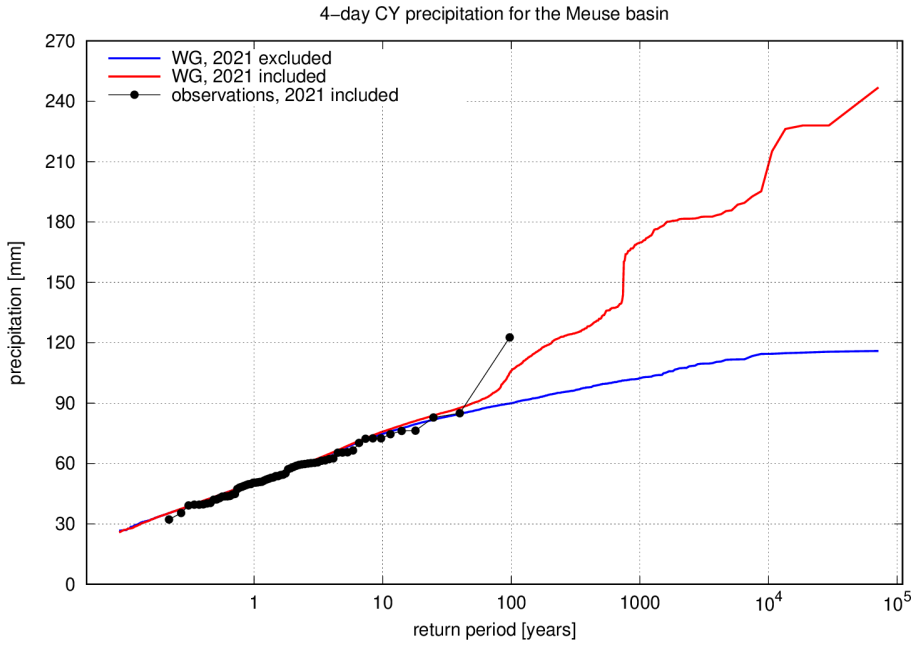

Figure 1.5: Annual precipitation extremes in the Meuse for a 4-day precipitation sum, in which the rainfall generator is shown with and without the extreme event of July 2021. The horizontal axis has a logarithmic scale.

Additionally, it is difficult to account for all features of climate change and their impact on precipitation extremes when using the current WG.

1.3 Research goal

The overarching cause of the potential unrealistic projection of rainfall extremes by the WG corresponding to the observed data as described in Section 1.2 is the limited length of the observation series and the presence or absence of outliers. This can be strongly mitigated by using a longer dataset as input of the WG. Unfortunately, daily observation data are only available over approximately 70 years, which is not enough to solve the problems. Instead, the possibility of using simulated weather from the SEAS5 dataset (ECMWF) and the RACMO dataset, which exceed the size of the observation record with order(s) of magnitude, in the GRADE procedure is analysed in this report. It is expected that by using these larger datasets as meteorological input of the WG or for the hydrological model of the GRADE model chain directly, the statistics of discharge extremes will improve.

The main goal of this report is to compare the synthetical datasets RACMO and SEAS5 datasets and their corresponding timeseries derived from the WG with the historical observations and the corresponding WG-timeseries, particularly in order to look at their performance to describe meteorological extremes. Beforehand the RACMO and SEAS5 datasets had to be bias-corrected.

Chapter 2

Methodology

In this chapter, the methodology used for the research is explained. Firstly, the different study areas are presented, after which the utilised datasets and their origins are being explained. Subsequently, these datasets are compared on their climatological similarity. After this comparison, a description of the bias corrections of the RACMO and SEAS5-datasets that have been performed is provided, followed up by a climatological comparison of the corrected datasets. Finally, theories used in order to obtain and present extremes are shortly explained.

2.1 Study areas

Three study areas that are relevant for flood protection in the Netherlands are used in this research. The Rhine, Meuse and Vecht, entering the Netherlands at respectively Lobith, Eijsden and Dalfsen[1] are three important rivers that are considered in Dutch flood protection. Their river basins, which capture precipitation and discharge a large part of this precipitation using the river, are the main study areas of this research.

2.1.1 Rhine basin

The Rhine basin, as presented in Figure 2.1, is a relatively large river basin of 185,000 km^2. In its entirety it covers parts of 9 countries, being Italy, Switzerland, Austria, Liechtenstein, Germany, France, Luxembourg, Belgium, and The Netherlands. Due to its large spatial extent, the Rhine basin contains large differences in topography, ranging from the high Alps in the South, to a flat topography in the Dutch Delta. Due to the size and differences in topography, there are large differences in annual precipitation within the basin, ranging from more than 2000 mm/year in the Alps to approximately 500 mm/year in the centre of the basin. Although the upper part of the river's discharge is dominated by snow melt from the Alpine region, the part of the river entering the Netherlands at Lobith is also strongly dependent on precipitation of the entire basin. In HBV the Rhine basin is divided into 134 sub-catchments, as shown in Figure 2.1.

[1]The Meuse and Vecht enter respectively at Borgharen and Gramsbergen, but the reference discharges hold for Borgharen/StPieter and Dalfsen.

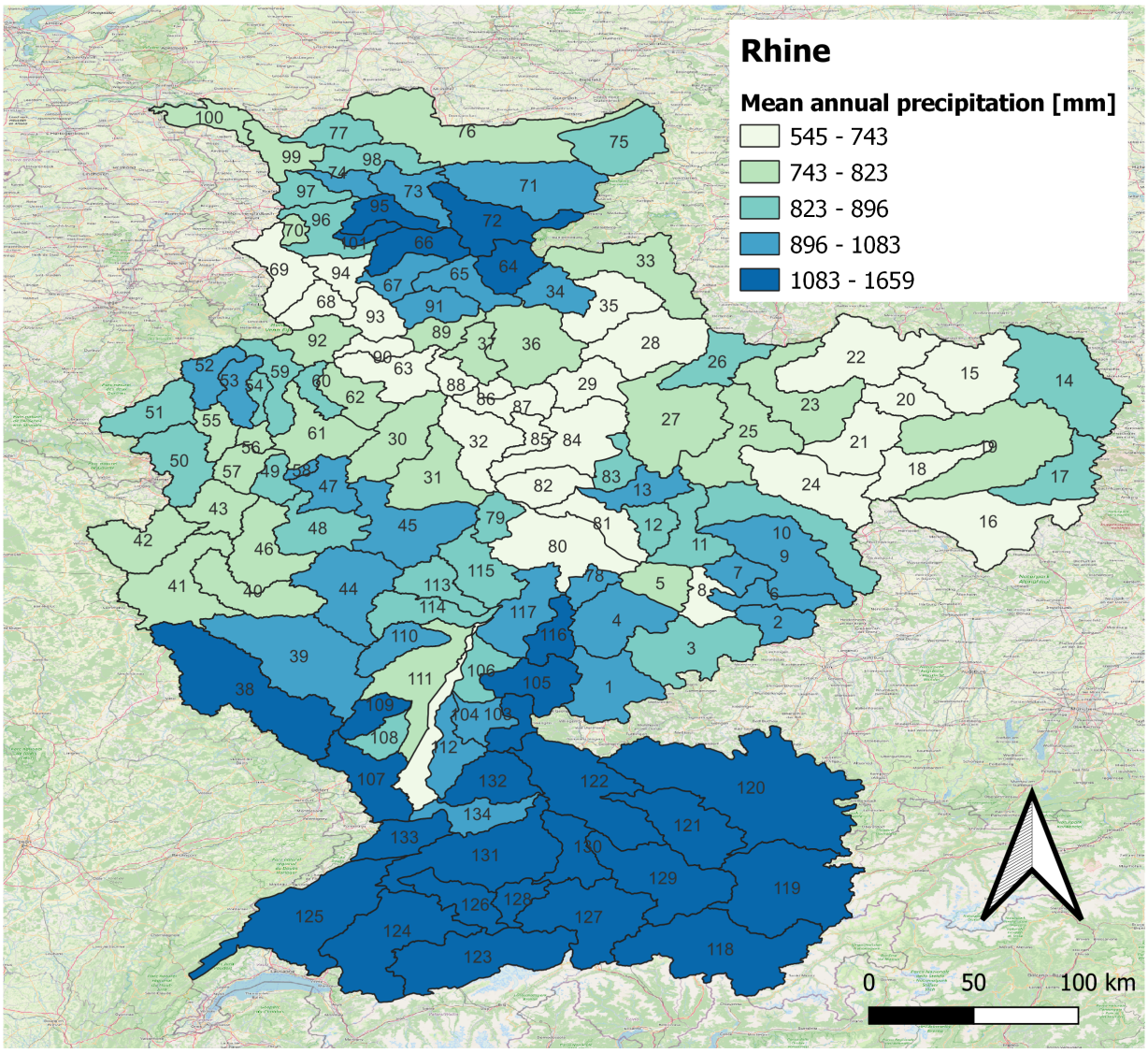

Figure 2.1: Overview of the Rhine basin. Annual average precipitation according to the observation-based passive dataset is presented per catchment. Catchment numbers are indicated in every catchment.

2.1.2 Meuse basin

The Meuse river has its origin in France and then flows through Belgium and the Netherlands. The river basin, as presented in Figure 2.2, also covers small parts of Luxembourg and Germany. The entire basin is approximately 36,000 km^2 and is therefore significantly smaller than the Rhine basin. There are considerable elevation differences in the Meuse basin, as it contains both a part of the Ardennes region as well as the Lorraine region. The topography of these two areas is very different however, which leads to a different response time of hydrological events to rainfall in these areas (the runoff generated in the Ardennes has a very direct response, contrary to the Lorraine region). Potential differences in rainfall due to relatively large spatial extent and topographical effects as well as differences in response time make the analysis of this rainfall-dominated river interesting in this research. In HBV the Meuse basin is divided into 15 sub-catchments.

2.1.3 Vecht basin

The Vecht basin, presented in Figure 2.3, is a relatively small river basin covering up parts of The Netherlands and Germany. The basin area is 3780 km^2, which is 10 times smaller than the Meuse basin and 50 times smaller than the Rhine basin. The Vecht basin does not have a large elevation difference and due to the small spatial extent, rainfall statistics (like the mean rainfall as presented in Figure 2.3) do not vary largely throughout the basin. Due to the low elevation of the basin, the river is largely dominated by rainfall. In HBV the entire Vecht basin is divided into 36 sub-catchments.

There are some differences between the basins that may influence the results of this research. Firstly, the large difference in spatial extent between the rivers may influence their hydrological response time, as well as the dominant type of precipitation extremes in the basins. Secondly, the size and topographical differences between the basins leads to large differences in rainfall variability within the basins. To account for potential differences between the basins, all three basins will be thoroughly investigated in the subsequent chapters of this research.

2.2 Data

As already stated in the Introduction, the goal of this study is to compare several different precipitation datasets and particularly their high extremes. The two reference datasets are the so-called passive dataset, which consists of observations, and the WG based on this dataset. The two synthetic datasets RACMO and SEAS5, and the corresponding results of the calculations of the WG using these synthetic datasets, will be intercompared as well as with the reference datasets. This section provides a small overview of all considered datasets.

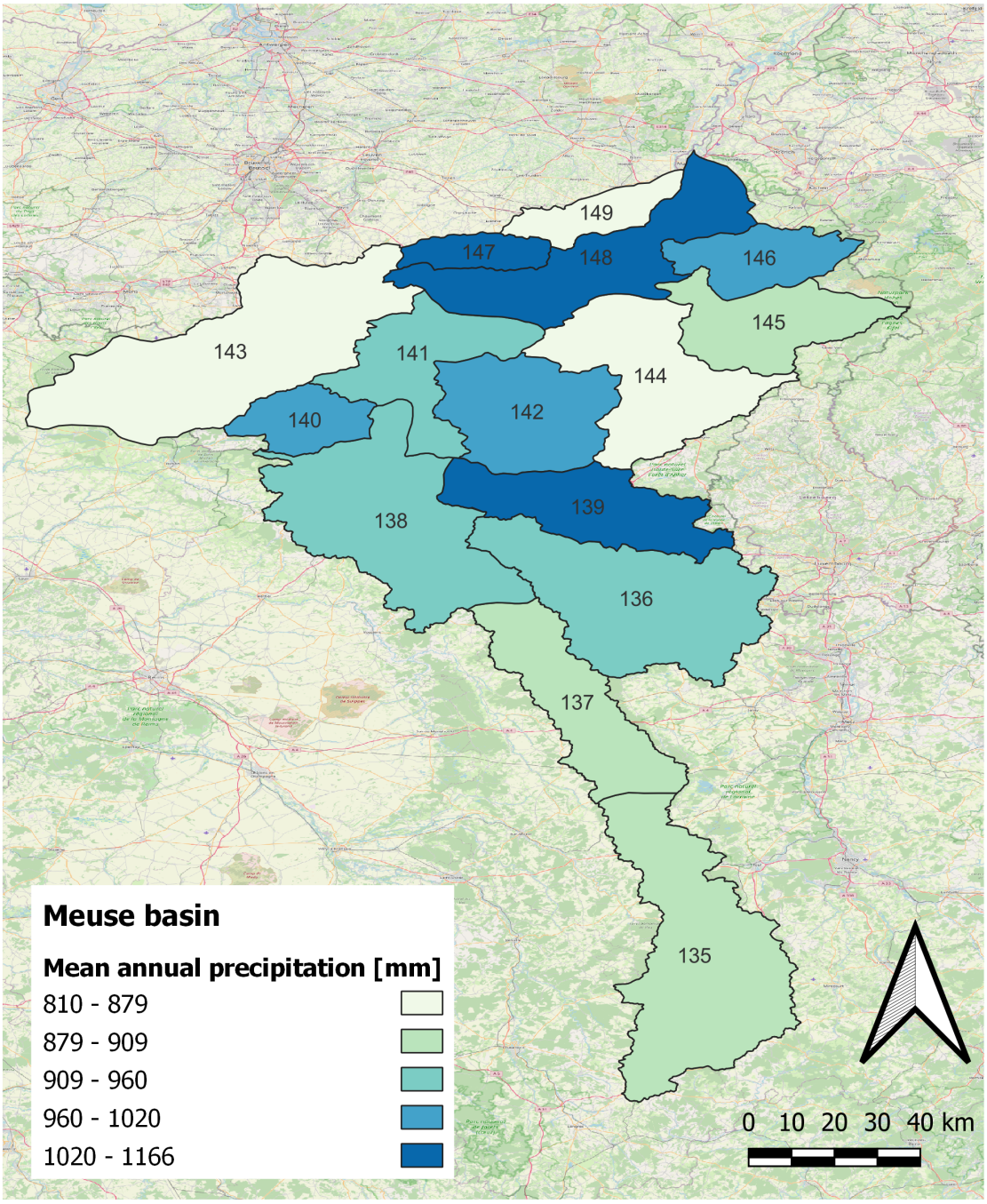

Figure 2.2: Overview of the Meuse basin. Annual average precipitation according to the observation-based passive dataset is presented per catchment. Catchment numbers are indicated in every catchment.

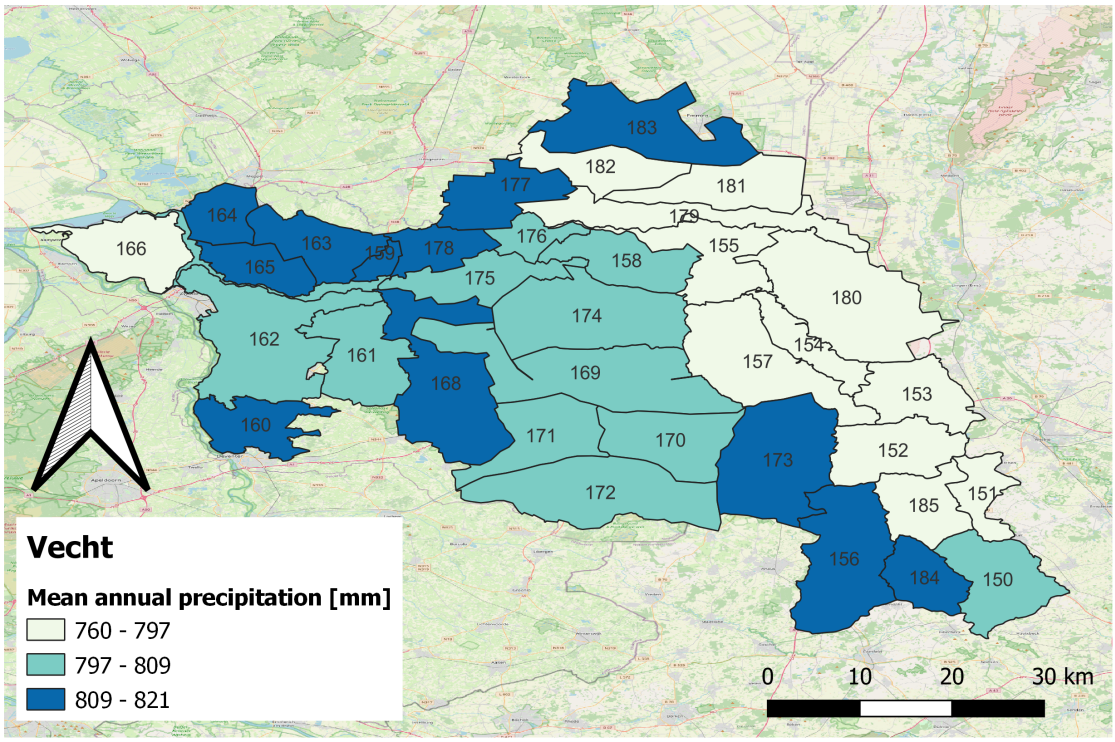

Figure 2.3: Overview of the Vecht basin. Annual average precipitation according to the observation-based passive dataset is presented per catchment. Catchment numbers are indicated in every catchment.

2.2.1 observation-based passive dataset and Weather Generator

The observational dataset used in this report is the so-called passive dataset[2]. For the Rhine and Vecht basin, precipitation of the observation-based passive dataset consists of the HYRAS precipitation dataset, constructed by the German Weather Service (Rauthe et al., 2013). The HYRAS dataset comprises around 6000 rain gauges that cover up the entire Rhine and Vecht basin, as shown in Figure 2.4. Gridded daily precipitation and temperature observations have been produced with a spatial resolution of 5x5 km^2. The observation-based passive dataset provides daily HYRAS precipitation data from 1950 onwards. Temperature and potential evaporation data of the observation-based passive dataset are derived from the EOBSv21e dataset (Cornes et al., 2018a).

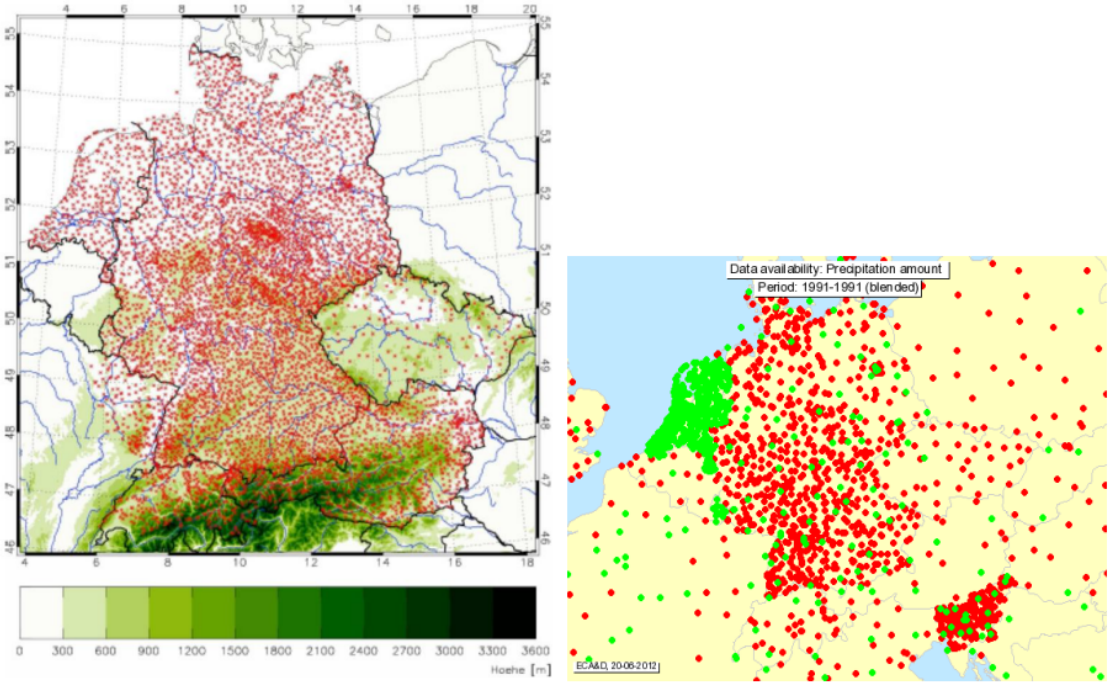

Figure 2.4: The stations used for the HYRAS dataset (left) and the EOBSv21e dataset (right). Green colour indicates public availability. Red stations are not publically available (Hegnauer et al., 2014).

Figure 2.4 shows that the HYRAS dataset does not cover the full Meuse basin. Therefore, for the Meuse a compound of different datasets in time is used. The precipitation dataset consists of EOBSv21e (Cornes et al., 2018a) data for the period of 1950-1960. Like HYRAS, EOBSv21e is a gridded dataset which contains precipitation, temperature and potential evaporation data. From 1961-2007 the observation-based passive dataset uses a dataset constructed by Buishand and Leander (2011), which contains daily precipitation for all 15 catchments of the Meuse and which were constructed from measured station data. Bouaziz et al. (2017) created high-resolution gridded observation-based daily

[2]The name 'passive' originates from earlier work (see e.g. Schmeits et al., 2014) where the data in the resampling step (the so-called active data) is not necessarily identical to the data that is used to transform the timeseries of historical dates back to meteorological data (the so-called passive data).

precipitation for the period 2008-2017, which is used for the last years of the observation-based passive precipitation dataset. Note that for the entire Meuse basin, EOBSv21e temperature and potential evaporation is used.

A more detailed description of all data sources used for the observation-based passive dataset for the Vecht and Rhine is found in Schmeits et al. (2014), for the Meuse basin, the individual sources can be consulted.

2.2.2 RACMO and WG-RACMO

RACMO (van Meijgaard et al., 2008) is a numerical weather prediction model developed by the KNMI and the Danish Meteorological Institute. Driven by EC-EARTH2.3[3] for its boundary conditions, the RACMO model generated 16 members with each member starting with a different initial state, resulting in different climate simulations. These weather data are available at a 12x12 km^2 resolution. All members are available for 1950-2100[4]. Many meteorological parameters are available at daily resolution; some (e.g. temperature and radiation) at 3-hourly, and precipitation is archived at hourly resolution.

Combination of the (independent) members from 1950 - 2021 results in a dataset of 1152 years that is representative for the same climate as the observation-based passive dataset, which will be used in this report.

Based on this RACMO dataset also a 50000-year timeseries is calculated using the WG finally called WG-RACMO.

2.2.3 SEAS5 and WG-SEAS5

SEAS5 (Johnson et al., 2019) is the most recent European Centre for Medium-Range Weather Forecasts (ECMWF) forecast system. Since 2017, it runs real-time seasonal, long-range forecasts, which consist of 51 members of 7 months of weather data consisting of amongst others precipitation, temperature and potential evaporation with a grid point resolution of 36x36 km^2. After several weeks, these forecasts are independent of their initial conditions[5]. For the period of 1981-2021, hindcasts have been made. For the months February, May, August and November 51 members of 7 months were generated and for the remaining months 25 members were generated. With adequate coupling of months (as described below), a large dataset of 8496 years of synthetic weather data has be constructed. To avoid dependency of initial conditions, during the coupling process, the first month of every member is not included in the generated dataset.

Coupling of all forecast members in order to create the large SEAS5 dataset is done by combining forecasts that differ half a calendar year in starting date. The coupling procedure is visualised in Table 2.2.3. After the coupling process, a dataset with almost 8500 years of synthetic weather remains, which is the SEAS5 dataset that is used in this report.

[3]These runs were done for the construction of the KNMI'14 climate scenario's. New runs (again 16 members) will be run in 2023 for the construction of the KNMI'23 climate scenario's.

[4]The RCP8.5 greenhouse gas concentration pathway was used from 2006 onwards.

[5]This holds for our latitudes; not for the tropical region, where the forecast skill is strongly determined by El-Nino.

Table 2.1: Combination of the individual SEAS5 ensemble members to construct a 8496 year timeseries.

First half year		**Second half year**
jan 1981-0	+	jul 1981-0
jan 1982-0	+	jul 1982-0
...		
jan 2021-0	+	jul 2021-0
jan 1981-1	+	jul 1981 -1
...		
jan 2021-24	+	jul 2021-24
feb 1981-0	+	aug 1981-0
...		
jun 2021-50	+	dec 2021-50

An advantage of this dataset compared to the observation-based passive dataset is the fact that it is available with a 6 hourly interval. For a detailed description of the SEAS5 forecast system is referred to Johnson et al. (2019).

Based on this SEAS5 dataset also a 50000-year timeseries is calculated using the WG finally called WG-SEAS5.

2.3 Climatological dataset comparison

In order to be applicable for GRADE, RACMO and SEAS5 must be climatologically comparable to the observation-based passive dataset. This section provides a dataset comparison for mean precipitation amounts and yearly average precipitation extreme for all catchments in the different basins.

2.3.1 Mean daily precipitation

In panels a of Figures 2.5 - 2.8, mean precipitation of RACMO, SEAS5 and the observation-based passive dataset are compared in a bar plot for all catchments in the different basins. RACMO and SEAS5 have been separated in the Figures for the Rhine basin for clearer visualisation. Daily average precipitation is plotted on the vertical axis, whereas the horizontal axis represents the catchment number for which the datasets are being compared (the location of the catchments is provided in Figure 2.1 - 2.3).

In general, for almost all catchments, the three datasets have comparable daily average precipitation amounts. This means that the synthetic datasets are capable of describing the mean precipitation in all different areas of the basins. As such, this climatological aspect is well covered in the synthetic datasets.

Looking into more detail, it should be noted that particularly the SEAS5 dataset has slightly more similar mean precipitation amounts to the observation-based passive dataset compared to the RACMO dataset. In the Meuse basin (Figure 2.7 a), RACMO overestimates precipitation in some catchments, whereas in the Vecht basin (Figure 2.8 a), a slight underestimation of daily average precipitation can be noticed. In the Rhine

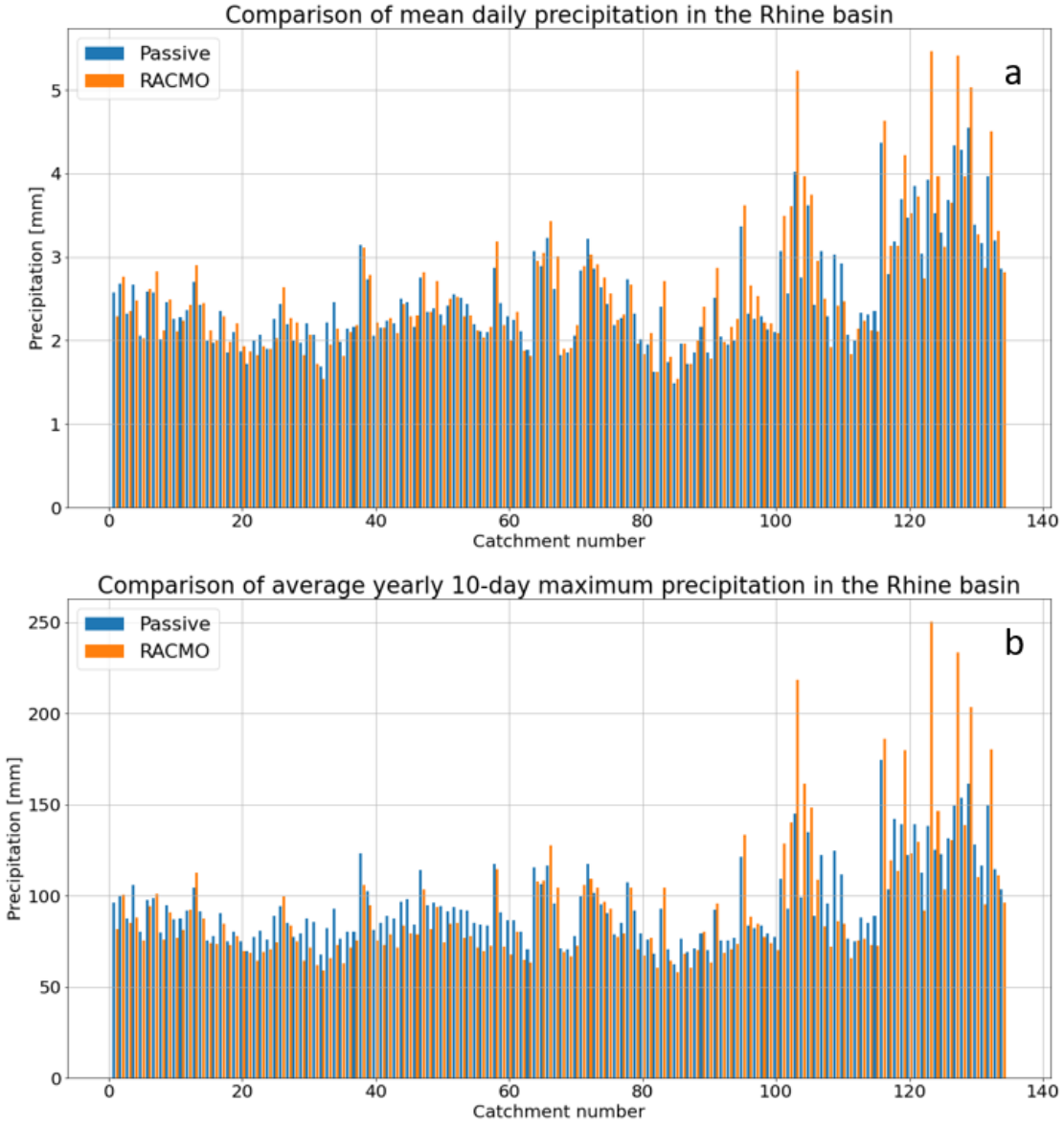

Figure 2.5: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Rhine basin for the observation-based passive dataset (blue) and RACMO (orange).

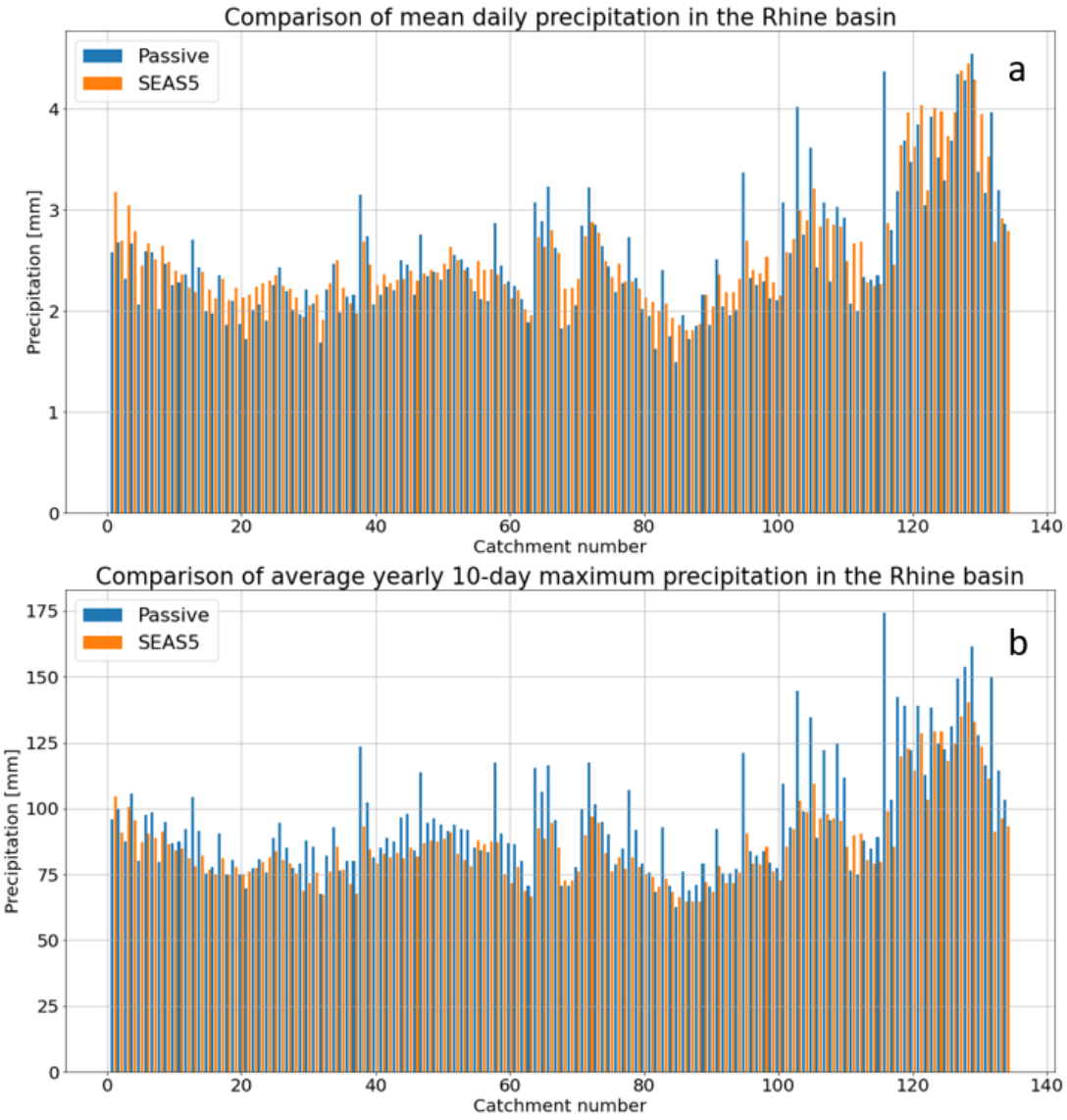

Figure 2.6: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Rhine basin for the observation-based passive dataset (blue) and SEAS5 (orange).

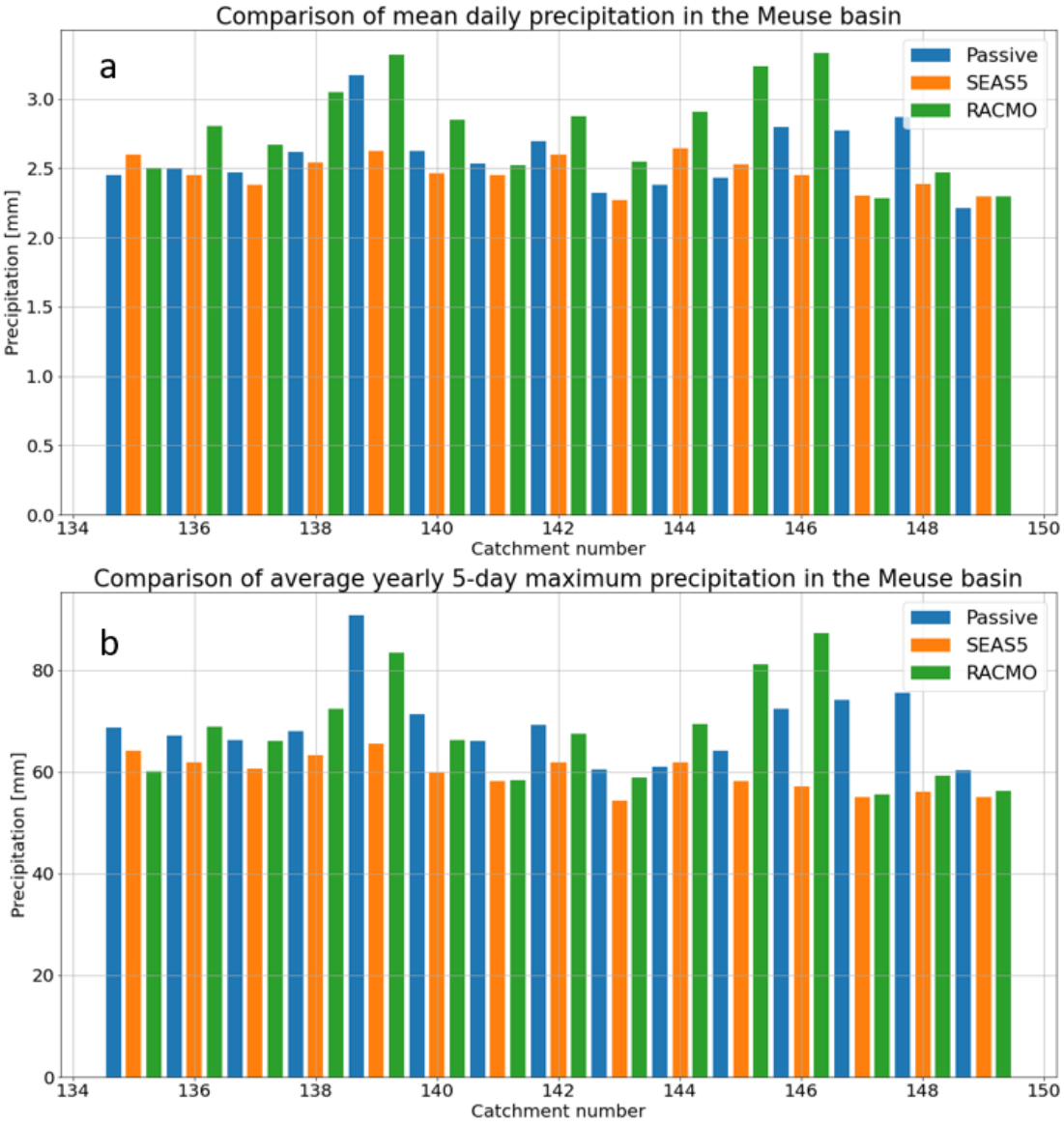

Figure 2.7: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Meuse basin for the observation-based passive dataset (blue), SEAS5 (orange) and RACMO (green).

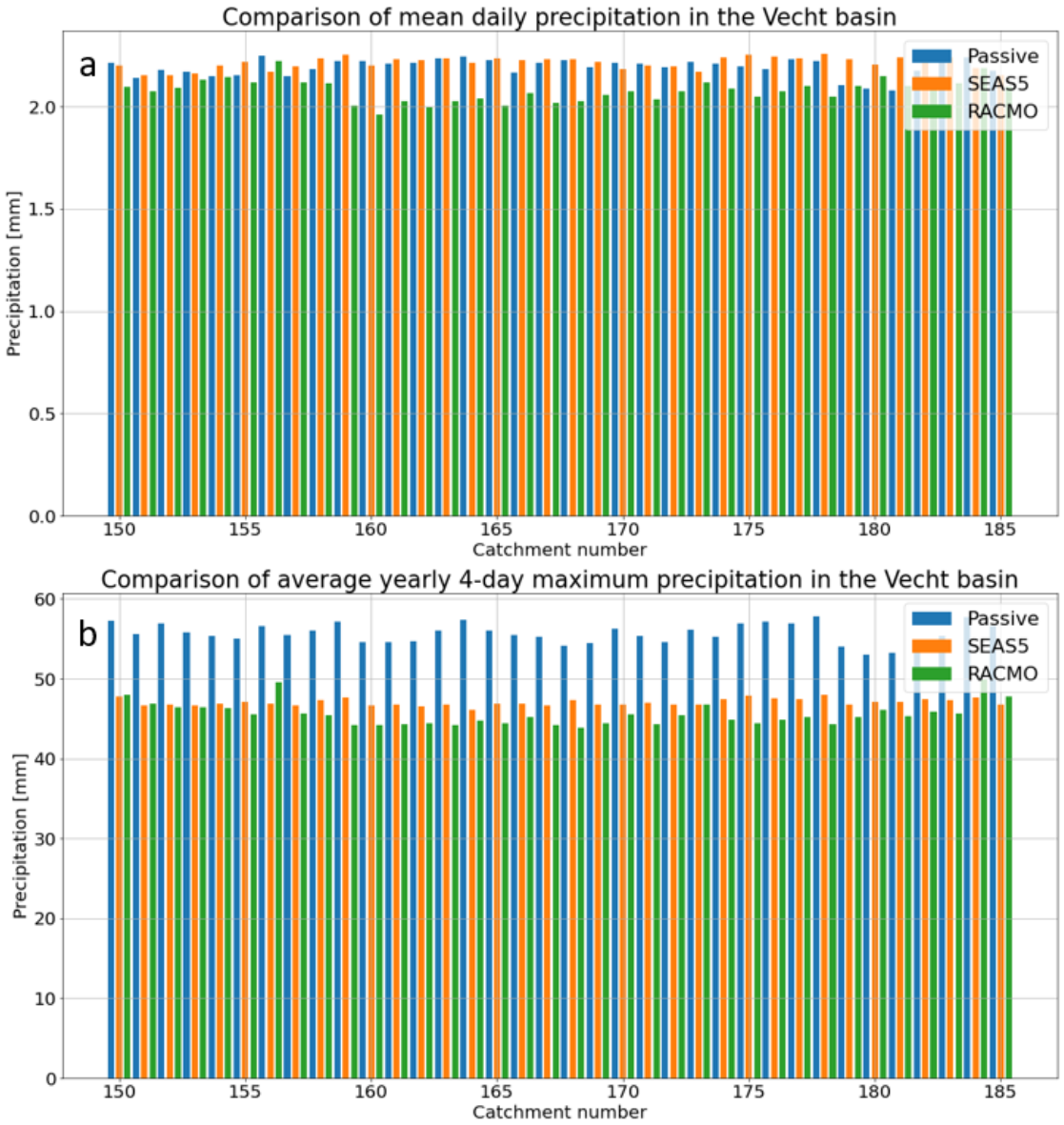

Figure 2.8: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Vecht basin for the observation-based passive dataset (blue), SEAS5 (orange) and RACMO (green).

basin (Figure 2.5 a and 2.6 a), both the SEAS5 and the RACMO dataset show over- and underestimation in different catchments.

Recapitulating, the general behaviour of all datasets regarding daily average precipitation is comparable and the differences between geographical regions are clearly projected by all datasets, which shows all datasets are able to describe this climatological aspect. However, the actual magnitude of mean precipitation may vary between the different datasets for different catchments.

2.3.2 Mean annual maximum precipitation

In order for the synthetic datasets to be reliable, the average of annual extremes for all catchments must be approximately equal for RACMO, SEAS5 and the observation-based passive dataset. This has been tested in panels b of Figures 2.5 - 2.8, where mean annual maximum precipitation is plotted on the vertical axis and the horizontal axis again represents the catchment number. Note that the multi-day sum differs for every basin. Due to amongst (others basin) size, the Rhine is expected to have a larger hydrological response time than the Meuse and Vecht. From experience, the basins are likely to respond to a sum of 10, 5 and 4 days for respectively the Rhine, Meuse and Vecht. However, the analysis has been performed for all multi-day sums in a range of 1-10 days, with similar conclusions as described below.

Similar to the precipitation means, the annual maxima of the synthetic datasets are comparable with the observation-based passive dataset for most catchments. This shows that although precipitation extremes clearly vary between catchments depending on their geographical location (e.g. the Alps compared to lower areas), the synthetic datasets are capable of describing these differences approximately equally well as the observation-based passive dataset. This shows the climatology of both synthetic datasets match that of the observation-based passive dataset.

It should be noted that although the general behaviour of the synthetic datasets and the observational dataset are comparable, their magnitudes may differ a bit. In other words, the climatic difference between different regions of the basins is described fairly well, but there is some difference in the actual precipitation amounts. In most catchments, there is a slight underestimation in annual maximum precipitation for both the SEAS5 and the RACMO dataset compared to the observation-based passive dataset. This behaviour is particularly clear in the Vecht basin. In some extraordinary cases (e.g. the RACMO dataset in the Alpine catchments in the Rhine basin), the synthetic datasets have larger annual maxima compared to the observation-based passive dataset.

Zooming in, the average annual maximum precipitation in the Vecht basin behaves approximately equal to the mean daily precipitation, with RACMO and SEAS5 both underestimating for all catchments. Note that the Vecht is a small basin and therefore the differences between each catchment are relatively small. In the Meuse basin, the largest outliers occur in the catchments that also showed the largest offset for mean daily precipitation. Note that in both basins the over- or underestimation is relatively small.

In the Rhine basin, catchments with a relatively large mean precipitation offset are also the biggest outliers for average annual maximum precipitation. Particularly the RACMO dataset has several large overestimations, which mostly occur in the Alpine region of the basin.

In general it is clear that the climatological differences in the basins, portrayed by the observation-based passive dataset, are well captured by both RACMO and SEAS5. However, the exact magnitude of both mean daily precipitation and average annual extreme precipitation may differ between different catchments. In other words, the datasets compare well qualitatively, but have some slight quantitative differences. In order to make the synthetic datasets even more reliable and useful for quantitative analysis, a simple bias correction is performed, which is common practice for model results such as the SEAS5 and RACMO dataset. An ideal bias correction would lead to even more comparable mean precipitation and average annual extremes. Because the magnitude differences only occur in few catchments, the correction is performed for every individual catchment.

2.4 Bias correction

Although all datasets are comparable regarding climatology, the absolute magnitude of several climate variables may differ per catchment. Although the number of catchments with a large offset is limited, a bias correction of both the SEAS5 and RACMO dataset could improve these synthetical datasets and make it useful for quantitative analysis. The bias correction used in this report is explained in this section.

To obtain an optimal bias correction, several aspects must be taken into account. Firstly, as different catchments require a different extent of scaling, the bias correction is performed on catchment level. Secondly, it is important to capture the difference in precipitation type between summer and winter in the bias correction. Generally, winter precipitation is dominated by large-scale precipitation, whereas higher temperatures in summer cause convective precipitation to occur. In order to capture these weather type varieties and be able to differentiate between summer and winter events in succeeding chapters of this report, the bias correction will be performed separately for all 12 calendar months.

The bias correction performed in this research is based on quantile mapping, i.e. every value is transformed to the value of observation-based passive dataset that has the same (cumulative) probability. For the temperature and potential evapotranspiration (pet) a normal distribution is fit to all days per calendar month, both for the observation-based passive data and the synthetic data.

For the precipitation, a 3-parameter distribution is fitted, given by:

$$R = \max[0, (1 - e^{c \log(1-p)})(a \log(1-p) + b)] \tag{2.1}$$

With R the precipitation amount, p the probability of exceedance and a, b and c fitting parameters. For small precipitation amounts, this function deviates from an exponential distribution, which leads to more accurate low precipitation amounts and therefore better corrected precipitation means.

Note that this correction is also performed for every catchment and for every individual month. The original datasets and corrections are shown in Figure 2.9 for the Rhine, with the upper panel for RACMO and the lower panel for SEAS5 (the correction of the other basins are given in Appendix A).

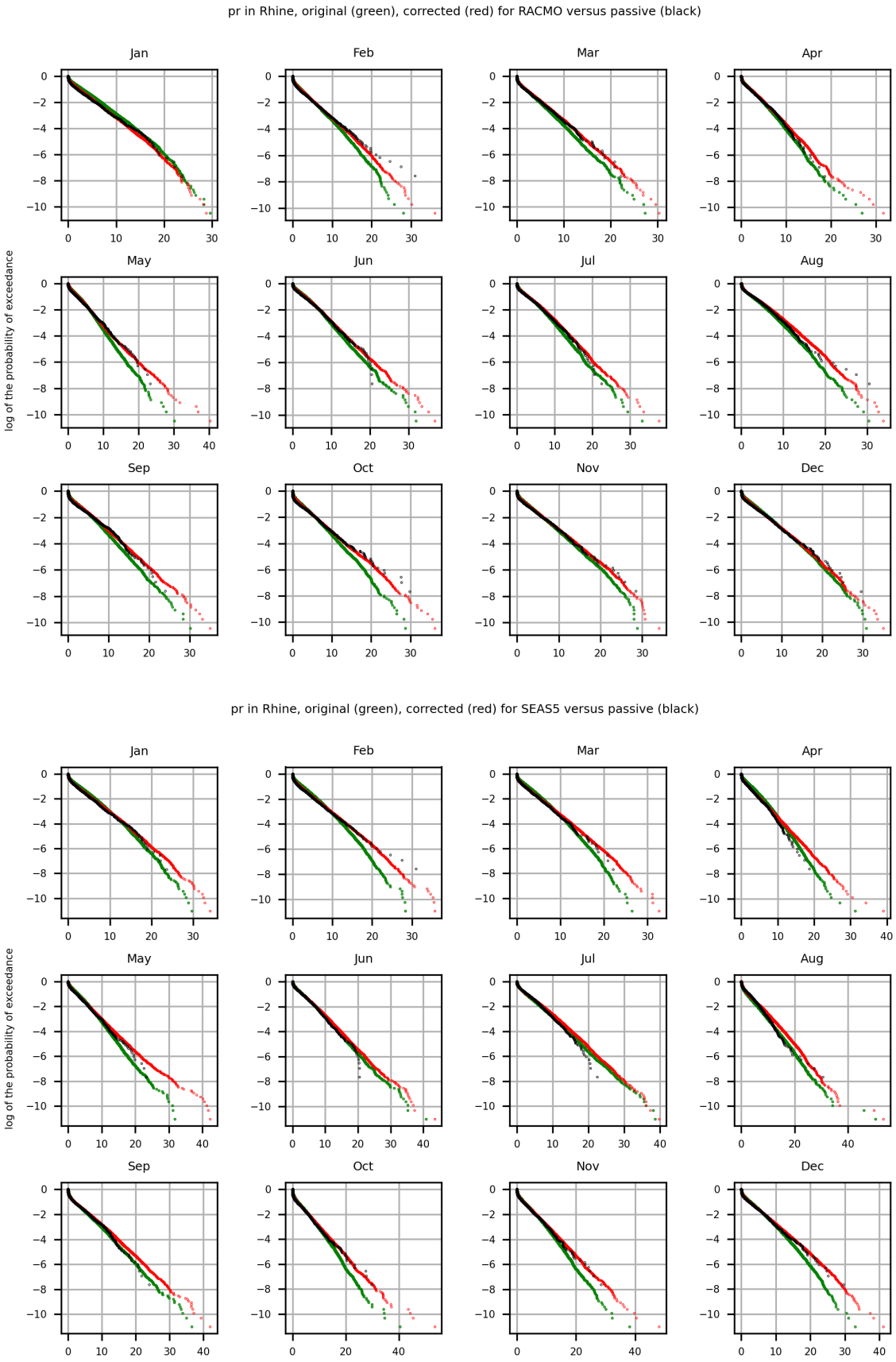

Figure 2.9: An overview of the bias correction for the Rhine basin for RACMO (upper) and SEAS5 (lower).

When looking at Figure 2.9, it is clear that the corrections work well. A comparison of climatological aspects of the corrected datasets and the observation-based passive dataset is provided in section 2.5.

Because the main focus of this report is the comparison of precipitation of RACMO, SEAS5 and passive (and corresponding WG), only the bias correction of precipitation data is explained in this Section. However, similar to precipitation, temperature and pet data have also been corrected. Both temperature (tas) and potential evapotranspiration (pet) have been corrected using a normal distribution. The results of these corrections are presented in Appendix A.

2.5 Corrected datasets

The results of the corrected RACMO and SEAS5 datasets are compared with the observation-based passive data in Figure 2.10 - 2.13, which represent respectively the Rhine, Meuse and Vecht basin. Panel a of all figures compares the mean daily precipitation in all catchments, whereas panel b shows a comparison of average annual maximum precipitation of the once-a-year multi-day sums. For the Rhine basin, RACMO and SEAS5 have been separated in the Figures to improve visualisation.

For all basins and for both the mean daily precipitation and the average annual maximum precipitation, the datasets are comparable to a very large extent.

To quantify the bias correction, the bias of the original and bias-corrected SEAS5 and RACMO compared to the observation-based passive dataset are presented in Table 2.5. Shown are the annual mean values as well as the 10-day, 5-day and 4-day annual maxima for the Rhine, Meuse and Vecht respectively, with the relative biases in the original and the bias-corrected datasets.

The table shows that in all three basins, for both datasets, the bias correction leads to a considerably better agreement with the observation-based passive dataset[6].

Table 2.2: Overview of the bias of the daily-mean and annual-maximum precipitation in the RACMO and SEAS5 datasets compared to the observation-based passive dataset before and after correction.

		Rhine		**Meuse**		**Vecht**	
		RACMO	SEAS5	RACMO	SEAS5	RACMO	SEAS5
Mean	passive	2.62 mm/day		2.56 mm/day		2.19 mm/day	
	bias original	2.12%	3.18%	8.28%	-2.80%	-4.81%	1.83%
	bias corrected	1.42%	1.00%	1.36%	0.72%	1.59%	0.87%
Max	passive	80.28 mm/10days		62.32 mm/5days		51.91 mm/4days	
	bias original	-0.46%	-3.35%	-2.63%	-10.23%	-16.09%	-12.62%
	bias corrected	1.43%	2.31%	-0.89%	-0.03%	2.80%	1.41%

Lastly, the annual cycle of all datasets are compared with the observation-based passive dataset, in order to see if the seasonal influence of all datasets is comparable. In

[6]The absolute value of the bias only increases for the 10-day Rhine annual maxima, but this change is not significant.

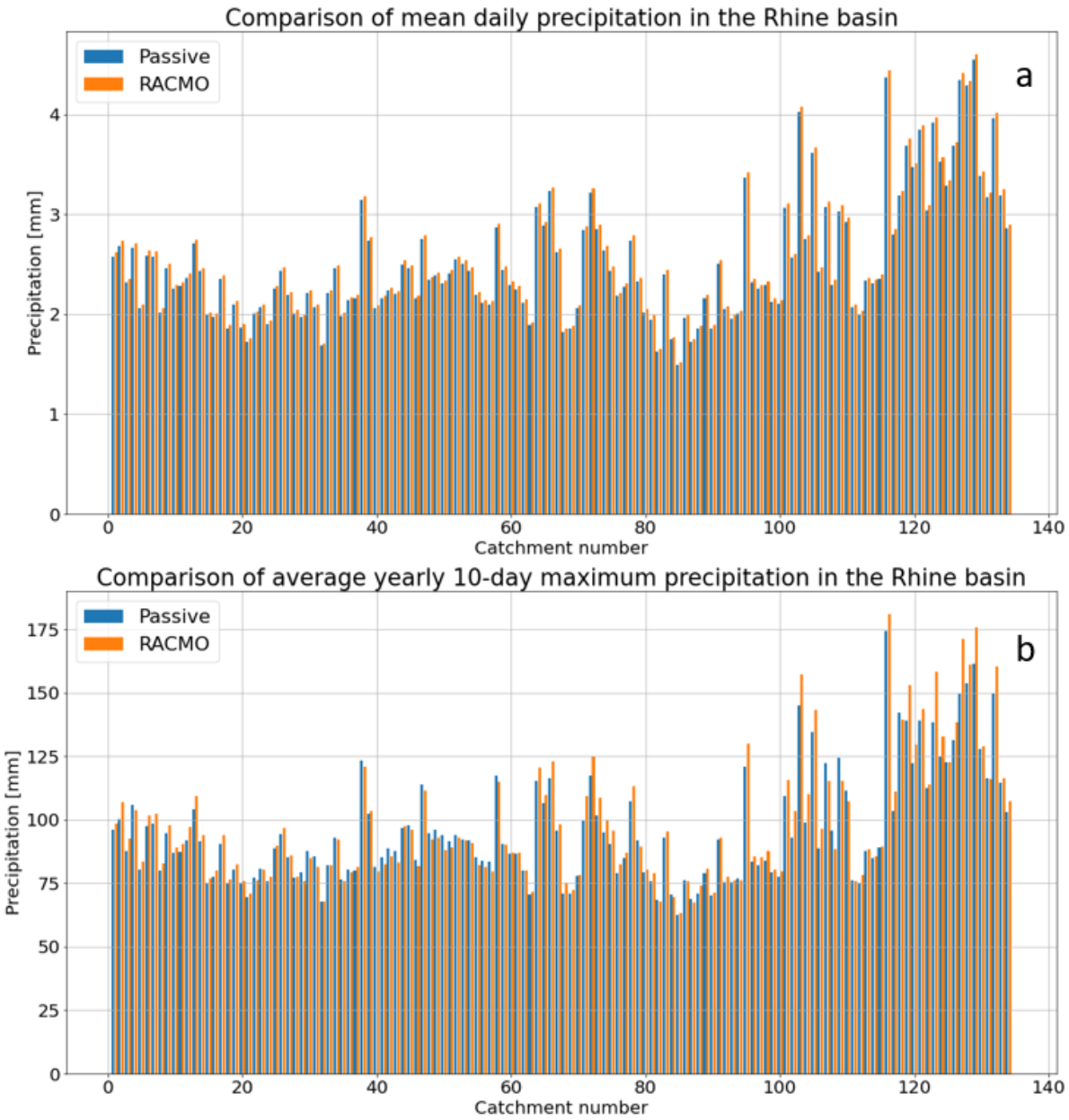

Figure 2.10: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Rhine basin for the observation-based passive dataset and corrected RACMO.

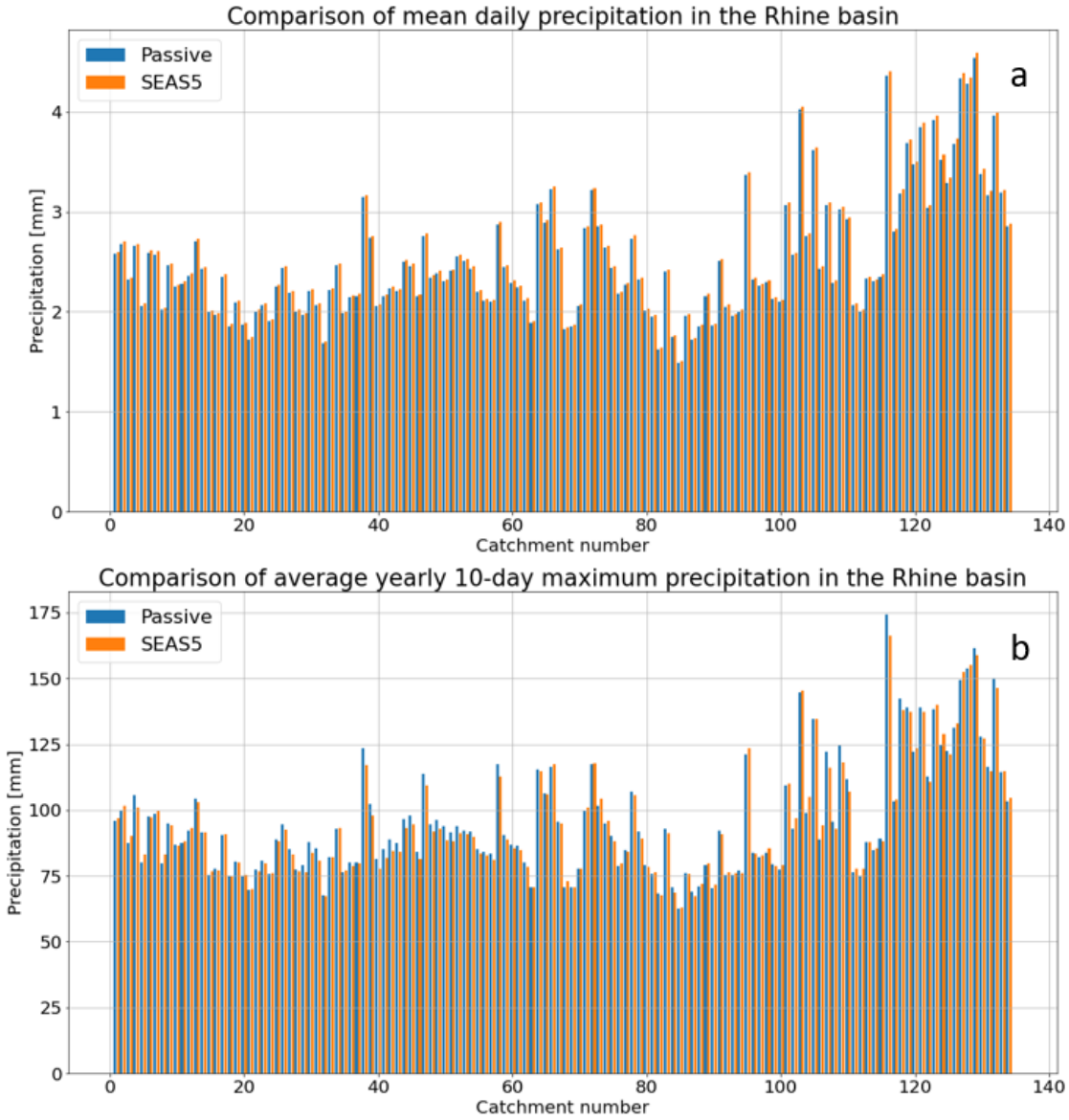

Figure 2.11: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Rhine basin for the observation-based passive dataset and corrected SEAS5.

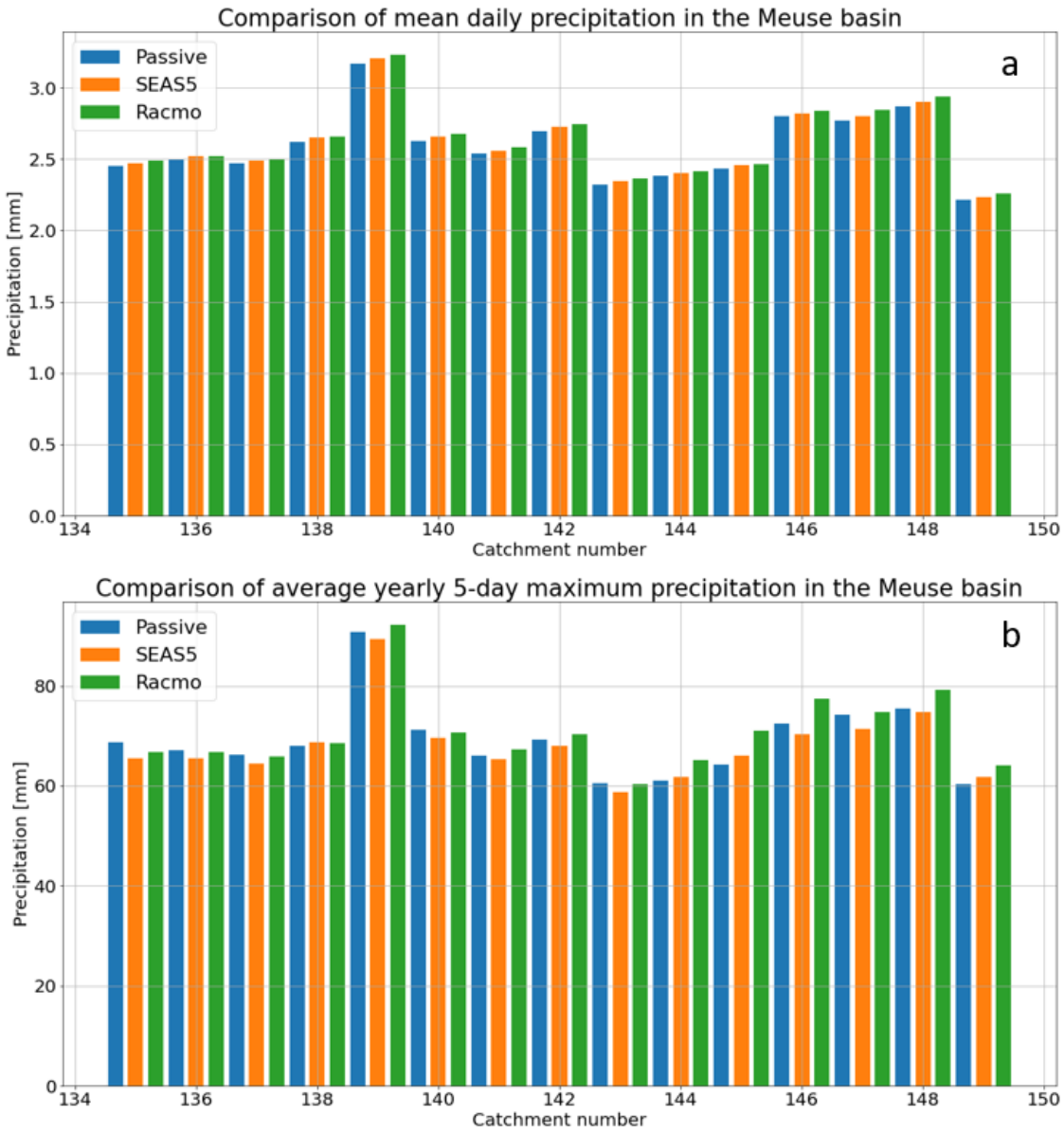

Figure 2.12: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Meuse basin for the observation-based passive dataset, corrected SEAS5 and corrected RACMO.

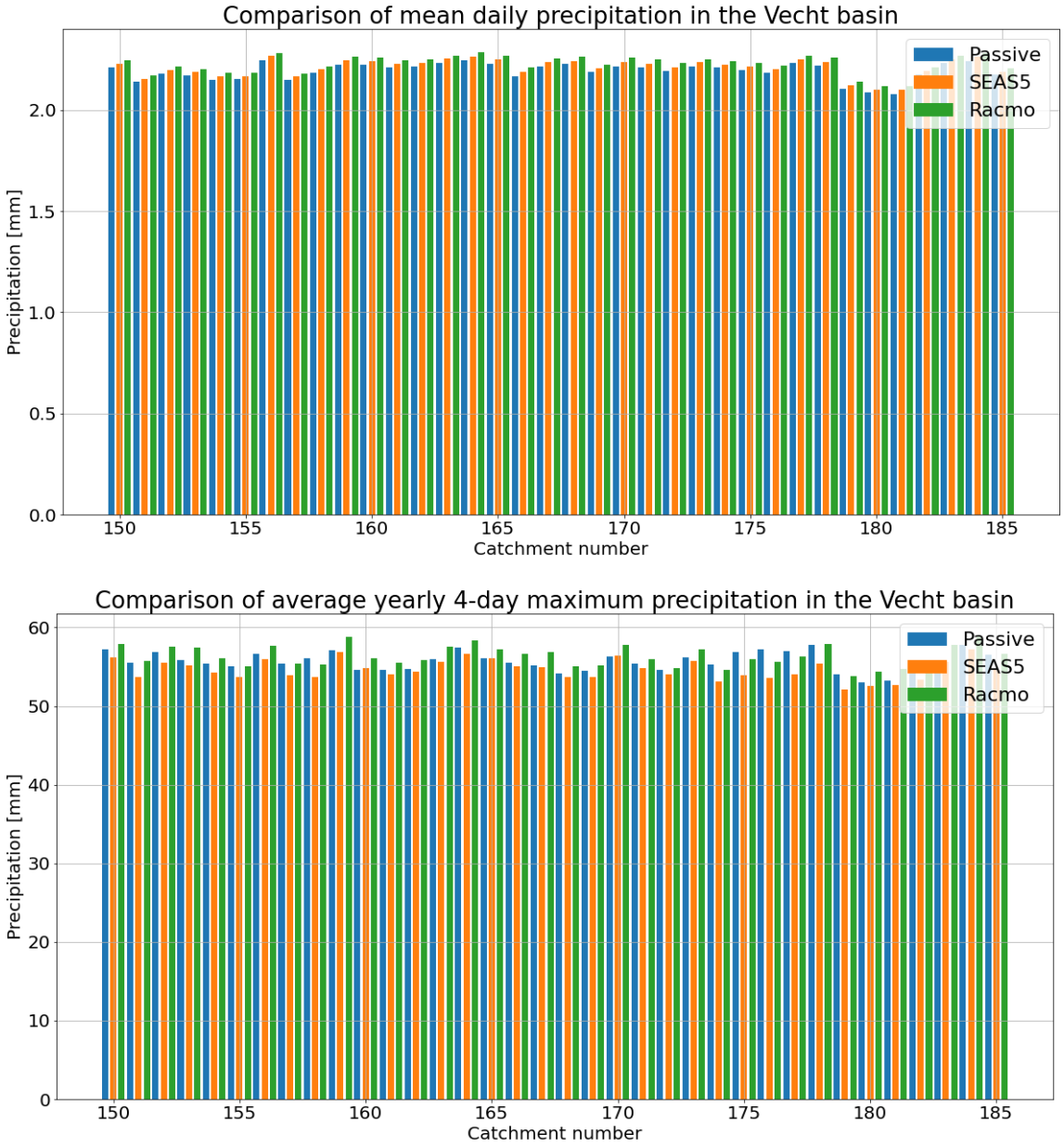

Figure 2.13: Comparison of a) mean daily precipitation and b) average annual precipitation maxima in the Vecht basin for the observation-based passive dataset, corrected SEAS5 and corrected RACMO.

Figure 2.14, the basin-averaged daily average precipitation of every day in the year is plotted for RACMO, SEAS5 and the observation-based passive dataset. The resulting graph is a good indication of the seasonal behaviour of mean precipitation in the basins. All panels show that the correction of SEAS5 and RACMO is largely congruent to the observation-based passive dataset considering seasonal behaviour. This indicates that the bias correction, which is based on monthly corrections, can account for seasonal effects as well. It should be noted that original RACMO compares a lot better to the observation-based passive dataset than original SEAS5.

Figure 2.15 shows the average of the monthly 1-day maximum precipitation in all basins and all datasets, both uncorrected and bias-corrected. It shows that the annual cycle of the extremes is strongly improved by the bias correction.

2.6 Theory and formulas

The main goal of this research is to analyse and compare the high extremes of RACMO and SEAS5 to the observation-based passive dataset and their corresponding results of the WG. To do this, a so called extreme value analysis is required. One method is to couple annual maxima to a return period (T), which is directly related to a probability of exceedance (P) according to Equation 2.2.

$$T = \frac{1}{P} \tag{2.2}$$

Note that the probability of exceedance and the non-exceedance (Q) probability are also related according to Equation 2.3.

$$P = 1 - Q \tag{2.3}$$

Extremes from a dataset can be coupled to the probability of Exceedance by ranking. With extremes ranked in ascending order, the probability of exceedance follows from the ranked extremes according to Equation 2.4.

$$P = \frac{m - 0.3}{n + 0.4} \tag{2.4}$$

with n being the total number of extreme events and m being the rank of a certain event. Using Equation 2.4, all events can be coupled to an exceedance probability and thereafter Equation 2.2 can be used to identify a corresponding return period[7]. In this report, extreme analysis is mainly performed by plotting the extremes and their corresponding return periods in so-called Gumbel plots, for which the Gumbel distribution is transformed into a straight line. For a detailed overview of the statistics of extreme analysis please consider Coles et al. (2001).

[7]Equation 2.4 implies that the largest event in n-year timeseries is plotted at a return period of T$\approx 1.43n$, which means that RACMO extends to return periods of 1650 years, and SEAS5 to 12,000 years.

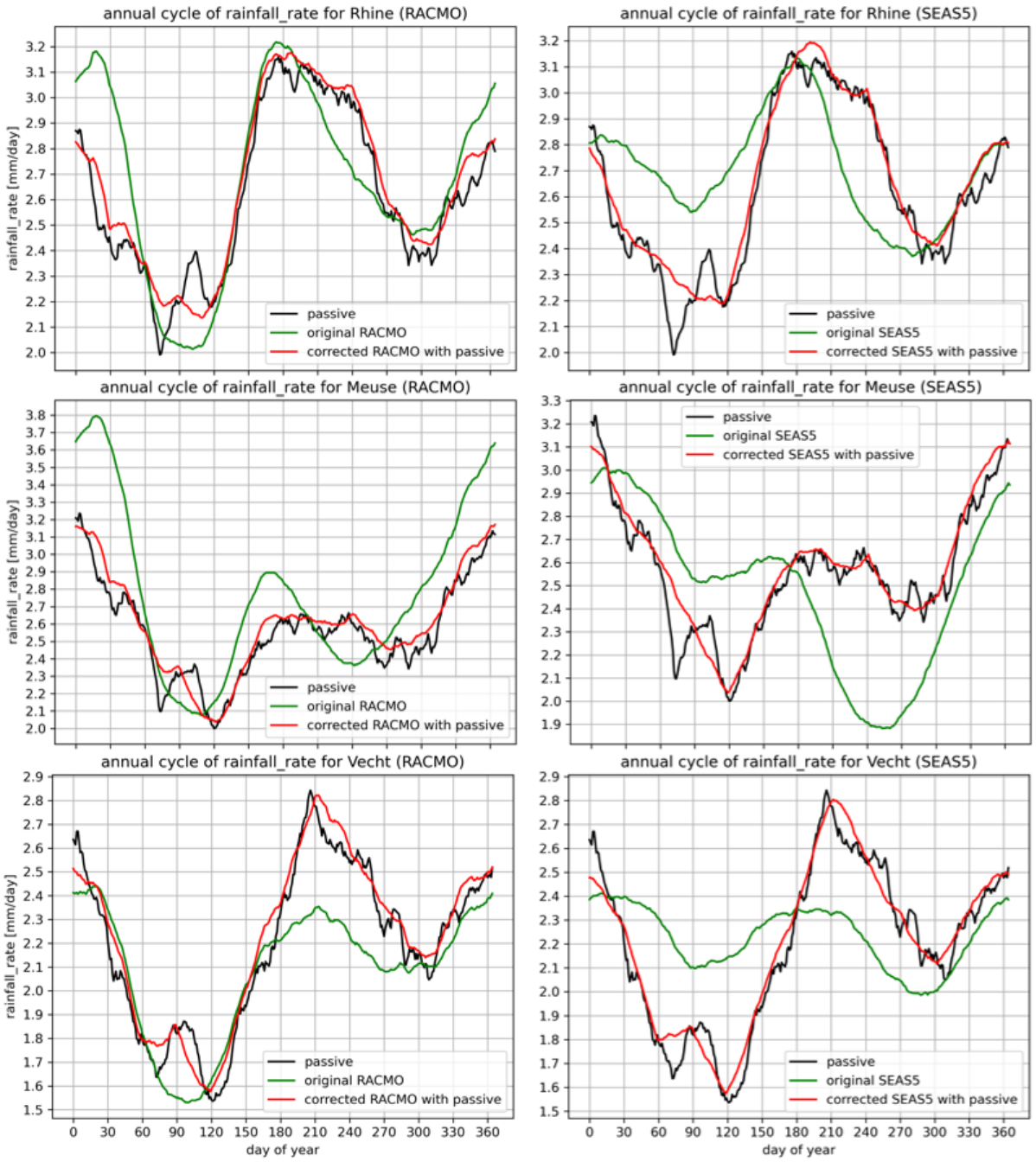

Figure 2.14: Annual cycle of the average basin-averaged daily precipitation for the Rhine, Meuse and Vecht.

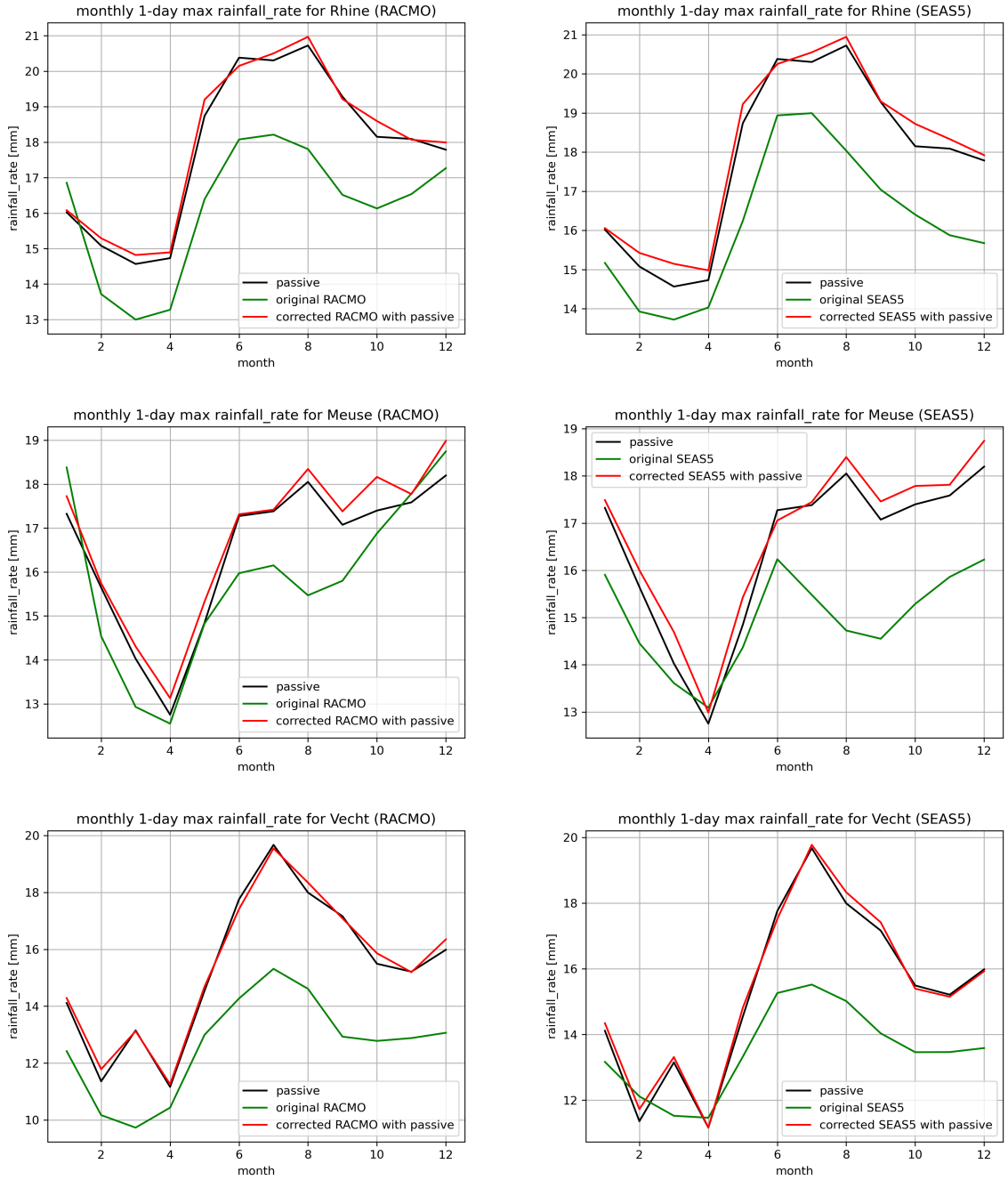

Figure 2.15: Annual cycle of average monthly-maximum daily precipitation for the Rhine, Meuse and Vecht.

Chapter 3

Results

This chapter provides an analysis of all datasets and their ability to describe extreme precipitation events. Firstly, dataset extremes are compared in an extreme value analysis. Subsequently, the ability of all datasets to describe spatial and temporal extent of summer extremes is described. In this chapter, the corrected RACMO and SEAS5 are considered for comparison, unless stated otherwise.

3.1 Extreme value analysis

In this section, precipitation extremes will be analysed with an extreme value analysis. More information on the method used is described in Section 2.6. Extreme values in this chapter are presented by plotting yearly maximum precipitation values against their corresponding return period. This is done for all considered basins and a variety of multi-day precipitation sums.

3.1.1 Rhine frequency analysis

Although runoff response to rainfall in catchments and basins is considered to be non-linear (e.g. Fortesa et al., 2020), it makes sense that discharge extremes of larger basins originate from high multi-day rainfall events. Based on earlier experience in GRADE (e.g. J. Beersma, personal communication, October 4, 2021), the expected hydrological response time of the Rhine basin is approximately 10 days. Since GRADE is particularly used for analysis of discharge extremes, the analysis of the 10-day rainfall sum therefore makes the most sense. In the following subsections, extreme rainfall of respectively RACMO and SEAS5 are therefore compared with the observation-based passive dataset and the WG in the Rhine basin for a 10-day sum. However, equivalent figures for a multi-day sum ranging from 1 to 10 days are presented in Appendix B.

RACMO

Figure 3.1 shows a frequency distribution for 5 different datasets, being original and corrected RACMO, the observation-based passive dataset and two sets of the WG based on the observation-based passive dataset and corrected RACMO for a 10-day sum. In all panels of the Figure, annual maximum 10-day sums of precipitation are plotted against

their statistical return period. From left to right, the maxima are determined for the entire calendar year (CY), the summer half year (SHY, April till September) and the winter half year (WHY, October till March). The latter two can therefore be used to identify how summer and winter events behave for the different datasets.

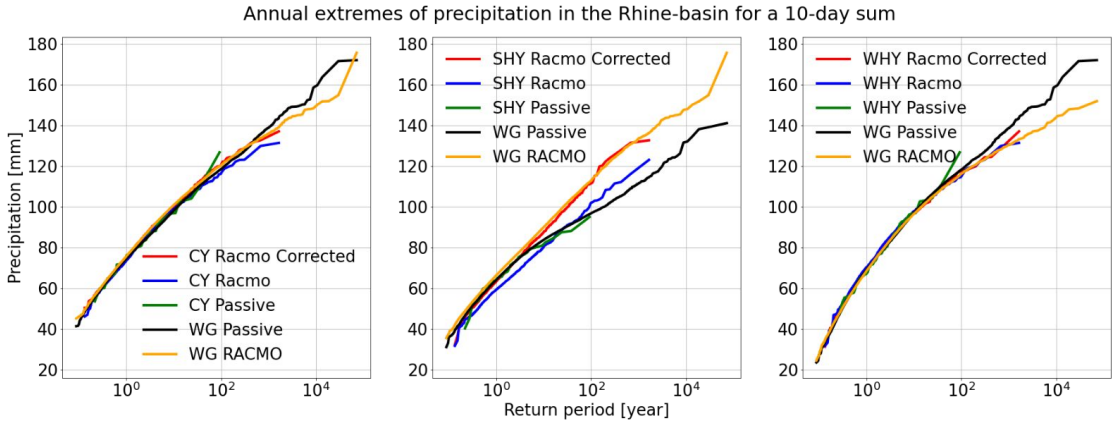

Figure 3.1: Precipitation extremes in the Rhine basin for a 10-day sum, for RACMO (corrected and uncorrected), the observation-based passive dataset, and the RACMO and passive-based WG. The vertical axis shows the 10-day precipitation magnitude in mm, the horizontal axis shows the corresponding return period to the extreme events. From left to right, the panels represent calendar year, summer half year and winter half year.

Extremes of all datasets look much alike for a 10-day precipitation sum in the Rhine basin as can be deduced from the CY panel (left). This firstly implies that the bias correction of the Rhine basin does have hardly any influence on the results of the precipitation extremes.

Secondly, the fact that all graphs in the CY panel are similar means that both RACMO and the WG likely provide realistic descriptions of the climatology of extreme 10-day precipitation. On the one hand it shows that the WG simulates long-duration multi-day sums (10-day in this example) well in the Rhine basin and the underestimation problem due to resampling as described in Section 1.2.1 has no significant impact. As described earlier, the flattening effect due to resampling is a lot less significant for longer-duration multi-day sums. On the other hand, the similarity between RACMO and the WG increases confidence both in the WG as in the RACMO results.

For large extremes (and high return periods) there is a slight difference between the datasets, with slightly higher WG extremes compared to RACMO. However, this could be the result of the low number of datapoints in this part of the graph (large return periods have few datapoints). The representation of the tail of the graph is thus a lot noisier, which may explain the differences between the datasets.

In Appendix B, figures equivalent to Figure 3.1 are depicted for short-period multi-day sums. With shorter-period multi-day sum, the differences between the CY maxima of all datasets increase. For the WG, the downward curving effect is clearly visible from a 5-day sum and shorter and is increasingly profoundly present for short-period multi-day precipitation sums. RACMO and SEAS5 on the other hand do not show this behaviour

and remain approximately linear (exponential considering the graph has a logarithmic scale). Although discharge extremes of the entire Rhine basin are likely to respond to large rainfall sums, exceptional events, or events in part of the basin, may respond to shorter-period multi-day precipitation sums, where the difference between RACMO and the WG is a lot more substantial. As such, potential underestimation of the WG for small multi-day sums must be considered.

The middle and right panel of Figure 3.1 show extremes based on SHY and WHY. As opposed to CY, particularly for SHY, the RACMO dataset does not resemble the observations and their corresponding WG result. Particularly the larger extremes of the RACMO dataset exceed the WG. Furthermore, the observation-based passive dataset curves down for large extremes, leading to lower corresponding WG values.

There is a large difference in extremity between SHY and WHY for the passive dataset and WG, whereas winter and summer events are approximately equally extreme for the datasets based on RACMO. This indicates that summer events are a lot more dominant in the RACMO dataset than in the WG and the observations. Potentially, this indicates that CY extremes involve a lot more summer events for RACMO compared to the WG. This contradicts the assumption that extreme precipitation events only occur in winter.

SEAS5

Figure 3.2 is similar to Figure 3.1, for SEAS5 instead of RACMO. Again, extremes are considered for different datasets and for respectively CY, SHY and WHY. Again, equivalent figures for a multi-day sum ranging from 1 to 10 days are presented in Appendix C. The results of the SEAS5 dataset are largely comparable to the RACMO dataset (Figure 3.1). This hints for the reliability of both datasets and strengthens presumptions of the previous paragraph. Both RACMO and SEAS5 show summer dominance, which increases the likeliness of an underestimation of summer extremes in the WG. Furthermore, the SEAS5 dataset also largely coincides with the WG for CY maxima with a 10-day sum, enforcing the argument that the WG works decently for these long-period multi-day sums.

Considering the short-period multi-day sums depicted in Appendix C, the SEAS5 dataset evidently does not curve down in Figures C.1-C.3, which emphasises the difference with the WG that does curve down. This difference is even more pronounced for SEAS5 than for RACMO due to its larger size. Consequently, particularly for short-period multi-day sums, SEAS5 and its corresponding WG result consist of higher precipitation extremes than the WG dataset based on the observation-based passive dataset. The latter may therefore likely underestimates precipitation amounts for larger extremes.

To summarise, the WG, RACMO and SEAS5 do not differ a lot in the Rhine basin considering a 10-day precipitation sum. On the one hand this means that for large multi-day sums in the Rhine basin, the WG does not seem to reach its methodical limits in the range of the return periods considered. On the other hand it speaks in favour of the usefulness of the RACMO and SEAS5 dataset. When considering shorter-period multi-day sums, the WG curves down for larger return periods, whereas the RACMO and SEAS5 dataset do not. Consequently, for shorter multi-day sums in the Rhine basin (5 days for RACMO, 8 days for SEAS5), the RACMO and SEAS5 dataset contain rainfall events of higher extremity than the WG, which hints on underestimation by the WG

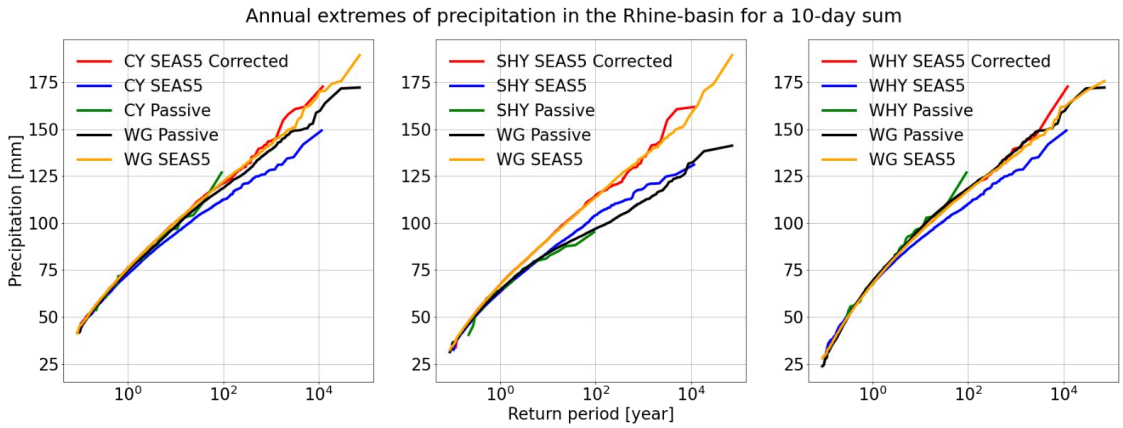

Figure 3.2: Precipitation extremes in the Rhine basin for a 10-day sum, for SEAS5 (corrected and uncorrected), the observation-based passive dataset and the SEAS5 and passive-based WG datasets. The vertical axis shows the 10-day precipitation magnitude in mm, the horizontal axis shows the corresponding return period to the extreme events. From left to right, the panels represent calendar year (CY), summer half year (SHY) and winter half year (WHY).

extremes in these cases. Additionally, the RACMO and SEAS5 datasets contain larger summer events than the WG. This could potentially mean that more extreme events happen in summer in these datasets, which contradicts the assumption that extreme precipitation mainly occurs during winter.

3.1.2 Meuse frequency analysis

This subsection provides an analysis of precipitation extremes in the Meuse basin for all considered datasets. Since the Meuse basin is considerably smaller than the Rhine basin, a shorter hydrological response time of 5 days is assumed in this report. As such, the main frequency analysis performed in this section is based on a 5-day precipitation sum. Equivalent figures for different multi-day sums are presented in Appendix B and Appendix C].

RACMO

Figure 3.3 shows a frequency distribution for all 5 datasets considered in this report. The Figure is typically equal to Figure 3.1 in Section 3.1.1, which can be considered for an extensive Figure description (note that instead of a 10-day sum, a 5-day sum is considered).

Similar to the results in the Rhine basin, the datasets considered in the CY panel of Figure 3.3 behave quite similar. As such, for a 5-day precipitation sum, RACMO and the WG likely contain fairly realistic 5-day extremes. However, note that for return periods around 10^2 years and larger, the WG starts to curve down and the difference with RACMO and corresponding WG increases. For very large extremes, the RACMO based WG contains higher 5-day precipitation events than the original WG, presumably due to underestimation of the WG based on the short observation-based passive dataset.

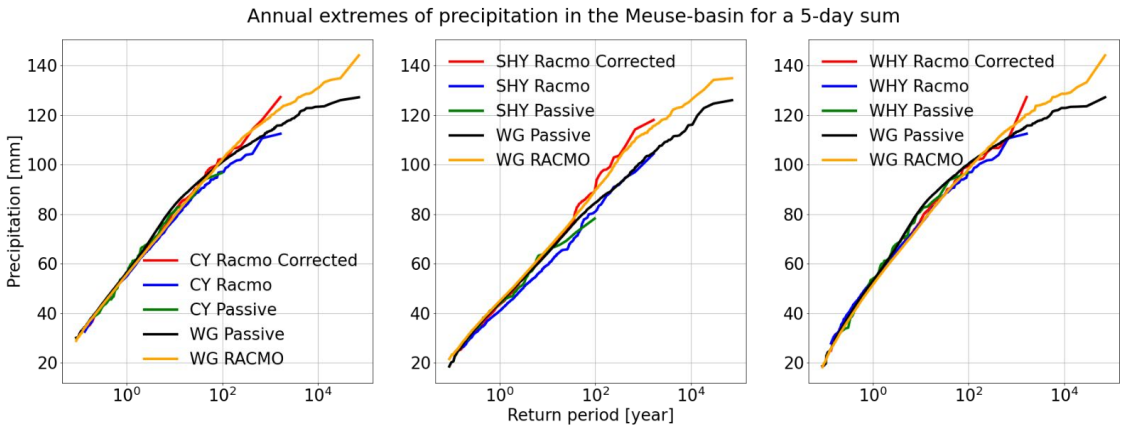

Figure 3.3: Precipitation extremes in the Meuse basin for a 5-day sum, for RACMO (corrected and uncorrected), the observation-based passive dataset and the RACMO and the passive-based WG. The vertical axis shows the 5-day precipitation magnitude in mm, the horizontal axis shows the corresponding return period to the extreme events. From left to right, the panels represent calendar year (CY), summer half year (SHY) and winter half year (WHY).

Although of smaller magnitude compared to the Rhine basin, there is a difference between RACMO (and corresponding WG) and the observation-based passive dataset and corresponding WG for SHY. Also, summer extremes in the Meuse basin are more dominant in RACMO (and its corresponding WG) than for the original WG.

Although having an approximately equal maximum event, WHY extremes of the WG are generally larger than SHY WG extremes. This difference between SHY and WHY is less distinctive for the RACMO dataset. Therefore, it is more likely that the CY extremes in the RACMO dataset consist of relatively more summer extremes than the WG.

In appendix B, an overview of all multi-day sums ranging from 1-10 is provided. Additionally, the results of a 2-day precipitation events are highlighted in Figure 3.4. Although the response time of the Meuse basin to precipitation events is generally around 5 days, an extreme event in July 2021 (Kreienkamp et al., 2021) has demonstrated that 2-day extremes may also lead to extreme water levels and consequently to flood hazard.

There is a substantial difference between the RACMO datasets and the original WG for large extremes as distinctly shown in all panels in Figure 3.4. As expected, the WG seems more prone to methodical artefacts for small multi-day precipitation sums. As such, RACMO contains a lot higher 2-day extremes than the WG, which likely underestimates 2-day precipitation extremes. The relative difference between the precipitation extremes of the datasets is considerably larger for 2-day events compared to 5-day events.

RACMO (and corresponding WG) has larger extremes than the WG for both SHY and WHY. However, the difference between the two datasets is larger for summer events. It therefore appears that particularly 2-day summer extremes are larger than primarily expected based on the WG data. This may also explain why the July 2021 event in the Meuse basin came as such a surprise.

Low representation of small multi-day summer extremes, or extremes in general, could potentially lead to problems in the Meuse basin. Although not yet hydrologically deter-

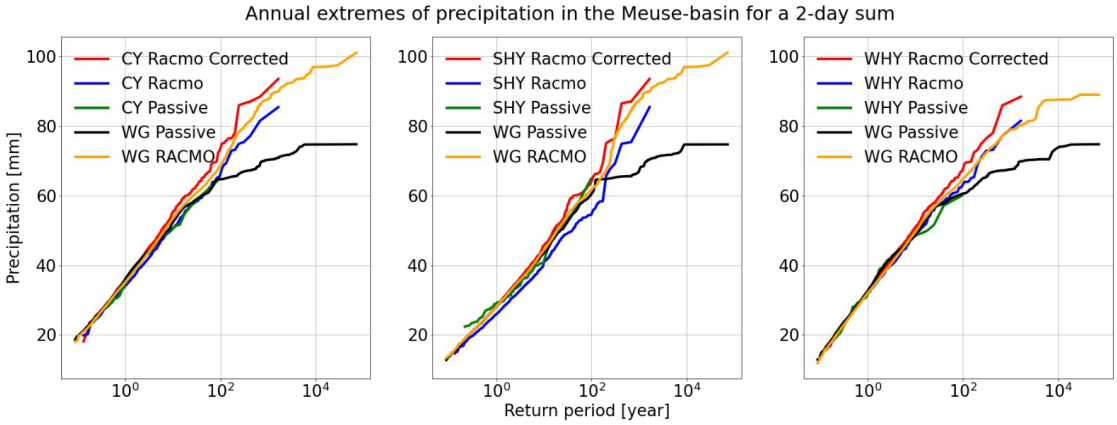

Figure 3.4: Precipitation extremes in the Meuse basin for a 2-day sum, for SEAS5 (corrected and uncorrected), the observation-based passive dataset and the SEAS5 and the passive-based WG. The vertical axis shows the 2-day precipitation magnitude in mm, the horizontal axis shows the corresponding return period to the extreme events. From left to right, the panels represent calendar year (CY), summer half year (SHY) and winter half year (WHY).

mined, higher precipitation extremes likely lead to higher runoff magnitudes and water levels. Based on the RACMO dataset, these hydrological extremes may (depending on the hydrological model) occur with higher probability compared to the WG dataset, on which current decision-making is based. Considering the apparent underestimation of short-period events by the current WG, this could lead to a false sense of security.

SEAS5

Figure 3.5 is similar to Figure 3.3, but instead of RACMO, SEAS5 is compared to the observational dataset and the WG.

The behaviour of the RACMO dataset and the SEAS5 dataset is much alike. However, given that the original SEAS5 dataset is longer, its difference with the WG is more distinctly visible for extremes with a return period larger than 1000 years. The increasing difference between the WG and SEAS5 strengthens the presumption that for large return periods, the WG may underestimate 5-day precipitation events for the Meuse.

Like the RACMO dataset, the SEAS5 dataset provides larger summer extremes than the WG for high return periods. Analysis of summer events of SEAS5 therefore leads to the same conclusions as for RACMO.

The 2-day precipitation extremes of SEAS5 are plotted in Figure C.1 of Appendix C. Although the main conclusions of the RACMO dataset also apply for the SEAS5 dataset, it is noteworthy that the difference between the WG and SEAS5 (and its corresponding WG) are considerably larger than for RACMO. Besides, the graphs (CY, SHY and WHY) all have an upwards bending tail, which causes the difference between the WG to increase even more for larger exceedance probabilities. Therefore, based on both RACMO and SEAS5, the WG likely underestimates 2-day precipitation events in the Meuse basin.

To summarise, in the Meuse basin all datasets are relatively comparable for events

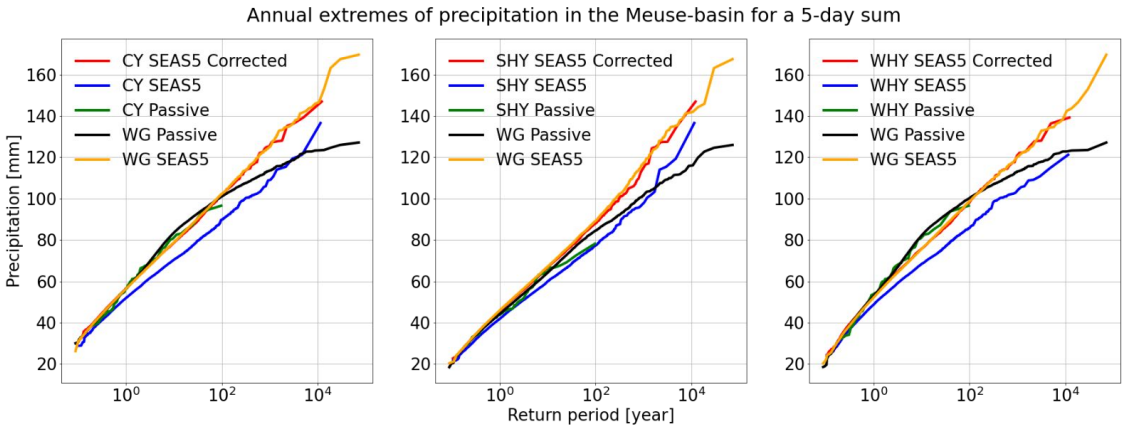

Figure 3.5: Precipitation extremes in the Meuse basin for a 5-day sum, for SEAS5 (corrected and uncorrected), the observation-based passive dataset and the SEAS5 and the passive-based WG. The vertical axis shows the 5-day precipitation magnitude in mm, the horizontal axis shows the corresponding return period to the extreme events. From left to right, the panels represent calendar year, summer half year and winter half year.

with a return period lower than 100 years. The original WG seems to underestimate larger precipitation extremes (for a 5-day sum and more extremely for a 2-day sum), with SEAS5 (and corresponding WG) containing considerably larger extremes than RACMO (and corresponding WG). The large difference between the original WG and those of RACMO and SEAS5 is predominantly visible for the SHY. It is therefore expected that the CY extremes consist of more summer events for RACMO and SEAS5 compared to the WG and that summer events may be more important in flood generation than suspected before, but more research on the hydrological response of the RACMO and SEAS5 dataset is therefore required.

3.1.3 Vecht frequency analysis

The Vecht is a relatively small basin and is therefore expected to have a shorter hydrological response time than the Meuse. However, the most likely hydrological response time is still unknown. As an indication, extremes of 4-day precipitation sums for all datasets are presented. However, an overview of extremes with other precipitation sums is provided in Appendix B and Appendix C.

RACMO

In Figure 3.6, extremes of RACMO are compared with extremes of the WG and the observation-based passive dataset. The Figure is typically equal to Figure 3.1 in Section 3.1.1, which can be considered for a more extensive Figure description (note that a 4-day precipitation sum is considered).

In the Rhine and Meuse basin, the RACMO dataset shows quite some similarities with the original WG and the observation-based passive dataset. Only for relatively large return periods, differences are visible. In the Vecht basin, these differences already occur for return periods of 10^2 years (the observation-based passive dataset that is used

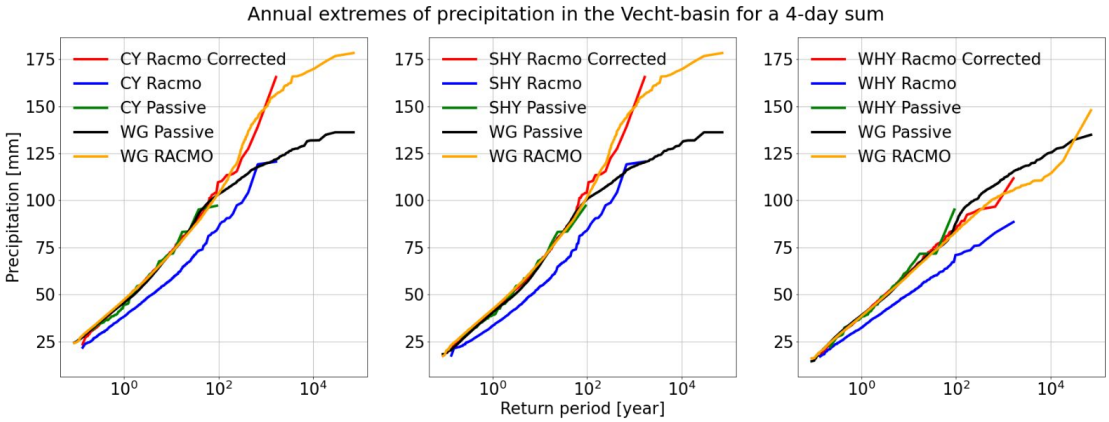

Figure 3.6: Precipitation extremes in the Vecht basin for a 4-day sum, for RACMO (corrected and uncorrected), the observation-based passive dataset and the RACMO and the passive-based WG. The vertical axis shows the 4-day precipitation magnitude in mm, the horizontal axis shows the corresponding return period to the extreme events. From left to right, the panels represent calendar year, summer half year and winter half year.

for the Vecht is 65 years in length). Again, the WG curves down, whereas the RACMO dataset (and corresponding WG) even show a slightly upwards curving bend. For a 4-day sum, the precipitation extremes in the WG are thus considerably lower compared to RACMO and its corresponding WG.

When comparing the CY panel to the SHY and WHY panel it is clear that the annual maxima of the RACMO dataset are largely dominated by summer events. Such dominance is not clear for the WG, where summer events and winter events are comparable. It is likely that the RACMO dataset is better at describing convective storms that often occur in summer, which can be very important extremes in small basins like the Vecht. This is further evaluated in Section 3.3 and 3.2. Depending on the hydrological result of the datasets, the difference between RACMO and the WG and potential underestimation of the WG, particularly for summer events, may as well cause a false sense of security regarding flood hazard to occur. Further investigation of the hydrological result of all datasets is therefore highly recommended.

In Appendix B, the Vecht basin extremes are presented for all other multi-day sums ranging from 1 to 10 days. Note that also for the Vecht basin, the relative difference between the WG and the RACMO datasets increases for shorter-period multi-day sums.

SEAS5

Figure 3.7 is similar to Figure 3.6, but instead of RACMO, SEAS5 is compared to the observational dataset and the WG.

The results of the SEAS5 dataset are largely comparable to those of the RACMO dataset. Since this dataset is larger, it contains even larger extremes, which makes the difference with the WG for very large return periods even more sizeable. The SEAS5 dataset also shows the importance of summer events in the Vecht basin, which strengthens the conclusions done based on the RACMO analysis.

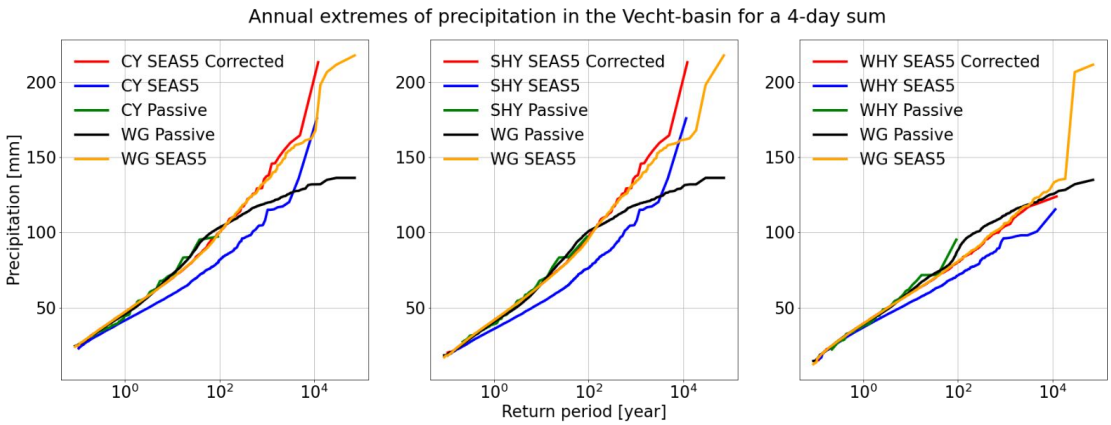

Figure 3.7: Precipitation extremes in the Vecht basin for a 4-day sum, for SEAS5 (corrected and uncorrected), the observation-based passive dataset and the SEAS5 and the passive-based WG. The vertical axis shows the 4-day precipitation magnitude in mm, the horizontal axis shows the corresponding return period to the extreme events. From left to right, the panels represent calendar year, summer half year and winter half year.

To summarise, the RACMO and SEAS5 datasets contain considerably higher extremes compared to the WG for the Vecht basin. Compared to these synthetic datasets, the WG largely underestimates summer dominance, which could potentially lead to underestimation of hydrological extremes and a potential false sense of security. Further research on the hydrological response of all datasets is therefore highly recommended. Note that WG underestimation and summer dominance is even clearer in the Vecht basin compared to the Rhine and Meuse basin.

3.2 Spatial distribution of precipitation

A significant advantage of particularly RACMO and to a lesser extent SEAS5 compared to the WG is the fact that spatial and temporal patterns are well described. In this section, several spatial visualisations of RACMO are presented.

In its resampling processs, the WG couples data by searching for similar events in the entire dataset. One of the disadvantages of this procedure is that different original precipitation events may be coupled. This may work well for basin average precipitation, but the problem is that spatial patterns become largely ambiguous using this method. As such, hydrological events that originate from certain precipitation events in the GRADE procedure can not be traced back to a valid individual event.

When not adjusted by the WG, the RACMO and SEAS5 datasets are largely continuous datasets (of considerable length). Although the coupling process of the datasets may lead to several jumps in the timeseries, in general, spatial patterns in these two datasets make great sense. In Figure 3.8, the spatial patterns of the largest summer events of the RACMO dataset are presented. Again, the 10, 5 and 4-day precipitation sums are used for respectively the Rhine, Meuse and Vecht basin. Summer events are used in this Figure, because spatial patterns are more explicit for convective events that occur in

this season. Note that only the RACMO dataset is considered in this section, but that the SEAS5 dataset is expected to show the same kind of results. Bear in mind that if RACMO and SEAS5 are resampled by the WG, these datasets would also provide a more random spatial pattern.

Figure 3.8 shows the spatial distribution of precipitation over the Rhine (a,b), Meuse (c,d) and Vecht (e,f). The spatial extent is presented as precipitation sum, but also as return period of the extreme event in every catchment. Particularly in the Rhine basin, where precipitation events may vary largely with geographic location in the catchment, this method of presenting precipitation allows for relative comparison between the catchments. Note that the three basins differ largely in size (see Section 2.1).

In all three basins, it is clear that the maximum summer event of the dataset only covers part of the entire basin. The extreme summer events are likely to be convective, covering only a small spatial extent and are therefore a lot more location specific. This is particularly clear in the Meuse and Vecht basin, where a large gradient is visible over the basins. This is important, because different parts of a basin may have a different influence on its hydrological response. In the Meuse basin for example, the Southern part of the basin has a slow response time to precipitation events, whereas in the Northern Ardennes, quick peak flows occur due to large precipitation events. The event type that is visualised in Figure 3.8 c and d, would thus lead to a rapid hydrological response, whereas an opposite precipitation distribution would lead to a much slower, more spread out hydrological response. This indicates the importance of capability of a dataset to describe the spatial extent of precipitation events. Such an improved spatio-temporal description is also expected to allow for a more accurate description of the (non-)coincidence of discharge peaks from the main river trunk and its tributaries in The Netherlands.

In general, in contrast to the WG, RACMO is capable of portraying the spatial distribution of precipitation events. This enables analysis of more location-specific precipitation events on the hydrological response in a catchment, for example, evaluation of precipitation event types that lead to flooding.

An important detail of RACMO and to lesser extent SEAS5, is the fact that these two datasets are available at respectively hourly and 6-hourly frequency. This means that certain extreme events can be studied with even higher detail. For example Figure 3.9 shows a link to an animation of the most extreme event of the Vecht, which shows large similarities with the events that occurred in the Meuse basin in July 2021. If multiple of these events are present in RACMO, this could potentially mean that such events may occur more frequently than we expect. The RACMO dataset provides the opportunity of detailed analysis of such events and their hydrological response.

Another striking element of the RACMO animation of the Vecht is the short duration of this most extreme event. Although a 4-day precipitation sum is currently considered the most relevant for floods for this basin, most of the precipitation seems to fall in 36 hours. Similar patterns are visible for the Meuse basin. In the next section, more elaboration on the temporal evolution of these summer events is provided.

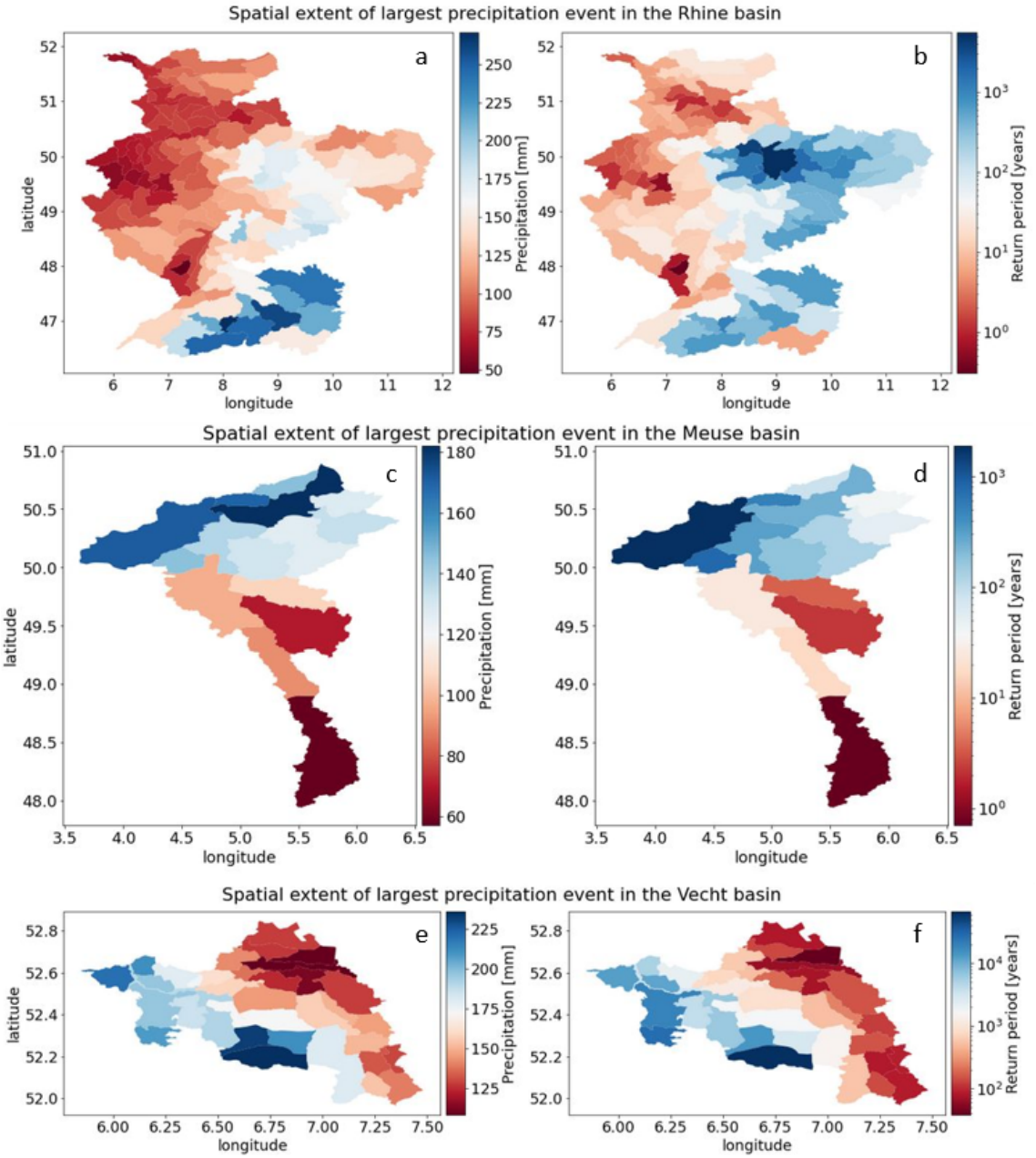

Figure 3.8: Spatial distribution of the largest RACMO multi-day basin-averaged precipitation summer events per basin. Figures a, c and e present absolute precipitation amounts of the event, Figures b,d and e represent the corresponding return period to enable intercomparison between catchments independent of climatic differences between these catchments.

Figure 3.9: Link to animation of the most extreme event in the Vecht. The Vecht basin is outlined in red. The event occurs on 7 September, at the end of the animation.

3.3 Temporal distribution of precipitation

In general, there is a difference between summer and winter precipitation. Where extreme winter events generally occur in large-scale depressions, summer events can also include weather systems with significant contributions of high-intensive convective (thunder)storms. This difference is visible both spatially and temporally.

The animation of Figure 3.9 showed that the maximum 4-day precipitation event in the Vecht basin seemed to fall in approximately 36 hours. Particularly in such a small sized basin, short extreme events may be important for the hydrological response of the catchment. Therefore, it is important to test whether datasets are capable of describing such events.

In this section, the temporal distributions of the most extreme summer events of the three basins are shown for the RACMO dataset with an hourly timescale. Besides, the RACMO dataset summer extremes are compared to the maximum summer event of 2000 years of WG (for simplicity of computation, the first 2000 years of the WG are used). Figure 3.10 a, b and c show the two largest multi-day summer events for respectively Rhine, Meuse and Vecht. To depict the characteristics of these events, they are portrayed with an hourly timescale. Since the bias correction has not been performed on hourly timescale, the uncorrected dataset is used in this Figure.

In the Rhine basin, the largest 10-day sum is represented by one or several relatively low precipitation peaks, which generally spread out over a few days. Due to the size of the basin, very extreme short and small-sized events are not dominant in the maximum 10-day sum summer events. Instead, events that spread out over a larger part of the basin and last a little longer, are more likely to result in basin extremes. However, for catchment-scale, these short convective events may be more important. In studies where catchment scale is required, the Rhine basin needs a dataset that can describe convective events well.

For the Meuse and particularly the Vecht basin, which are significantly smaller in size than the Rhine basin, summer events seem to be more dominated by extremes of 1 or 2 days with extreme precipitation amounts. Compared to the basin size, short extreme events can spread out over a relatively large part of the basin and can therefore become dominant summer events. This illustrates the importance of a dataset that describes

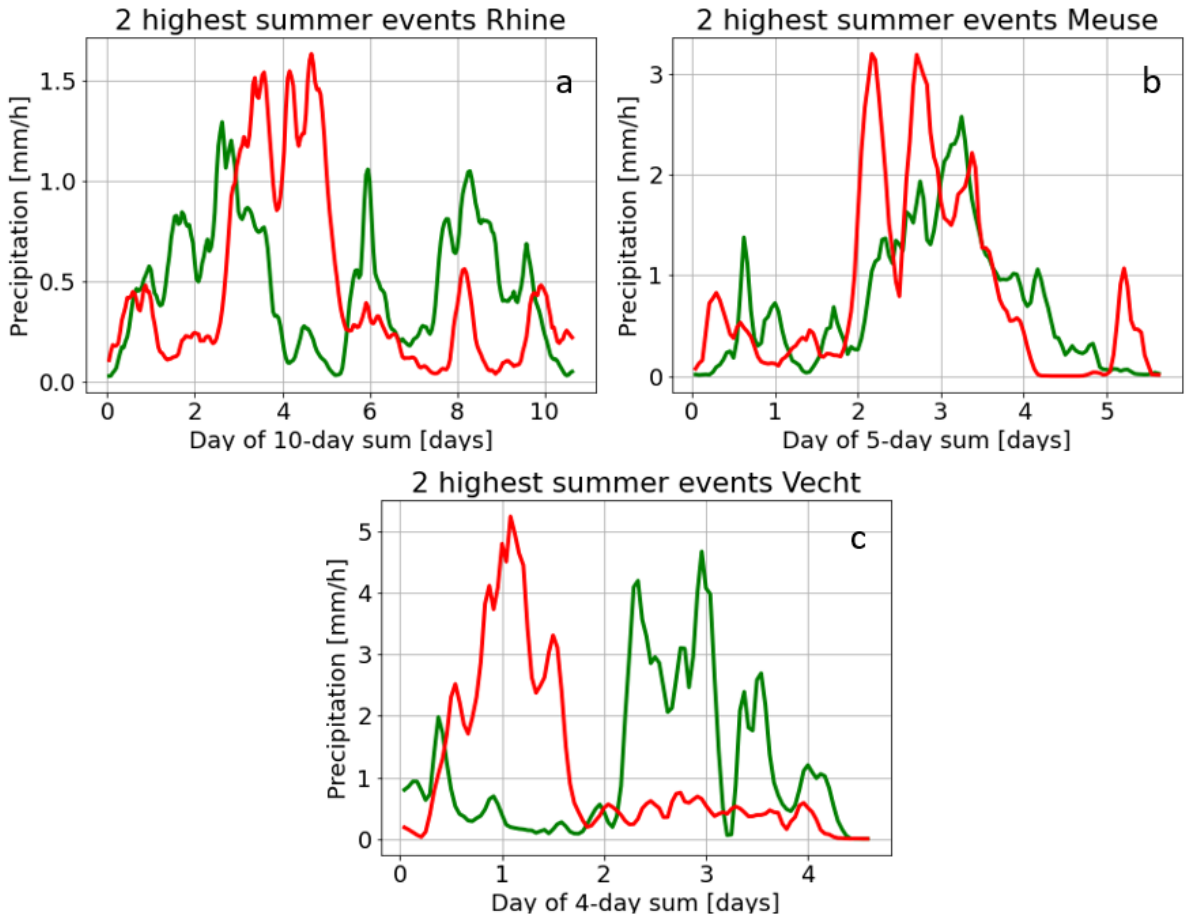

Figure 3.10: Visualisation of the two highest multi-day summer events in every considered basin. The green line represents the maximum event and the red line the second largest event. The horizontal axis shows the day of the multi-day sum corresponding to the event. The vertical axis shows precipitation amounts in mm/h. Note that all graphs are uncorrected RACMO data.

such convective events well.

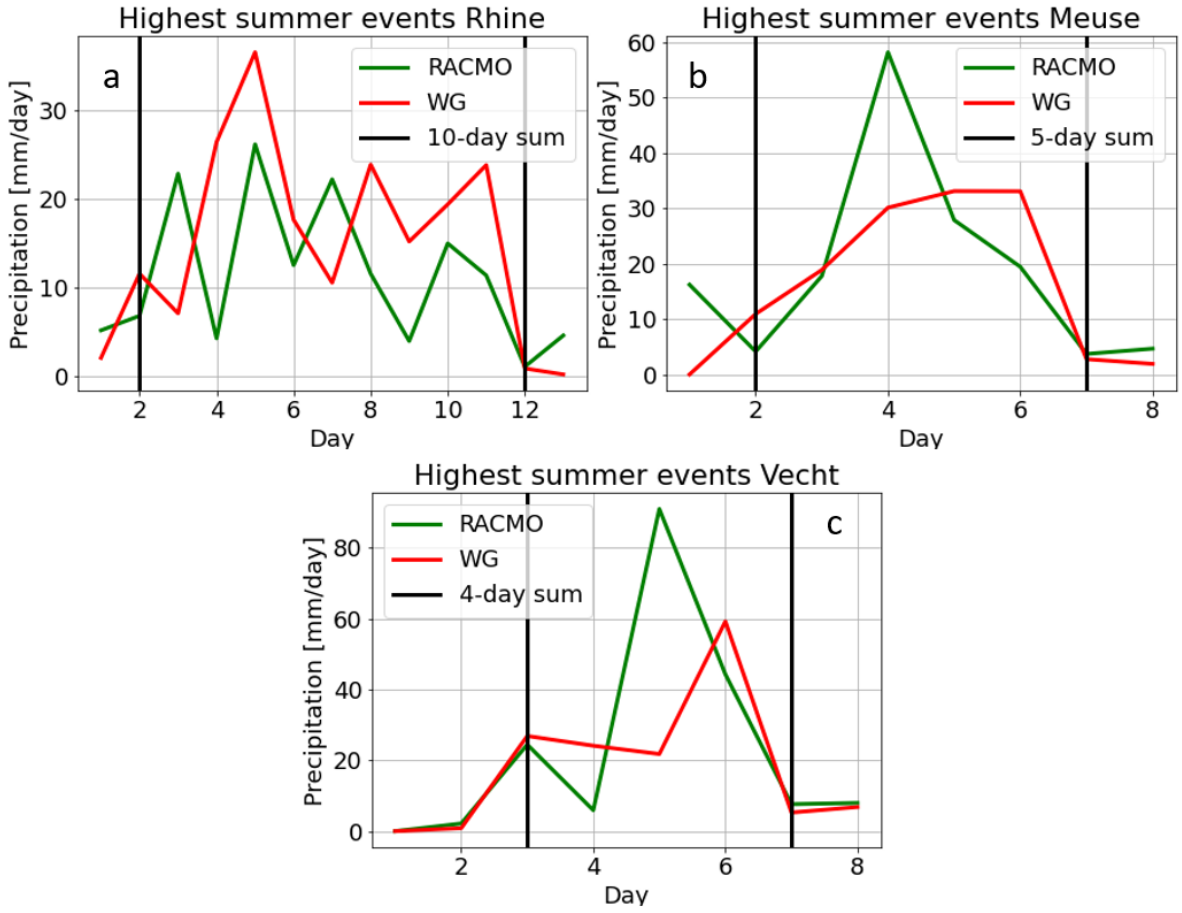

Figure 3.11: Comparison of the largest summer extremes of RACMO (green) and the WG (red). The vertical black lines show the start and end of the multi-day period that is considered in every catchment. The vertical axis shows precipitation in mm/day.

In Figure 3.11 a,b and c, RACMO's most extreme summer event is compared with the WG for all basins. Note that the corrected RACMO dataset is used here, with a daily timescale. Particularly for the Vecht and the Meuse basin, the peaks of the RACMO dataset are more than 30 percent higher than the WG peaks. In general, peak magnitude of the WG is limited to the largest peak in the source data. Since this dataset only consists of 65 years, it is likely that it does not contain very extreme (convective) peaks. This way, particularly for the Vecht and Meuse basin, the WG is therefore not able to capture the nature of the summer extremes, which consist of shorter events with large peaks. The RACMO dataset (and also the SEAS5 dataset) is able to simulate such type of convective events, as shown by the green line in Figure 3.11 b and c. This is another advantage of using RACMO (and SEAS5) over the WG[1].

[1]Note that we do not explore the capability of RACMO and/or SEAS5 in representing strong convective situations. It is likely that - although RACMO has a resolution of 12km - these models are too coarse to simulate convective situations adequately. A possible solution might be to use non-hydrostatic

To summarise, the short, extreme characteristics of summer events seem to be well described by RACMO (and expectedly also SEAS5). As the WG is based on resampling of the the observation-based passive dataset, it does not contain these convective extremes. The WG can not describe short events of extreme magnitude, which may lead to inaccurate weather descriptions, particularly in the Vecht and Meuse basin. This meteorological differences between RACMO and WG could potentially lead to different hydrological responses as well.

3.4 Dry spells

Besides extreme precipitation amounts, also dry spells are important. Here we present a preliminary investigation of long dry periods. We use the minimum 60-day precipitation over the whole basin as a measure for the extremity of dry periods. Figure 3.12 shows that RACMO shows a good representation of dry spells, which even improves further after bias correction.

models (e.g. HARMONIE) to downscale RACMO events.

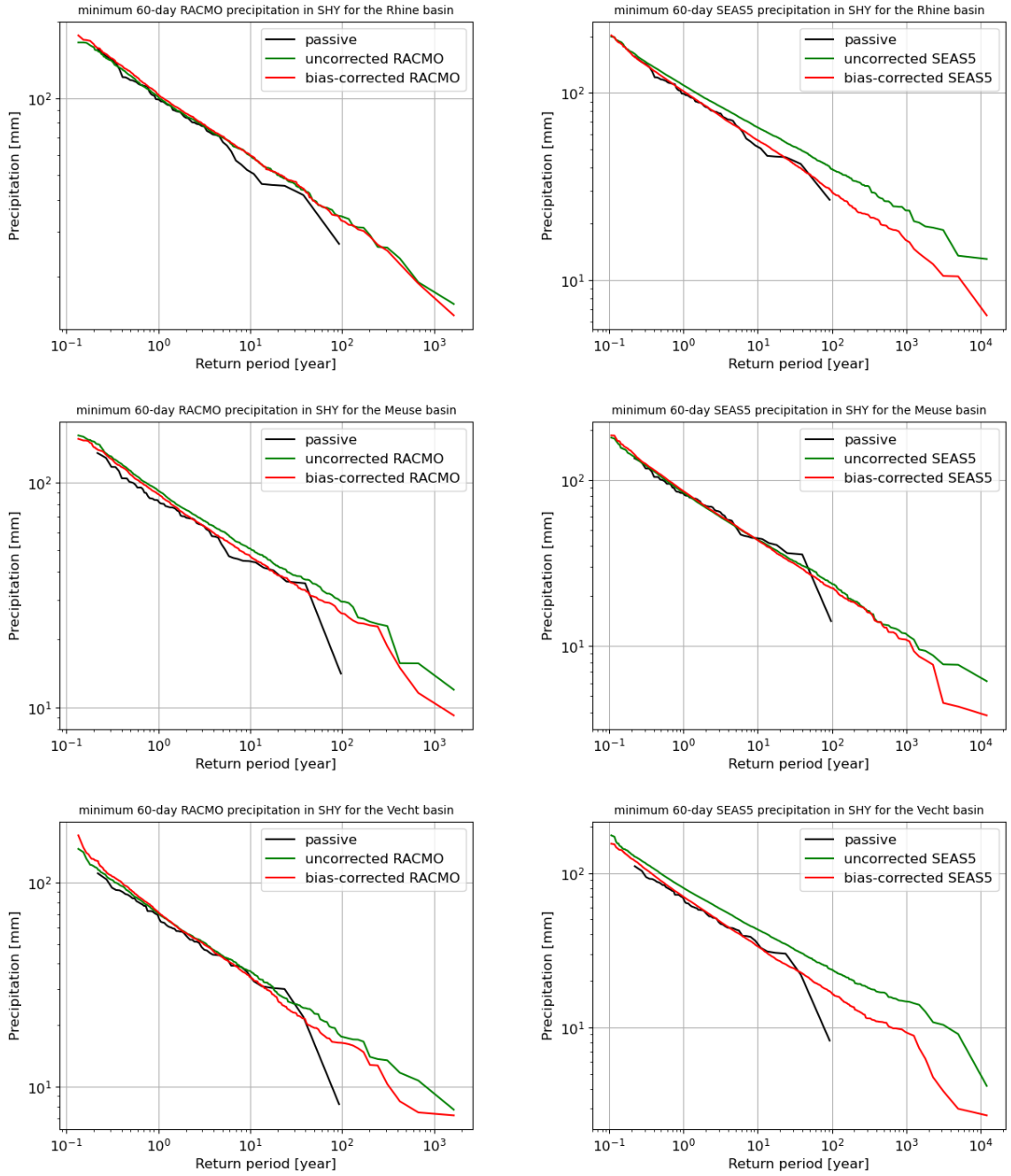

Figure 3.12: Gumbel plots of the minimum 60-day cumulative precipitation over the whole basin for RACMO (left) and SEAS5 (right) for the Rhine, Meuse and Vecht, compared with the observation-based passive dataset (black). The logarithmic vertical axis shows precipitation in mm/day.

Chapter 4

Discussion

This chapter provides a short overview of the uncertainties of the methodology used in this study, as well as the basic uncertainties of the extreme analysis. Furthermore, the first hydrological results are presented.

4.1 Uncertainties

There are several uncertainties that may influence the results of this study. In this section, the uncertainties of the datasets as well as the correction are briefly addressed. Besides, comments are made about the general uncertainties of extreme analysis.

4.1.1 Data

As described in Section 2.2, there are several datasets used in this study. A detailed overview of their accuracy is provided in their source documents. However, in this subsection, several general comments about data uncertainty are made.

The observation-based passive dataset is used as reference dataset in this study and is also used for correction of RACMO and SEAS5. Reliability of this dataset is therefore essential. In general, the reliability of data depends on the density of stations used as source of the gridded data product. In case of the observation-based passive dataset, HYRAS stations are used, which are densely available over the Rhine and Vecht basin to provide a reliable dataset. Note that the observation-based passive dataset in the Meuse basin has several different sources. Theoretically this means that for the Meuse basin, there is a difference in data accuracy for different time periods. However, all three source datasets consist of a large number of weather stations, implying that the passive dataset is reliable for the Meuse basin as well.

For a detailed overview of the reliability of RACMO and SEAS5, their source documents can be considered (van Meijgaard et al., 2008; Johnson et al., 2019). It should be noted here that the resolution of both datasets is considerably smaller than the resolution of the observation-based passive dataset. This could be one of the reasons why the difference in precipitation magnitude between RACMO and SEAS5 and the observation-based passive dataset is largest in the Alpine regions, as elevation effects may be better captured by the latter dataset. However, both RACMO and SEAS5 have been corrected

for bias using the observation-based passive dataset, which likely filters some of these uncertainties.

Furthermore, coupling of half years in the SEAS5 dataset may have led to several discontinuities in the entire timeseries. Regarding the fact that SEAS5 is a very long timeseries, chances are non-negligible that several of the CY, SHY or WHY multi-day maxima are found within the coupling range, allthough the probability of having a maximum in both ranges to be combined will be much smaller[1]. However, due to the relative size of every SEAS5 member (6 months) compared to the maximum multi-day sum (10 days), it is expected that this effect is limited. Note that due to the resampling process, the WG may have a discontinuity on every day and therefore the SEAS5 alternative is a lot more continuous. Furthermore, the results have shown that RACMO and SEAS5 look a lot alike. Since RACMO is a lot more continuous than SEAS5 and the conclusions of the two datasets are similar, it seems that the coupling process is not of large influence on the study results.

4.1.2 Correction

Both RACMO and SEAS5 have been corrected in this study. This is common procedure in many climate studies. The bias correction influences RACMO and SEAS5 in several ways. Both averages and extremes are altered, as well as the timing of precipitation. However, in the vast majority of catchments, the correction affects the statistics with less than 20 percent.

The choice of making a monthly bias correction is justified by the importance of seasonal influence in this study. The type of precipitation events may differ largely between winter and summer. A monthly bias correction accounts for these differences. This study has shown the importance of distinguishing between summer and winter events, which explains the use of a monthly bias correction.

Furthermore, the choice was made to provide a separate bias correction for every sub-catchment of the Rhine, Meuse and Vecht, rather than for the whole basins.

We emphasize that the bias correction leaves the meteorological situations that lead to (extreme) precipitation unaltered, both in space and time. The fact that correction of daily data leads to a correct climatology of multi-day extremes as well underlines the quality of both RACMO and SEAS5 in reproducing the temporal and spatial correlations of (extreme) meteorological events.

In Figures 3.1 - 3.7, both the original and corrected RACMO and SEAS5 have been plotted. In many graphs, there is no large difference between the two, showing the bias correction does not have a large influence on extreme representation in many cases. In some other cases, the uncorrected SEAS5 and RACMO extremes are slightly lower in magnitude than the corrected extremes. However, the climatological behaviour (linear line, no down-curving, same relationship between summer and winter, that is shown in Figures 3.1 - 3.7) is also visible for the uncorrected datasets. Consequently, the main

[1] A 10-day extreme has a probability of 10/182=6% to contain a concatenation-point. The probability that the concatenation leads to an extra 10-day-extreme is much smaller, as the extremes are only a small fraction of the whole dataset. Multi-day extremes may be slightly under-estimated as a small part will be cut by the end/start of the 182-day period, and will be combined with less extreme days and consequently is no longer extreme.

qualitative conclusions of this report are not affected by the bias correction.

4.1.3 Extreme analysis

One of the major complications in extreme analysis is the lack of reference. Like stated before, decent observation data only exist for approximately 70 years, meaning none of the return periods larger than 100 years can be verified. Outside this range, it is impossible to say with certainty which of the datasets considered in this study approaches reality the best.

Although neither RACMO, nor SEAS5 can be verified as realistic, the comparison with the WG in this study is still very useful. Methodically, the WG does not provide realistic results for large return periods (particularly for smaller multi-day sums). Underestimation of large extremes as often presented in Figures 3.1 - 3.7, is therefore evident. RACMO and SEAS5 show more linear behaviour in these plots and are independent of the methodical underestimation as for the WG. Besides, these datasets can be used as proxy for the magnitude of underestimation of the WG. Added to this, the fact that RACMO and SEAS5 are largely similar to the observation-based passive dataset for its range adds to the reliability of these datasets. In short, RACMO and SEAS5 seem more likely to approach reality than the WG, particularly for large return periods and small multi-day sums.

Lastly, whenever the WG is used to elongate the RACMO and SEAS5 timeseries, the methodical shortcomings of the WG also apply. Overestimation of the SEAS5 WG (section 1.2.2) is clearly represented in Figure 3.7 for WHY. The WG result of SEAS5 is considerably higher than the original SEAS5 dataset. In the resampling process, the extreme daily event (see WHY in Figure C.7 in Appendix C) is resampled too often, leading to the large 4-day precipitation amounts of Figure 3.7. A clear example of methodical underestimation (downwards curving of the graph) is visible for the RACMO based WG in Figure 3.4. Do note, however that these problems now occur for return periods of much lower frequency due to the length of the source datasets RACMO and SEAS5. As such, use of the WG based on RACMO and SEAS5 expectedly leads to a dataset less prone to the methodical shortcomings of the WG. Note, that if the WG of RACMO and SEAS5 is used, climate change effects and spatial or temporal patterns are more ambiguous due to the resampling process. Therefore, depending on the goal of a study, it must be well considered whether the original RACMO and SEAS5 or their corresponding WG results are used.

4.2 Preliminary hydrological results

The main focus of this research is the comparison of precipitation of the WG with RACMO and SEAS5 alternatives. For further evaluation in how far the results shown above are relevant for the final results in GRADE, it is very important that the meteorological results are converted to hydrological results (i.e. discharges). During the time scope of this research, there was no time to perform a detailed hydrological analysis. However, some preliminary results of the Meuse basin are provided here.

The hydrological results presented in this report are the results of the HBV model

of the Meuse, run by Deltares (Hegnauer, 2022). In Figure 4.1, discharge extremes at Borgharen are presented, with annual maximum discharges on the vertical axis and their corresponding return period on the horizontal axis. The observations of the Meuse are presented in black, whereas the observation-based WG, the RACMO and the SEAS5 results are presented in respectively orange, blue and green.

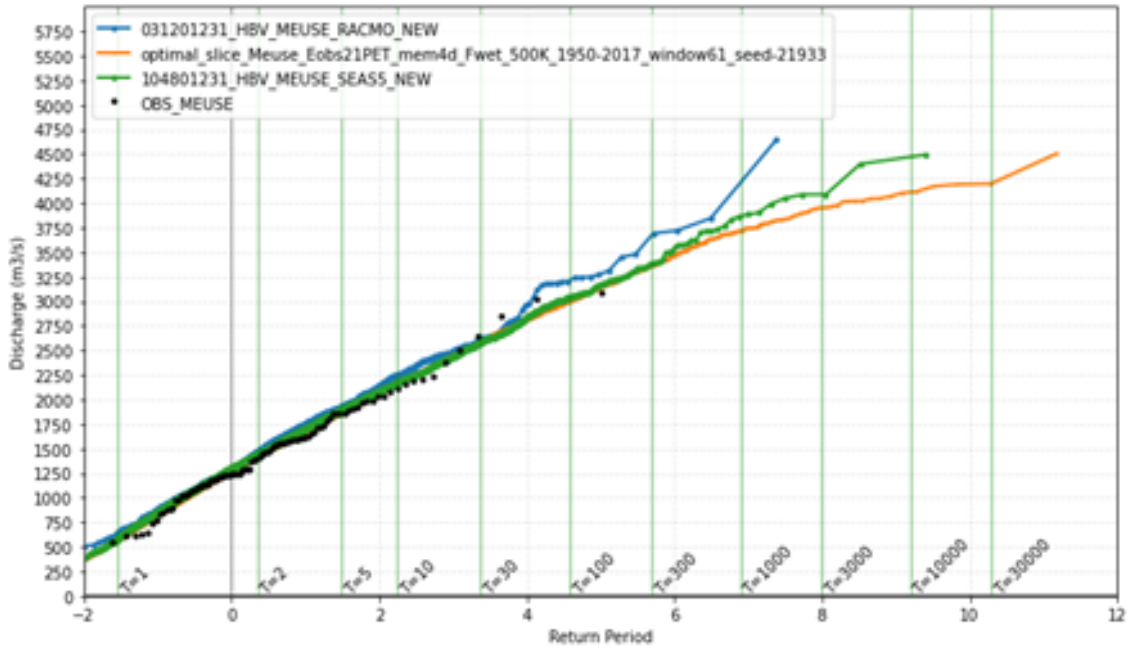

Figure 4.1: Discharge extremes in the Meuse basin, in which RACMO, SEAS5 and the WG are compared with observed discharges (Hegnauer, 2022).

For small extremes with low return periods, all datasets compare well and fit well to the observations. However, in the regime where observations lack, the discharge results differ per source dataset. Like for the precipitation statistics, the WG discharge result is bending downwards for large return periods. The results of both RACMO and SEAS5 do not bend downwards, which means these discharge datasets contain higher extremes than the WG-based discharge dataset. A first visual inspection shows a factor 50 difference in return period between RACMO and the WG result and a factor 5 between SEAS5 and the WG. As such, the difference between the datasets for precipitation is therefore also reflected in the discharge result in the Meuse basin.

These preliminary results of discharge extremes indicate the importance of further research on the hydrological response of SEAS5 and RACMO. This should include Figures like Figure 4.1 for the Rhine and Meuse basin, as well as a thorough analysis of the nature of the extreme events that occur and their timing.

4.3 Further research

There are several different studies that could potentially follow up this exploratory research. Several of them are briefly addressed in this section.

The most important follow-up research, particularly for Rijkswaterstaat, would be an analysis of the hydrological result of RACMO and SEAS5 in the Rhine, Vecht and Meuse. The results from this study suggest that particularly for the smaller basins, where small multi-day precipitation sums can be important in flood generation, precipitation extremes are likely underestimated. This implies that water levels and runoff will likely have higher extremes as well. The non-linearity of a catchment makes the exact influence of precipitation on streamflow hard to predict, which means thorough analysis of the hydrological results of RACMO and SEAS5 essential. It is particularly interesting to test the hydrological response on summer precipitation events, to see whether summer events are more likely to occur than currently expected. A potential way to study this is to compare hydrological results of summer events from RACMO and SEAS5 to the events that occurred in the Meuse basin in 2021.

Furthermore, an interesting follow-up research is to test other synthetic weather datasets to check the reliability of RACMO and SEAS5. One way to do this is by co-operating with various other countries and comparing their seasonal forecasting systems or climate models with RACMO, SEAS5 and the observation-based passive dataset, provided that these datasets are long enough. Another possibility is to extend the RACMO dataset by the new runs that are delivered in 2023. In theory, RACMO runs have been provided up to 2100, enabling creation of a larger precipitation dataset. This dataset can be used to verify conclusions and analyse the influence of climate change on extreme precipitation and potentially extreme discharges.

Lastly, this research has shown that summer events are more dominant in the RACMO and SEAS5 dataset compared to the WG and the observation-based passive dataset. An interesting starting point is to see whether the choice of the dataset affects the relative contributions to the annual precipitation maxima of summer and winter extremes respectively. This could potentially change the current perception of extreme precipitation events (and corresponding flood hazards) mainly occurring during winter. Furthermore, extra investigation of the hydrological result of precipitation extremes and particularly whether they occur in summer or winter may give insight in whether future flood events can also be caused by summer convective events. This could also provide a change of perception.

Chapter 5

Conclusion

The weather generator as currently used in GRADE shows several methodical imperfections, potentially leading to either under- or overestimation of large precipitation extremes. A potential solution is to replace the presently used source dataset (passive) with longer, synthetical datasets like RACMO and SEAS5. The goal of this study is to compare these synthetical datasets and their corresponding timeseries derived from the weather generator with the current source dataset and its corresponding timeseries derived from the weather generator (WG) in order to look on their performance to describe meteorological extremes. Beforehand the RACMO and SEAS5 datasets had to be bias-corrected .

Comparison of RACMO, SEAS5 and passive for several main climate parameters showed that climatological differences in the Rhine, Meuse and Vecht basin are well captured by RACMO and SEAS5. A quantile-mapping bias correction has been executed, in which RACMO and SEAS5 were translated towards the climatology of the passive dataset. The correction was performed per calendar month and for every individual catchment, leading to even better climatological similarities without overfitting.

Extreme value analyses show comparable extreme precipitation and return periods for 10-day precipitation sums of RACMO, SEAS5, passive and the WG in the Rhine basin. For multi-day sums smaller than 5 days, RACMO and SEAS5 (and their corresponding WG) contain higher precipitation events than the original WG, which hints on underestimation of the WG for such events. Similar underestimation might occur in the WG for the Meuse basin for a 5-day precipitation sum, for events with a return period larger than 100 years. Analysis of 2-day extremes, which can cause flood hazard as proven in July 2021, shows even larger differences between the WG and RACMO and SEAS5, as well as their corresponding WG result. Accordingly, the Vecht basins shows large underestimation of the WG for precipitation extremes compared to RACMO and SEAS5 and their corresponding WG result. In general, RACMO and SEAS5 and their corresponding WG either provide equal or higher magnitude precipitation extremes compared to the WG based on passive.

A comparison between the summer events of all datasets showed considerably larger summer precipitation in RACMO and SEAS5 compared to passive and corresponding WG for all basins, with the Vecht in particular. Comparing summer and winter showed increased summer dominance regarding calendar year maxima. This indicates that the

assumption of extreme precipitation events mainly occurring during winter may have to be reconsidered.

Analysis of the temporal and spatial spread of summer precipitation extremes in the RACMO dataset shows the ability of RACMO to describe intensive summer events. Contrary to RACMO and SEAS5, the WG is not capable of simulating the extreme daily peaks that characterise these events. RACMO and SEAS5 therefore provide a better description of short, intensive precipitation events and allow for spatial analysis of precipitation extremes.

Bibliography

Jules J. Beersma and T. Adri Buishand. Drought in the Netherlands – Regional frequency analysis versus time series simulation. *Journal of Hydrology*, 347(3):332–346, 2007. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2007.09.042. URL https://www.sciencedirect.com/science/article/pii/S0022169407005082.

Laurène Bouaziz, Markus Hrachowitz, Jaap Schellekens, Albrecht Werts, and Hubert Savenije. What controls the very quick runoff response in the Meuse basin? In *EGU General Assembly Conference Abstracts*, page 14266, 2017.

T.A. Buishand. Estimation of a large quantile of the distribution of multi-day seasonal maximum rainfall: the value of stochastic simulation of long-duration sequences. *Climate Research*, 34:185,194, 2007. doi: https://doi.org/10.3354/cr00701. URL https://www.int-res.com/abstracts/cr/v34/n3/p185-194/.

TA Buishand and R Leander. Rainfall generator for the Meuse basin. Technical report, KNMI, 2011.

Stuart Coles, Joanna Bawa, Lesley Trenner, and Pat Dorazio. *An introduction to statistical modeling of extreme values*, volume 208. Springer, 2001.

Richard C. Cornes, Gerard van der Schrier, Else J. M. van den Besselaar, and Philip D. Jones. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres*, 123(17):9391–9409, 2018a. doi: https://doi.org/10.1029/2017JD028200. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017JD028200.

Richard C. Cornes, Gerard van der Schrier, Else J. M. van den Besselaar, and Philip D. Jones. An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres*, 123(17):9391–9409, 2018b. doi: https://doi.org/10.1029/2017JD028200. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017JD028200.

Yigzaw Desta, Haddush Goitom, and Gebremeskel Aregay. Investigation of runoff response to land use/land cover change on the case of Aynalem catchment, North of Ethiopia. *Journal of African Earth Sciences*, 153:130–143, 2019. ISSN 1464-343X. doi: https://doi.org/10.1016/j.jafrearsci.2019.02.025. URL https://www.sciencedirect.com/science/article/pii/S1464343X19300585.

Josep Fortesa, Jérôme Latron, Julián García-Comendador, Miquel Tomàs-Burguera, Aleix Calsamiglia, Joan Estrany, et al. Driving factors of non-linearity rainfall-runoff

relationships at different time scales in small Mediterranean-climate catchments. In *EGU General Assembly Conference Abstracts*, page 3127, 2020.

M Hegnauer, JJ Beersma, HFP Van den Boogaard, TA Buishand, and RH Passchier. Generator of rainfall and discharge extremes (GRADE) for the Rhine and Meuse basins. *Final report of GRADE*, 2:1209424–004, 2014.

Mark Hegnauer. Memo: Results of the analysis done on the different options for a weather generator for the Meuse. Technical report, Deltares, March 2022.

S. J. Johnson, T. N. Stockdale, L. Ferranti, M. A. Balmaseda, F. Molteni, L. Magnusson, S. Tietsche, D. Decremer, A. Weisheimer, G. Balsamo, S. P. E. Keeley, K. Mogensen, H. Zuo, and B. M. Monge-Sanz. SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12(3):1087–1117, 2019. doi: 10.5194/gmd-12-1087-2019. URL https://gmd.copernicus.org/articles/12/1087/2019/.

Frank Kreienkamp, Sjoukje Y Philip, Jordis S Tradowsky, Sarah F Kew, Philip Lorenz, Julie Arrighi, Alexandre Belleflamme, Thomas Bettmann, Steven Caluwaerts, Steven C Chan, et al. Rapid attribution of heavy rainfall events leading to the severe flooding in Western Europe during July 2021. Technical report, Universiteit Gent, 2021.

Benjamin B. Mirus and Keith Loague. How runoff begins (and ends): Characterizing hydrologic response at the catchment scale. *Water Resources Research*, 49(5): 2987–3006, 2013. doi: https://doi.org/10.1002/wrcr.20218. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/wrcr.20218.

Monika Rauthe, Heiko Steiner, Ulf Riediger, Alex Mazurkiewicz, Annegret Gratzki, et al. A Central European precipitation climatology–Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, 22(3): 235–256, 2013.

Maurice J Schmeits, Erwin LA Wolters, Jules J Beersma, and T Adri Buishand. Rainfall generator for the Rhine basin: Description of simulations using gridded precipitation datasets and uncertainty analysis, 2014.

E. van Meijgaard, L.H. van Ulft, W.J. van de Berg, F. C. Bosveld, B.J.J.M. van den Hurk, G. Lenderink, and A.P. Siebesma. The knmi regional atmospheric climate model racmo version 2.1. Technical report, KNMI, 2008. URL https://cdn.knmi.nl/knmi/pdf/bibliotheek/knmipubTR/TR302.pdf.

Appendices

Appendix A

Bias corrections

In this section, the figures of bias corrections that are not presented in the main part of this report will be presented. First, the correction of the precipitation for the Meuse and Vecht basin are presented, followed up by a bias correction for temperature and potential evaporation for all three basins (Rhine, Meuse, Vecht).

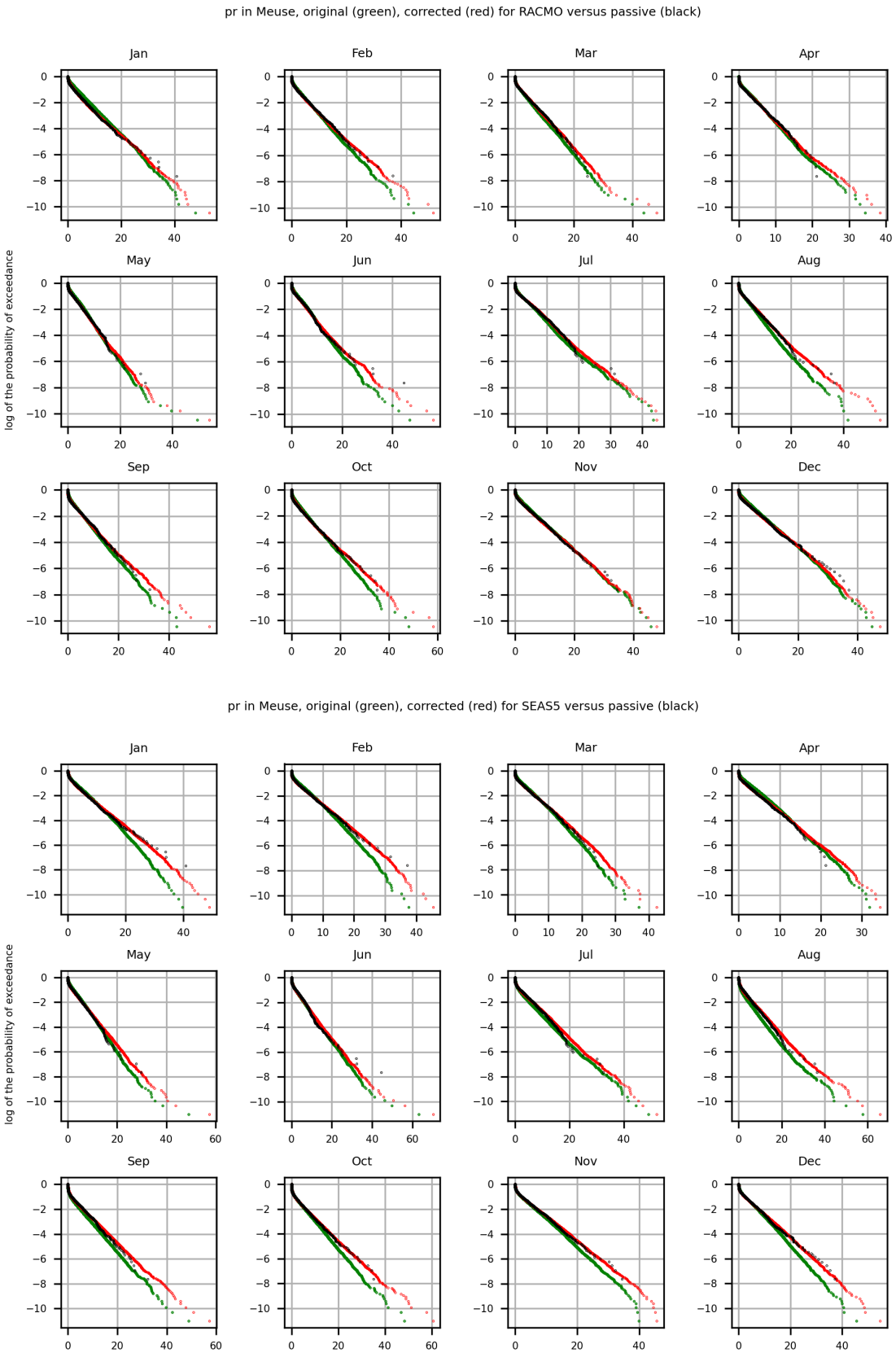

Figure A.1: Correction of precipitation in the Meuse basin.

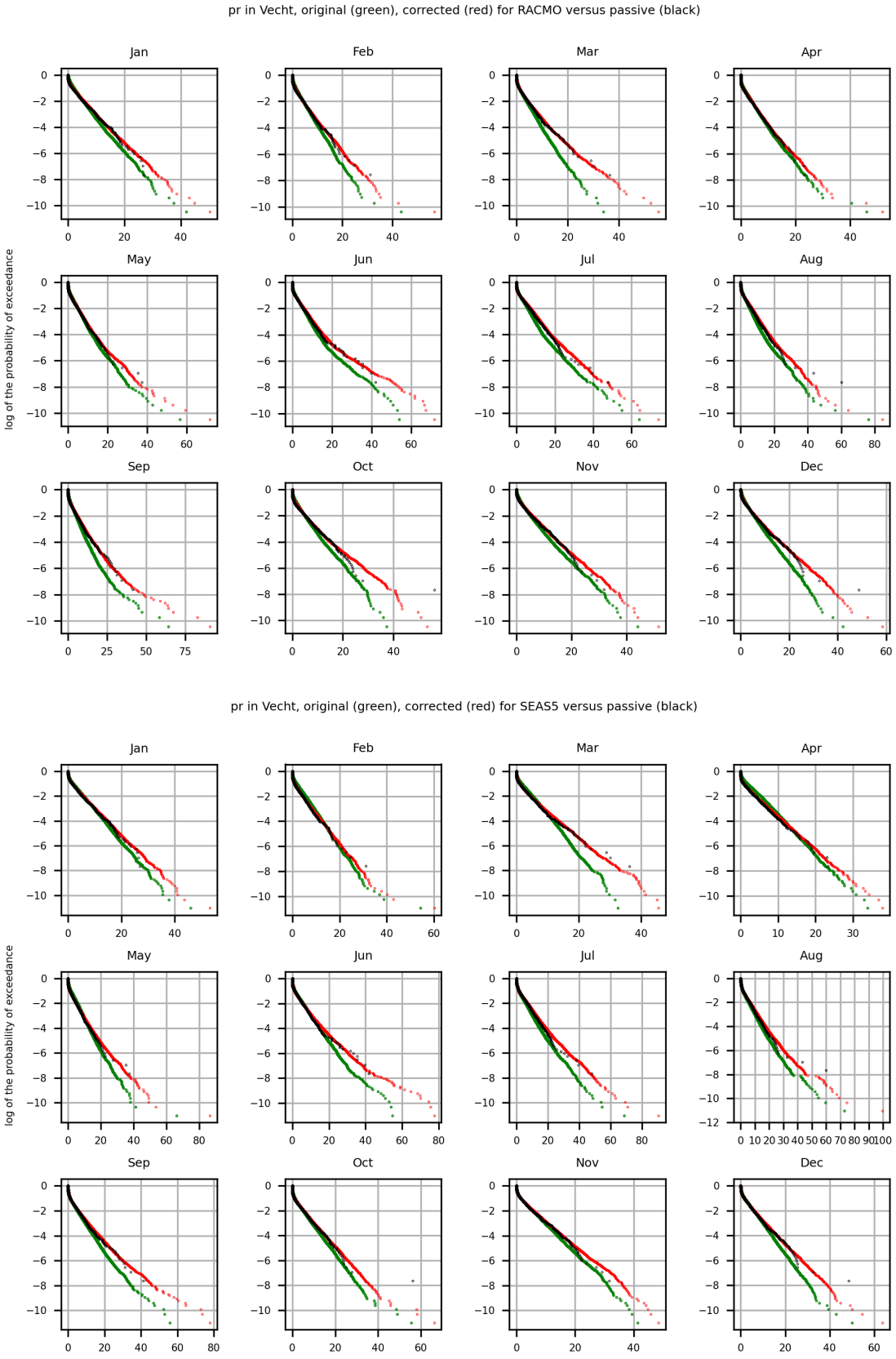

Figure A.2: Correction of precipitation in the Vecht basin.

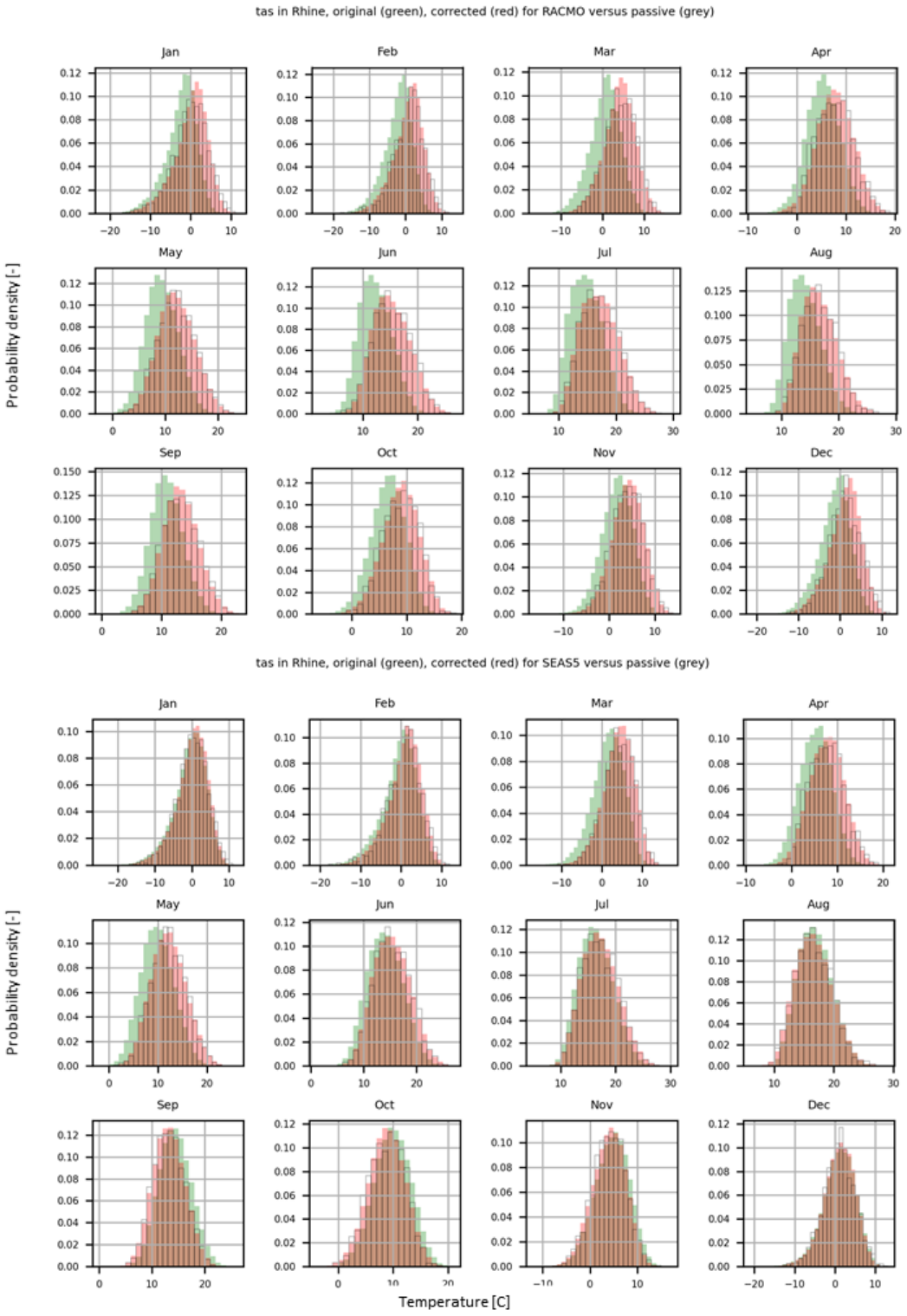

Figure A.3: Correction of temperature in the Rhine basin.

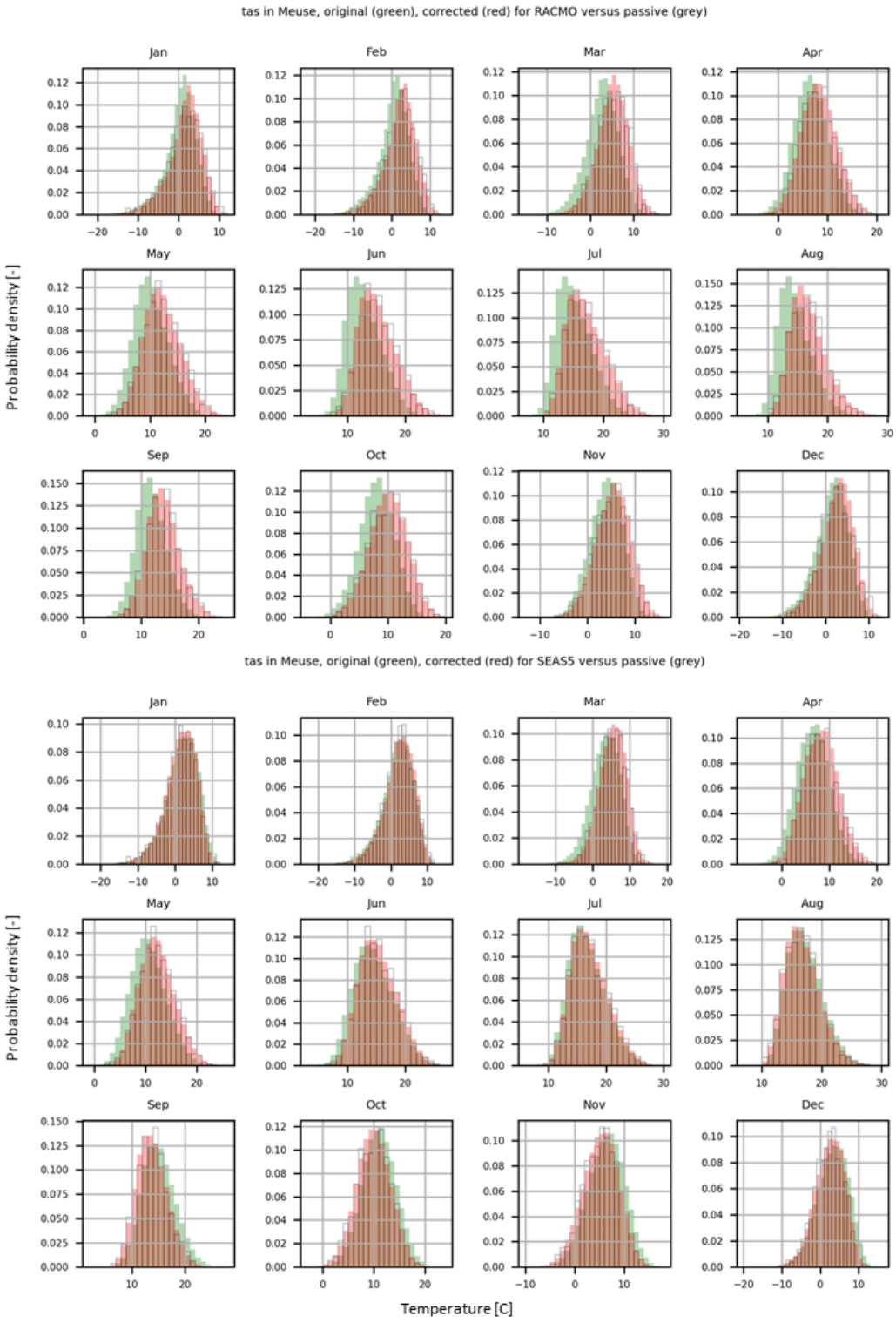

Figure A.4: Correction of temperature in the Meuse basin.

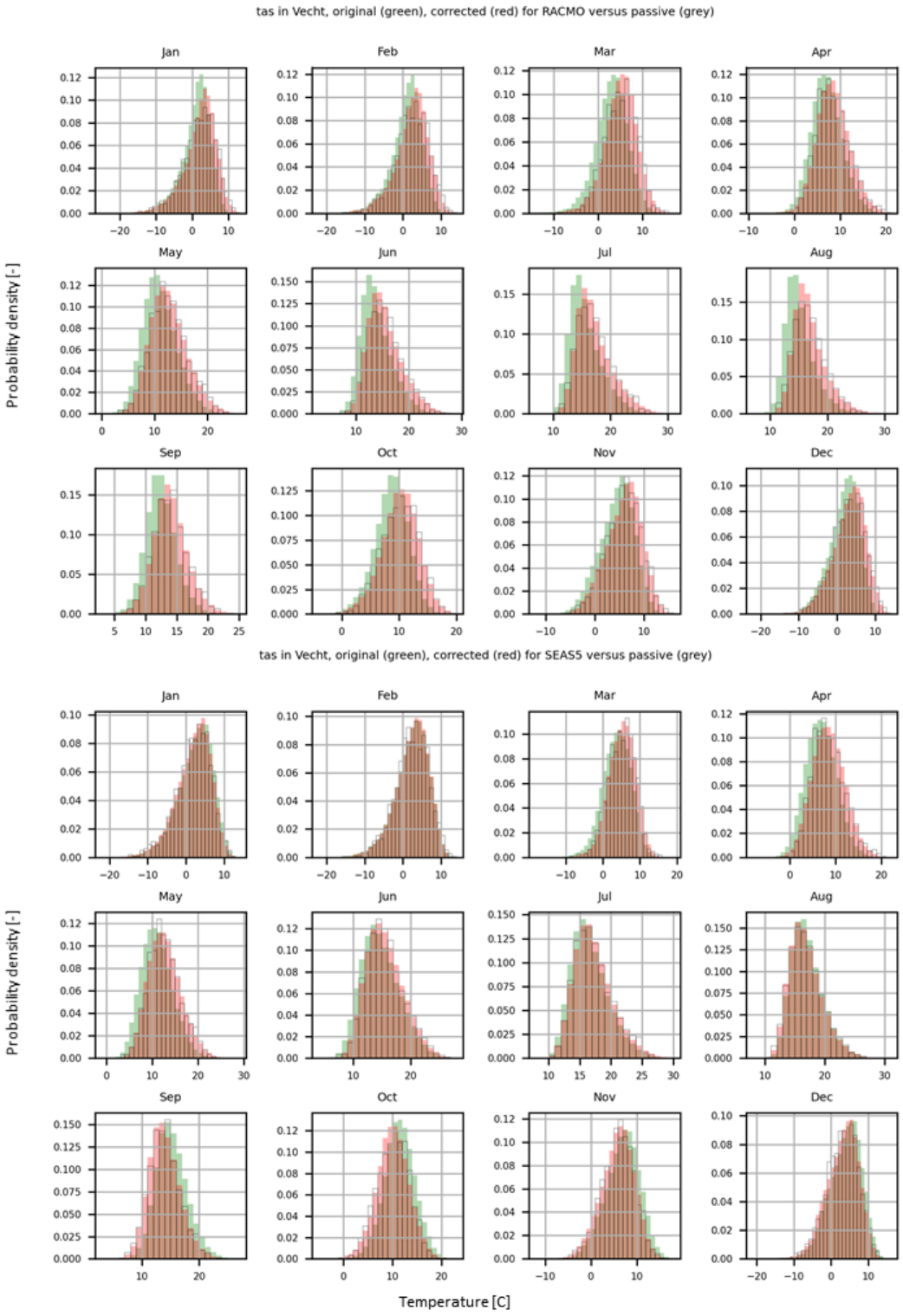

Figure A.5: Correction of temperature in the Vecht basin.

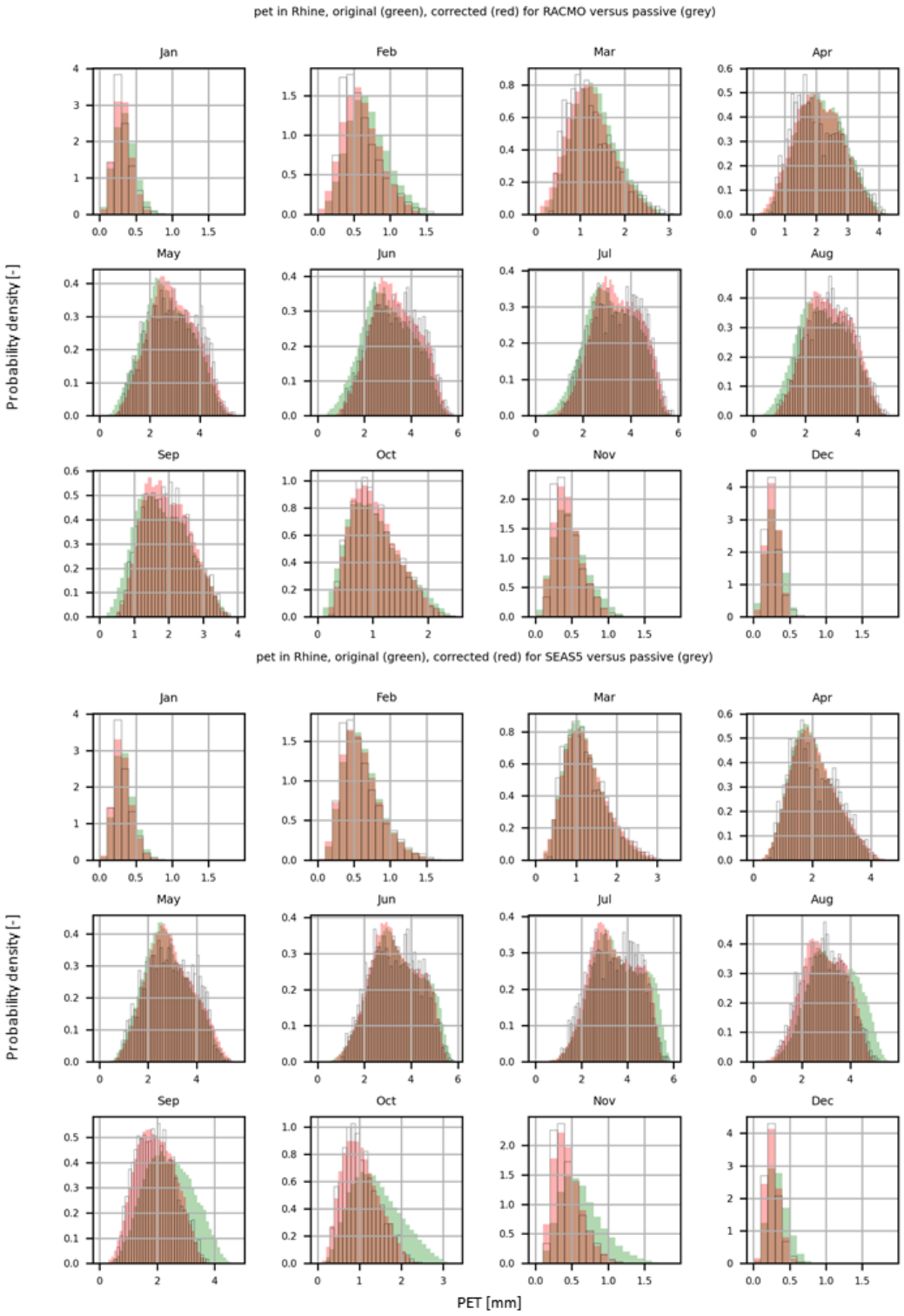

Figure A.6: Correction of potential evaporation in the Rhine basin.

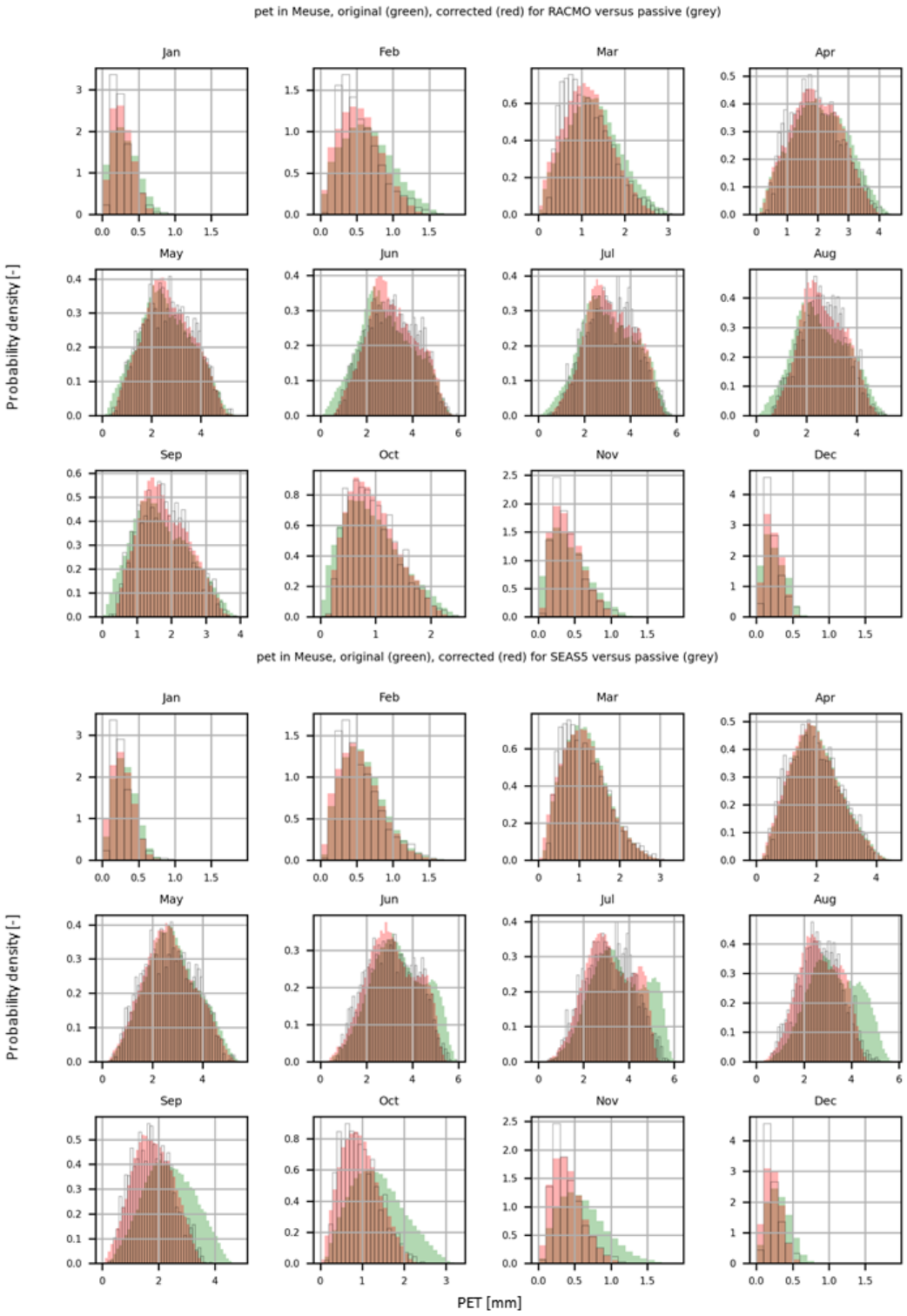

Figure A.7: Correction of potential evaporation in the Meuse basin.

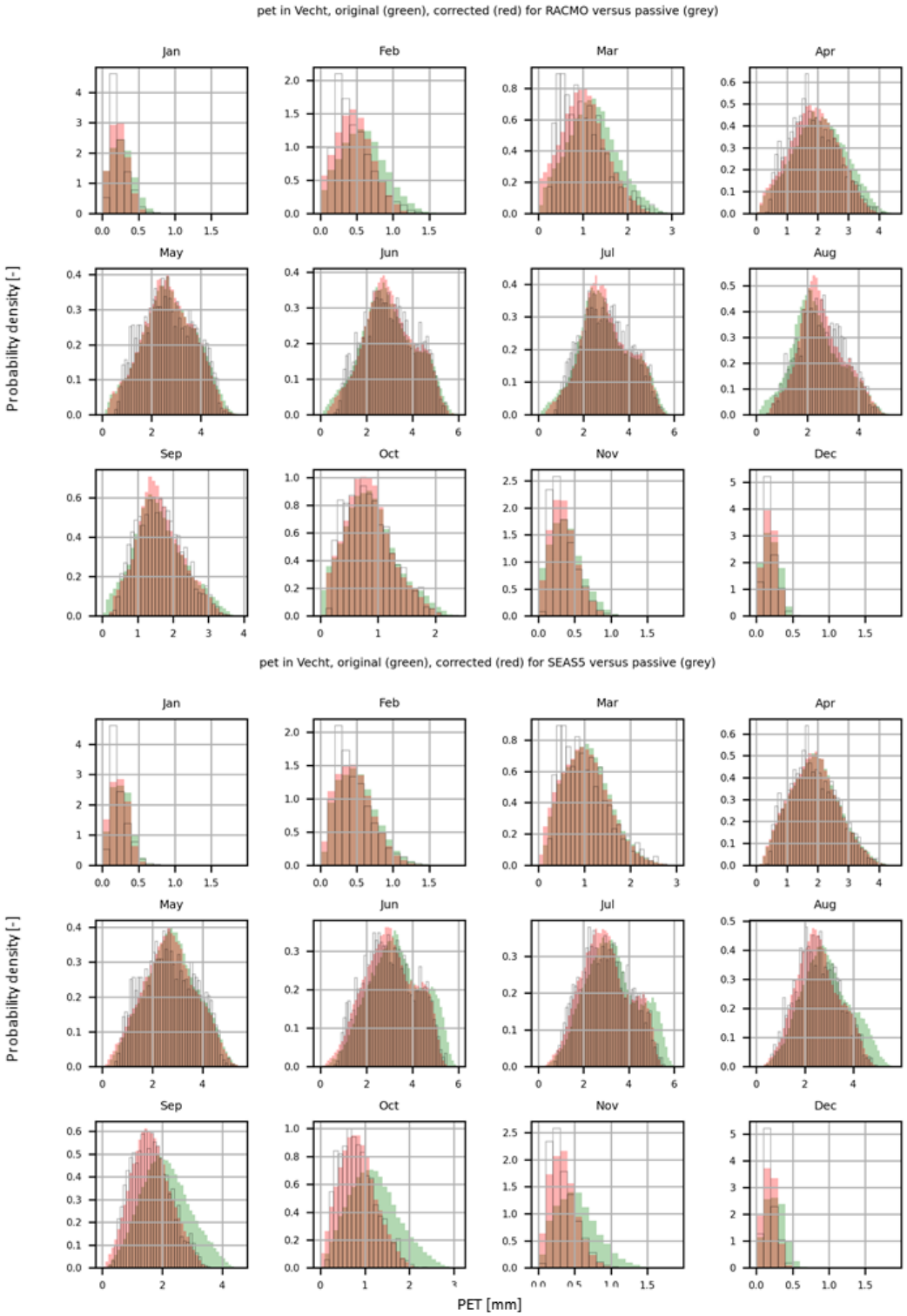

Figure A.8: Correction of potential evaporation in the Vecht basin.

Appendix B

RACMO extremes

This appendix will provide an overview of all annual maximum multi-day precipitation sums of precipitation ranging from 1 to 10 days, in respectively the Rhine, Meuse and Vecht basin. Like in section 3.1, three panels are presented in every figure, which, from left to right, show results based on timeseries of calendar year (CY), summer half year (SHY) and winter half year (WHY). In these figures RACMO (both corrected and uncorrected) and the WG based on the corrected RACMO timeseries are compared to the observation-based passive dataset (and its corresponding timeseries derived from the WG).

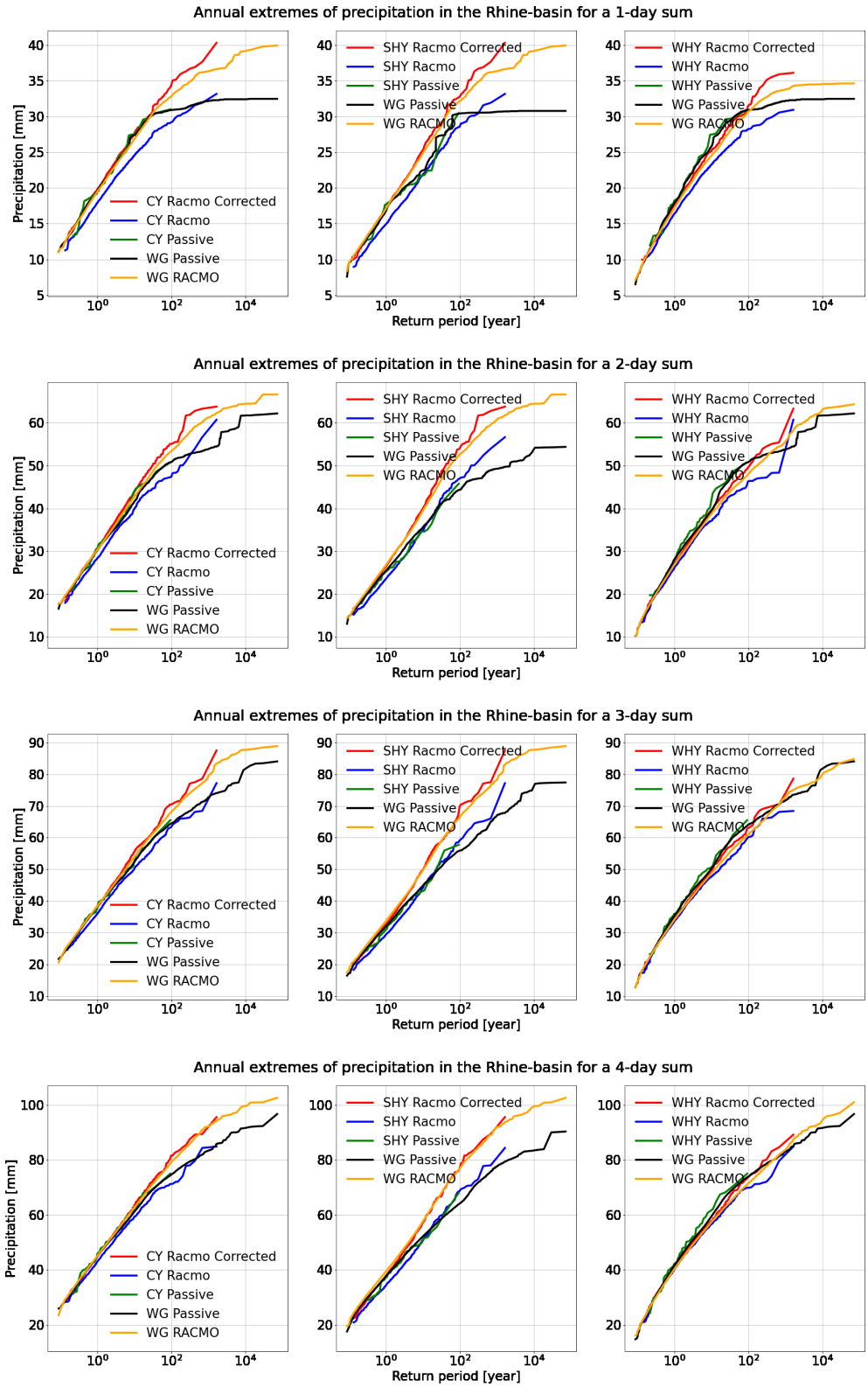

Figure B.1: Rhine 1-4 day precipitation sums.

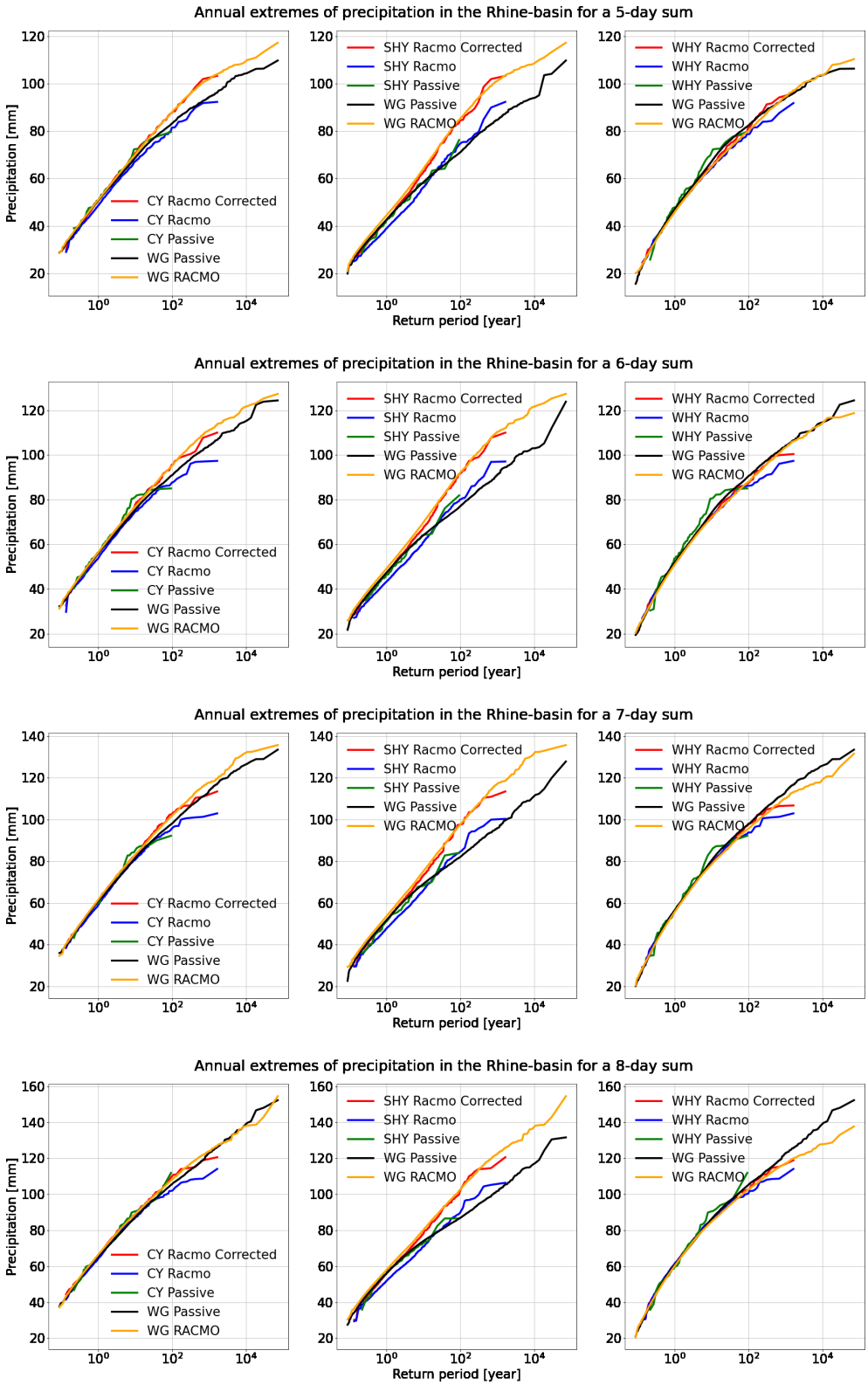

Figure B.2: Rhine 5-8 day precipitation sums.

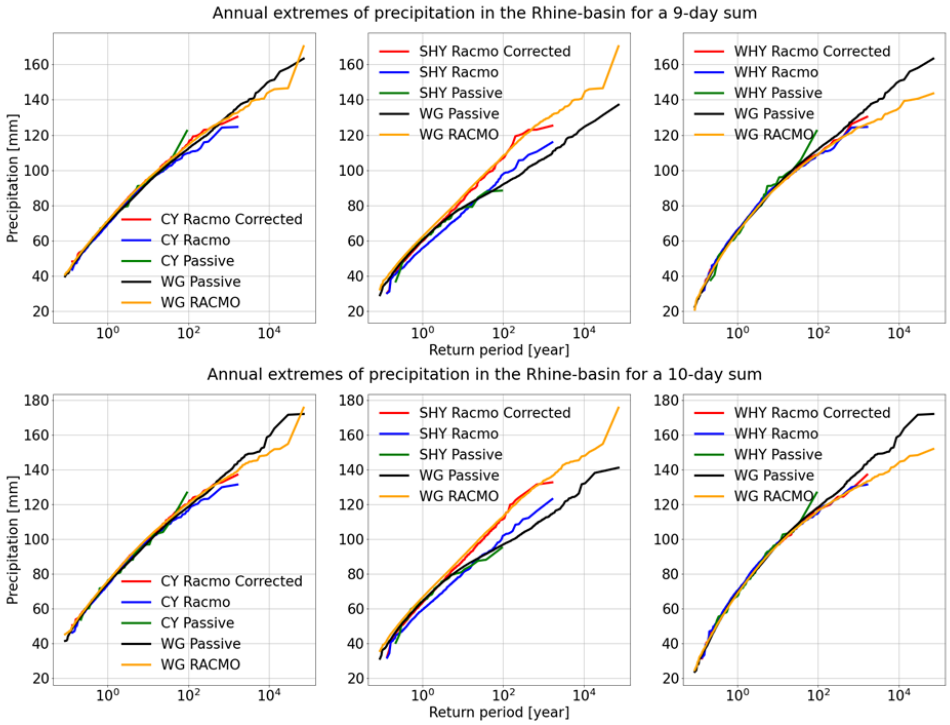

Figure B.3: Rhine 9 and 10 day precipitation sums.

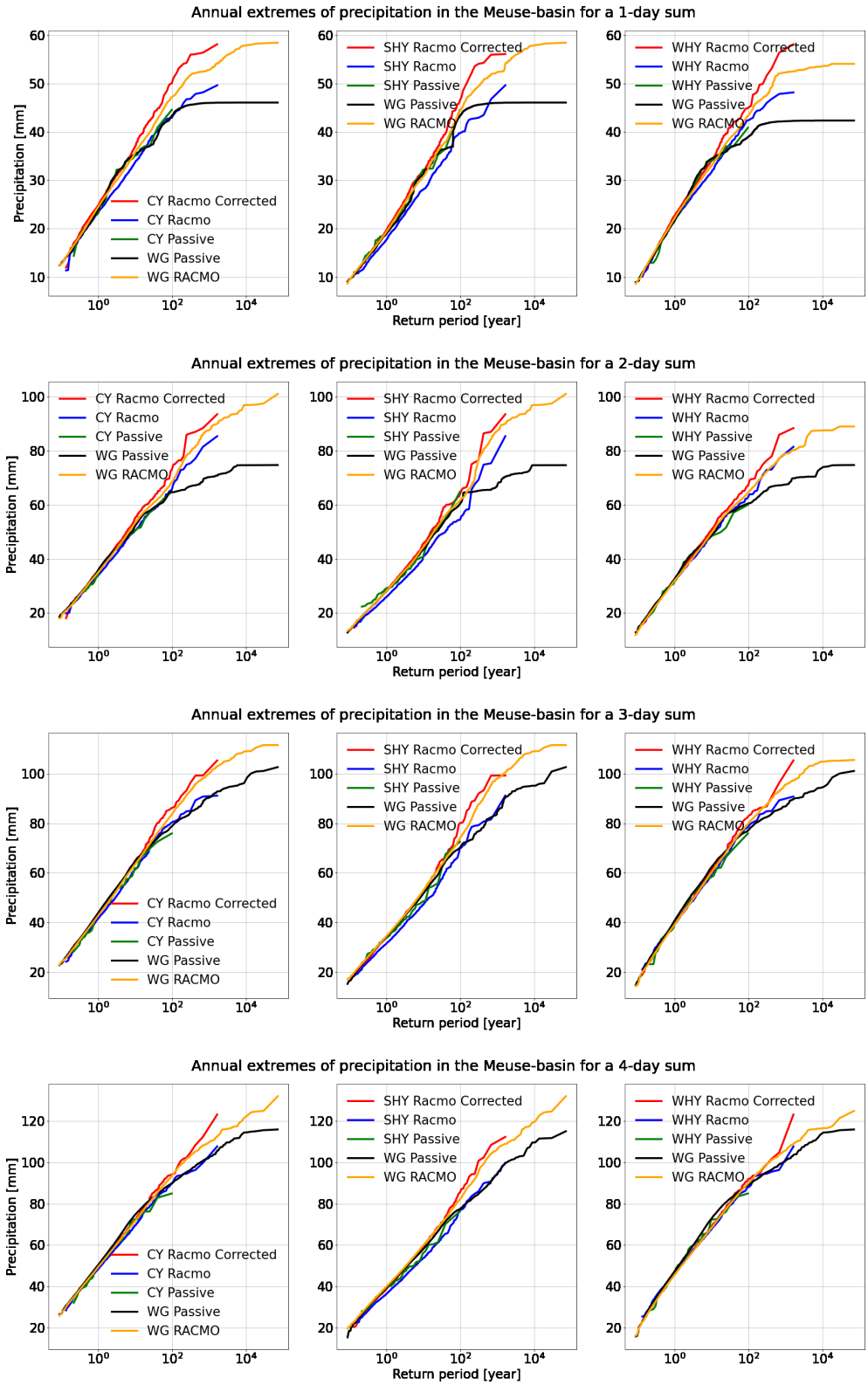

Figure B.4: Meuse 1-4 day precipitation sums.

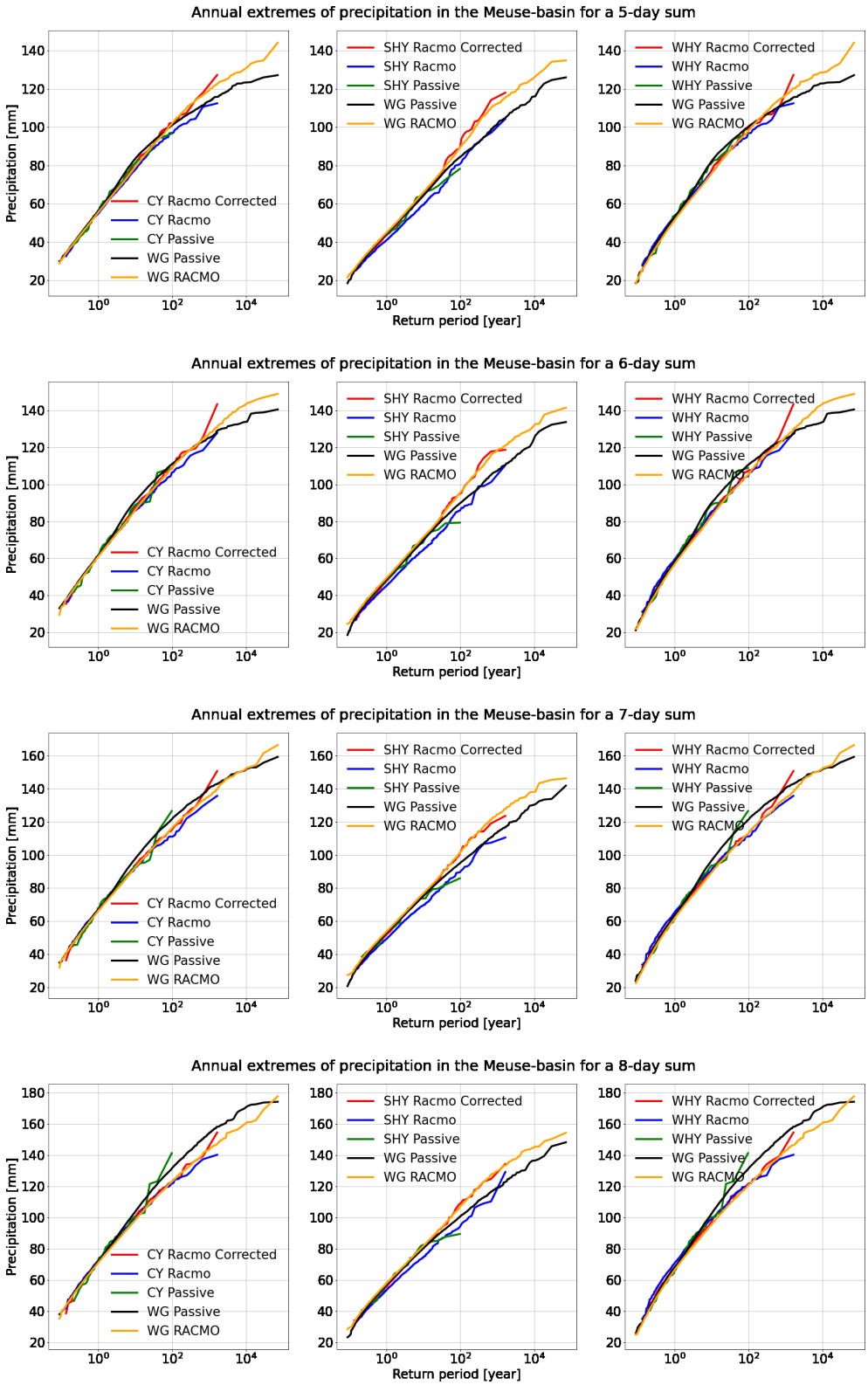

Figure B.5: Meuse 5-8 day precipitation sum

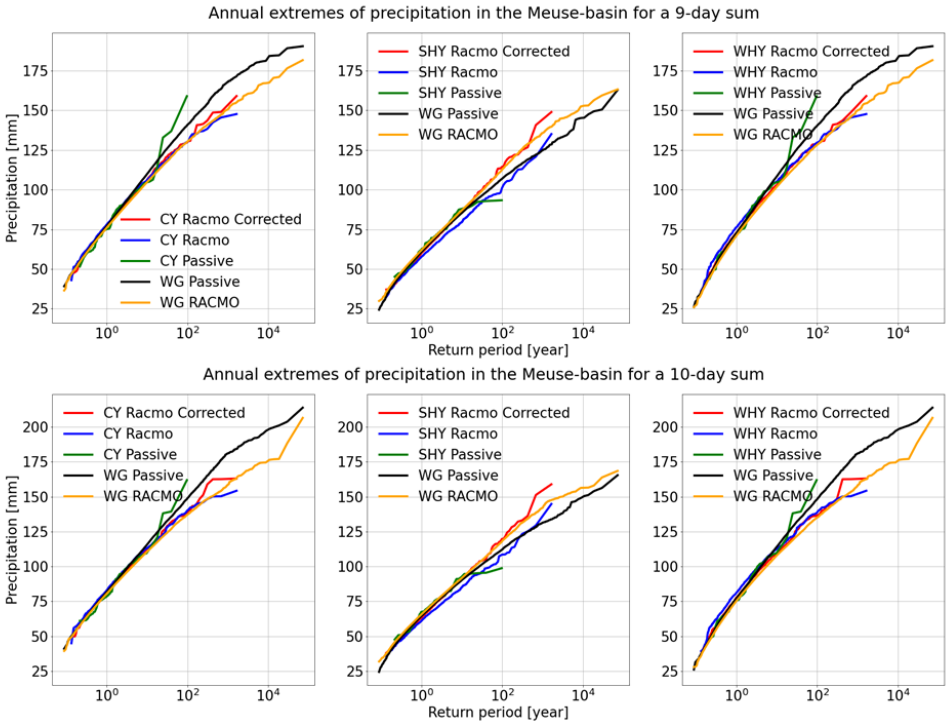

Figure B.6: Meuse 9 and 10 day precipitation sum

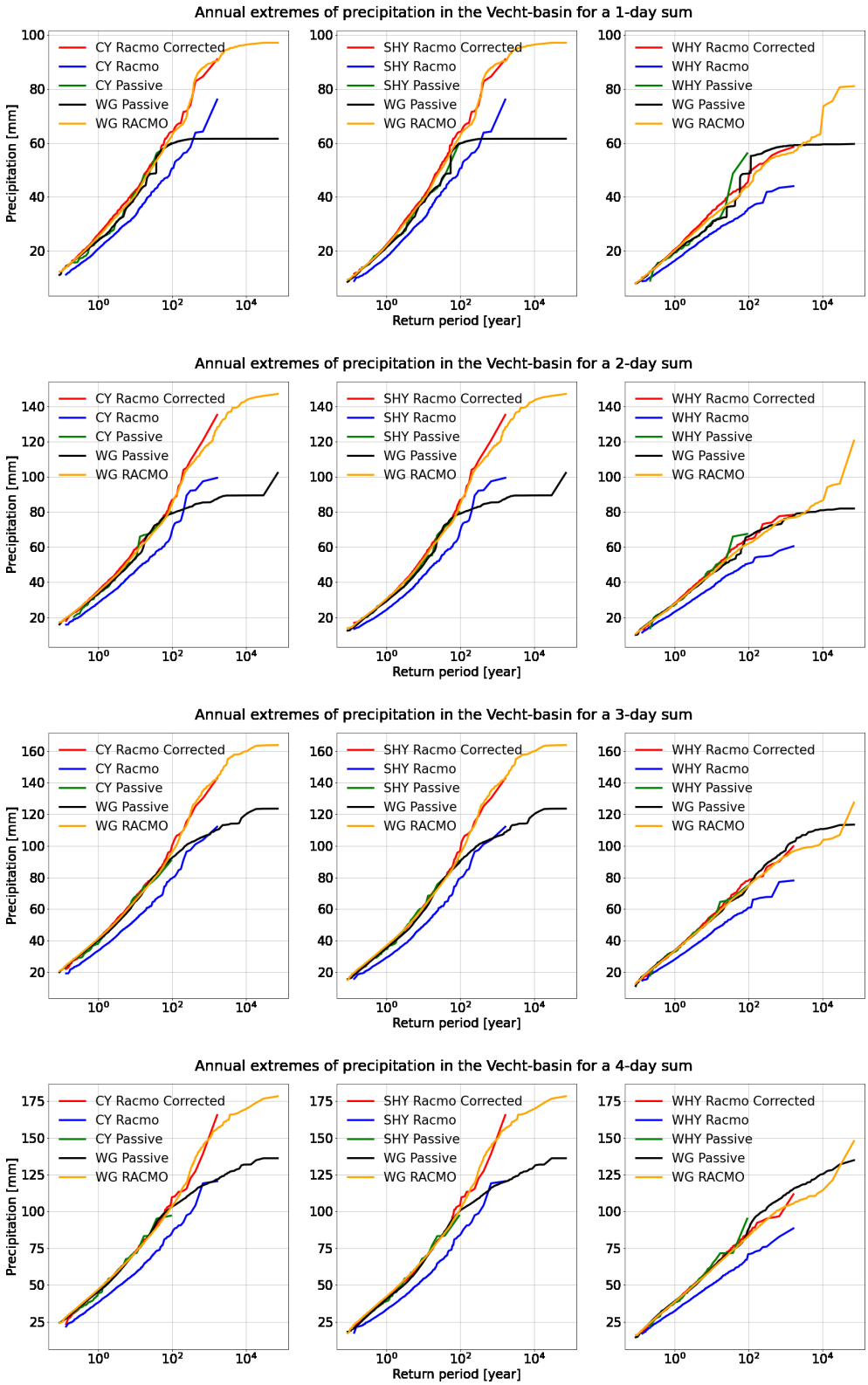

Figure B.7: Vecht 1-4 day precipitation sums.

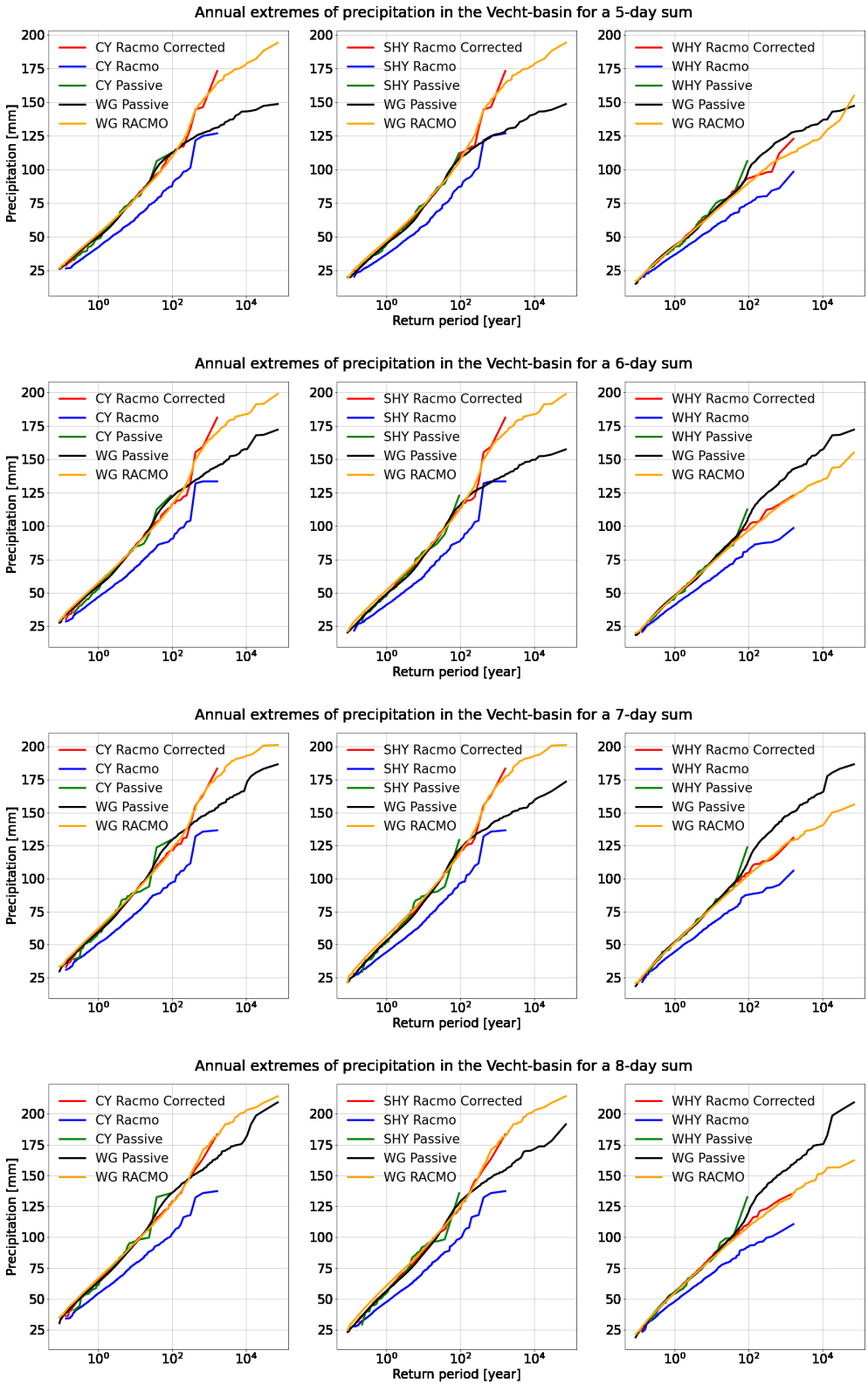

Figure B.8: Vecht 5-8 day precipitation sums.

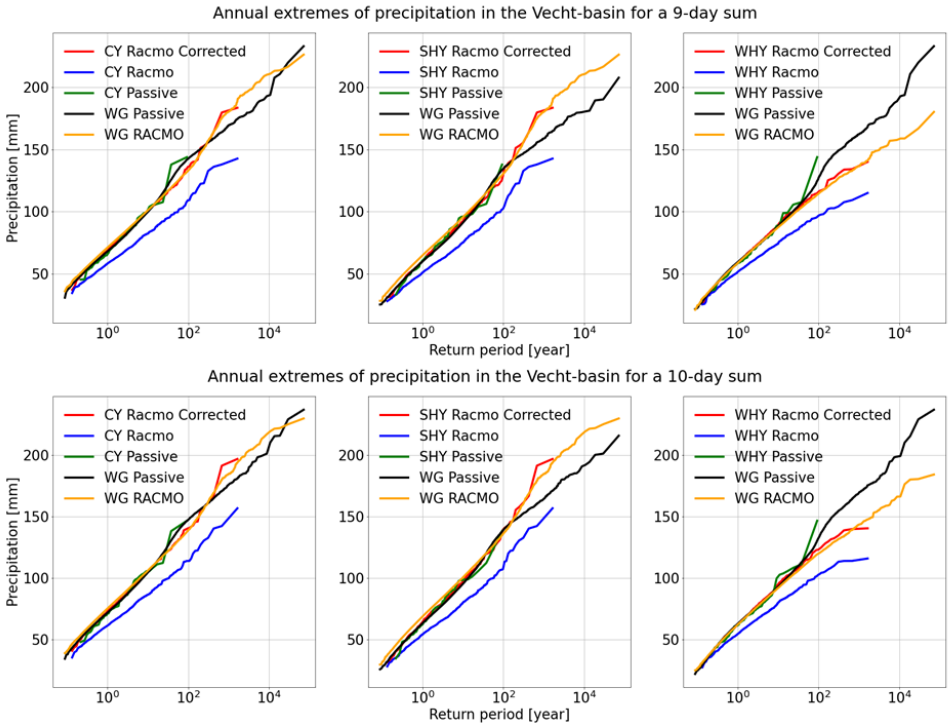

Figure B.9: Vecht 9 and 10 day precipitation sums.

Appendix C

SEAS5 extremes

This appendix will provide an overview of all annual maximum multi-day precipitation sums of precipitation ranging from 1 to 10 days, in respectively the Rhine, Meuse and Vecht basin. Like in section 3.1, three panels are presented in every figure, which, from left to right, show results based on timeseries of calendar year (CY), summer half year (SHY) and winter half year (WHY). In these figures SEAS5 (both corrected and uncorrected) and the WG based on the corrected SEAS5 timeseries are compared to the observation-based passive dataset (and its corresponding timeseries derived from the WG).

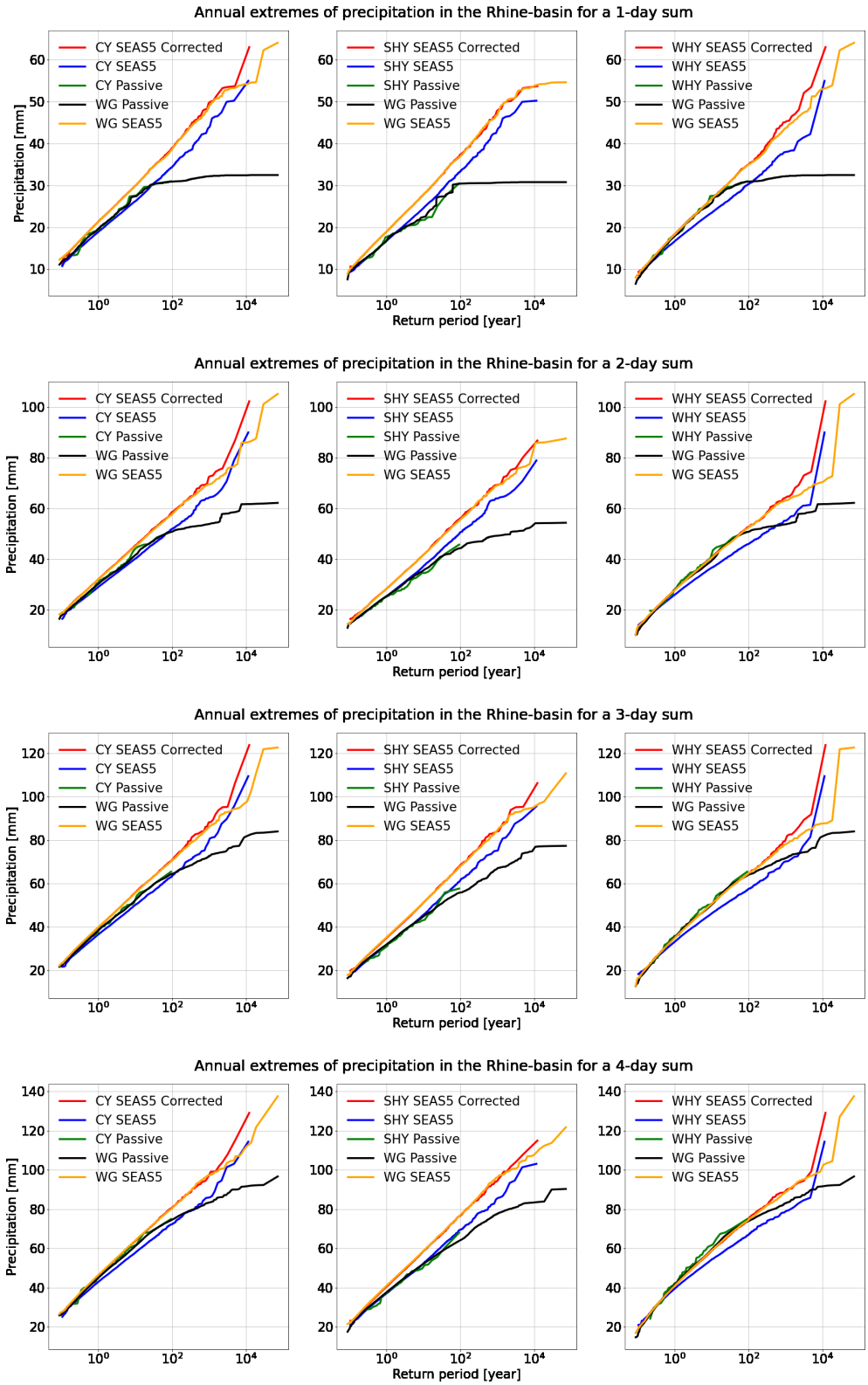

Figure C.1: Rhine 1-4 day precipitation sums.

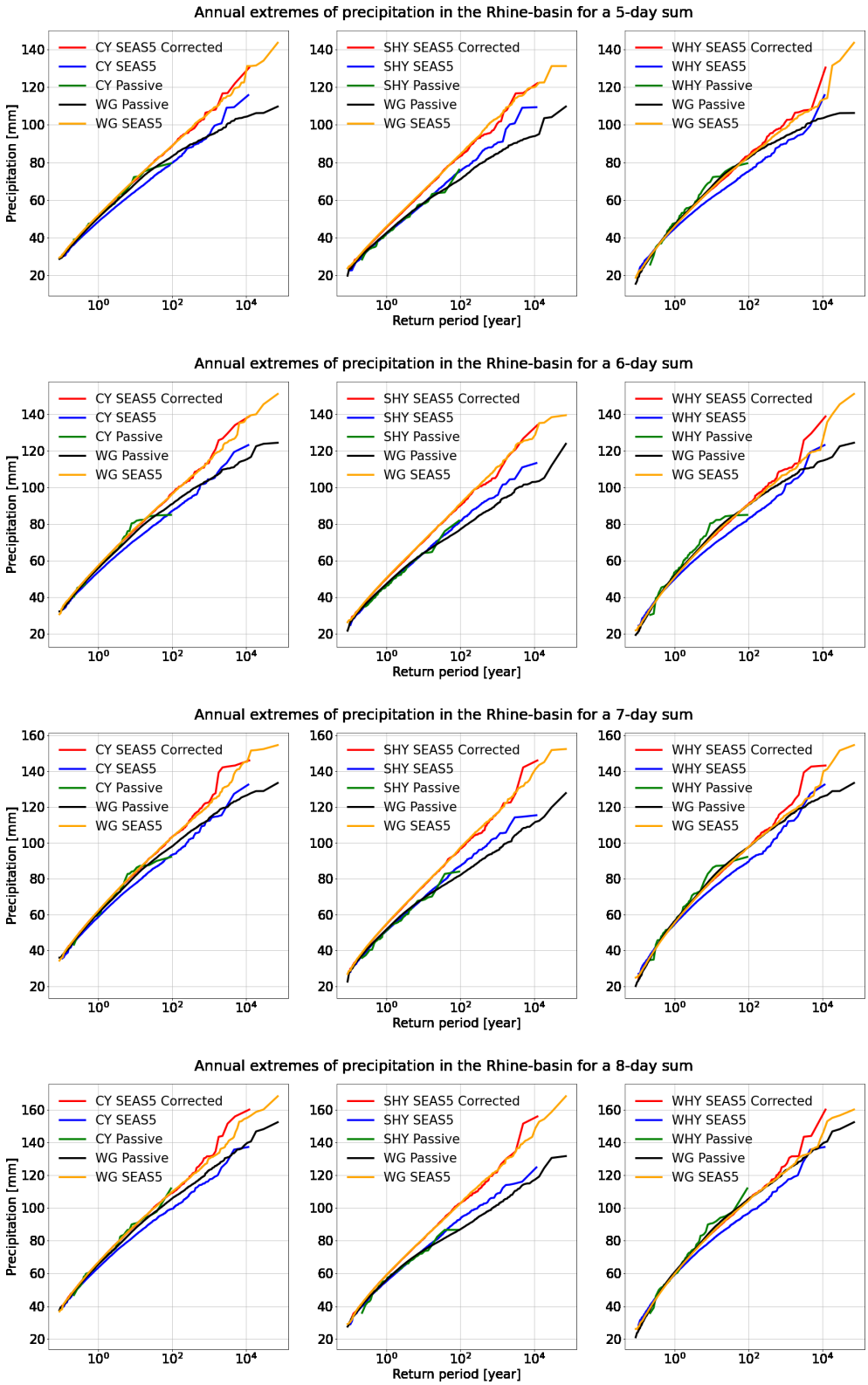

Figure C.2: Rhine 5-8 day precipitation sums.

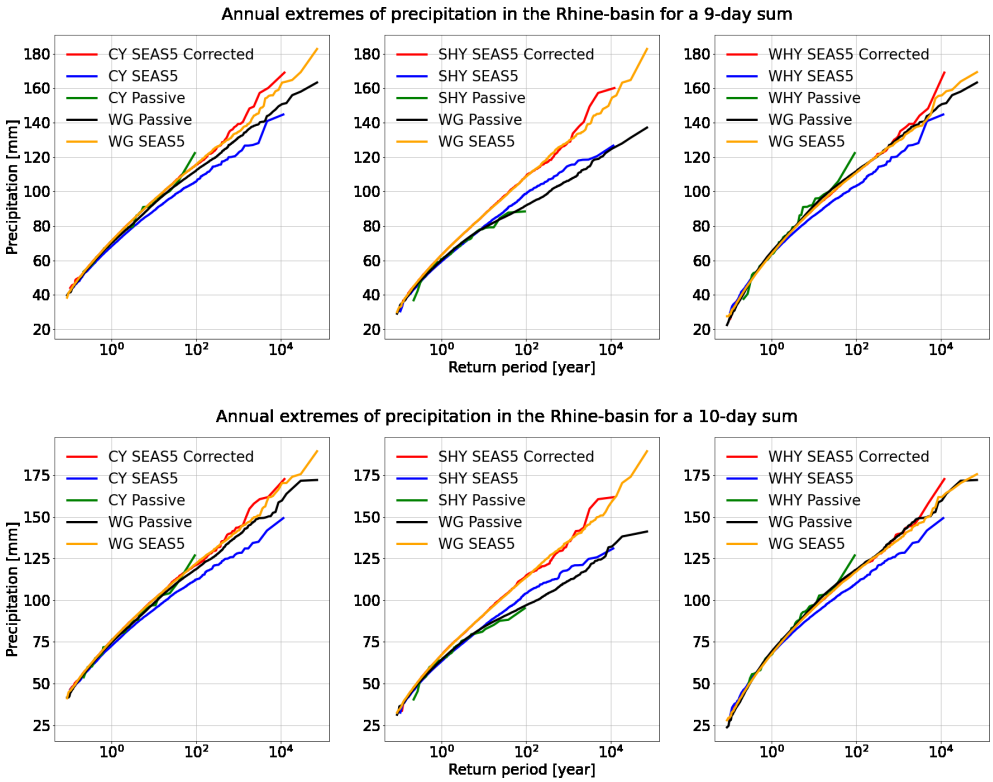

Figure C.3: Rhine 9 and 10 day precipitation sums.

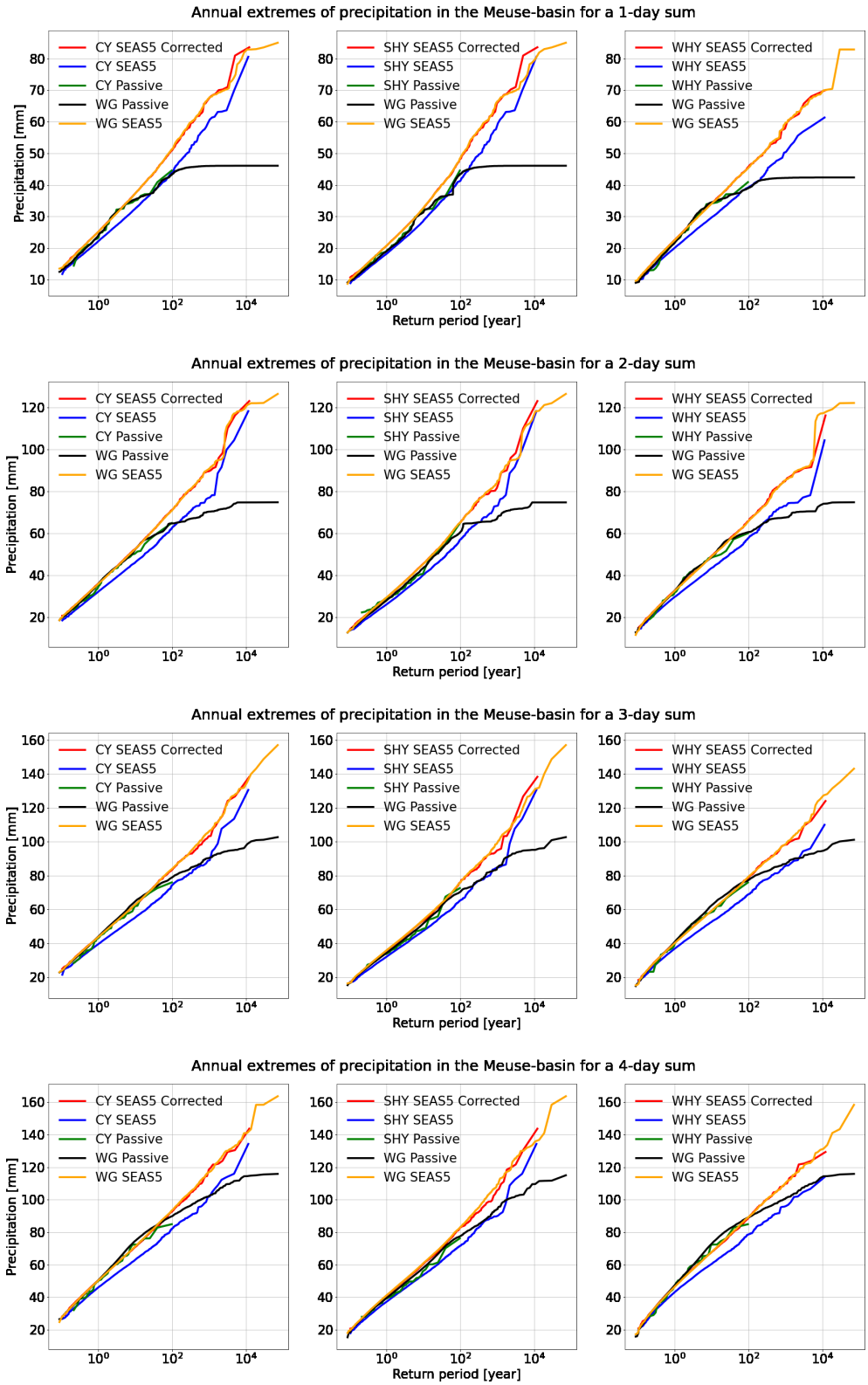

Figure C.4: Meuse 1-4 day precipitation sums.

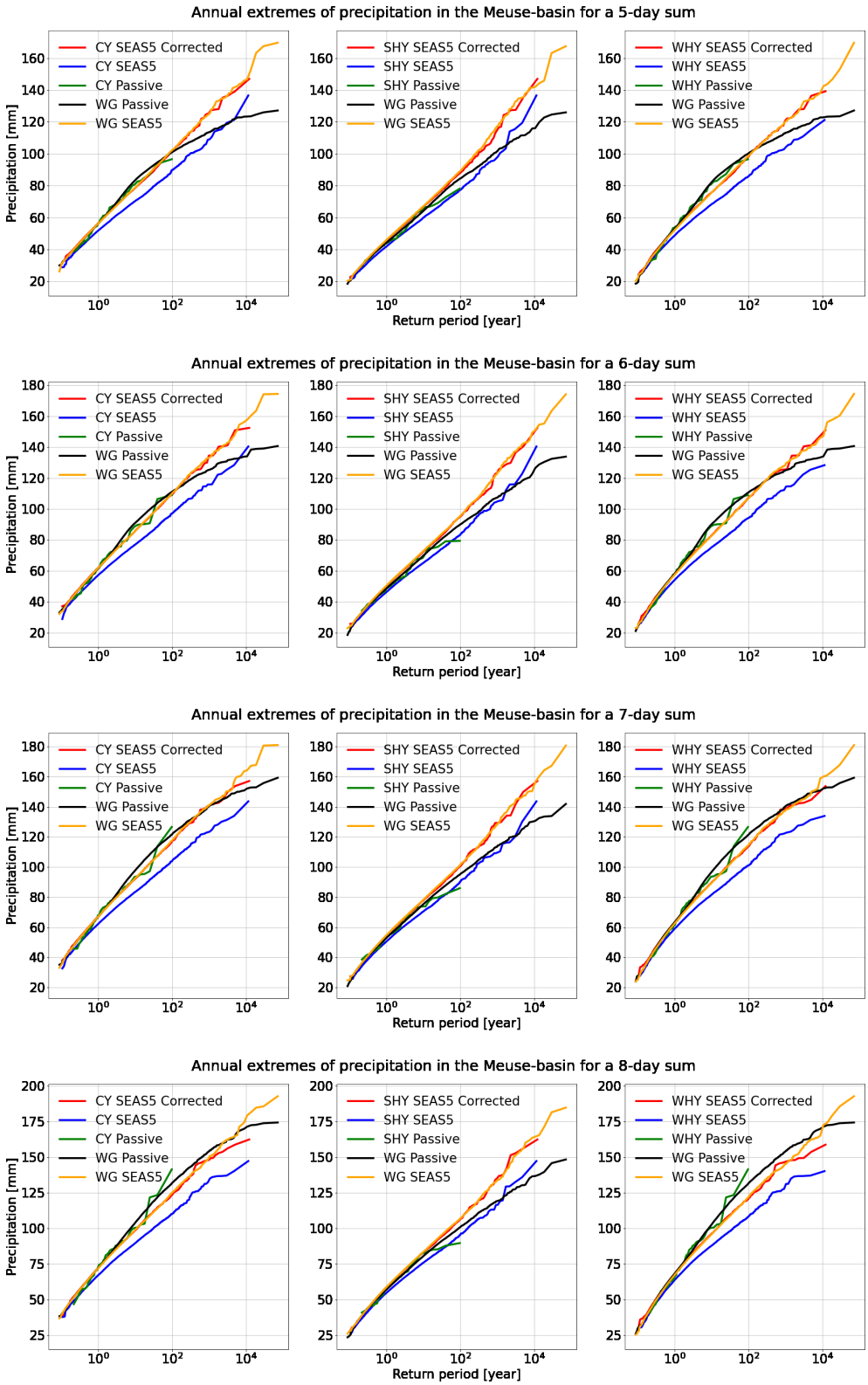

Figure C.5: Meuse 5-8 day precipitation sums.

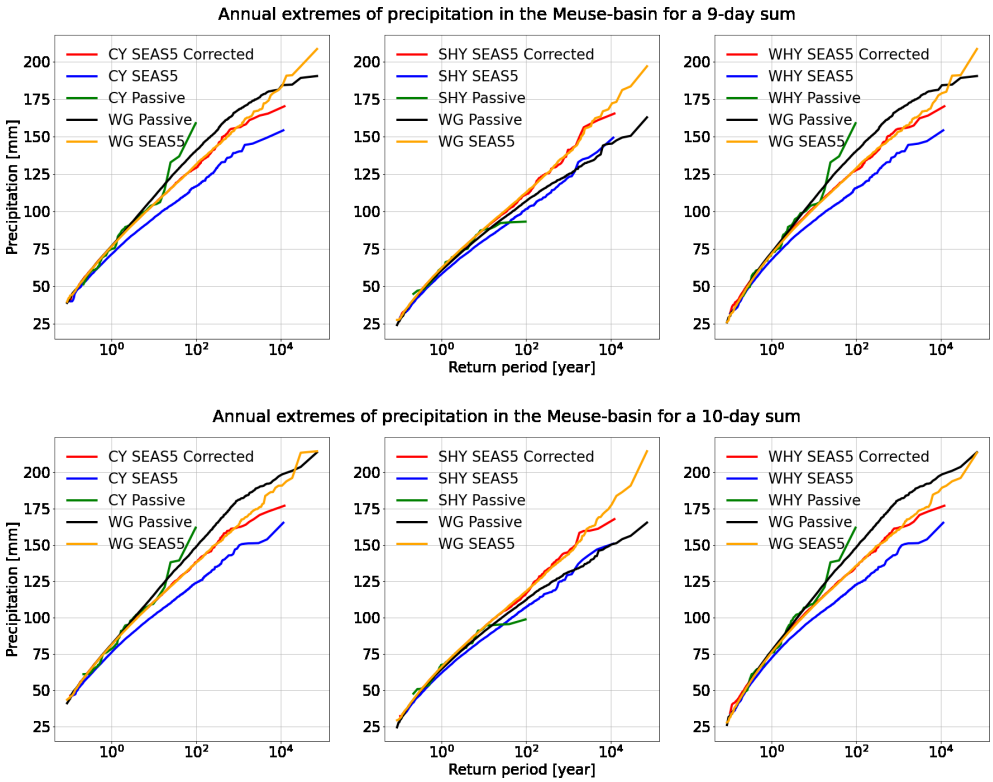

Figure C.6: Meuse 9 and 10 day precipitation sums.

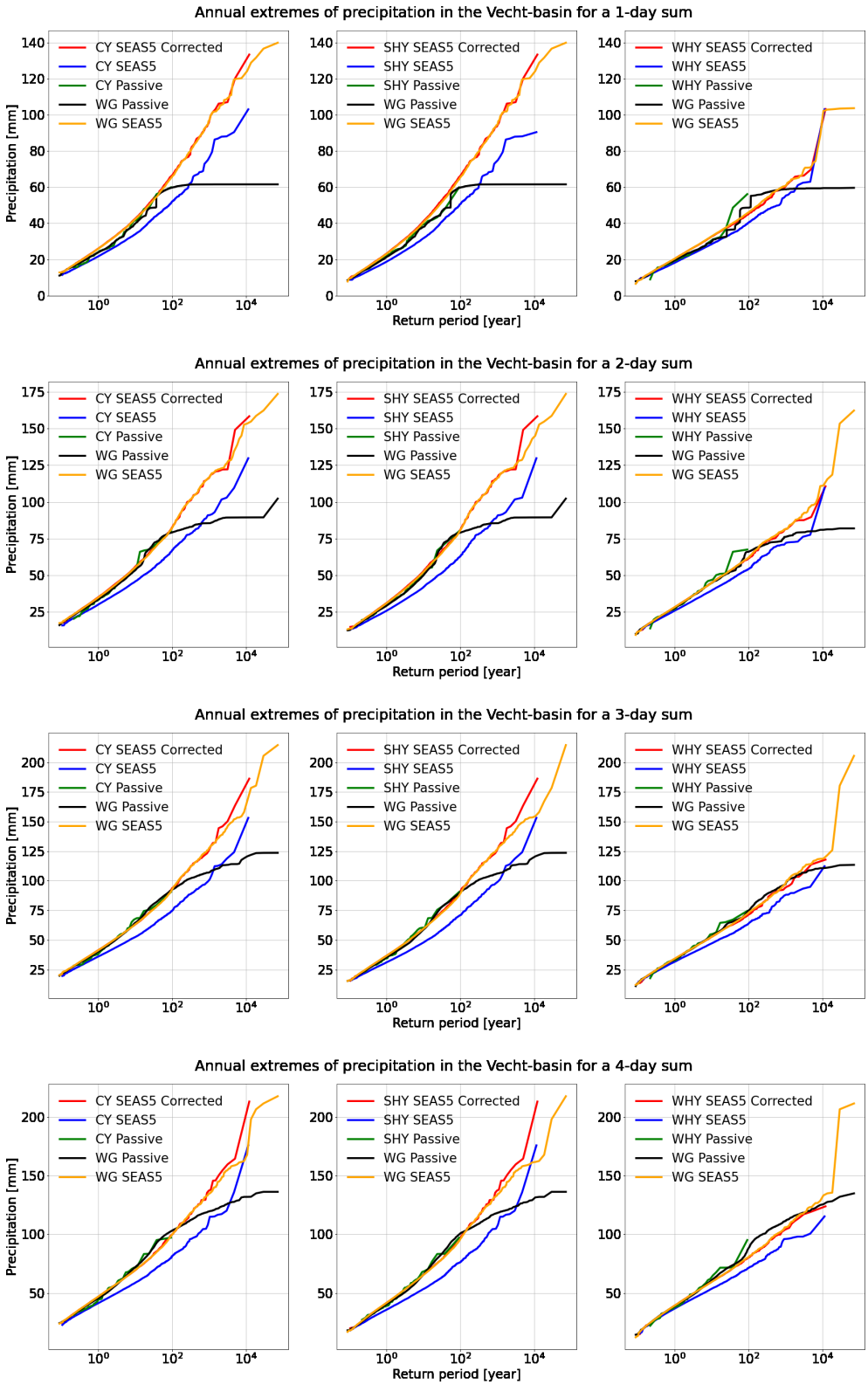

Figure C.7: Vecht 1-4 day precipitation sums.

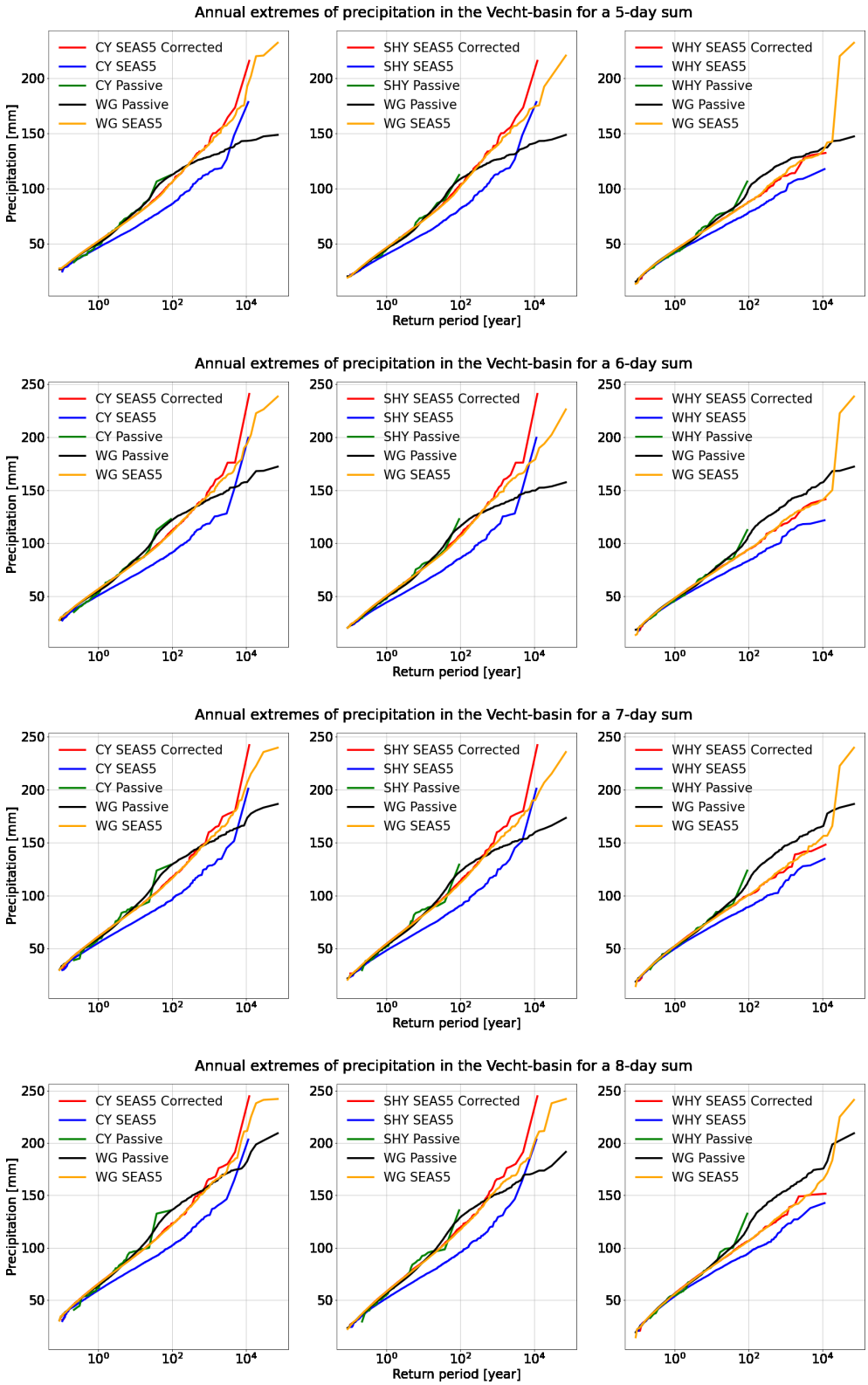

Figure C.8: Vecht 5-8 day precipitation sums.

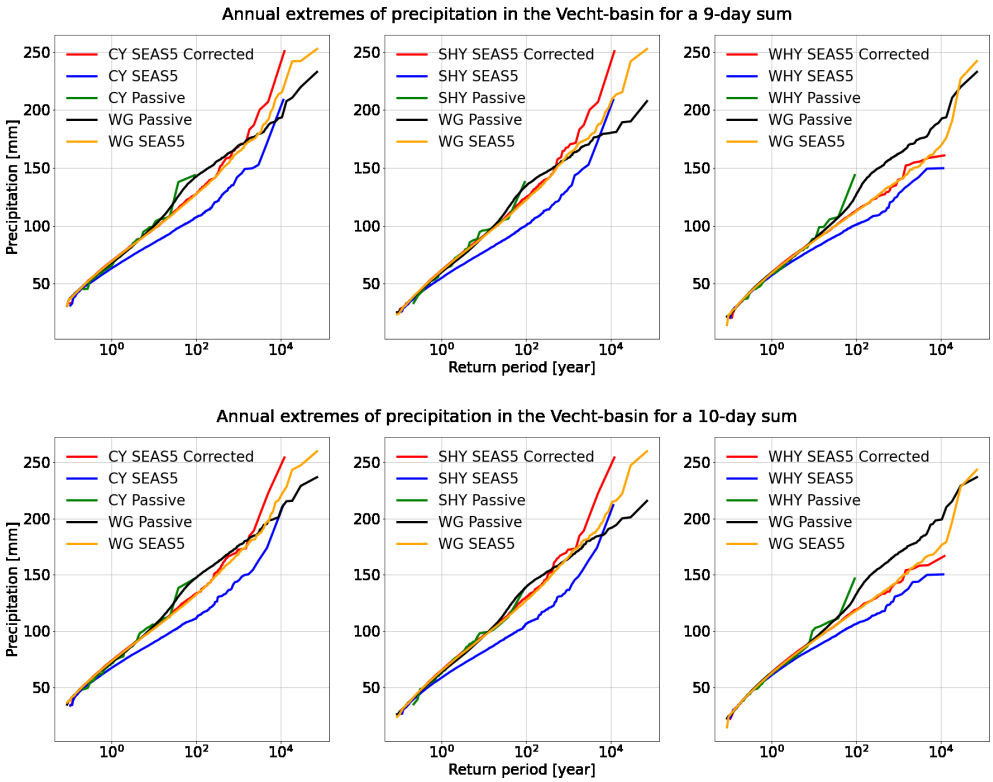

Figure C.9: Vecht 9 and 10 day precipitation sums.

Royal Netherlands Meteorological Institute

PO Box 201 | NL-3730 AE De Bilt
Netherlands | www.knmi.nl