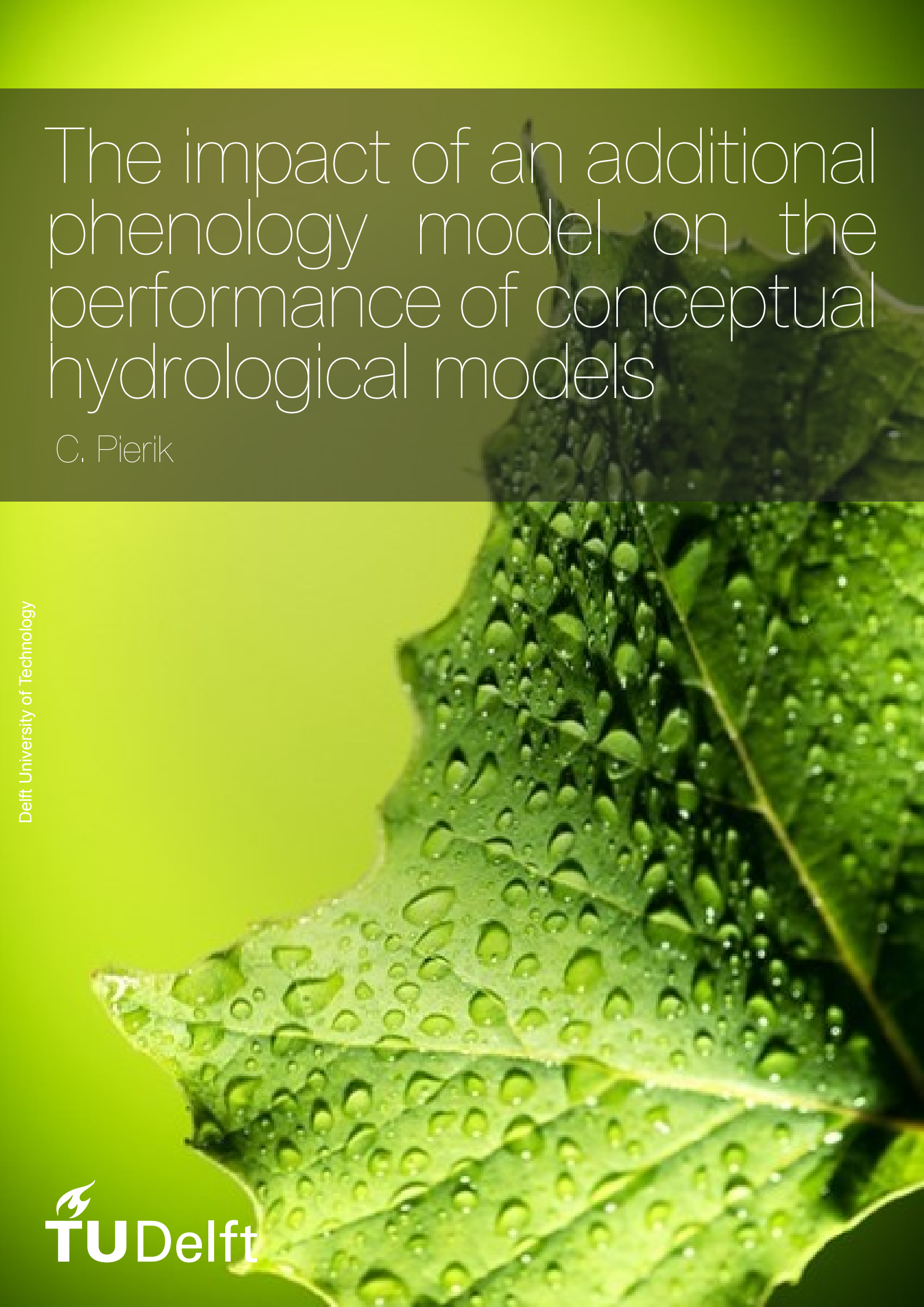
The impact of an additional phenology model on the performance of conceptual hydrological models

C. Pierik



Delft University of Technology

TU Delft

The impact of an additional phenology model on the performance of conceptual hydrological models

by

C. Pierik

Student Name	Student Number
Casper Pierik	4454960

Assessment commitee:	Markus Hrachowitz
	Miriam Coenders
	Riccardo Toarmina
Supervisor:	Markus Hrachowitz
Faculty:	Faculty of Civil engineering and Geo-sciences, Delft

Cover: Transpiring leaf. Retrieved from Stein Tree Service (Service 2020)

TU Delft

Abstract

This thesis investigates the impact of a phenology model on conceptual hydrologic model. In conventional conceptual hydrologic models the evapotranspiration is partitioned into evaporation and transpiration by a combination of the potential evaporation and the availability of water. This way the seasonal differences in vegetation dynamics are not taken into account. These vegetation dynamics represent the presence of transpiring leaves in summer, changing to an almost complete absence of transpiration in winter for some vegetation types. Including information on these vegetation dynamics could potentially improve the way streamflow and evaporation is modelled by conceptual hydrologic models.

To investigate the impact of a phenology model on conceptual hydrologic models the FLEX-model is adjusted with a phenology model based on the 'K_c-ET_o' approach of the Food and Agricultural Organisation (FAO) of the UN. With the conventional FLEX-model and the FAO-adjusted FLEX-model, 28 catchments from the USA are simulated in terms of streamflow and transpiration. The 28 catchments differ in terms dominant vegetation type and climate indices. The conventional model and the adjusted model are calibrated with streamflow observations and transpiration data from NASA Global Land Data Assimilation System (GLDAS) both separately and together. The conventional FLEX-model and the FAO-adjusted FLEX-model are compared to each other in order to determine under what climate conditions and for which vegetation types the phenology adjustment improves the performance of the model in terms of the ability to simulate streamflow and transpiration and the predictive capacity of the model.

The streamflow simulation of the FLEX-model does not improve by the FAO-adjustment. The FAO adjustment does cause an improvement in the mean seasonal sum of the discharge in the spring months, which is structurally underestimated by the conventional model due to an overestimation in spring of the transpiration caused by the lack of information on vegetation dynamics. However the transpiration simulation of the FLEX-model does improve drastically by the FAO-adjustment. The large overestimation of the transpiration in early spring by the conventional FLEX-model is solved by the FAO-adjustment, improving the NS- and NSlog-efficiencies for all catchments, regardless of their vegetation type of climate conditions.

Preface

I am very proud to present to you my thesis on 'The impact of an additional phenology model on the performance of conceptual hydrological models'. I have learnt a lot during this research and I am very thankful to all who have supported me during this process. I am very grateful to Markus, for his enthusiasm, his encouragements and creative ideas and to Ricciardo and Miriam for joining the thesis commitee and providing me with their constructive critiques.

Lastly I want to thank my friends and roommates, in particular Geert and Coen-Jan, who motivated me during our collective study sessions at the campus, and my girlfriend and family for their ongoing support.

C. Pierik
Delft, October 2022

Contents

1 Introduction

Fresh water is one of the most important natural resources of the earth and therefore hydrology is a very important field of study. It describes the never-ending transport of water from the atmosphere to the earth's surface and back to the atmosphere again. This hydrologic cycle is a very complex system. The water is present in the system in various forms, such as snow, evaporation and liquid water and on top of that, there are a lot of different ways of water storage, such as rivers, groundwater, the atmosphere and the interception zone.

This hydrologic cycle is too complex to be modelled completely accurately. However, there is a variety of ways to model the hydrological cycle so that a meaningful understanding of what is happening in terms of water transport and storage is given. In the field of hydrologic models, you could identify two competing philosophies of most process-based models. On the one side of the spectrum, there is the description of the small-scale processes which are then integrated numerically into the catchment scale, and on the other side of the spectrum, the system can be represented in a spatially lumped model (Hrachowitz and Clark 2017).

One of these ways of hydrological modelling more on the spatially lumped side of the earlier mentioned spectrum is the conceptual hydrological model. A conceptual hydrologic model describes the main hydrologic processes with simple mathematical equations. The main hydrological processes may include processes such as evapotranspiration, surface storage, percolation, snowmelt, baseflow, and runoff, depending on the modeller's choice (Aghakouchak and Habib 2010). It is a more simple way of modelling compared to the more physically based models from a mathematical perspective. Conceptual models can be implemented as lumped, where the forcing data and the hydrologic response is representing the whole catchment or as (semi-)distributed, where the catchment is gridded or subdivided into multiple land cover or topography categories for example (Kumar, Samaniego, and Emad 2010)(Gao et al. 2014)(Fenicia et al. 2016).

A conceptual hydrologic model can be very useful for multiple reasons. The model can describe the hydrological processes that are present in the catchment of interest. This can give a better understanding of what is happening in the catchment in terms of water fluxes and storage in a mathematical way. On top of that, the model can be used for predicting the system. Floods, droughts, groundwater recharge and land-atmosphere exchange can be predicted with a hydrologic model which can be of great importance in the field of water resource planning, flood defence and water management strategies (Hrachowitz and Clark 2017). Also, the model can be used for testing certain changes in the system. The model can for example

show the impact of certain human activity in the catchment on the hydrological cycle. All these applications show the importance of well-performing hydrologic models.

Terrestrial evapotranspiration is a large part of the water cycle. Around 60% of all precipitation on land returns to the atmosphere through terrestrial evapotranspiration (Taikan and Shinjiro 2006). Evapotranspiration is the sum of the evaporation and the transpiration, which is the water produced by vegetation through photosynthesis. The vegetation that is present in a catchment has a large impact on the evapotranspiration. The ratio of transpiration to total terrestrial evapotranspiration is estimated to be between 61% and 77%, according to (Good, Noone, and Bowen 2015). However, the uncertainty on this ratio is large, since it is also depending on the physiology and structures of the vegetation (Pan et al. 2020). This means that the accurate partitioning of transpiration and other evaporation is a complex problem and that there is a lot of improvement in the hydrologic models that are representing these processes to be made. The conventional conceptual hydrologic models use in most cases a fixed partitioning of evaporation and transpiration. This way of modelling does not include the seasonal changes of the vegetation that influence the amount of transpiration that occurs.

There are many approaches present in the current literature on how to model the partitioning of evaporation and transpiration. Examples of this are Shuttleworth–Wallace, ENWATBAL, Cupid-DPEVAP, SWEAT, TSEB, FAO dual Kc model and HYDRUS-1D (Kool et al. 2014). However, the combination of a conceptual hydrological model and a sophisticated integration of the partitioning of evaporation and transpiration has not been sufficiently investigated. In the research of Okay Mert, three methods have been developed to establish the mentioned combination (Mert 2021). A method based on the FAO dual Kc model mentioned before, a method based on the research of Jarvis, which uses a relation between temperature and transpiration (Jarvis and Mcnaughton 1986), and finally a combination of the two methods. The methods were combined with the GR4J model and the FLEX model, both simple lumped conceptual hydrologic models. The models were tested on two catchments and the performance of the conventional models was compared with the models that had the adjustment for the partitioning of evaporation and transpiration.

1.1. Problem statement

Conceptual hydrologic models give a greatly simplified representation of the hydrologic system and the models often make a lot of assumptions. One of these assumptions regards the partitioning of evapotranspiration into interception evaporation (E_i) and transpiration (E_t). The models assume a constant ratio between these two fluxes throughout the year and with this assumption disregard the impact of seasonal changes on the vegetation. This could lead to larger uncertainties in the simulations and predictions of the water fluxes in hydrological systems. As mentioned before there are a few methods developed in the research of Okay Mert. Although the first results of testing this principle turned out to be promising, it does not fully answer questions on how well this principle works and under what conditions in terms of vegetation and climate conditions the adjustment can have a positive impact on the model performance and in which part of the performance.

1.2. Research objective

In this thesis we used information about plant phenology to modify a simple lumped conceptual hydrological model. This modification is based on the FAO dual Kc model and will be implemented in the conventional lumped FLEX model. Both the conventional model and the adjusted model are tested in a large set of catchments with different climate conditions and vegetation types. We calibrated the model on streamflow and transpiration separately and

a combination of the two. After this, We compared the performance of the conventional and adjusted model regarding predictive capacity and the general ability of the model to simulate streamflow and transpiration correctly. The main research question that will be answered in this thesis is:

Under what climate and land cover conditions can a phenology adjustment improve the performance of a conceptual hydrologic model?

Where the performance of a conceptual model is subdivided into two parts:

1. The general modelling of hydrological processes such as evapotranspiration and streamflow in a catchment
2. The predictive capacity of the model

1.3. Relevance

The first part of the performance of a conceptual hydrologic model is important because it gives a better understanding of the hydrological processes in a catchment. If the streamflow in a catchment is modelled properly, but the evaporation and transpiration is deviating a lot from the data, the model does not give a full understanding of what is happening in the catchment. The second part is important because this enables people to better anticipate problems regarding floods, droughts and other water resource planning issues. What follows from this, is that this research will give a better foundation in determining whether this way of modelling is suitable for a certain set of circumstances and catchment properties.

2

Methodology

This chapter describes the experimental set-up and the model which we used to answer the research question. First, We explain the study area from which catchments are used for the analysis. Why we will clarify why we chose this area and what strategy we used to select a set of catchments. This will give an idea of where the FAO model adjustment this research tests gives an improvement in the model performance. After that, we will elucidate the data that will be used in the model. This data contains the forcing data that drives the model, the streamflow and transpiration data that is used for calibrating the model and evaluating its performance and finally, the Leaf Area Index (LAI) data which is used in the integration of vegetation phenology information in the FAO model adjustment. Also, we will describe the model setup itself, which conceptual hydrologic model we chose and why, and how this model works both with the FAO model adjustment and in its conventional setup.

2.1. Study area

This study extracted catchment from the CAMELS database of the United States of America (Newman et al. 2014). This is one of the most well-known databases with daily time series of meteorological forcing data and streamflow. The data set contains 671 catchments in the contiguous United States (CONUS) minimally impacted by human activities. The catchments are small to medium-sized with a median basin size of 336 km^2. The meteorological forcing data are the daily time series from gridded meteorological data sets (Daymet, Maurer and NLDAS). We chose the CAMELS data set because in the USA there is a wide range of different vegetation types and climate conditions over the country. The CAMELS dataset also provides a simulation with the coupled Snow-17 and SAC-SMA system, which is a conceptual air-temperature-index-based snow accumulation and ablation model (Anderson 1973). This can provide an indication or benchmark for model performance and gives attributes of the catchments, which helps in the catchment selection.

2.1.1. Catchment selection

Out of the CAMELS dataset, we selected 28 catchments. In this part, the different criteria that we used in the selection procedure will be presented.

First of all, we selected catchments with four different dominant land cover types: deciduous broadleaf forest, evergreen needleleaf forest, grassland and mixed forest. The expectations for these land cover types in terms of the impact of phenology on the hydrological processes go from large for deciduous broadleaf forest to small for evergreen needleleaf forest, with

grassland and mixed forest in between. This is because deciduous broadleaf forest loses their leaves in fall which would eliminate the transpiration during fall and winter and this is not the case for evergreen needleleaf forest and grassland. The catchments are selected to have a dominant land cover fraction of at least 0.85, to ensure the impact of the vegetation type is significant.

In terms of the size of the catchments, we selected relatively small catchments with sizes not larger than 130 km^2. Smaller catchments are in general more homogeneous in there terms of their meteorological forcing data, so the uncertainty induced by using a lumped model becomes smaller.

Another requirement we considered is the difference in LAI and the green vegetation fraction (GVF) throughout the year. the GVF is the fraction of ground covered by green vegetation (Newman et al. 2014). The difference in LAI and GVF throughout the year refers to the difference between the maximum and minimum monthly LAI and GVF. The CAMELS dataset uses a 1 km Moderate Resolution Imaging Spectroradiometer (MODIS) data product which gives monthly values between 2002 and 2014 (Newman et al. 2014). One of these products is the maximum difference in LAI. Large values indicate a large difference in transpiration throughout the year and therefore potentially a large impact of the model adjustment this research is testing. For all four different dominant land cover types, we selected catchments that have a variable difference in LAI when the LAI in summer is compared to the LAI in winter. This way we can show the impact of vegetation phenology on the performance of both the conventional model and the FAO-adjusted model.

Finally, we selected catchments based on climatic indices: both aridity and seasonality. Aridity is the ratio between the annual potential evaporation and precipitation and is described by Equation 2.1 (Budyko, Miller, and Miller 1974).

$$Aridity = \frac{\overline{E_p}}{P} \tag{2.1}$$

Seasonality of precipitation describes to what extent the precipitation is spread over the year. For this research, we used the seasonality from the CAMELS dataset, which is described in (Woods 2009). A seasonality close to zero indicates uniform precipitation throughout the year. Positive values indicate precipitation peaks in summer and negative values indicate precipitation peaks in winter.

Because the system input in terms of water is determined by the rainfall, it is interesting to see what changes in the quantity and timing of this input do with the system. Aridity indicates the quantity of the rainfall input and seasonality links to the timing of the rainfall. When the timing of rainfall is matching the periods with high temperatures the amount of transpiration could be boosted, which could lead to a larger impact by the FAO model adjustment. To be able to show the impact of these indices on the model performance, we chose catchments with deciduous broadleaf forest as their dominant land cover type with both relatively low values and relatively high values for both aridity and seasonality.

An overview of the selected catchments is given in Table 2.1 and Figures 2.1 and 2.2.

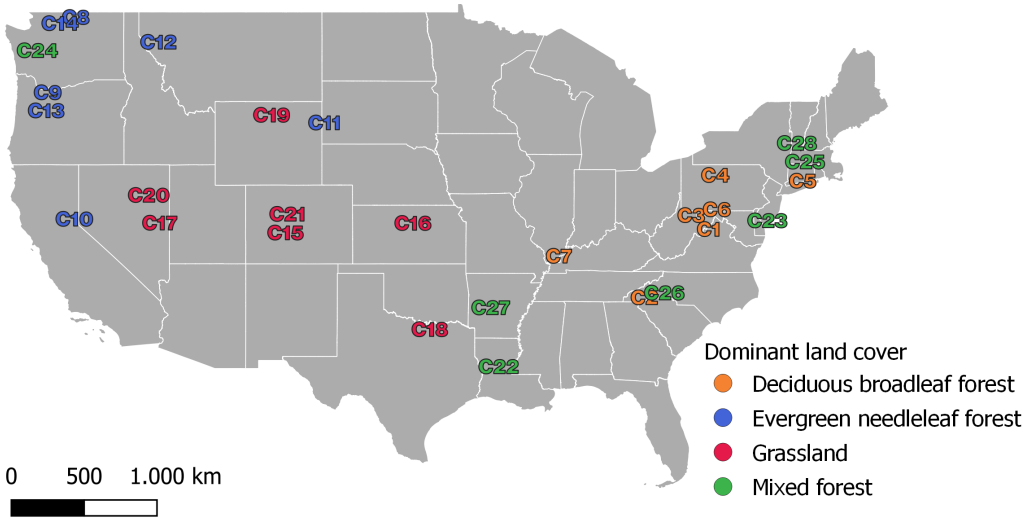


Figure 2.1: Catchment locations with the catchment name. The colors show the dominant vegetation type in the catchments

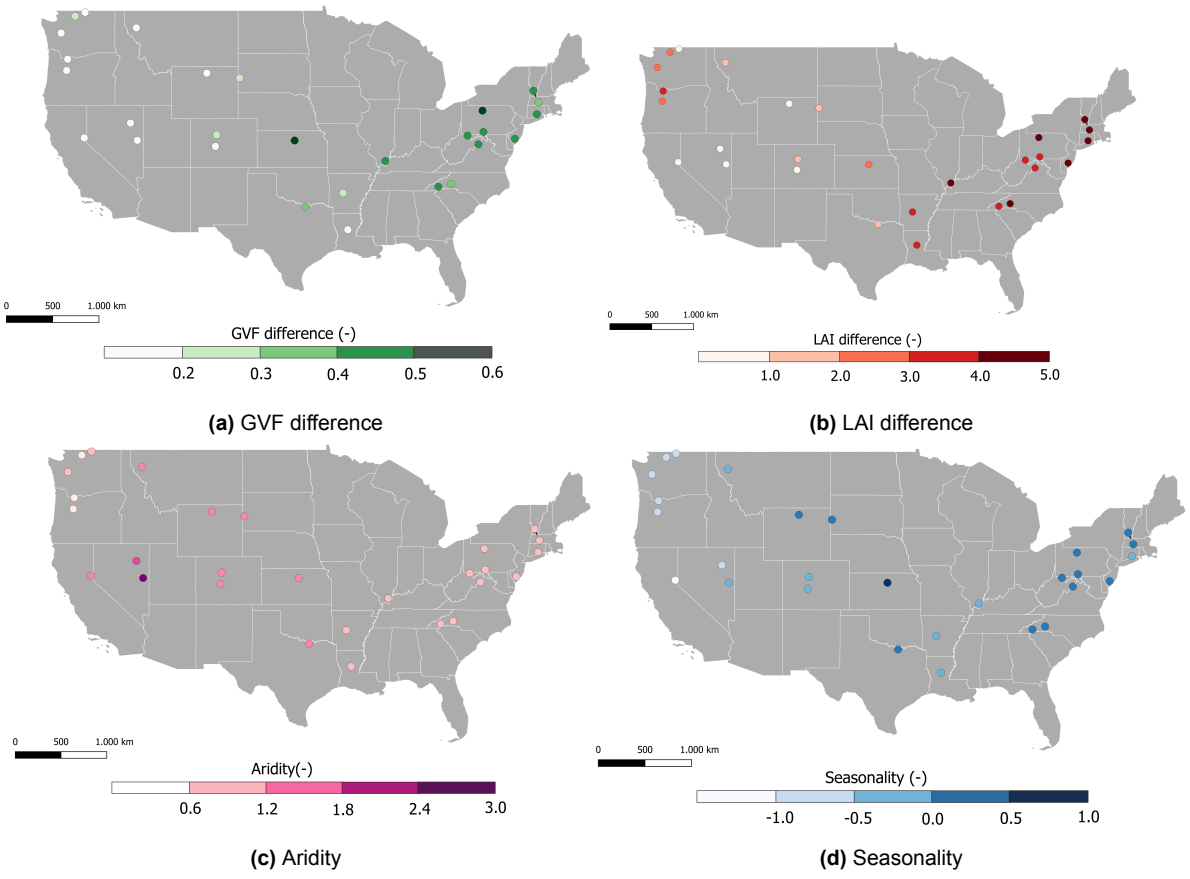


Figure 2.2: Locations of the selected catchments with their climate indices. Figure a shows the green vegetation fraction, Figure b shows the leaf area index, Figure c shows the aridity and Figure d shows the seasonality of the catchments.

Catchment	Gauge ID	Max LAI	LAI difference	Max GVF	GVF difference	Dom. land cover frac.	Dom. land cover	Seasonality	Aridity	Area (km2)
C1	03076600	4,12	3,55	0,83	0,47	1,00	Deciduous Broadleaf Forest	0,07	0,67	127,40
C2	03455500	4,43	3,76	0,86	0,43	0,67	Deciduous Broadleaf Forest	0,06	0,59	72,95
C3	01596500	4,47	3,91	0,85	0,48	1,00	Deciduous Broadleaf Forest	0,10	0,74	124,65
C4	01542810	5,27	4,74	0,90	0,55	1,00	Deciduous Broadleaf Forest	0,18	0,66	13,62
C5	01195100	5,47	4,83	0,91	0,50	0,98	Deciduous Broadleaf Forest	0,00	0,64	14,84
C6	01613050	4,19	3,67	0,83	0,47	0,80	Deciduous Broadleaf Forest	0,12	0,79	27,90
C7	03384450	5,04	4,45	0,88	0,48	0,81	Deciduous Broadleaf Forest	-0,07	0,74	111,13
C8	12447390	1,20	0,79	0,49	0,18	0,86	Evergreen Needleleaf Forest	-0,60	0,86	58,10
C9	14138900	4,85	3,02	0,91	0,19	1,00	Evergreen Needleleaf Forest	-0,79	0,24	21,68
C10	10343500	2,07	0,96	0,67	0,16	1,00	Evergreen Needleleaf Forest	-1,13	1,10	27,60
C11	06408700	2,65	1,65	0,74	0,21	1,00	Evergreen Needleleaf Forest	0,49	1,28	20,80
C12	12374250	2,88	1,93	0,76	0,20	0,99	Evergreen Needleleaf Forest	-0,24	1,07	50,79
C13	14141500	4,22	2,55	0,87	0,17	0,98	Evergreen Needleleaf Forest	-0,77	0,29	59,87
C14	12178100	2,84	2,12	0,72	0,30	0,88	Evergreen Needleleaf Forest	-0,74	0,33	69,68
C15	07083000	0,79	0,53	0,33	0,15	1,00	Grasslands	-0,17	1,36	60,76
C16	06879650	2,35	2,08	0,70	0,52	1,00	Grasslands	0,67	1,01	11,51
C17	10244950	0,76	0,44	0,33	0,11	1,00	Grasslands	-0,34	2,61	28,20
C18	08050800	1,89	1,50	0,61	0,31	1,00	Grasslands	0,08	1,11	100,93
C19	06278300	0,87	0,64	0,38	0,19	1,00	Grasslands	0,04	1,32	58,85
C20	10316500	0,99	0,80	0,39	0,24	1,00	Grasslands	-0,55	1,55	64,78
C21	09066200	1,61	1,22	0,53	0,26	0,99	Grasslands	-0,29	1,40	16,10
C22	07373000	4,72	3,09	0,85	0,19	1,00	Mixed Forests	-0,20	0,80	131,18
C23	01411300	5,35	4,57	0,89	0,44	0,92	Mixed Forests	0,05	0,80	79,32
C24	12073500	3,87	2,44	0,86	0,18	0,94	Mixed Forests	-0,85	0,71	15,67
C25	01162500	5,40	4,15	0,88	0,33	1,00	Mixed Forests	0,05	0,60	49,71
C26	02143040	5,16	4,25	0,89	0,39	1,00	Mixed Forests	0,01	0,69	66,48
C27	07362587	5,31	3,96	0,87	0,29	1,00	Mixed Forests	-0,07	0,75	69,77
C28	01170100	5,38	4,62	0,90	0,46	0,86	Mixed Forests	0,07	0,58	106,99

Table 2.1: Overview of the catchments with their properties

2.2. Data

2.2.1. CAMELS dataset

Forcing data

The CAMELS database contains an extension to the daily meteorological forcing from three data sets: Daymet (Thornton et al. 2014), Maurer (Livneh et al. 2013), and NLDAS (Xia et al. 2012). For this research, we used the forcing data from Daymet. We used This dataset also for the calculation of the climate indices the database provides. The Daymet forcing data contains precipitation, shortwave radiation, maximum air temperature, minimum air temperature and water vapor pressure. The precipitation goes directly into the model. We used the maximum and minimum air temperature to calculate the mean temperature, which goes directly into the model. All data has a spatial resolution of 1 km by 1 km and is averaged over the area of the catchment, making it suitable for a lumped model.

Potential evaporation

Finally, we used the shortwave radiation and water vapor pressure together with temperature data to calculate the potential evaporation. This is done using the Priestley-Taylor formulation presented in Equation 2.2, calibrated for each catchment separately. This is a method related to the Penman-Monteith equation, with the adjustment that it needs less observational data, which is compensated with a coefficient (α) which is found through calibration. The Priestley-Taylor formulation used is:

$$E_p = \alpha \frac{\Delta(R_n - G)}{\rho_w * \lambda * (\Delta + \gamma)} \tag{2.2}$$

Where E_p is the potential evaporation; Δ (kPa $\circ C^{-1}$) is the slope of the saturation vapor pressure-temperature relationship; R_n (MJ m^{-2} day^{-1}) is the net radiation which is estimated with the day of the year, the Daymet variables and some equations from the FAO report (Allen et al. 1998). G (MJ m^{-2} day^{-1}) is the soil heat flux (assumed to be zero in this case), ρ_w is the density of water (kg/m^3), λ is Volumetric latent heat of vaporization which is 2453 MJ m^{-3}. γ

(kPa $\circ C^{-1}$) is the psychrometric constant and α (-) is the P–T coefficient, which is calibrated. The value for α will be high for catchments with more arid climate conditions and low for catchments with more humid climate conditions.

Streamflow
The daily streamflow measurements are retrieved from the United States Geological Survey (USGS). The USGS developed and maintains a database with over 9000 stream gages called Geospatial Attributes of Gages for Evaluating Streamflow (GAGES-II). The streamflow data has a temporal resolution of one day. (Newman et al. 2015)

2.2.2. Transpiration

To be able to evaluate the partitioning of transpiration and evaporation by the model, transpiration data is needed. In this study we retrieved this data from The NASA Global Land Data Assimilation System (GLDAS). System version 2 has two components: GLDAS-2.0 and GLDAS-2.1. The difference between these two components is that GLDAS-2.0 is forced with only meteorological forcing input data and GLDAS-2.1 is forced with a combination of both model and observation data (Rui and Beaudoing 2021). GLDAS-2.1 is forced with Global Data Assimilation System atmospheric analysis fields (GDAS) (Derber, Parrish, and Lord 1991), the disaggregated Global Precipitation Climatology Project (GPCP) V1.3 Daily Analysis precipitation fields (Adler et al. 2003), and the Air Force Weather Agency's AGRicultural METeorological modeling system (AGRMET) radiation fields. This dataset produces land surface states and fluxes at different spatial and temporal resolutions, ranging from 0.25° to 1.0° and 3 hourly to monthly data (Rodell et al. 2004)(Rui and Beaudoing 2021). From the two components we chose to use GLDAS-2.1 for the transpiration data in this study, because this in this dataset the more advanced modelling strategy is used. The GLDAS-2.1 provides a whole list of modeled fluxes from radiation and heat fluxes to water fluxes. In this study we used the transpiration data of GLDAS-2.1 from the Noah land surface model, which is provided in W/m^2 at a spatial resolution of 0.25°. We then averaged the GLDAS-2.1 transpiration over the area of the catchments and transformed the data to mm/day with a conversion of 1 W/m^2 = 0.0353 mm/day (Allen et al. 1998). The temporal resolution of 3 hours, is averaged to a daily resolution before going into the model. The Noah Land Surface Model uses a Jarvis scheme(Jarvis and Mcnaughton 1986) or Ball-Berry scheme(Ball, Woodrow, and Berry 1987) to calculate transpiration. This method also uses LAI data as input (Niu et al. 2011). Because there is no observational data on transpiration available, and the transpiration product of for example the Noah Land Surface Model is the best option around, we justified the choice to use this data in the calibration of the model.

2.2.3. Leaf area index

The LAI variable indicates the number of equivalent leave layers per unit of ground area. We retrieved the LAI in this study from the MODIS sensors of the Terra and Aqua satellites of NASA referred to as MCD15A3H. The level-4 MODIS product provides data for LAI and Photosynthetically Active Radiation (FPAR) in combination with a look-up table. The MODIS algorithm uses as input the spectral information of the MODIS red (648 nm) and near-infrared (NIR,858 nm) with their uncertainties in combination with the vegetation structural type and sun-sensor geometry. For each pixel modeled and observed bidirectional reflectance factors (BRFs) are compared for a given set of conditions that are expected of a given biome type(Ranga Myneni 2015). For a more detailed description of the MODIS data, the research of Ranga Myenini (2015) can consulted.

This is a remote sensing product with a spatial resolution of 500 m and a temporal resolution of 4 days(Ranga Myneni 2015). As with the other data the LAI is averaged over the whole area of all the catchments separately. As for the temporal resolution the LAI data is resampled to three-week data. This is done because the data can be very noisy which makes it inconvenient to determine the changes in seasons in a semi-automatic way.

2.3. Hydrologic model

In this study we aim at increasing the performance of a conceptual hydrologic model with a phenology integration. A phenology integration means taking information about vegetation phenology and deriving a partitioning of the evapotranspiration from this. For the conventional hydrologic model we chose the lumped FLEX-model(FLEX^{C}) (Gao et al. 2014) and for the phenology integration we chose a method based on the research of the Food and Agriculture Organization (FAO) (Allen et al. 1998). This adjusted model will be referred to as (FLEX^{FAO}). In this part, we describe the conventional hydrologic model together with the method for phenology integration and how the model is calibrated and evaluated.

2.3.1. FLEX-model

For testing the phenology integration we chose the lumped FLEX model. The Flex (Flux Exchange) hydrological model or (FLEX^{C}) is a lumped conceptual model that consists of five reservoirs that interact with each other. (FLEX^{C}) is based on the HBV model, which is one of the most widely used and successful conceptual hydrologic models in the field (Bergström 1992). We chose this model, because of its simple structure, which makes it easier to implement the phenology integration and also emphasizes the adjustments more. A complex model has more uncertainty in its structure and more elements that could be impacted by any adjustments.

As mentioned before (FLEX^{C}) is built out of five reservoirs: The interception reservoir (S_i), which represents the rainfall intercepted by the canopy, the snow reservoir (S_w), which is used when temperatures are below a certain threshold to simulate snowfall, the unsaturated soil reservoir (S_u), which is the water stored in the unsaturated zone of the soil and runoff-reservoirs representing the formation of runoff in both a slow reacting reservoir (S_s) and a fast-reacting reservoir (S_f). An overview of the model structure is presented in Figure 2.3, where all the buckets represent the reservoirs, the arrows represent the fluxes between the reservoirs, and the red symbols represent the parameters that are calibrated in the model. The model starts with the forcing data as input. This input is precipitation (P) and temperature (T) data. All reservoirs will be enlightened in more depth in this chapter and the equations that are used to generate the fluxes between the reservoirs will be explained. All equations are retrieved from (Fenicia et al. 2006) and (Gao et al. 2014)

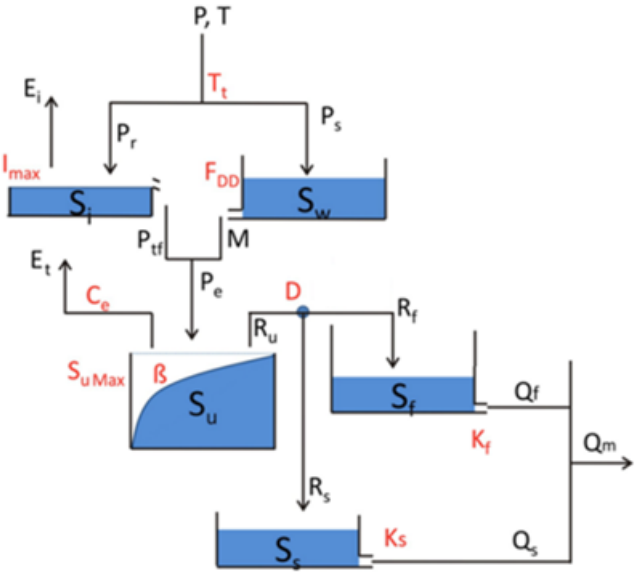


Figure 2.3: Structure of (FLEX^{C}) (Gao et al. 2014)

Interception reservoir

As mentioned before the input of the model consists of P and T. P is the water input that is used to fill the reservoirs and T is the temperature used to determine the partitioning of the precipitation in the interception reservoir and the snow reservoir. Where the threshold between this partitioning is set, is determined by the threshold temperature (T_t), which is calibrated in the model. If the temperature is above T_t, the precipitation is assumed to be rainfall (P_r) and will be used to fill S_i. P_r will fill S_i until the maximum interception storage (I_{max}). When S_i is filled the water will leave S_i as excess rainfall (P_tf). This is represented in Equation 2.3. The water will also leave the interception reservoir as interception evaporation (E_i), which is determined by the potential evaporation (E_p) and the storage. This is done with Equation 2.4.

$$P_{tf} = \begin{cases} 0; & \text{if } S_i < I_{max} \\ P_r; & \text{if } S_i = I_{max} \end{cases} \tag{2.3}$$

$$E_i = \begin{cases} E_p; & \text{if } S_i > 0 \\ 0; & \text{if } S_i = 0 \end{cases} \tag{2.4}$$

The water balance of S_i is given by equation 2.5:

$$\frac{dS_i}{dt} = P_r - E_i - P_{tf} \tag{2.5}$$

Snow reservoir

If the temperature is below T_t, the precipitation is assumed to be snowfall(P_s) and will fill S_w. From S_w the water will leave the reservoir in the shape of snowmelt (M) in mm/day. This flux is only induced when the temperature is above T_t again. The quantity of the flux is determined by Equation 2.6 with the temperature in combination with the degree day factor (F_{DD}). The degree-day factor is defined as the amount of water that is melted per day and per Celsius degree above T_t. This value will be found by calibration.

$$M = F_{DD}(T - T_t) \qquad \text{if} \quad T > T_t \tag{2.6}$$

The water balance of S_w is given by equation 2.7:

$$\frac{dS_w}{dt} = P_s M \tag{2.7}$$

Unsaturated reservoir
S_u is filled with the effective precipitation (P_e), which is the combination of P_{tf} and M, reduced by the runoff coefficient (C_r). The runoff coefficient is determined by equation 2.8:

$$C_r = \frac{1}{1 + \exp\left(\frac{\frac{-S_u}{S_{uMax}+\frac{1}{2}}}{\beta}\right)} \tag{2.8}$$

Where S_u is the soil moisture content, S_{uMax} is the maximum soil moisture capacity in the root zone and β is the parameter describing the spatial process heterogeneity in the catchment. Parameter β is found through calibration. From S_u water is going out through the transpiration (E_t). E_t is calculated with Equation 2.9, with the remainder of E_p when E_i is subtracted, and a constraint of S_u, S_{uMax} and C_e where C_e indicates the fraction of S_{uMax} above which the actual evaporation is equal to potential evaporation.

$$E_t = (E_p - E_i) \min\left(1, \frac{S_u}{S_{uMax} C_e}\right) \tag{2.9}$$

Another flow out of S_u is the generated flow during rainfall events (R_u). R_u is determined by P_e and C_r, with equation 2.10

$$R_u = P_e C_r \tag{2.10}$$

The water balance of the unsaturated reservoir is then given by equation 2.11

$$\frac{dS_u}{dt} = P_e(1 - C_e) - E_t \tag{2.11}$$

Fast and slow reacting reservoir
From S_u the runoff is generated which is stored in the surface runoff reservoirs. The surface runoff reservoirs are divided into a fast-reacting reservoir (S_f) and a slow-reacting reservoir (S_s) taking into account the reaction time of the system between the storm and peak flow. This division is determined by the splitter functions in Equation 2.13, where R_f is the flow into the fast-response routine, R_s indicates the flow into the groundwater reservoir and D is a splitter to separate recharge from preferential flow, which will be determined through calibration.

$$R_f = R_u D \tag{2.12}$$

$$R_s = R_u(1 - D) \tag{2.13}$$

How much of S_f and S_s then can flow out to runoff is determined by K_f, which is the timescale of the fast runoff, and K_s, which is the timescales of the slow runoff. Both will be determined by calibration and then used in Equations 2.14 and 2.16 to calculate the fast runoff (Q_f) and the slow runoff (Q_s), which added together is the total runoff (Q_m)

$$Q_f = \frac{S_f}{K_f} \tag{2.14}$$

$$Q_s = \frac{S_s}{K_s} \tag{2.15}$$

$$Q_m = Q_f + Q_s \tag{2.16}$$

The water balances of both the fast-reacting reservoir and the slow-reacting reservoir are presented in Equations 2.17 and 2.18

$$\frac{dS_f}{dt} = R_f - Q_f \tag{2.17}$$

$$\frac{dS_s}{dt} = R_s - Q_s \tag{2.18}$$

2.3.2. Phenology integration

In the FLEX model evapotranspiration is the function of potential evapotranspiration and water availability in the soil. When the actual soil moisture content is too large, evapotranspiration will be limited by the climate conditions represented by the potential evaporation. The other way around, when the soil moisture becomes low, evapotranspiration will be limited by the water supply. In Equation 2.9 is visible how this process works in (FLEX^{C}). Also in other conceptual hydrologic models, evapotranspiration is mostly calculated as a function of both potential evaporation and soil moisture deficiency (Liu et al. 2017). Information about plant phenology is missing in this method. Through calibration, this might be compensated partly, so that the model at least still simulates the streamflow of the system properly, but the disadvantage is that the division between how much of the potential evapotranspiration is attributed to E_i and how much to E_t is simulated less accurately. This has the consequence that the interaction between the buckets in the model is significantly different because E_i is a flux out of the interception reservoir and E_t is a flux out of the unsaturated soil reservoir.

The method to integrate phenology into the FLEX-model is based on the research of the FAO (Allen et al. 1998). In this research, guidelines are presented for reference and crop evapotranspiration determination through meteorological data and crop coefficients. To determine the reference crop evaporation multiple methods are developed. One of these methods is called the 'K_c-ET_o' approach. In this approach, crop-evapotranspiration is determined by the reference crop evapotranspiration and the crop coefficient (K_c). K_c is then used to find the crop-evapotranspiration through Equation 2.19, where ET_c is the crop-evapotranspiration and ET_0 the reference crop-evapotranspiration.

$$ET_c = K_c * ET_0 \tag{2.19}$$

According to (Allen et al. 1998) K_c represents four different characteristics that influence the relationship between actual and potential evapotranspiration of the crops:

- The crop height in relationship to the aerodynamic resistance
- Albedo of both the crops and the surface regarding the net-radiation
- The canopy resistance
- Evaporation from soil

K_c can be found empirically through a look-up table after which a K_c-curve can be constructed. This is a curve that is divided into four crop development stages: the initial stage, the crop development stage, the mid-season stage, and the late-season stage. For the initial stage and the mid-season stage, the K_c values are determined through a look-up table, where the values are then adjusted to reflect the wetting frequency of soil surface and climate conditions. An example of this curve is shown in Figure 2.4

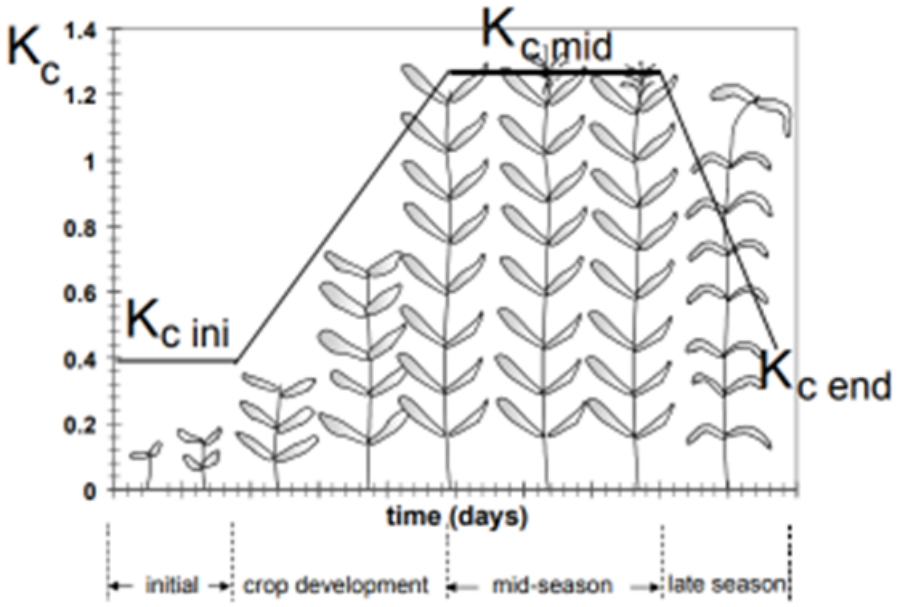


Figure 2.4: K_c curve for different growing stages (Allen et al. 1998)

This approach is developed for crops with one growing season and not for vegetation like forests. Also, this approach demands detailed information about the local circumstances regarding vegetation. Therefore, we used a variation of this approach to partition E_p into evaporation and transpiration. The method consists of three steps.

First, we determined the timing of the development stages by the LAI curve. Because the LAI gives direct information on the vegetation's activity in terms of transpiration the LAI is suitable for determining when the vegetation is starting to grow leaves and when the vegetation is releasing the leaves. Also in the case of a type of vegetation that holds its leaves throughout the year the amount of transpiration varies and with that the partitioning of the E_p. When there are sudden changes in the LAI curve, we interpreted that as a transition to a new development phase and a timestamp is then made that can go into the model.

After that the K_v-curve will be constructed by the FLEX^{FAO} model, where K_v is the vegetation coefficient, which is equivalent to K_c in the 'K$_c$-ET$_o$' approach. However, the notation K_v does not imply that this is a crop coefficient, which is not the case in the catchments in this research. To construct the K_v-curve, the values for K_v are determined by the FLEX^{FAO} model. This is done by calibrating the model on K_v in the initial stage ($K_{v,ini}$) and K_v in the mid-season ($K_{v,mid}$). The $K_{v,ini}$ will range between 0 and 0.5 and the $K_{v,mid}$ will range between 0.5 and 1. Between these stages, K_v will be linearly interpolated between $K_{v,ini}$ and $K_{v,mid}$ and between $K_{v,mid}$ and $K_{v,ini}$ of the next season leading to the construction of the K_v curve.

Finally, the found K_v curve can be used to partition E_p. This is done by modifying the part in the conventional FLEX model that calculated the E_i and E_t in a way that includes the partitioning of evaporation and transpiration by K_v. This means that Equation 2.4 will become Equation 2.20, and Equation 2.9 will become Equation 2.21, where t is time.

$$E_i = \begin{cases} E_p * (1 - K_v(t)); & \text{if } S_i > 0 \\ 0; & \text{if } S_i = 0 \end{cases} \tag{2.20}$$

$$E_t = (E_p * K_v(t)) \min\left(1, \frac{S_u}{S_{uMax}C_e}\right) \tag{2.21}$$

2.3.3. Model calibration and evaluation

A conceptual hydrologic model is a combination between a data-driven model and a physically based model, which means that in the formulations that describe the hydrologic processes, parameters are present that can not be measured by fieldwork realistically. To solve this problem the model needs to be calibrated. By calibration values are found for the parameters that can not be measured by fieldwork, that give the best performances of the model.

To calibrate the models we used a Monte Carlo sampling strategy. With this method, a set of 250.000 simulations with parameter samples out of a uniform parameter distribution within a certain range are carried out. These ranges are presented in Table 2.2. For every simulation, we determined the performance of that simulation over a calibration period of three years between January 1, 2005, and December 31, 2008. Also, we used a spin-off period between 1 and 1.5 years to fill the reservoirs. We determined The performance by simulating the streamflow and the transpiration and comparing it to the observational streamflow data from the CAMELS dataset and the GLDAS-2.1 transpiration data respectively with objective functions. Two objective functions are selected for this research proposed by (Nash and Sutcliffe 1970), shown in Equation 2.22 and Equation 2.23, where $Q_{s,i}$ is the simulated streamflow and $Q_{o,i}$ the observed streamflow from the CAMELS dataset, described in Chapter 2.2.1. The Nash-Sutcliffe efficiency (NS) in Equation 2.22 puts more emphasis on how well the model fits the data in the high flows and the logarithmic Nash-Sutcliffe efficiency (NSlog) in Equation 2.23 puts more emphasis on how well the model fits the data in the low flows.

$$N_{Q,NS} = 1 - \frac{\sum_{i=1}^{n}(Q_{s,i} - Q_{o,i})^2}{\sum_{i=1}^{n}(Q_{s,i} - \overline{Q}_{o,i})^2} \tag{2.22}$$

$$N_{Q,NSlog} = 1 - \frac{\sum_{i=1}^{n}(log(Q_{s,i}) - log(Q_{o,i}))^2}{\sum_{i=1}^{n}(log(Q_{s,i}) - \overline{log(Q_{o,i})})^2} \tag{2.23}$$

Both objective functions are also applied on the transpiration flux of the model, leading to Equation 2.24 and Equation 2.25, where $E_{t,s,i}$ is the simulated transpiration and $E_{t,o,i}$ the transpiration data from the GLDAS-2.1 dataset described in Chapter 2.2.2.

$$N_{E_t,NS} = 1 - \frac{\sum_{i=1}^{n}(E_{t,s,i} - E_{t,o,i})^2}{\sum_{i=1}^{n}(E_{t,s,i} - \overline{E}_{t,o,i})^2} \tag{2.24}$$

$$N_{E_t,NSlog} = 1 - \frac{\sum_{i=1}^{n}(log(E_{t,s,i}) - log(E_{t,o,i}))^2}{\sum_{i=1}^{n}(log(E_{t,s,i}) - \overline{log(E_{t,o,i})})^2} \tag{2.25}$$

With these objective functions, the Euclidean distance can be calculated. The Euclidean distance combines multiple objective functions so that the best performing samples can be selected considering multiple objective functions. We calibrated each catchment in three different ways:

- Only on streamflow, using Equation 2.26.
- Only on transpiration, using Equation 2.27.

- On streamflow and transpiration, using Equation 2.28.

$$D_{e,Q} = \sqrt{(1 - N_{Q,NS})^2 + (1 - N_{Q,NSlog})^2} \tag{2.26}$$

$$D_{e,E_t} = \sqrt{(1 - N_{E_t,NS})^2 + (1 - N_{E_t,NSlog})^2} \tag{2.27}$$

$$D_{e,Q,E_t} = \sqrt{(1 - N_{Q,NS})^2 + (1 - N_{Q,NSlog})^2 + (1 - N_{E_t,NS})^2 + (1 - N_{E_t,NSlog})^2} \tag{2.28}$$

For every calibration, we considered the 100 samples with the lowest value for D_e to be the best behavioral samples and we used them for evaluation from January 1, 2008, to December 31, 2011. This way the predictive capacity of the model can be evaluated. For all three ways of calibration, we evaluated both simulated streamflow and simulated transpiration. This way a broad overview of the impact of the FAO adjustment on the different hydrologic processes can be presented.

Parameter	Dimension	Description	Calibration range
F_{DD}	L/T*Temp	Degree day factor	(1,8)
T_t	Temp	Threshold temperature for snow	(-2.5,2.5)
I_{max}	L	maximum storage capacity of S_i	(0.1,5)
S_{uMax}	L	Maximum storage capacity of S_u	(50,1000)
C_e	-	Coefficient for constraint on E_p	(0.25,1)
β	-	Coefficient of shape of C_r	(0.01,5)
D	-	Coefficient of partitioning of runoff	(0,1)
K_f	T	Time scale for Q_f	(1,20)
K_s	T	Time scale for Q_s	(10,100)
$K_{v,ini}$	-	vegetation coefficient in initial stage	(0,0.5)
$K_{v,mid}$	-	vegetation coefficient in mid-season stage	(0.5,1)

Table 2.2: Calibrated parameter ranges

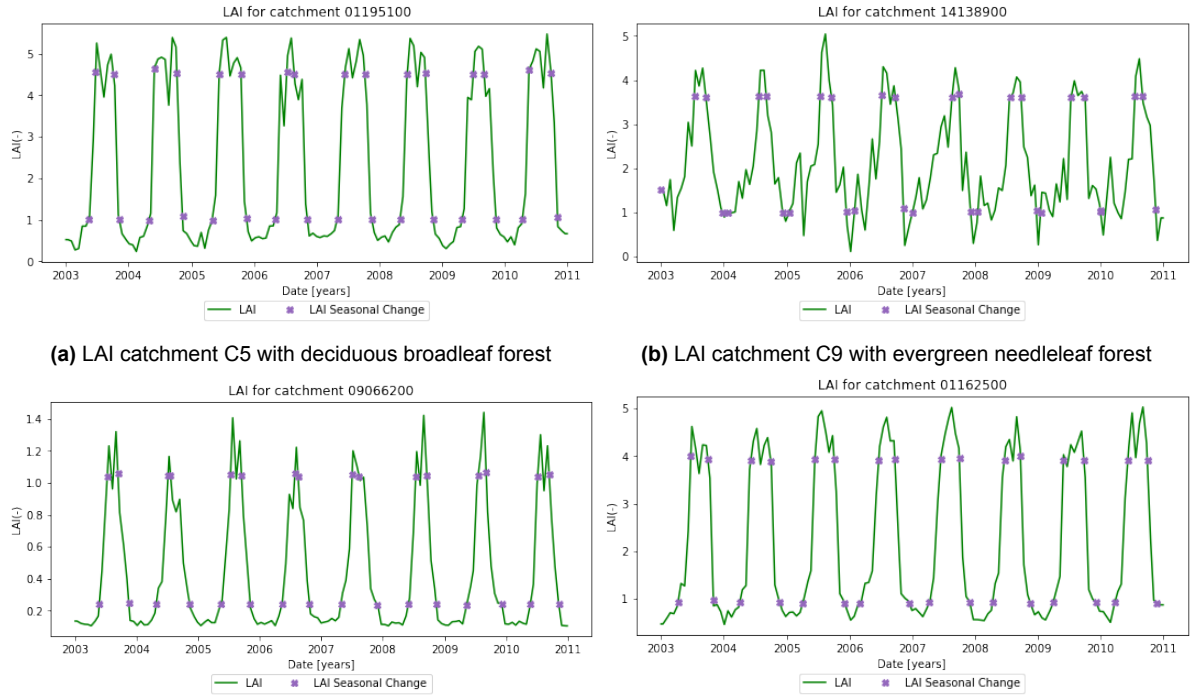
3

Results

In this chapter, the results are presented. First, we present the results of the analysis of the LAI which we used for the determination of the growing stages of the vegetation. Then, we show the results of the model simulations. This is to give an impression of how well the models work and what the differences are between the different calibration strategies. After that, we give a more detailed description of the model performances. These results are subdivided into the three different calibration strategies described in Chapter 2.3.3. For every calibration strategy, the results are subdivided into the results for the streamflow and the transpiration. For both the streamflow and the transpiration, we then present the conventional FLEX model ($FLEX^{C}$) and the FLEX model with the FAO adjustment ($FLEX^{FAO}$) for comparison.

3.1. Determination growing stages

The growing stages are determined with the MODIS-derived Leaf Area Index. In Figure 3.1, the LAI curve is presented for one catchment for every vegetation type. This way an indication is given about the seasonal variation of the LAI in catchments with different vegetation types. The green line shows the average LAI in the catchment and the purple dots show the boundaries between the different growing stages, which are used in the model. The other LAI curves with the timestamps for the growing stages can be found in Appendix A.



(a) LAI catchment C5 with deciduous broadleaf forest

(b) LAI catchment C9 with evergreen needleleaf forest

(c) LAI catchment C21 with grassland

(d) LAI catchment C25 with mixed forest

Figure 3.1: Leaf Area Index (LAI) for four catchments. The purple dots indicate the transition of the vegetation's development stages. The periods between the top purple dots indicate the dormant seasons. The periods between the bottom purple dots indicate the peak seasons.

To give an indication of what dates can be used as boundaries for the different growing stages when no LAI data is present a range of the dates is shown in Table 3.1. For every dominant vegetation type the table shows a range of the dates that we used in this research as boundaries for the different growing stages.

Growing stage	Deciduous broadleaf forest	Evergreen needleleaf forest	Grassland	Mixed forest
1	Feb 3rd - May 13th	Feb 15th - May 30th	Mar 1st - Jun 18th	Feb 16th - Apr 26th
2	May 19th - Jul 16th	Jun 19th - Jul 18th	Apr 23th - Jul 18th	Apr 22th - Jun 27th
3	Aug 26th - Oct 16th	Aug 7th - Oct 24th	Aug 4th - Oct 10th	Aug 23th - Oct 23th
4	Nov 8th - Dec 22th	Nov 3th - Dec 23th	Oct 29th - Dec 18th	Nov 10th - Dec 26th

Table 3.1: Date ranges for the start of the four growing stages for catchments with deciduous broadleaf forest, evergreen needleleaf forest, grassland, and mixed forest respectively.

3.2. Model simulations

In this chapter we will show the streamflow and transpiration simulations of four catchments out of the whole batch of catchments. Not all catchments will be enlightened in detail, because of the large amount of catchment we examined in this research. Instead for all four different land cover types that were taken into account in this thesis, we selected one catchment and present them in this chapter. We did this to show some of the patterns that can be recognised in the simulations. We will not go into detail about all the specific particularities of the streamflow and transpiration simulations for all catchments individually, because the scope of this thesis is more focused on the performance of the models on a broader scale. All the streamflow and transpiration simulations of the catchments individually can be found in Appendix B and C if the reader is interested.

3.2.1. Streamflow

Some of the simulations of the streamflow of $FLEX^C$ and $FLEX^{FAO}$ are presented in the Figures 3.2 - 3.5. For every vegetation type one catchment is presented in this subsection. The streamflow simulations of the other catchments can be found in Appendix B. Every catchment simulation in the figures consists of three parts. The top graph shows the simulated streamflow and the observed streamflow for the whole simulation, and the bottom two parts show the simulated streamflow and the observed streamflow for the first year of the calibration period and the first year of the evaluation period. The colors in the figure indicate which calibration strategy was used, and the linestyles indicate whether $FLEX^C$ or $FLEX^{FAO}$ was used for the simulation. A pattern that comes back in most simulations regardless of their landcover type is that the models that used only streamflow data and the models that used both streamflow data and transpiration data for calibration, which are represented by the red and green lines, show a reasonably good correlation with the observed streamflow data. The models that used only transpiration data for calibration, which is represented by the blue line, show little or very little correlation with the observed streamflow data. For the catchment shown in this chapter the things that stand out besides the characteristics mentioned above will be laid out. The catchments give typical responses in comparison to the whole batch of catchments.

In Figure 3.2 the streamflow simulation of catchment C5 is visible. The landcover in this catchment is dominated by deciduous broadleaf forest. The baseflow is matched closely throughout the whole simulation both for the calibration period and the evaluation period and the peakflows are matched occasionally. From this graph not a clear distinction between the models calibrated on streamflow and the models calibrated on streamflow and transpiration can be seen.

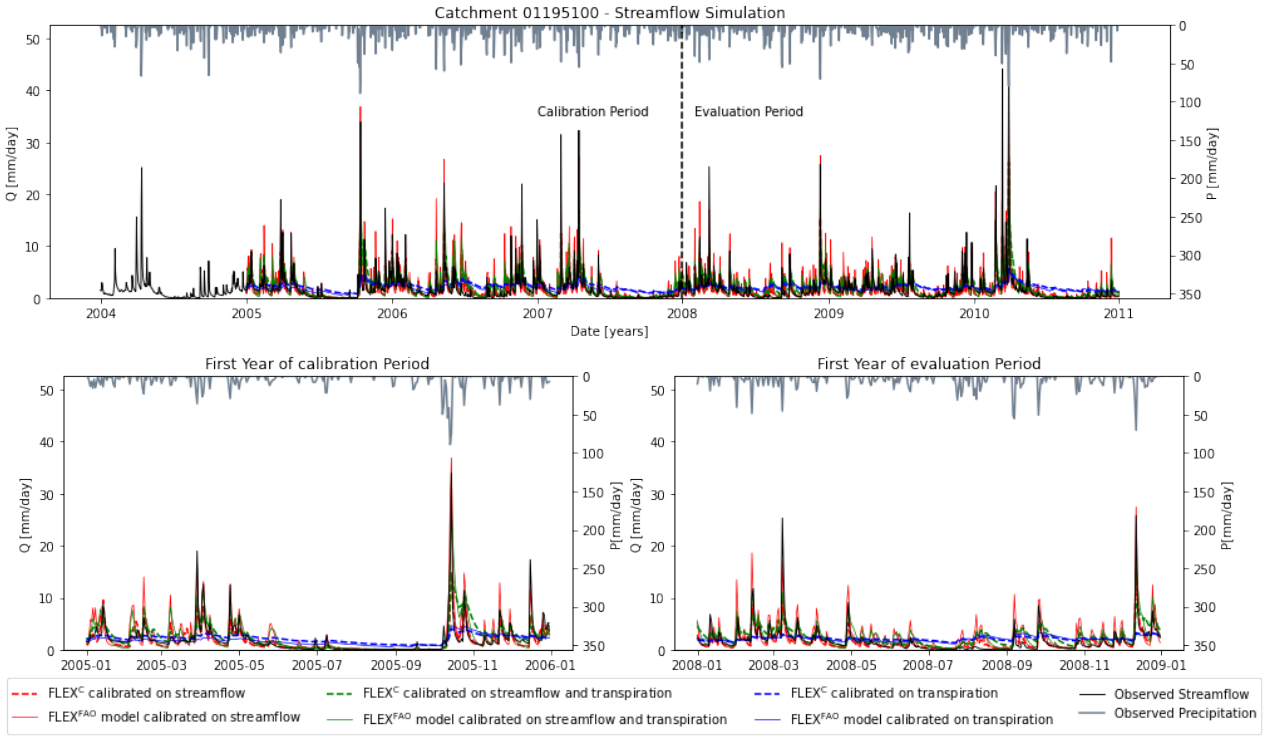


Figure 3.2: Catchment C5 streamflow simulation. Landcover dominated by deciduous broadleaf forest

In Figure 3.3 the streamflow of catchment C9 is presented. The landcover in this catchment is dominated by evergreen needleleaf forest. In the figure it is visible that similar to catchment

C5 in Figure 3.2 that the models calibrated on only transpiration give relatively poor streamflow responses compared to the observed streamflow. The baseflow is modelled reasonably well, but the response in the peakflow is not sufficient. The models that calibrated on streamflow or streamflow and transpiration are both performing better, however the models that calibrated on both streamflow and transpiration still miss some of the peaks in their response.

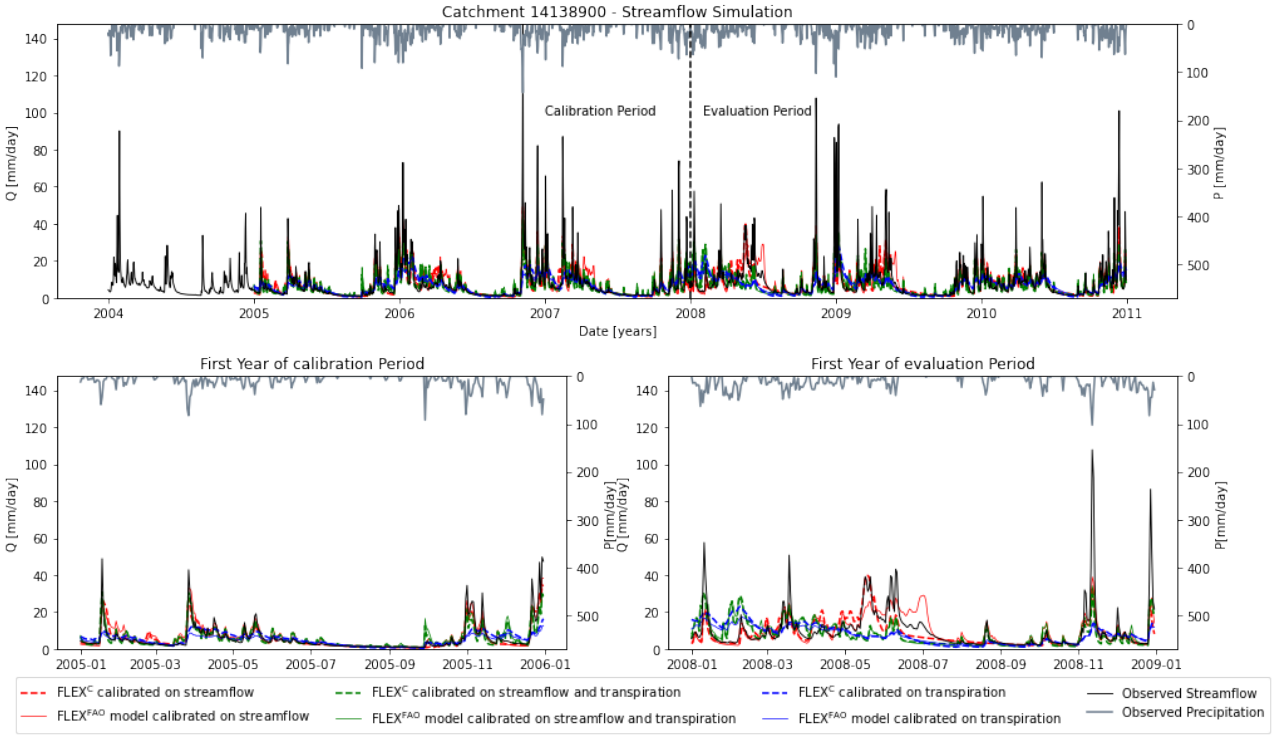


Figure 3.3: Catchment C9 streamflow simulation. Landcover dominated by evergreen needleleaf forest

In Figure 3.4 the streamflow simulation of catchment C21 is visible, which has a dominant landcover of grassland. In this catchment the river has a clear seasonal pattern with a peakflow in late spring or early summer. What stands out in this simulation is that the FLEX^{C} calibrated on transpiration is underestimating in the peakflows and the FLEX^{FAO} calibrated on transpiration is overestimating in the spring. FLEX^{C} calibrated on streamflow and transpiration gives a good simulation of the peakflows, but FLEX^{FAO} calibrated on streamflow and transpiration underestimates the peakflows similarly to FLEX^{FAO} calibrated on transpiration.

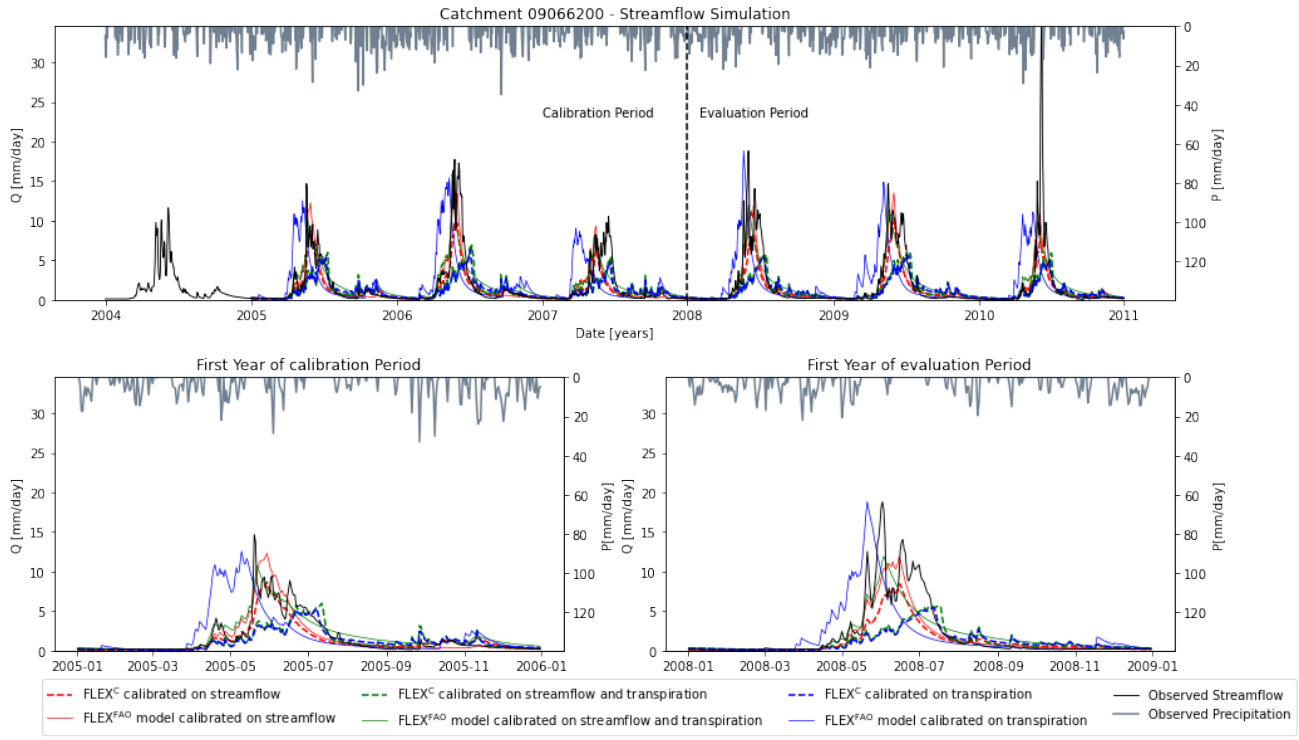


Figure 3.4: Catchment C21 streamflow simulation. Landcover dominated by grassland

Catchment 25 is dominated by mixed forest. The streamflow simulation is presented in Figure 3.5. Different from the catchments earlier shown in this chapter the $FLEX^{FAO}$ calibrated on transpiration already shows better responses to the peakflows, even though no streamflow observations were used for calibration. If this is also the case for other catchments with mixed forests will be shown in Chapter 3.4.

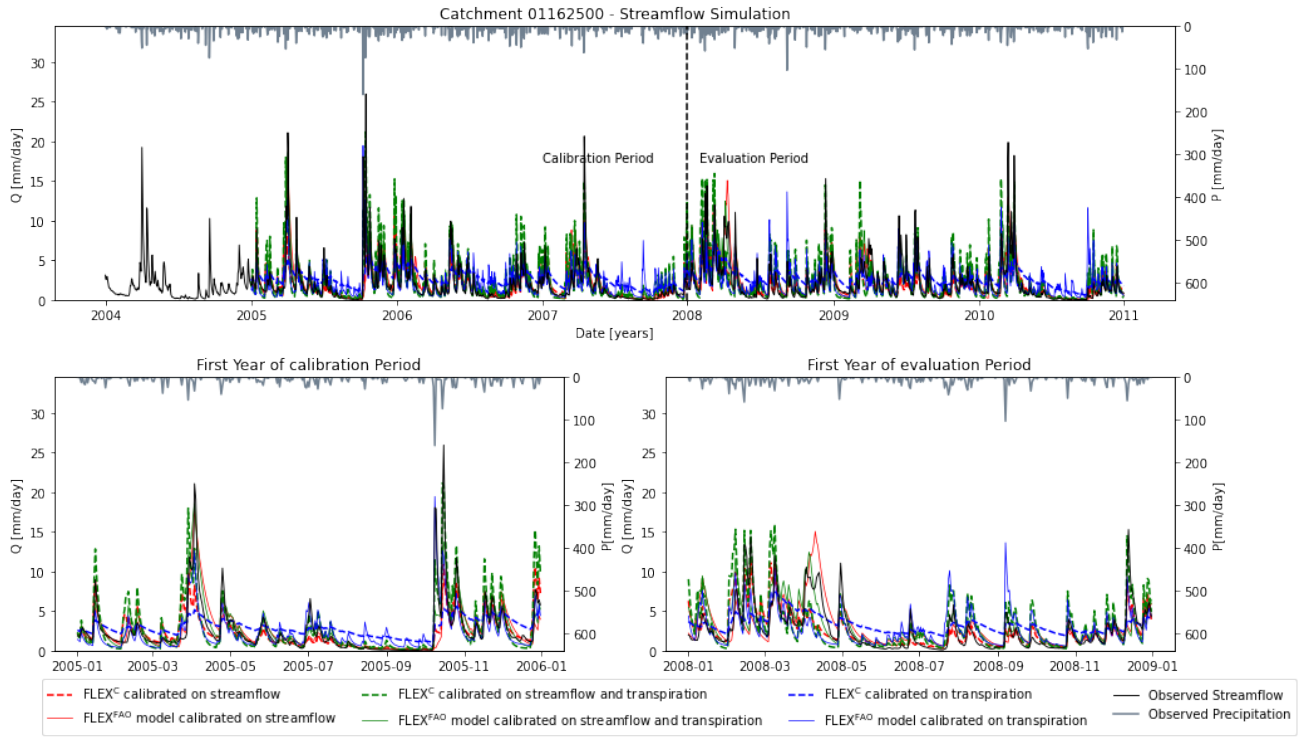


Figure 3.5: Catchment C25 streamflow simulation. Landcover dominated by mixed forest

3.2.2. Transpiration

We also evaluated the transpiration simulations in this study. The results of the transpiration simulations give a more consistent image of the impact of different calibration strategies and the FAO adjustment. The transpiration simulations of the same four catchments as in Chapter 3.2.1 are presented in Figures 3.6, 3.7, 3.8 and 3.9. The results are resampled to a 7 day mean so that the seasonal pattern becomes more clear and the results are less 'noisy'. The transpiration simulations of the other catchments are presented in Appendix C. In the figures the GLDAS21 transpiration data is presented by the black line. The red lines represent the $FLEX^{C}$ models where the linestyles distinguish between different calibration strategies. The blue lines show the $FLEX^{FAO}$ models also with different linestyles for the three calibration strategies. These colors show that for all three calibration strategies the $FLEX^{FAO}$ models present a much better simulation of the transpiration. Especially in the winter and early spring, when the transpiration is increasing mostly due the growing potential evaporation in this period, the $FLEX^{C}$ models overestimate the transpiration. What also can be noticed in the figures is that $FLEX^{C}$ calibrated on streamflow gives a transpiration that drops to zero occasionally in the summer months, when the GLDAS21 transpiration is peaking. However the $FLEX^{FAO}$ models have this behaviour to a much lesser extend, and follow the GLDAS21 data more accurately, even though this data was not used in the calibration.

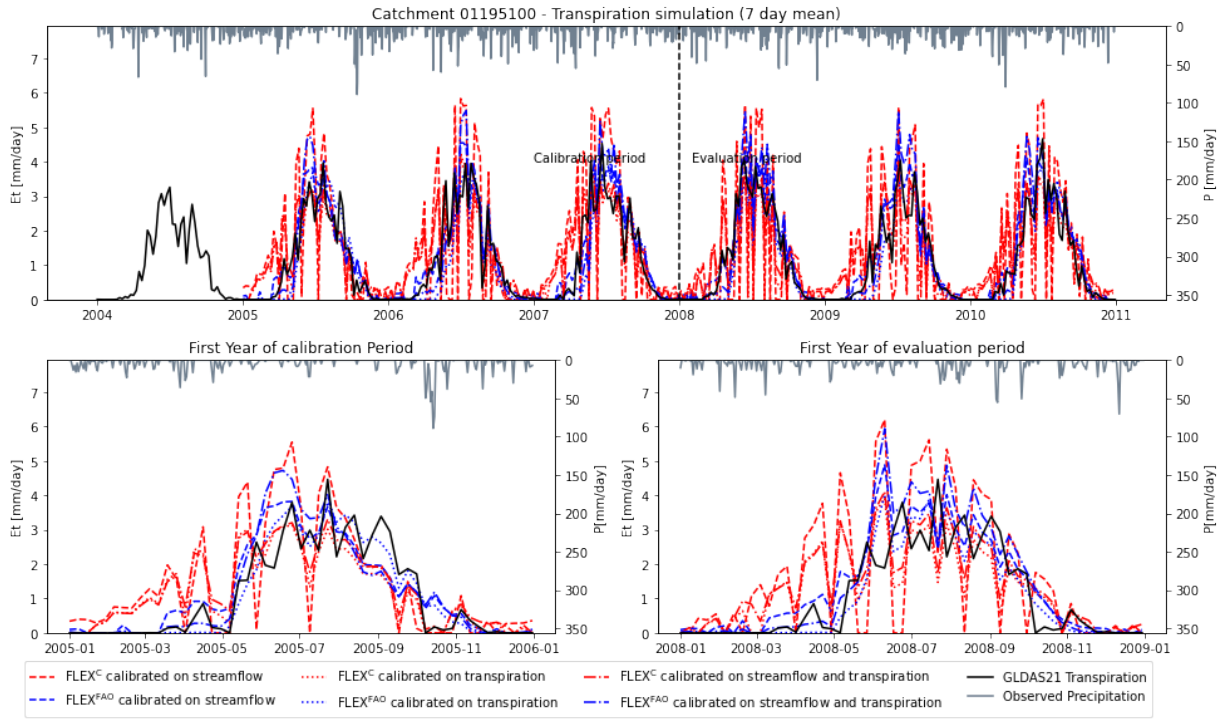


Figure 3.6: Catchment C2 transpiration simulation. Landcover dominated by deciduous broadleaf forest

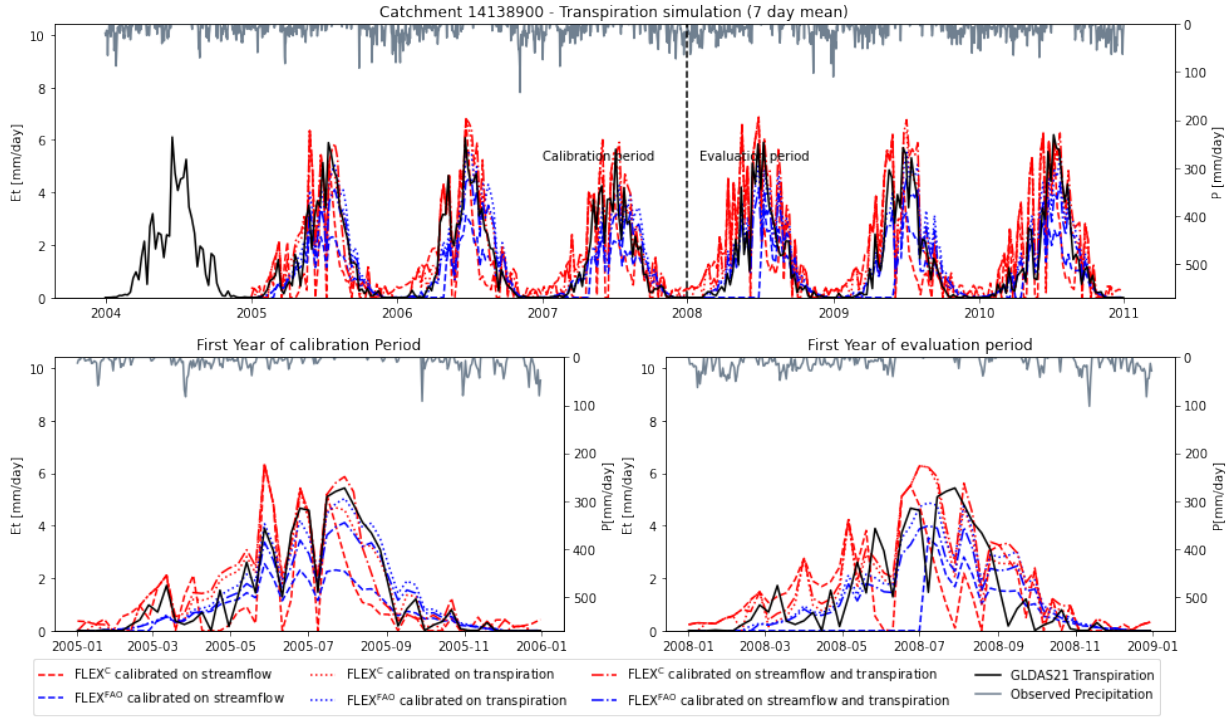


Figure 3.7: Catchment C9 transpiration simulation. Landcover dominated by evergreen needleleaf forest

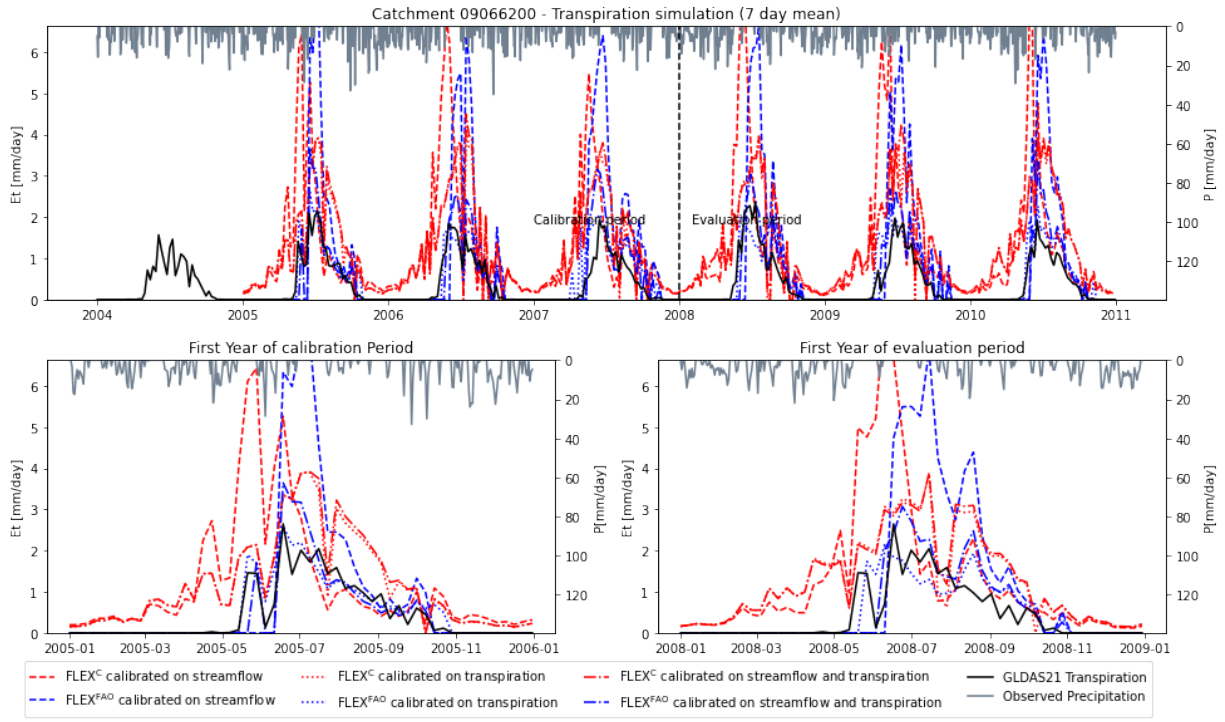


Figure 3.8: Catchment C9 transpiration simulation. Landcover dominated by grassland

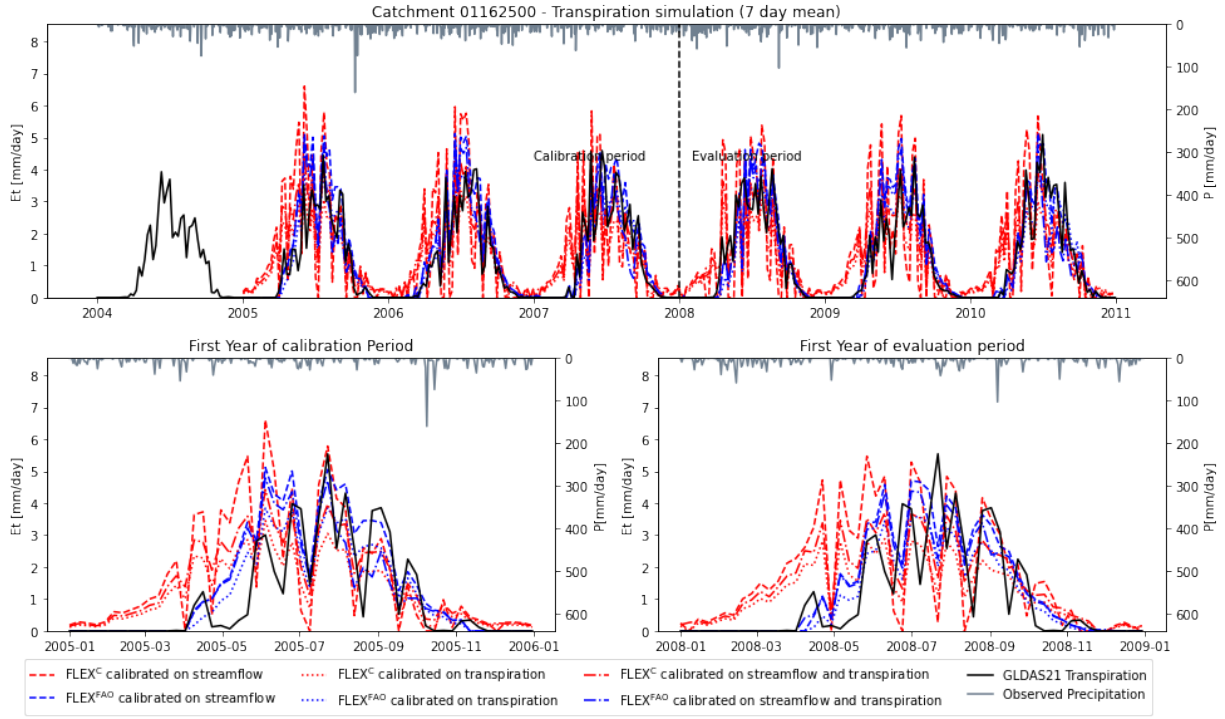


Figure 3.9: Catchment C9 transpiration simulation. Landcover dominated by mixed forest

3.3. Model performance when calibrated on streamflow

In this research we used three different calibration strategies. FLEX^{C} and FLEX^{FAO} have been calibrated on streamflow, on transpiration and on both streamflow and transpiration sep-

arately. In this subsection the results are presented for $FLEX^{C}$ and $FLEX^{FAO}$ that have been calibrated on streamflow only. The results are separated into the streamflow simulation and the transpiration simulation. First the performance of the models is shown for the streamflow simulations of all catchments and then the performance of the transpiration simulations of all catchments.

3.3.1. Streamflow

The Nash-Sutcliffe (NS) efficiencies and logarithmic Nash-Sutcliffe (NSlog) efficiencies of the streamflow simulations of $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on streamflow only are presented in the boxplots in Figure 3.10. The performance of the model is shown in terms of the ability to simulate both low flows and high flows, and on the ability to predict future fluxes in the model. The NS-efficiency emphasizes more on high flows and the NSlog-efficiency emphasizes more on low flows, and the ability to predict future fluxes in the model is presented by showing the objective functions for the calibration period and the evaluation period separately. In accordance with this, Figure 3.10 shows four graphs. The top figure shows the NS-efficiency in the calibration period. The second graph shows the NS-efficiency in the evaluation period. The third figure shows the NSlog-efficiency in the calibration period and finally the bottom figure shows the NSlog-efficiency in the evaluation period. For all catchment, which are shown by the x-axes, $FLEX^{C}$ and $FLEX^{FAO}$ are presented next to each other. The colors indicate the dominant vegetation type in the catchments.

Figure 3.10 shows that for most catchments the performance of the models is reasonably good. The y-axes only show the boxplots with a objective function above -0.4. We did this to put more emphasis on the well performing catchments, with the disadvantage that some of the catchment with very low scores are ignored. Catchments score low on some or all parts of the performance and are therefore ignored are C11, C12 and C17. Furthermore the graphs show that for all catchments the NSlog-efficiency is very similar or higher than the NS-efficiency, which means both $FLEX^{C}$ and $FLEX^{FAO}$ are most times better in simulating the low flows than the high flows. Also most catchments show similar values for the objective functions in the calibration period and the objective functions in the evaluation period, which means that in general $FLEX^{C}$ and $FLEX^{FAO}$ have a reasonable predictive capacity, where there is not a noticeable difference in performance between $FLEX^{C}$ and $FLEX^{FAO}$.

The impact of the FAO adjustment does vary. For the catchments with deciduous broadleaf forest and mixed forest $FLEX^{C}$ and $FLEX^{FAO}$ give similar results or even a decrease in performance. For catchments with evergreen needleleaf forest increase of decrease varies a lot between the catchment, with some catchments improving in performance and some catchments worsening in performance. The catchments with grassland are improving in NS-efficiency, however the NSlog-efficiency is staying similar or even decreasing in some cases. The performance comparison is shown in more detail in Figure 3.11 and 3.12.

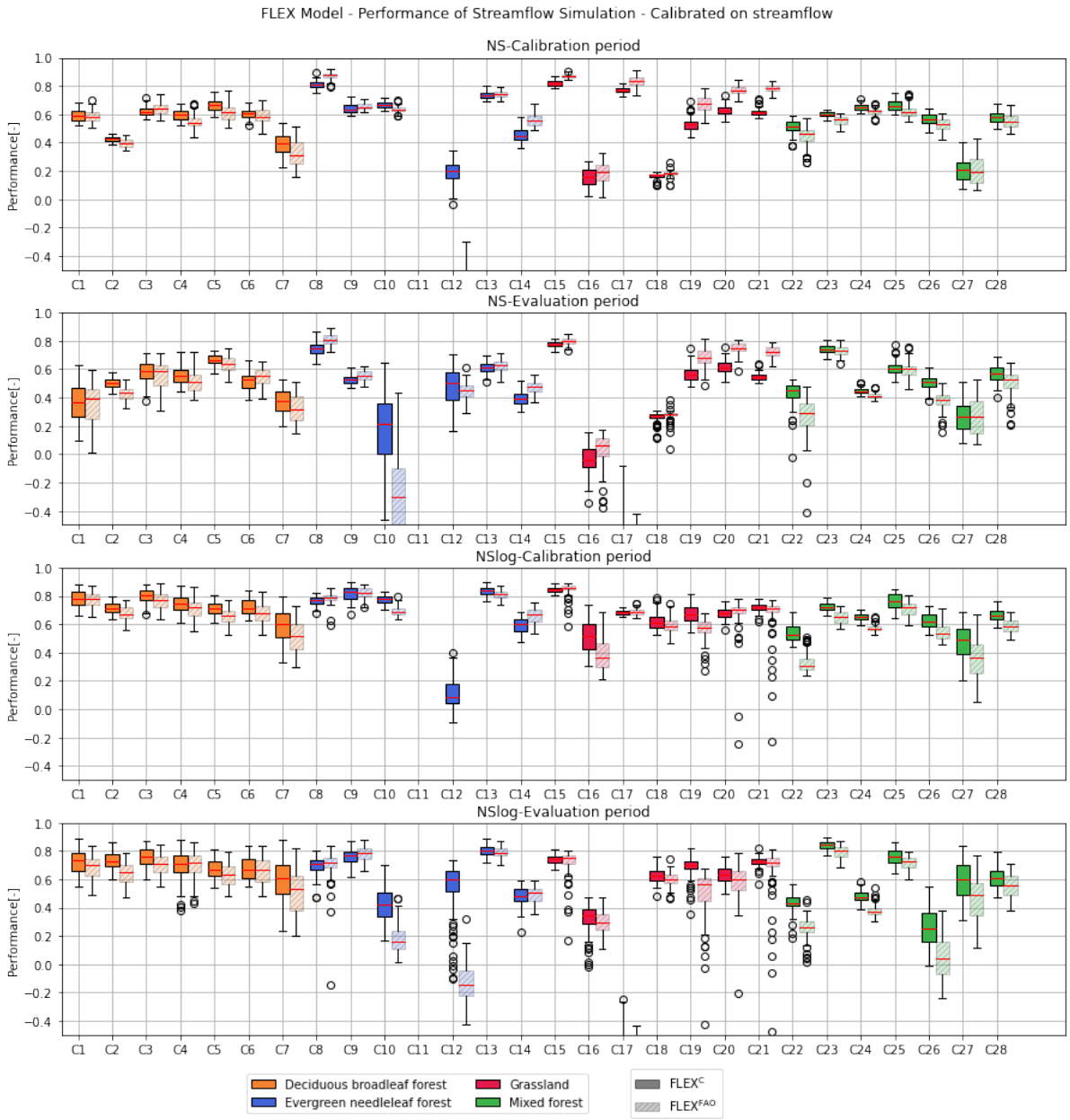


Figure 3.10: Performance of the streamflow simulation of $FLEX^C$ and $FLEX^{FAO}$ when calibrated on streamflow, presented next to each other for every catchment. The four graphs separate the performance into the NS and NSlog objective function and the performance in the calibration period and the evaluation period. The colors show the dominant vegetation type in all the catchments. The x-axes show all the catchments that have been simulated.

Figure 3.11 shows how much the NS-efficiency and the NSlog-efficiency increased or decreased after the FAO adjustment. For every catchment the ΔNS and ΔNSlog are presented next to each other. The colors indicate the dominant vegetation type.

The figure shows that for most catchments the efficiencies stay the same or decrease slightly. For catchments with mixed forest is visible that the NSlog-efficiency decreases more than the NS-efficiency, with between 0.04 and 0.22 for the NSlog-efficiency and between 0.01 and 0.05 for the NS-efficiency. The NS-efficiency of the catchments with grassland does increase structurally between 0.01 and 0.16, while this increase is not there for the NSlog-efficiency for

the same catchments.

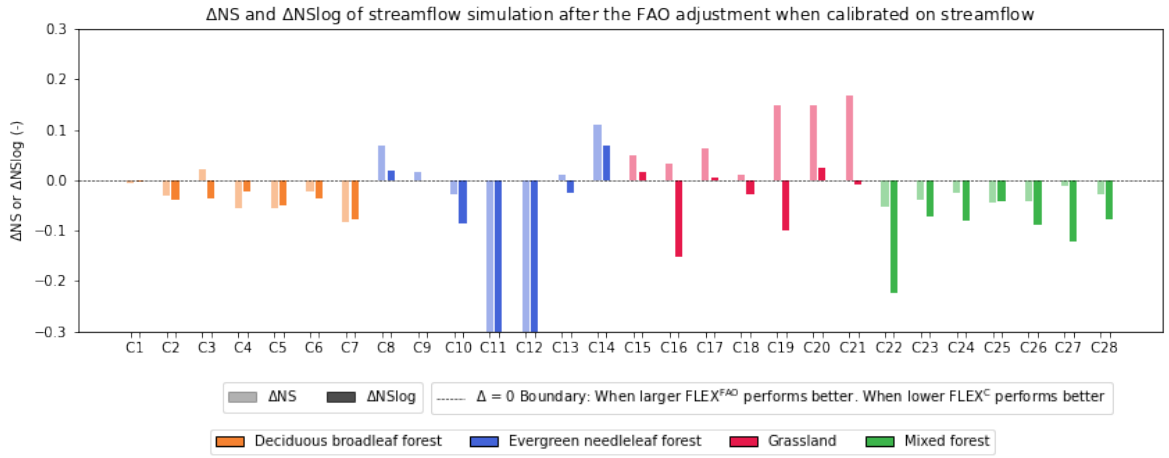


Figure 3.11: Performance comparison of the streamflow simulation between $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on streamflow. The figure shows the increase or decrease of the NS-efficiency and the NSlog-efficiency after the FAO adjustment for every catchment respectively. The colors indicate the dominant vegetation type.

To be able to show a potential relationship between the impact of the FAO adjustment and some key climate indices, ΔNS and ΔNSlog are also plotted against some climate indices in Figure 3.12. In Figure 3.12 the performance of streamflow simulation $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on streamflow is compared. The figures show how much the performance is increased or decreased after the FAO adjustment. This increase or decrease is presented with respect to three different climate indices. The first two figures show how the performance comparison relates to the maximum difference in LAI throughout the year. The middle two figures show the performance comparison in relationship with the aridity and the final figures in relationship with seasonality. The top row shows the NS-efficiency and the bottom row the NSlog-efficiency. The dot marker and the cross marker represent the performance in the calibration period and the evaluation period respectively. The colors indicate the dominant vegetation type in the catchments.

In the first two figures is visible how the catchments with deciduous broadleaf forest and mixed forest have a higher difference in LAI than the catchments with grassland and evergreen needleleaf forest. This is expected due to the fact evergreen needleleaf forests and grassland is not impacted by phenology as much. For ΔNS the improvement by the FAO adjustment seems higher for catchments with low LAI differences throughout the year. Both the aridity and seasonality are not clearly related to the vegetation types in the catchments. The figures confirm that catchments with grassland are improving more than other dominant vegetation types. Furthermore the figures don't show any clear relationships between the climate indices and the NS- or NSlog-efficiencies.

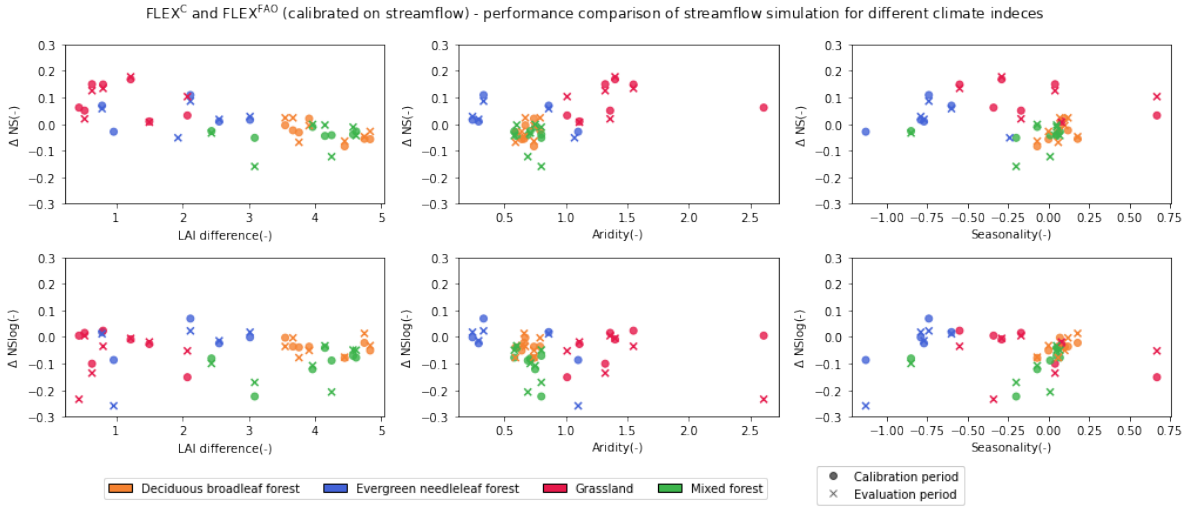


Figure 3.12: Performance comparison between FLEX^{C} and FLEX^{FAO}. The top figures show the comparison of the NS-efficiency and the bottom figures show the comparison of the NSlog-efficiency. The y-axes show how much the NS/NSlog-efficiency is increased or decreased after the FAO adjustment. The dot marker shows the performance in the calibration period and the cross marker the performance in the evaluation period. The colors indicate the dominant vegetation type in the catchment. The x-axes show different climate indices.

In Figure 3.13 the mean seasonal sums of the streamflow for the observed streamflow, FLEX^{C} and FLEX^{FAO} are presented for all catchment. All four graphs show one season, where the top graph shows the spring, second graph the summer, third graph the fall and the bottom graph the winter. The colors show the dominant vegetation type of the catchments.

The figure shows that in the spring the streamflow is slightly overestimated by FLEX^{FAO} between 2 and 51 mm/day, and slightly underestimated by FLEX^{C} between 20 and 68 mm/day for all catchments with deciduous broadleaf forest and some of the catchments grassland and mixed forest. For catchments with evergreen needleleaf forest FLEX^{FAO} is also overestimating the streamflow in spring with between 70 and 113 mm/day, however this is not for all catchments, FLEX^{C} is relatively closely matching the observed streamflow. In summer it can be noticed that FLEX^{C} and FLEX^{FAO} are underestimating the streamflow for the well performing catchments C15, C19, C20 and C21 with grassland, where FLEX^{FAO} is closer to the observed streamflow than FLEX^{C}. For the other seasons there are no clear patterns visible in the over- or underestimation by the streamflow simulation of FLEX^{C} and FLEX^{FAO} in relationship to their dominant vegetation type.

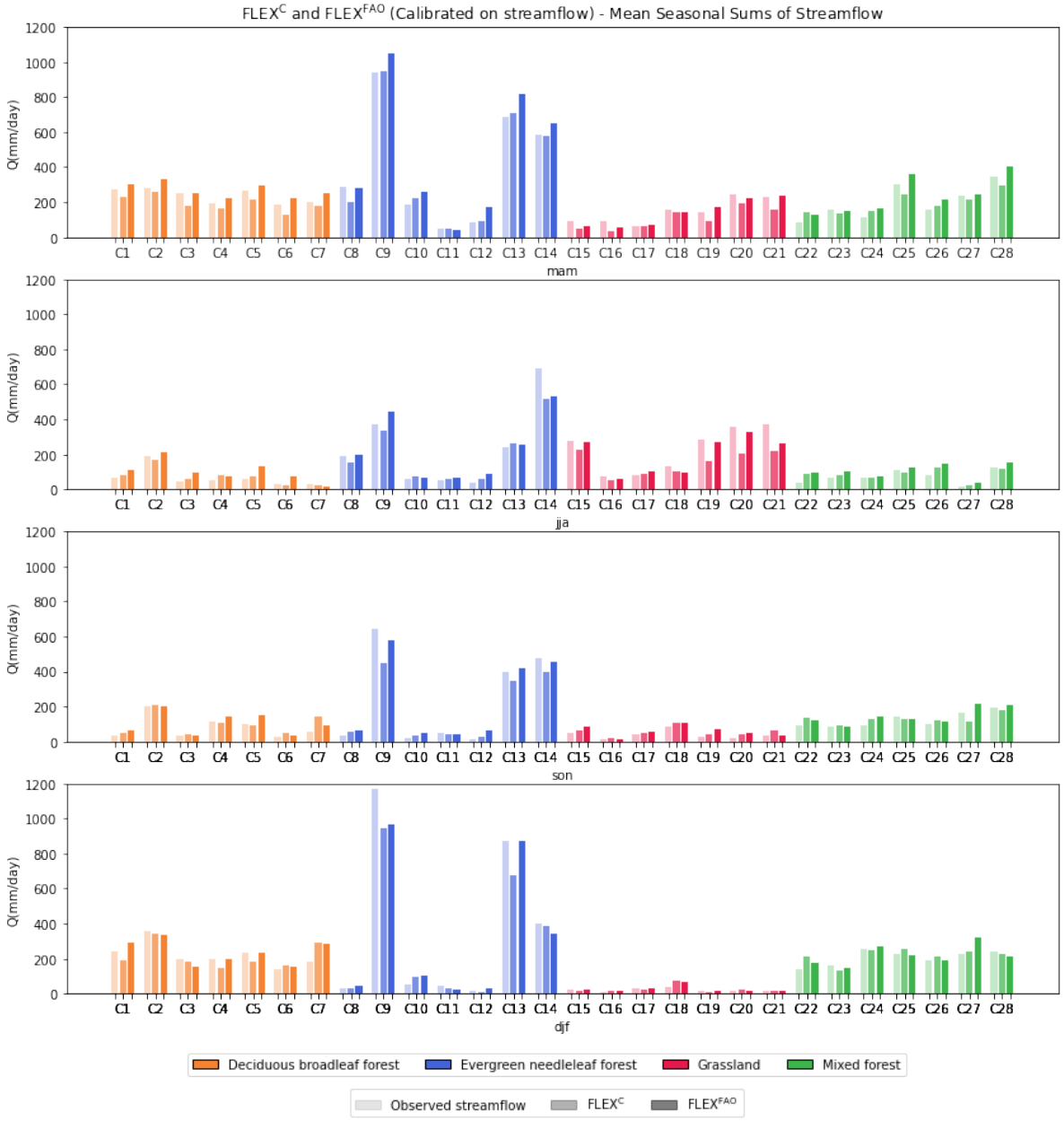


Figure 3.13: Mean seasonal sums of the streamflow for the observed streamflow, FLEX^{C} and FLEX^{FAO}. The seasons are indicated by mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. The colors show the dominant vegetation type in the catchments, which are indicated on the x-axes. For every catchment, the observed streamflow, FLEX^{C} and FLEX^{FAO} are presented in that order.

3.3.2. Transpiration

The Nash-Sutcliffe (NS) efficiencies and logarithmic Nash-Sutcliffe (NSlog) efficiencies of $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on streamflow are presented in the boxplots in Figure 3.14. The performance of the model is shown in terms of the ability to simulate both low flows and high flows, and on the ability to predict future fluxes in the model. The NS-efficiency focuses more on high flows and the NSlog-efficiency focuses more on low flows, and the ability to predict future fluxes in the model is presented by showing the objective functions for the calibration period and the evaluation period separately. In accordance with this, Figure 3.14 shows four graphs. The top figure shows the NS-efficiency in the calibration period. The second graph shows the NS-efficiency in the evaluation period. The third figure shows the NSlog-efficiency in the calibration period and finally the bottom figure shows the NSlog-efficiency in the evaluation period. For all catchment, which are shown by the x-axis, $FLEX^{C}$ and $FLEX^{FAO}$ are presented next to each other. The colors indicate the dominant vegetation type in the catchments.

The y-axes show a range between -2 and 1, which is a large range for objective functions. Even with this large range some catchments still lie outside the range. We chose this range so that the relatively well performing catchments, which are most relevant for analysis are also still visualized. Catchments that have efficiencies even below 2 can be interpreted as being very disbehavioural, which are catchments C15, C17, C19, C20 and C21 in the NS-efficiency and catchments C18, C26 and C27 in the NSlog efficiency of the $FLEX^{C}$ simulation. With this observation it must be noted that the efficiencies are calculated with the simulated transpiration in combination with the GLDAS-21 transpiration, while this data is not used in the model calibration or in any way in the model. However, even with this noted, the figure shows that for almost all catchments especially the NSlog-efficiencies of $FLEX^{C}$ are very low with values even below zero. However in this figure it is already visible that the performance drastically improves after the FAO adjustment for almost all catchments, regardless of their dominant vegetation type and even when $FLEX^{C}$ was performing very badly. It can be seen that $FLEX^{C}$ simulates the NS-efficiency better for catchments with deciduous broadleaf forest and mixed forests in comparison with the NSlog-efficiency, while for grassland it is the other way around. For catchments with evergreen needleleaf forest it varies. The efficiencies in the calibration period and the evaluation period perform very similar for all catchments, which confirms that the models both have a good predictive capacity.

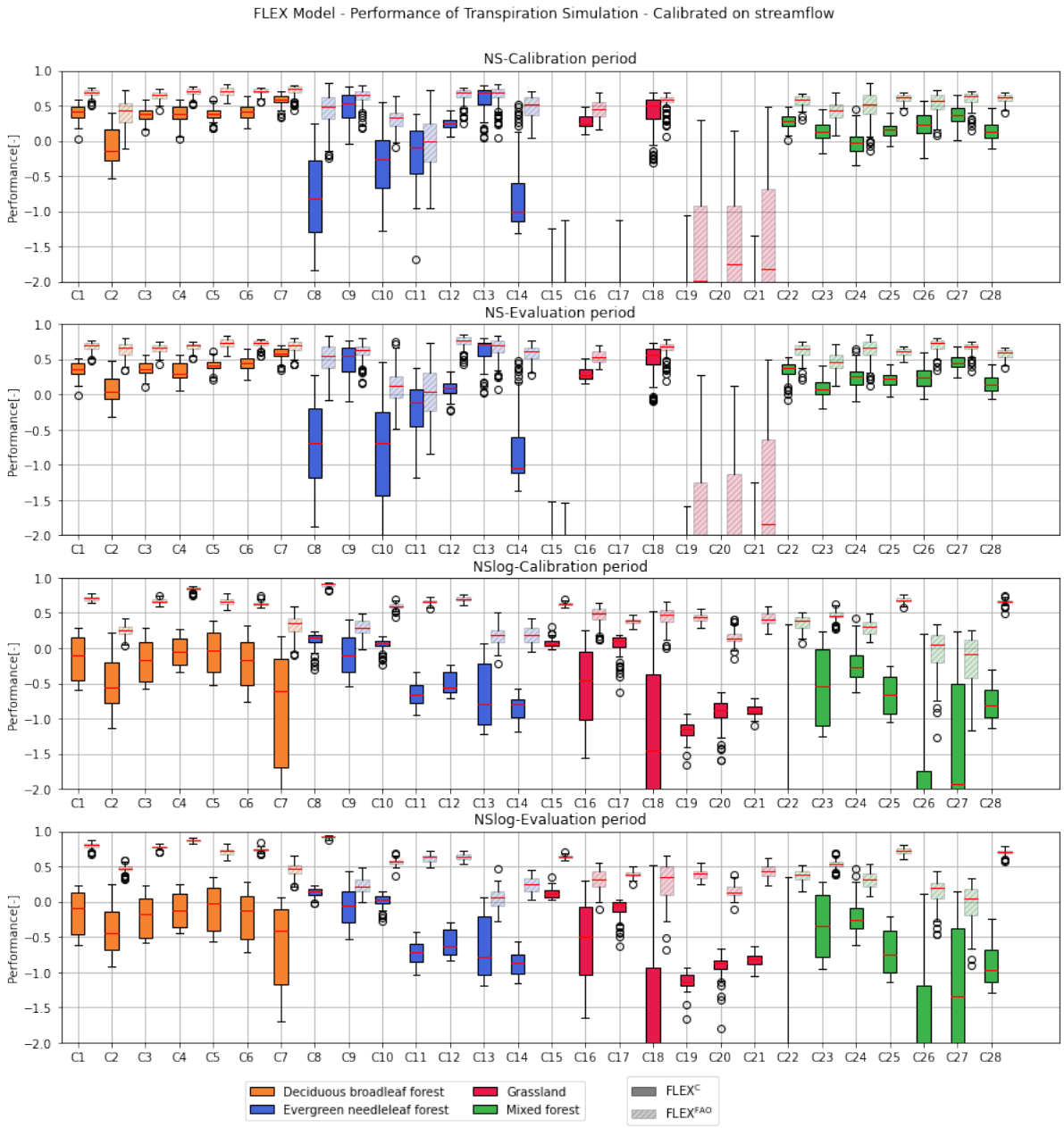


Figure 3.14: Performance of the transpiration simulation of $FLEX^C$ and $FLEX^{FAO}$ when calibrated on streamflow, presented next to each other for every catchment. The four graphs separate the performance into the NS and NSlog objective function and the performance in the calibration period and the evaluation period. The colors show the dominant vegetation type in all the catchments. The x-axes show all the catchments that have been simulated.

Figure 3.15 shows how much the NS-efficiency and the NSlog-efficiency increased or decreased after the FAO adjustment. For every catchment the ΔNS and ΔNSlog are presented next to each other. The colors indicate the dominant vegetation type.

The figure shows that for most catchments the efficiencies increase drastically. What stands out is that the improvement of the NSlog-efficiency is higher than the improvement of the NS-efficiency for catchments with deciduous broadleaf forest and mixed forest, while this is not necessarily the case for catchments with evergreen needleleaf forest or grassland. For the catchments with deciduous broadleaf forest the increase is between 0.15 and 0.60 for NS-

efficiency and between 0.70 and 0.90 for the NSlog-efficiency. For catchments with mixed forest the increase is between 0.30 and 0.50 for NS-efficiency and between 1.6 and 5.0 for the NSlog-efficiency. An important sidenote to this graph is that the catchment with the largest improvement are also catchment with very low efficiencies by FLEX^{C}. This in combination with FLEX^{FAO} giving similar performances as the other catchments, makes a large improvement.

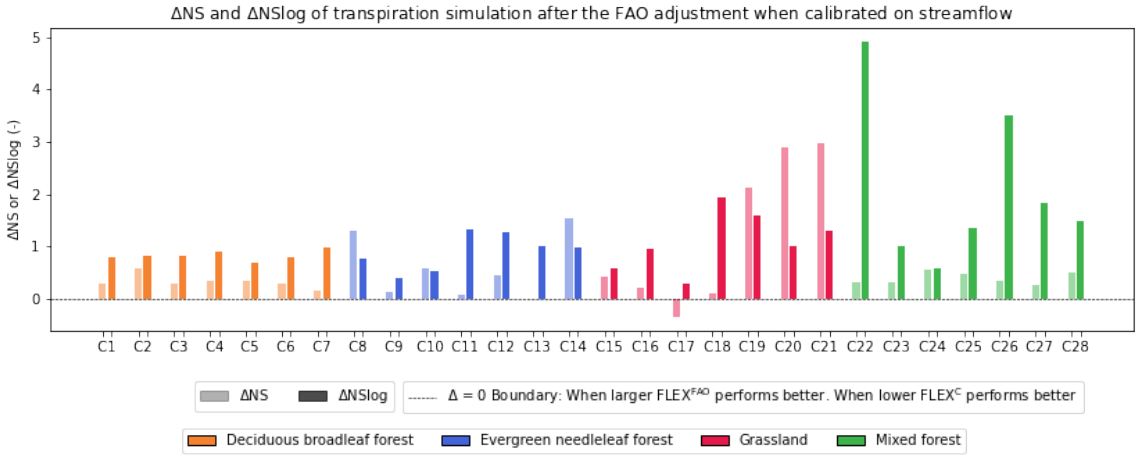


Figure 3.15: Performance comparison of the transpiration simulation between FLEX^{C} and FLEX^{FAO} calibrated on streamflow. The figure shows the increase or decrease of the NS-efficiency and the NSlog-efficiency after the FAO adjustment for every catchment respectively. The colors indicate the dominant vegetation type.

In Figure 3.16 the performance of the transpiration simulation FLEX^{C} and FLEX^{FAO} calibrated on streamflow is compared. The figures show how much the performance is increased or decreased after the FAO adjustment. This increase or decrease is presented with respect to three different climate indices. The first two figures show how the performance comparison relates to the maximum difference in LAI throughout the year. The middle two figures show the performance comparison in relationship with the aridity and the final figures in relationship with seasonality. The top row shows the NS-efficiency and the bottom row the NSlog-efficiency. The dot marker and the cross marker represent the performance in the calibration period and the evaluation period respectively. The colors show the dominant vegetation type in the catchments.

In the first two figures is visible how the catchments with deciduous broadleaf forest and mixed forest have a higher difference in LAI than the catchments with grassland and evergreen needleleaf forest. This is expected due to the fact evergreen needleleaf forests and grassland is not impacted by phenology as much. The both the aridity and the seasonality are not clearly related to the vegetation types in the catchments. The figures confirm that the NS-efficiency and the NSlog-efficiency greatly improves after the FAO adjustment. The first most left two figures show that the NS-efficiency improves more for catchments with low LAI difference and the NSlog-efficiency improves more for all catchments approximately the same amount, with some outliers at the catchments with mixed forest and high LAI difference. In the most right two figures there is not a clear relation visible between seasonality and the improvement of the performance after the FAO adjustment.

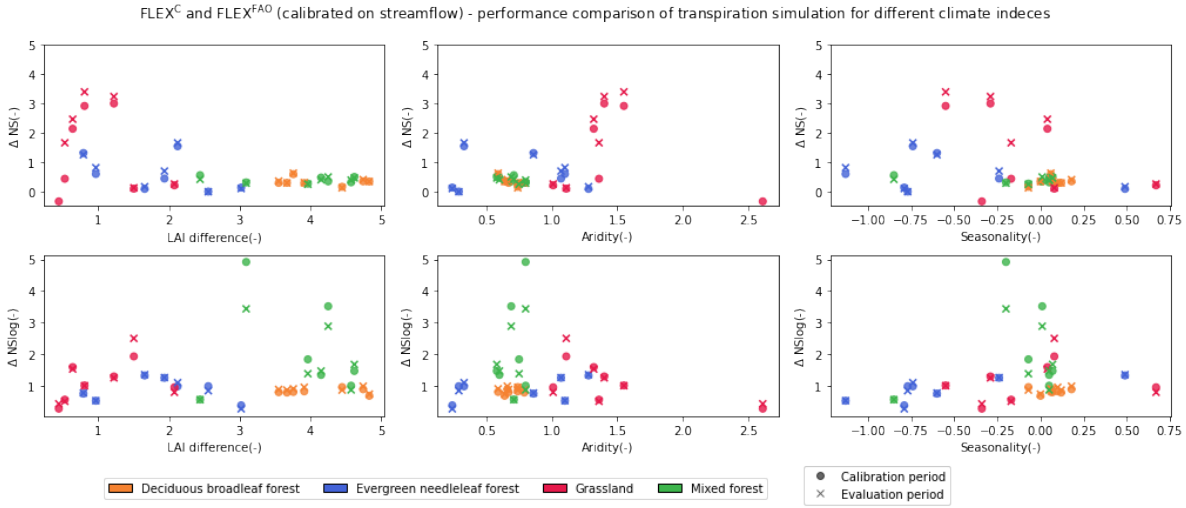


Figure 3.16: Performance comparison between FLEX^{C} and FLEX^{FAO}. The top figures show the comparison of the NS-efficiency and the bottom figures show the comparison of the NSlog-efficiency. The y-axes show how much the NS/NSlog-efficiency is increased or decreased after the FAO adjustment. The dot marker shows the performance in the calibration period and the cross marker the performance in the evaluation period. The colors indicate the dominant vegetation type in the catchment. The x-axes show different climate indices.

In Figure 3.17 the mean seasonal sums of the transpiration for the GLDAS21-transpiration, FLEX^{C} and FLEX^{FAO} are presented of all catchment. All four graphs show one season, where the top graph shows the spring, second graph the summer, third graph the fall and the bottom graph the winter. The colors show the dominant vegetation type of the catchments.

The figure shows that in the spring the mean seasonal sum of the transpiration is greatly overestimated by FLEX^{C} for all catchments. For catchments with deciduous broadleaf forest this overestimation varies between 29 and 198 mm/day, while for FLEX^{FAO} it varies between an underestimation of 48 mm/day and an overestimation of 55 mm/day. For catchments with evergreen needleleaf forest the overestimation of FLEX^{C} varies between 55 and 215 mm/day, while for FLEX^{FAO} it varies between an underestimation of 80 mm/day and an overestimation of 83 mm/day. For catchments with grassland the overestimation of FLEX^{C} varies between 7 and 180 mm/day, while for FLEX^{FAO} it varies between an underestimation of 8 mm/day and an overestimation of 69 mm/day. Finally for catchments with mixed forest the mean seasonal sum of the transpiration of FLEX^{C} compared to the GLDAS21-transpiration varies between an underestimation of 8 mm/day and an overestimation 137 mm/day, while for FLEX^{FAO} it varies between an overestimation of 3 mm/day and an overestimation of 88 mm/day. In the other seasons the mean seasonal sum of the transpiration simulated by FLEX^{C} is much closer to the GLDAS21-transpiration. In the winter months FLEX^{C} also overestimates the the mean seasonal sum of the transpiration. This overestimation is small when compared to the other seasons, however relatively large compared to the GLDAS-21 transpiration data.

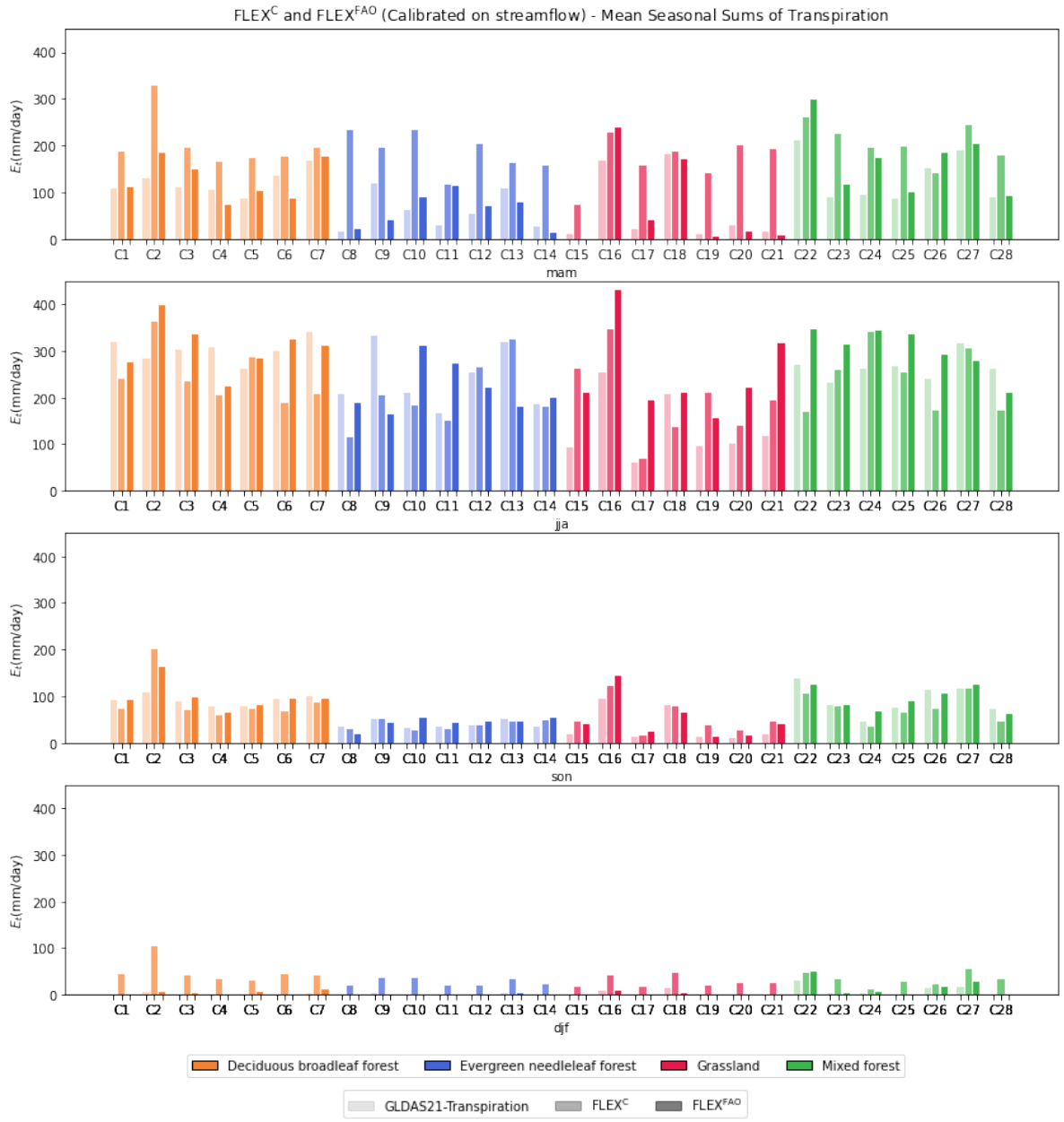


Figure 3.17: Mean seasonal sums of the transpiration for the GLDAS21-transpiration, FLEX^{C} and FLEX^{FAO}. The seasons are indicated by mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. The colors show the dominant vegetation type in the catchments, which are indicated on the x-axes. For every catchment, the GLDAS21-transpiration, FLEX^{C}, and FLEX^{FAO} are presented in that order.

3.4. Model performance when calibrated on transpiration

In this subsection the results will be presented for FLEX^{C} and FLEX^{FAO} that are calibrated on transpiration only. The results are separated into the streamflow simulation and the transpiration simulation. Because the overall performance of the streamflow simulations when the models are calibrated on transpiration only is very low logically, these results are not so relevant for the purposes of this research. Therefore these results are not shown in this chapter. For the reader that is still interested the results are moved to Appendix D. The results of the performance of the transpiration simulation is presented in this chapter.

3.4.1. Transpiration

The Nash-Sutcliffe (NS) efficiencies and logarithmic Nash-Sutcliffe (NSlog) efficiencies of $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on transpiration are presented in the boxplots in Figure 3.18. The performance of the model is shown in terms of the ability to simulate both low flows and high flows, and on the ability to predict future fluxes in the model. The NS-efficiency focuses more on high flows and the NSlog-efficiency focuses more on low flows, and the ability to predict future fluxes in the model is presented by showing the objective functions for the calibration period and the evaluation period separately. In accordance with this, Figure 3.18 shows four graphs. The top figure shows the NS-efficiency in the calibration period. The second graph shows the NS-efficiency in the evaluation period. The third figure shows the NSlog-efficiency in the calibration period and finally the bottom figure shows the NSlog-efficiency in the evaluation period. For all catchment, which are shown by the x-axes, $FLEX^{C}$ and $FLEX^{FAO}$ are presented next to each other. The colors indicate the dominant vegetation type in the catchments.

The y-axes only show the boxplots with a objective function above -0.4. We did this to put more emphasis on the well performing catchments, with the disadvantage that some of the catchment where $FLEX^{C}$ has low efficiencies are ignored. These catchments are C15, C19, C20 and C21. With this observation it must be noted that the efficiencies are only low for the $FLEX^{C}$. $FLEX^{FAO}$ however gives much higher performances, which confirms an improvement in performance of the FAO adjustment. It can be seen that $FLEX^{C}$ simulates the NS-efficiency better for all catchments in comparison with the NSlog-efficiency. The efficiencies in the calibration period and the evaluation period perform very similar for all catchments, which confirms that the models both have a good predictive capacity. What also stands out is that the range of the boxplots is very small. The performance comparison is shown in more detail in Figures 3.19 and 3.20.

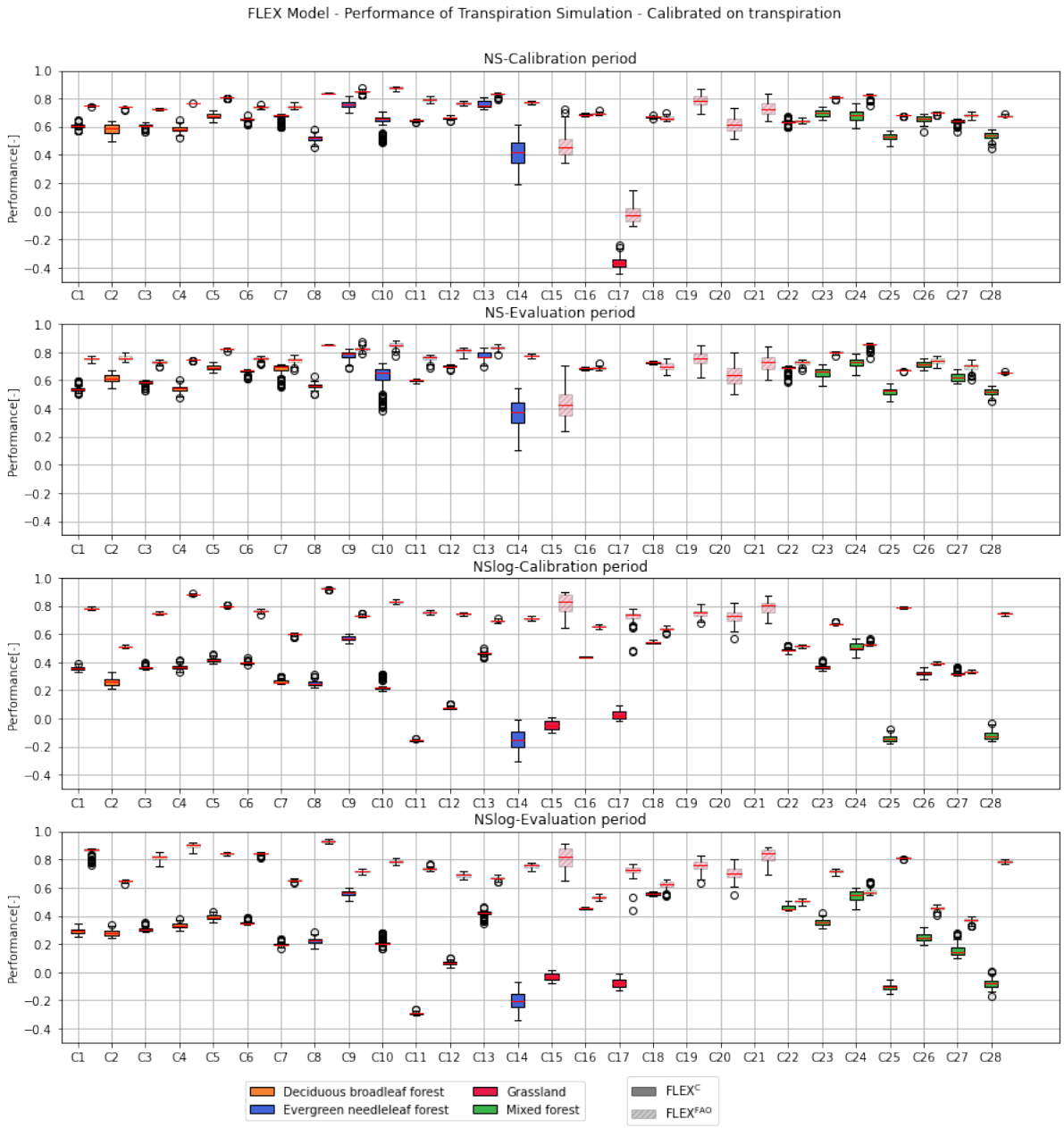


Figure 3.18: Performance of the transpiration simulation of $FLEX^C$ and $FLEX^{FAO}$ when calibrated on transpiration, presented next to each other for every catchment. The four graphs separate the performance into the NS and NSlog objective function and the performance in the calibration period and the evaluation period. The colors show the dominant vegetation type in all the catchments. The x-axes show all the catchments that have been simulated.

Figure 3.19 shows how much the NS-efficiency and the NSlog-efficiency increased or decreased after the FAO adjustment. For every catchment, the ΔNS and ΔNSlog are presented next to each other. The colors indicate the dominant vegetation type.

The figure shows that for most catchments the efficiencies increase drastically. What stands out is that the improvement of the NSlog-efficiency is higher than the improvement of the NS-efficiency for catchments with deciduous broadleaf forest and evergreen needleleaf forest, while this is not necessarily the case for catchments with mixed forest or grassland. For the catchments with deciduous broadleaf forest, the increase is between 0.06 and 0.18 for the NS-

efficiency and between 0.25 and 0.51 for the NSlog-efficiency. For catchments with evergreen needleleaf forest, the increase is between 0.09 and 0.35 for the NS-efficiency and between 0.16 and 0.91 for the NSlog-efficiency. For catchments with grassland, the increase is between 0.0 and 2.12 for the NS-efficiency and between 0.22 and 1.78 for the NSlog-efficiency. For catchments with mixed forest, the increase is between 0.01 and 0.15 for the NS-efficiency and between 0.03 and 0.93 for the NSlog-efficiency. An important sidenote to this graph is that the catchments with the largest improvement are also catchments with very low efficiencies by $FLEX^C$. This in combination with $FLEX^{FAO}$ giving similar performances as the other catchments, makes a large improvement.

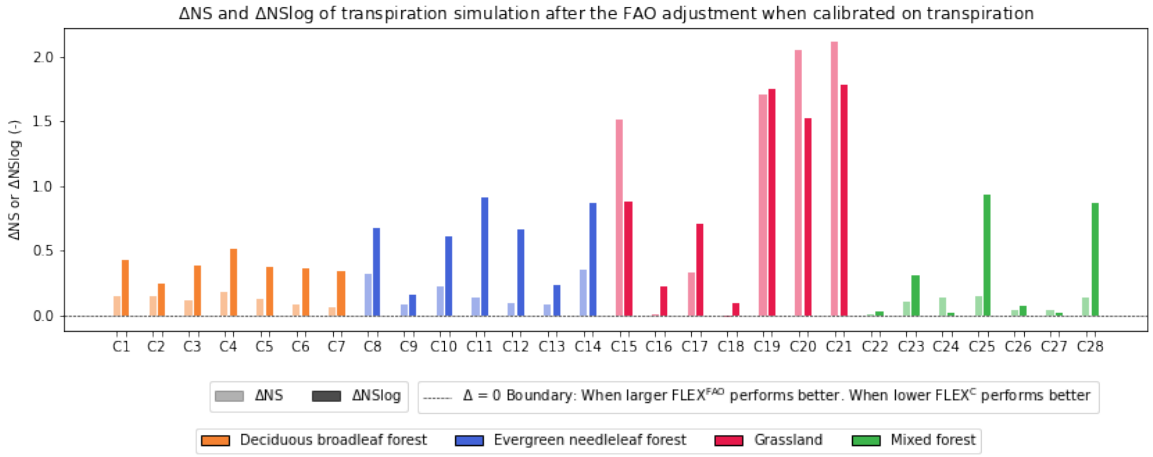


Figure 3.19: Performance comparison of the transpiration simulation between $FLEX^C$ and $FLEX^{FAO}$ calibrated on transpiration. The figure shows the increase or decrease of the NS-efficiency and the NSlog-efficiency after the FAO adjustment for every catchment respectively. The colors indicate the dominant vegetation type.

In Figure 3.20 the performance of the transpiration simulation $FLEX^C$ and $FLEX^{FAO}$ calibrated on transpiration is compared. The figures show how much the performance is increased or decreased after the FAO adjustment. This increase or decrease is presented with respect to three different climate indices. The first two figures show how the performance comparison relates to the maximum difference in LAI throughout the year. The middle two figures show the performance comparison in relationship with the aridity and the final figures in relationship with seasonality. The top row shows the NS-efficiency and the bottom row the NSlog-efficiency. The dot marker and the cross marker represent the performance in the calibration period and the evaluation period respectively. The colors show the dominant vegetation type in the catchments.

The figures confirm that the NS-efficiency and the NSlog-efficiency greatly improve after the FAO adjustment. The first two figures show that the efficiencies improve more for catchments with low LAI difference. Catchments with a relatively high aridity have high values for ΔNS and ΔNSlog, however this is only in the case of grassland, and not for catchments with similar values for aridity but other dominant vegetation types. Also in this figure, the observation that the catchments with a high ΔNS or ΔNSlog are also the catchments with a low overall performance by the $FLEX^C$ needs to be taken into account.

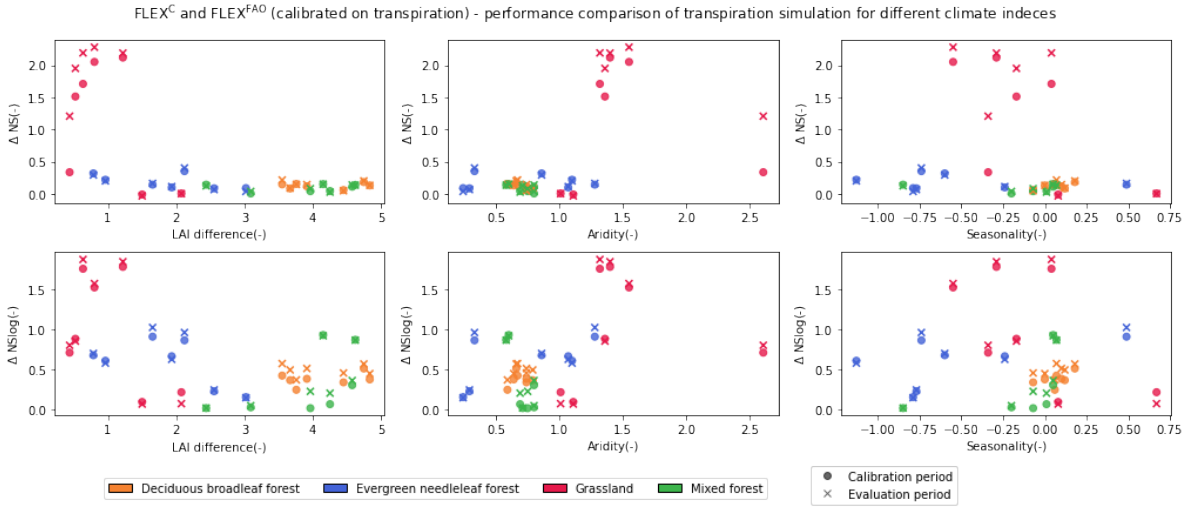


Figure 3.20: Performance comparison between FLEX^{C} and FLEX^{FAO}. The top figures show the comparison of the NS-efficiency and the bottom figures show the comparison of the NSlog-efficiency. The y-axes show how much the NS/NSlog-efficiency is increased or decreased after the FAO adjustment. The dot marker shows the performance in the calibration period and the cross marker the performance in the evaluation period. The colors indicate the dominant vegetation type in the catchment. The x-axes show different climate indices.

In Figure 3.21 the mean seasonal sums of the transpiration for the GLDAS21-transpiration, FLEX^{C} and FLEX^{FAO} are presented of all catchments. All four graphs show one season, where the top graph shows the spring, the second graph the summer, the third graph the fall, and the bottom graph the winter. The colors show the dominant vegetation type of the catchments.

The figure shows that in the spring the mean seasonal sum of the transpiration is greatly overestimated by FLEX^{C} for all catchments. For catchments with deciduous broadleaf forest, this overestimation varies between 6 and 103 mm/day, while for FLEX^{FAO} it varies between an underestimation of 54 mm/day and an overestimation of 8 mm/day. For catchments with evergreen needleleaf forest, the overestimation of FLEX^{C} varies between 61 and 89 mm/day, while for FLEX^{FAO} it varies between an underestimation of 8 mm/day and an overestimation of 24 mm/day. For catchments with grassland, it varies between an underestimation of 17 and an overestimation of 64 mm/day for FLEX^{C}, while for FLEX^{FAO} it varies between an underestimation of 19 mm/day and an overestimation of 35 mm/day. Finally, for catchments with mixed forest, the comparison of the mean seasonal sum of the transpiration FLEX^{C} in spring varies between an underestimation of 21 mm/day and an overestimation of 81 mm/day, while for FLEX^{FAO} it varies between an underestimation of 28 mm/day and an overestimation of 31 mm/day. In the other seasons, the mean seasonal sum of the transpiration simulated by FLEX^{C} is much closer to the GLDAS21-transpiration. In the winter months FLEX^{C} also overestimates the mean seasonal sum of the transpiration. This overestimation is small when compared to the other seasons, but relatively large compared to the GLDAS21-transpiration data.

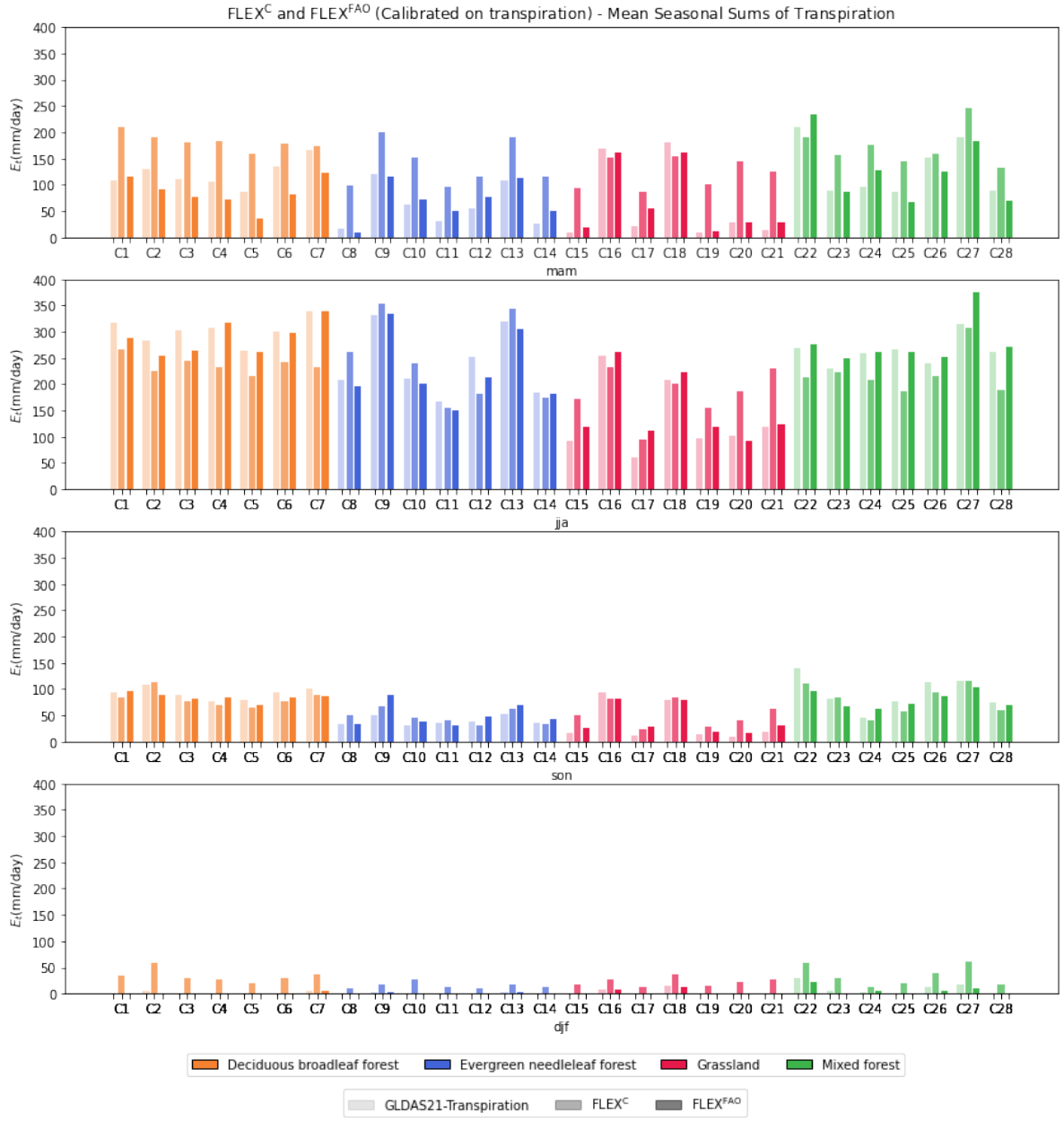


Figure 3.21: Mean seasonal sums of the transpiration for the GLDAS21-transpiration, FLEX^{C} and FLEX^{FAO}. The seasons are indicated by mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. The colors show the dominant vegetation type in the catchments, which are indicated on the x-axes. For every catchment, the GLDAS21-transpiration, FLEX^{C}, and FLEX^{FAO} are presented in that order.

3.5. Model performance when calibrated on streamflow and transpiration

In this research we used three different calibration strategies. FLEX^{C} and FLEX^{FAO} have been calibrated on streamflow, on transpiration and on both streamflow and transpiration separately. In this subsection the results will be presented for FLEX^{C} and FLEX^{FAO} that have been calibrated on streamflow and transpiration. The results will be separated into the streamflow simulation and the transpiration simulation. First the performance of the models will be shown

for the streamflow simulations of all catchments and then the performance of the transpiration simulations of all catchments.

3.5.1. Streamflow

The NS-efficiencies and logarithmic NSlog-efficiencies of the streamflow simulations of $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on streamflow and transpiration are presented in the boxplots in Figure 3.22. The performance of the model is shown in terms of the ability to simulate both low flows and high flows, and on the ability to predict future fluxes in the model. The NS-efficiency emphasizes more on high flows and the NSlog-efficiency emphasizes more on low flows, and the ability to predict future fluxes in the model is presented by showing the objective functions for the calibration period and the evaluation period separately. In accordance with this, Figure 3.22 shows four graphs. The top figure shows the NS-efficiency in the calibration period. The second graph shows the NS-efficiency in the evaluation period. The third figure shows the NSlog-efficiency in the calibration period and finally the bottom figure shows the NSlog-efficiency in the evaluation period. For all catchment, which are shown by the x-axis, $FLEX^{C}$ and $FLEX^{FAO}$ are presented next to each other. The colors indicate the dominant vegetation type in the catchments.

Figure 3.22 shows that for most catchments the performance of the models is reasonably good. The efficiencies are slightly lower than the models that are only calibrated on streamflow, between 0.05 and 0.2 for most catchments for the NS-efficiency and between 0.0 and 0.1 for the NSlog-efficiency. The y-axes only show the boxplots with an objective function above -0.4. We did this to put more emphasis on the well-performing catchments, with the disadvantage that some of the catchments with very low scores are ignored. Catchments that score low on some or all parts of the performance and are therefore ignored are C11, C12 and C17. Furthermore, the graphs show that for all catchments the NSlog-efficiency is very similar or higher than the NS-efficiency, which means both $FLEX^{C}$ and $FLEX^{FAO}$ are most times better in simulating the low flows than the high flows. Also, most catchments show similar values for the objective functions in the calibration period and the objective functions in the evaluation period, which means that in general $FLEX^{C}$ and $FLEX^{FAO}$ have a reasonable predictive capacity, where there is not a noticeable difference in performance between $FLEX^{C}$ and $FLEX^{FAO}$.

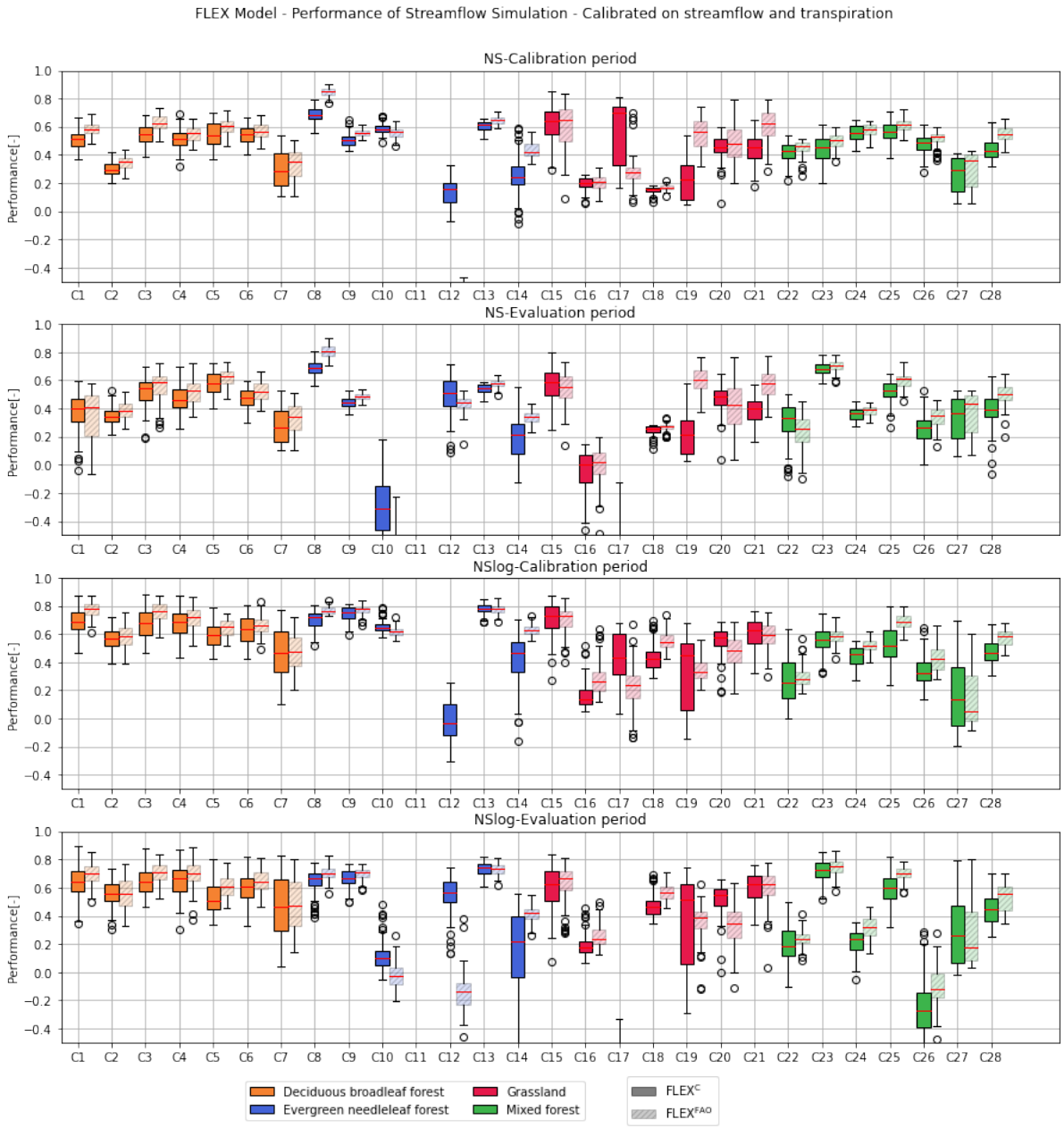


Figure 3.22: Performance of the streamflow simulation of $FLEX^C$ and $FLEX^{FAO}$ when calibrated on streamflow and transpiration, presented next to each other for every catchment. The four graphs separate the performance into the NS and NSlog objective function and the performance in the calibration period and the evaluation period. The colors show the dominant vegetation type in all the catchments. The x-axes show all the catchments that have been simulated.

Figure 3.23 shows how much the NS-efficiency and the NSlog-efficiency increased or decreased after the FAO adjustment. For every catchment, the ΔNS and ΔNSlog are presented next to each other. The colors indicate the dominant vegetation type.

The figure shows that for most catchments the efficiencies increase for most catchments. For catchments with deciduous broadleaf forest the NS-efficiency increases between 0.01 and 0.08 and the NSlog-efficiency between 0.01 and 0.095. For the catchments with evergreen needleleaf forest most catchments also increase for both efficiencies, except for C11 and C12, which will be ignored due to their low overall performances visible in Figure 3.22. Catchment

C17 will be discarded for the same reason. With this in mind the NS-efficiencies vary between a decrease of 0.02 and an increase of 0.18. The NSlog-efficiencies of the catchments with evergreen needleleaf forest vary between a decrease of 0.03 and an increase of 0.16. The increase of the NS-efficiencies of the catchments with grassland vary between 0.005 and 0.34. The NSlog-efficiencies do decrease in some cases varying between a decrease of 0.11 and an increase of 0.13. The increase of the NS-efficiencies of the catchments with mixed forest vary between 0.02 and 0.11. The NSlog-efficiencies of these catchments vary between a decrease of 0.09 and an increase of 0.17.

In general, there is no clear connection visible between how much the performances of the models increase and the dominant vegetation type. The catchments with deciduous broadleaf forest and mixed forest do seem more consistent in their performances and improvement of performance.

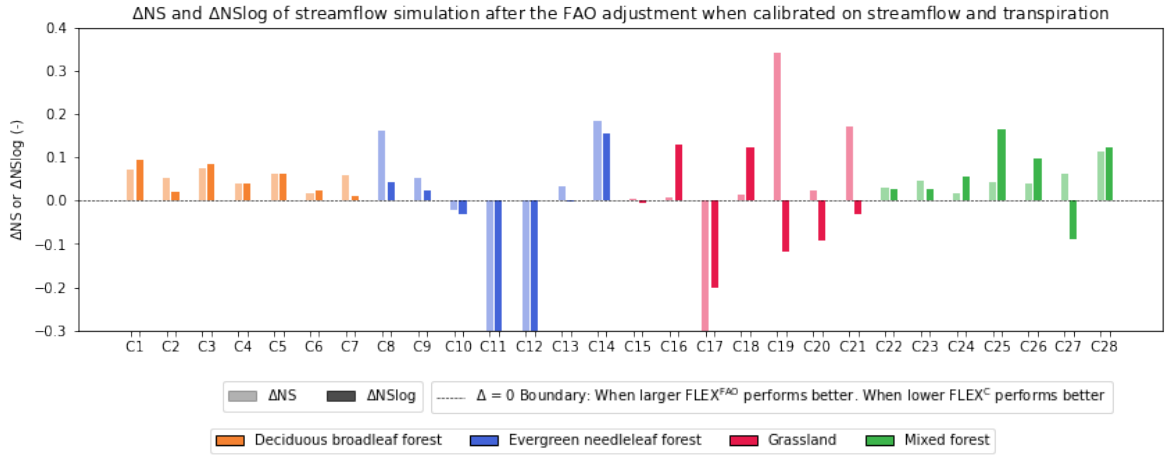


Figure 3.23: Performance comparison of the streamflow simulation between $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on streamflow and transpiration. The figure shows the increase or decrease of the NS-efficiency and the NSlog-efficiency after the FAO adjustment for every catchment respectively. The colors indicate the dominant vegetation type.

To be able to show a potential relationship between the impact of the FAO adjustment and some key climate indices, ΔNS, and ΔNSlog are also plotted against some climate indices in Figure 3.24. In Figure 3.24 the performance of streamflow simulation $FLEX^{C}$ and $FLEX^{FAO}$ calibrated on streamflow is compared. The figures show how much the performance increased or decreased after the FAO adjustment. This increase or decrease is presented in relationship to three different climate indices. The first two figures show how the performance comparison relates to the maximum difference in LAI throughout the year. The middle two figures show the performance comparison in relationship with the aridity and the final figures in relationship with seasonality. The top row shows the NS-efficiency and the bottom row the NSlog-efficiency. The dot marker and the cross marker represent the performance in the calibration period and the evaluation period respectively. The colors show the dominant vegetation type in the catchments.

In the first two figures is visible how the catchments with deciduous broadleaf forest and mixed forest have a higher difference in LAI than the catchments with grassland and evergreen needleleaf forest. This is expected due to the fact evergreen needleleaf forests and grassland are not impacted by phenology as much. Furthermore, the figures don't show any clear relationships between the climate indices and the NS- or NSlog-efficiencies.

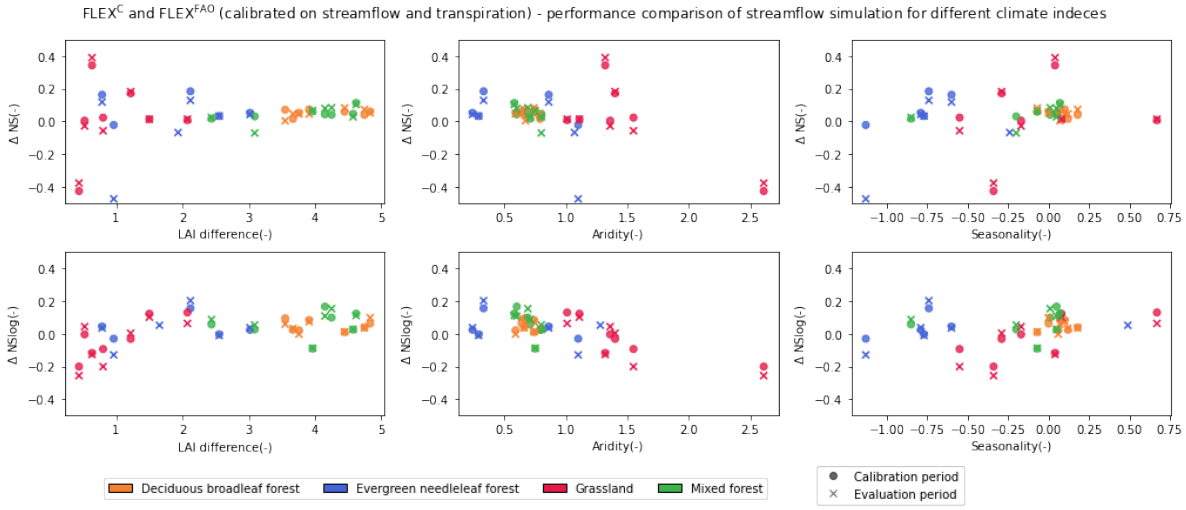


Figure 3.24: Performance comparison between FLEX^{C} and FLEX^{FAO}. The top figures show the comparison of the NS-efficiency and the bottom figures show the comparison of the NSlog-efficiency. The y-axes show how much the NS/NSlog-efficiency is increased or decreased after the FAO adjustment. The dot marker shows the performance in the calibration period and the cross marker the performance in the evaluation period. The colors indicate the dominant vegetation type in the catchment. The x-axes show different climate indices.

In Figure 3.25 the mean seasonal sums of the streamflow for the observed streamflow, FLEX^{C} and FLEX^{FAO} are presented for all catchments. All four graphs show one season, where the top graph shows the spring, the second graph the summer, the third graph the fall, and the bottom graph the winter. The markers separate the efficiencies in the calibration period and the evaluation period. The colors show the dominant vegetation type of the catchments.

In the figure is visible that overall both FLEX^{C} and FLEX^{FAO} match the mean seasonal sums of the observed streamflow quite well throughout the different seasons. However, no clear consistencies on whether the models are overestimating or underestimating are present in the different seasons. With small differences FLEX^{C} and FLEX^{FAO} are in some cases overestimating and in some cases underestimating the mean seasonal sums of the observed streamflow.

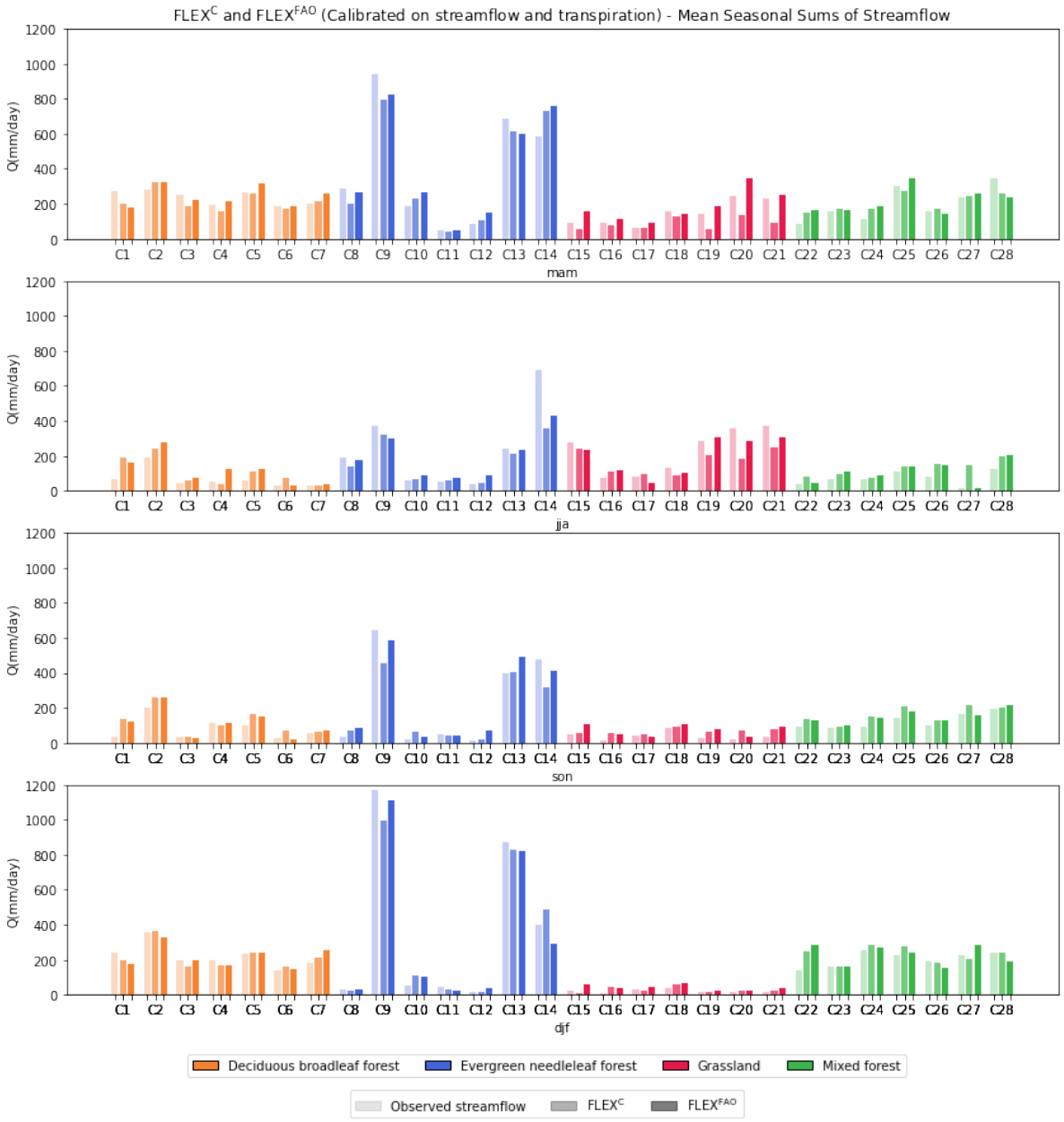


Figure 3.25: Mean seasonal sums of the streamflow for the observed streamflow, FLEX^{C} and FLEX^{FAO}. The seasons are indicated by mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. The colors show the dominant vegetation type in the catchments, which are indicated on the x-axes. For every catchment, the observed streamflow, FLEX^{C} and FLEX^{FAO} are presented in that order.

3.5.2. Transpiration

The Nash-Sutcliffe (NS) efficiencies and logarithmic Nash-Sutcliffe (NSlog) efficiencies of FLEX^{C} and FLEX^{FAO} calibrated on streamflow and transpiration are presented in the boxplots in Figure 3.26. The performance of the model is shown in terms of the ability to simulate both low flows and high flows, and the ability to predict future fluxes in the model. The NS-efficiency focuses more on high flows and the NSlog-efficiency focuses more on low flows, and the ability to predict future fluxes in the model is presented by showing the objective functions for the calibration period and the evaluation period separately. In accordance with this, Fig-

ure 3.26 shows four graphs. The top figure shows the NS-efficiency in the calibration period. The second graph shows the NS-efficiency in the evaluation period. The third figure shows the NSlog-efficiency in the calibration period and finally, the bottom figure shows the NSlog-efficiency in the evaluation period. For all catchments, which are shown by the x-axes, $FLEX^{C}$ and $FLEX^{FAO}$ are presented next to each other. The colors indicate the dominant vegetation type in the catchments.

The y-axes only show the boxplots with an objective function above -0.4. We did this to put more emphasis on the well-performing catchments, with the disadvantage that some of the catchments modeled by $FLEX^{C}$ have low efficiencies and are ignored. These catchments are C11, C15, C17, C19, C20, and C21. With this observation, it must be noted that the efficiencies are only low for the $FLEX^{C}$. $FLEX^{FAO}$ however gives much higher performances, which confirms an improvement of the FAO adjustment. It can be seen that $FLEX^{C}$ simulates the NS-efficiency better for catchments with deciduous broadleaf forests and mixed forests in comparison with the NSlog-efficiency. The efficiencies in the calibration period and the evaluation period perform very similarly for all catchments, which confirms that the models both have a good predictive capacity. The performance comparison is shown in more detail in Figure 3.27 and 3.28.

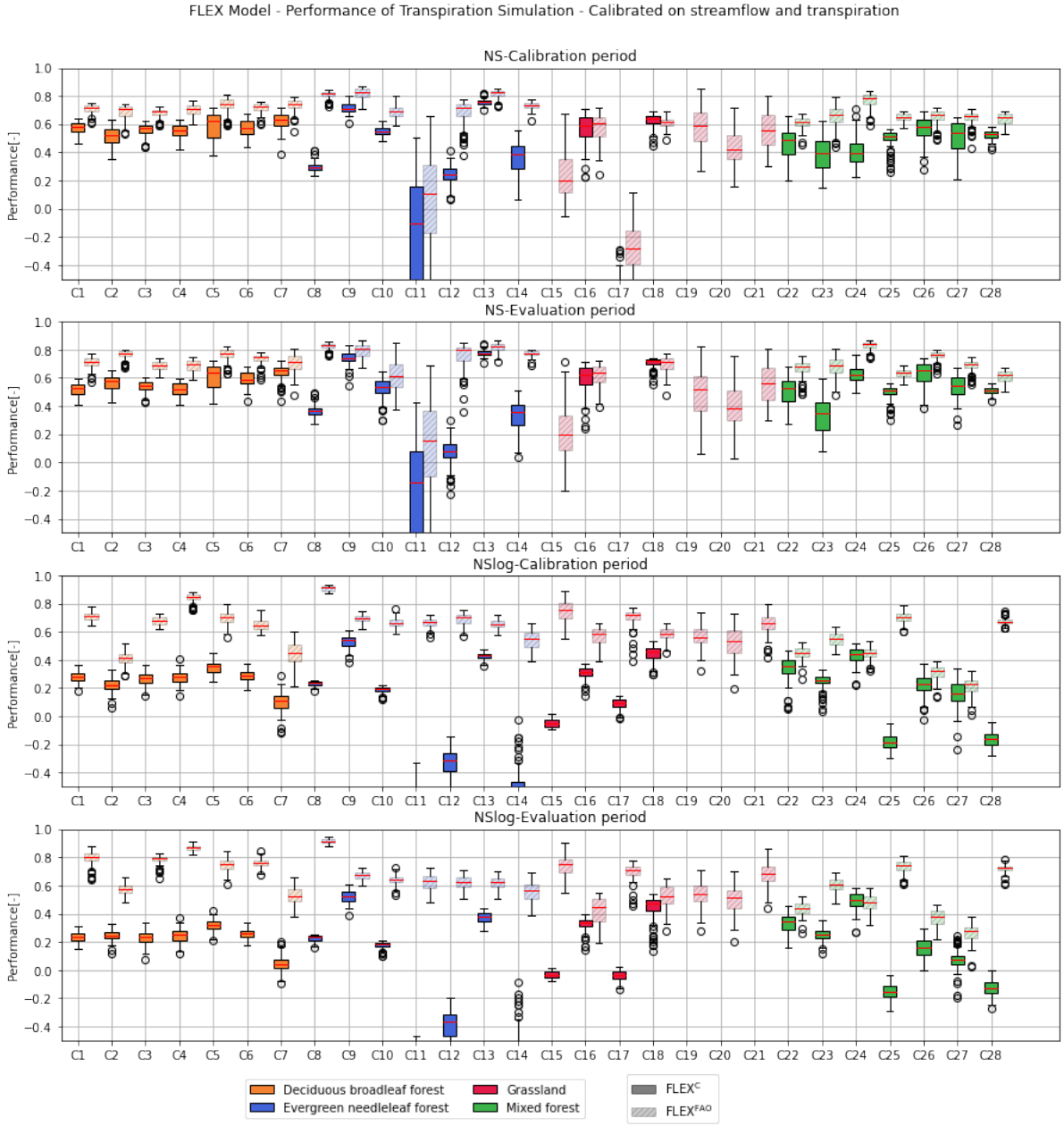


Figure 3.26: Performance of the transpiration simulation of $FLEX^{C}$ and $FLEX^{FAO}$ when calibrated on streamflow and transpiration, presented next to each other for every catchment. The four graphs separate the performance into the NS and NSlog objective function and the performance in the calibration period and the evaluation period. The colors show the dominant vegetation type in all the catchments. The x-axes show all the catchments that have been simulated.

Figure 3.27 shows how much the NS-efficiency and the NSlog-efficiency increased or decreased after the FAO adjustment. For every catchment, the ΔNS and ΔNSlog are presented next to each other. The colors indicate the dominant vegetation type.

The figure shows that for most catchments the efficiencies increase drastically. What stands out is that the improvement of the NSlog-efficiency is higher than the improvement of the NS-efficiency for catchments with deciduous broadleaf forest and evergreen needleleaf forest, while this is not necessarily the case for catchments with mixed forest or grassland. For the catchments with deciduous broadleaf forest, the increase is between 0.12 and 0.18 for the

NS-efficiency and between 0.19 and 0.56 for the NSlog-efficiency. For catchments with mixed forest, the increase is between 0.09 and 0.38 for the NS-efficiency and between 0.01 and 0.89 for the NSlog-efficiency. An important sidenote to this graph is that the catchments with the largest improvement are also catchments with very low efficiencies by FLEX^{C}. This in combination with FLEX^{FAO} giving similar performances as the other catchments, makes a large improvement.

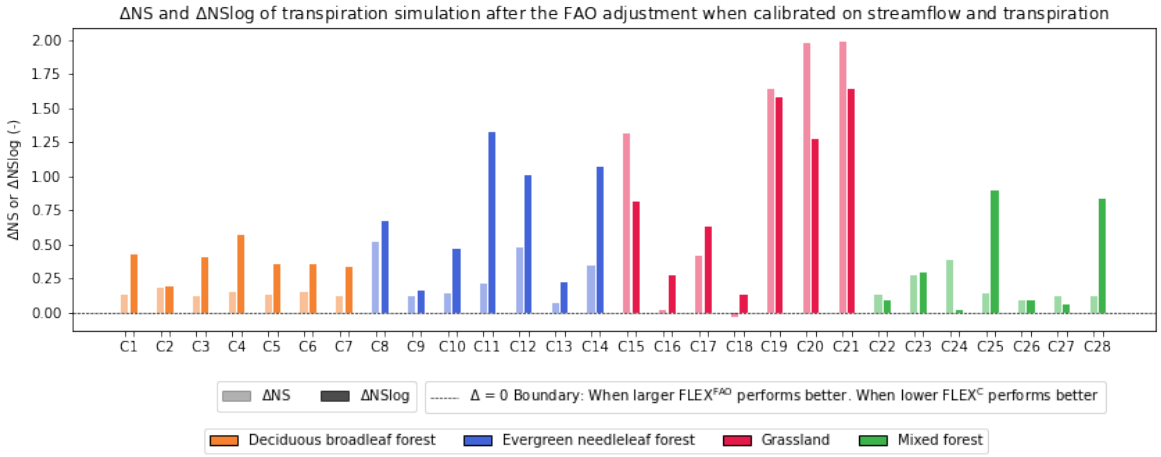


Figure 3.27: Performance comparison of the transpiration simulation between FLEX^{C} and FLEX^{FAO} calibrated on streamflow and transpiration. The figure shows the increase or decrease of the NS-efficiency and the NSlog-efficiency after the FAO adjustment for every catchment respectively. The colors indicate the dominant vegetation type.

In Figure 3.28 the performance of the transpiration simulation FLEX^{C} and FLEX^{FAO} calibrated on streamflow and transpiration is compared. The figures show how much the performance increased or decreased after the FAO adjustment. This increase or decrease is presented in relationship with three different climate indices. The first two figures show how the performance comparison relates to the maximum difference in LAI throughout the year. The middle two figures show the performance comparison in relationship with the aridity and the final figures in relationship with seasonality. The top row shows the NS-efficiency and the bottom row the NSlog-efficiency. The dot marker and the cross marker represent the performance in the calibration period and the evaluation period respectively. The colors show the dominant vegetation type in the catchments.

The figures confirm that the NS-efficiency and the NSlog-efficiency greatly improves after the FAO adjustment. The first two figures show that the efficiencies improve more for catchments with low LAI difference and high aridity. Also in this figure the observation that the catchments with a high ΔNS or ΔNSlog are also the catchments with a low overall performance by the FLEX^{C} needs to be taken into account.

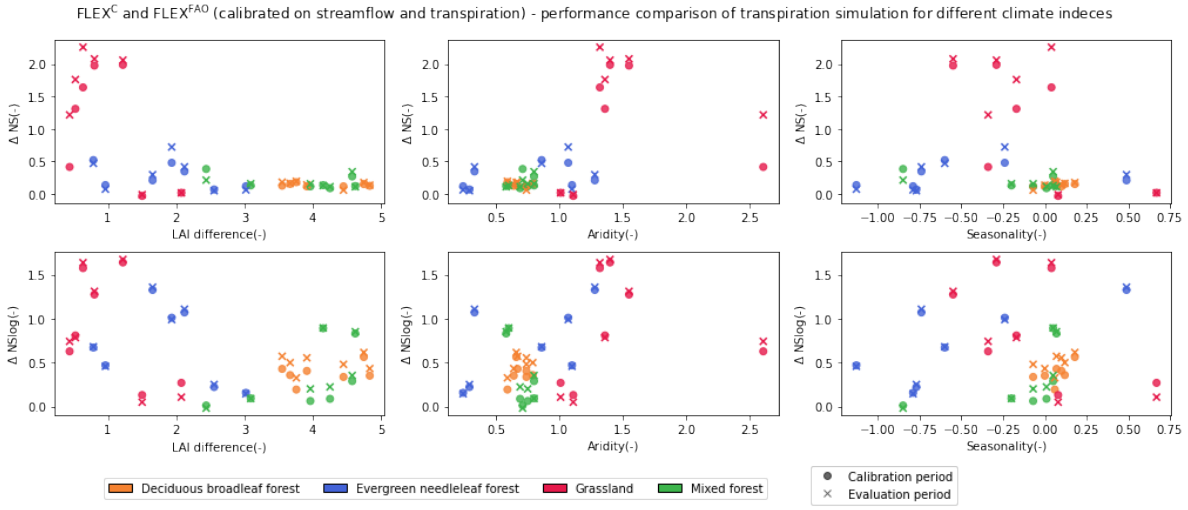


Figure 3.28: Performance comparison between FLEX^{C} and FLEX^{FAO}. The top figures show the comparison of the NS-efficiency and the bottom figures show the comparison of the NSlog-efficiency. The y-axes show how much the NS/NSlog-efficiency is increased or decreased after the FAO adjustment. The dot marker shows the performance in the calibration period and the cross marker the performance in the evaluation period. The colors indicate the dominant vegetation type in the catchment. The x-axes show different climate indices.

In Figure 3.29 the mean seasonal sums of the transpiration for the GLDAS21-transpiration, FLEX^{C} and FLEX^{FAO} are presented of all catchments. All four graphs show one season, where the top graph shows the spring, the second graph the summer, the third graph the fall, and the bottom graph the winter. The colors show the dominant vegetation type of the catchments.

The figure shows that in the spring the mean seasonal sum of the transpiration is greatly overestimated by FLEX^{C} for all catchments. For catchments with deciduous broadleaf forest, this overestimation varies between 50 and 138 mm/day, while for FLEX^{FAO} it varies between an underestimation of 41 mm/day and an overestimation of 31 mm/day. For catchments with evergreen needleleaf forest, the overestimation of FLEX^{C} varies between 79 and 176 mm/day, while for FLEX^{FAO} it varies between an underestimation of 24 mm/day and an overestimation of 52 mm/day. For catchments with grassland, the overestimation of FLEX^{C} varies between 21 and 110 mm/day, while for FLEX^{FAO} it varies between an underestimation of 24 mm/day and an overestimation of 34 mm/day. Finally for catchments with mixed forest the overestimation of FLEX^{C} varies between 56 mm/day and 132 mm/day, while for FLEX^{FAO} it varies between an underestimation of 8 mm/day and an overestimation of 64 mm/day. In the other seasons, the mean seasonal sum of the transpiration simulated by FLEX^{C} is much closer to the GLDAS21-transpiration. In the winter months FLEX^{C} also overestimates the mean seasonal sum of the transpiration. This overestimation is small when compared to the other seasons, but relatively large compared to the GLDAS21-transpiration data.

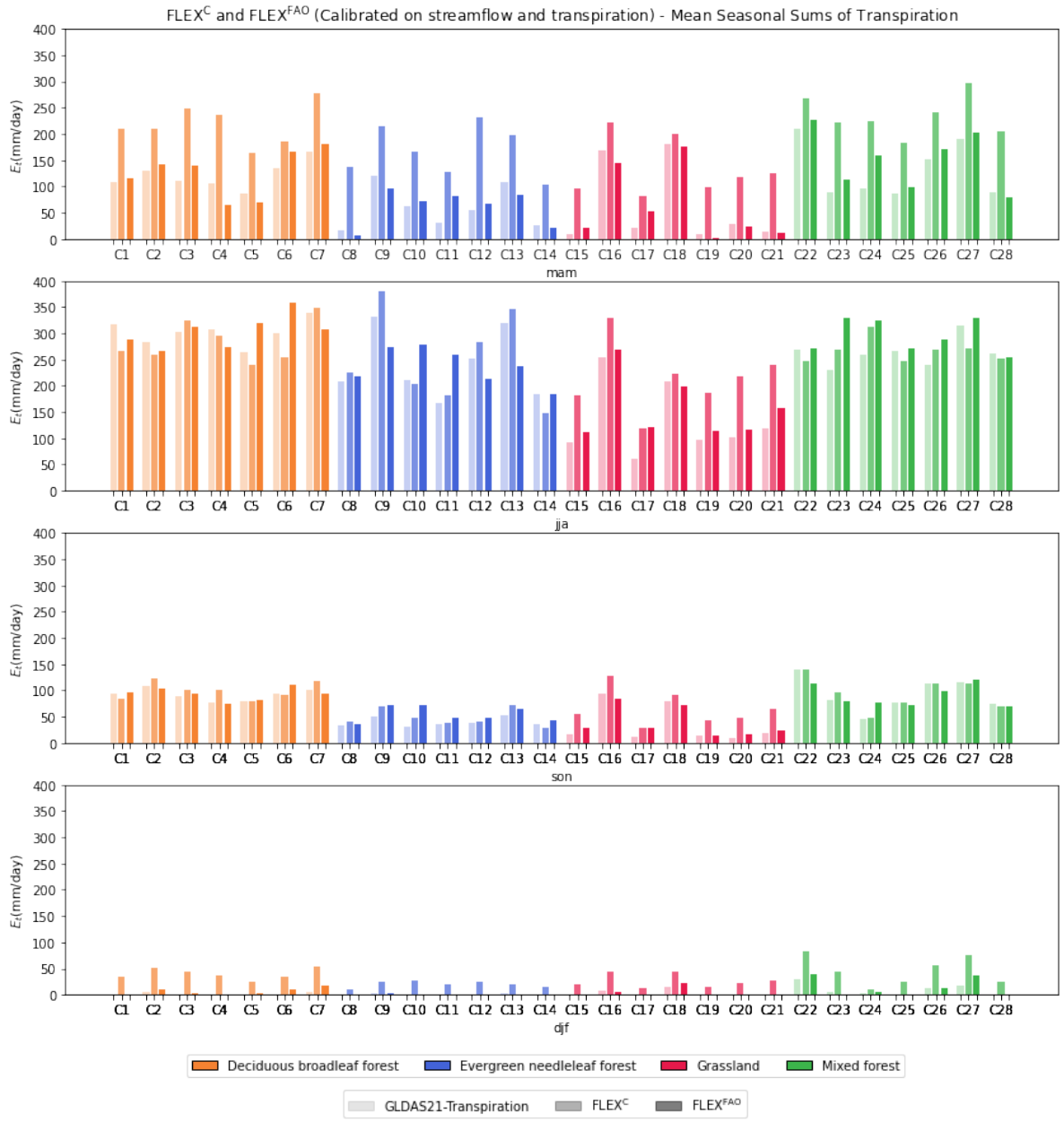


Figure 3.29: Mean seasonal sums of the transpiration for the GLDAS21-transpiration, FLEX^{C} and FLEX^{FAO}. The seasons are indicated by mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. The colors show the dominant vegetation type in the catchments, which are indicated on the x-axes. For every catchment, the GLDAS21-transpiration, FLEX^{C}, and FLEX^{FAO} are presented in that order.

4

Discussion

In FLEX^{C} and many other lumped conceptual hydrologic models, the total evaporation is separated into evaporation and transpiration in a way that does not account for phenology or vegetation dynamics. In FLEX^{C} evaporation is a flux going out of the interception reservoir as interception evaporation and transpiration is a flux going out of the unsaturated soil reservoir as transpiration. The partitioning of these two fluxes is determined by first quantifying the interception evaporation with the potential evaporation and the model input, and then using the remainder as transpiration from the unsaturated soil reservoir. While this way of modeling is very suitable for both fast and high-quality streamflow modeling (Gao et al. 2014), the evapotranspiration fluxes are not always adequate. Neglecting the vegetation dynamics causes the model to not include the fact that deciduous broadleaf forests for example are leafless in the winter months practically leading to no transpiration in these months. On the other hand in the summer months, transpiration would have a much more dominant role in the total evapotranspiration due to the presence of a developed canopy, with many transpiring leaves.

The FAO adjustment which we used in this research does take these vegetation dynamics into account. This chapter will discuss what the impact of this FAO adjustment is on the performance of the streamflow simulation and the transpiration simulation of FLEX^{C} and FLEX^{FAO}, based on the results. The performance that is discussed in this chapter is separated into the general ability to simulate streamflow and transpiration and the predictive capacity of the model. First, the streamflow simulation of both models will be discussed, and then the transpiration simulation of both models, comparing different calibration strategies with each other.

4.1. Streamflow simulation

In general, both FLEX^{C} and FLEX^{FAO} can model the observed streamflow quite well, with higher performances for the NSlog-efficiencies than the NS-efficiencies. This means that the models are both better at simulating the low flows than the peak flows. The reasons for this are difficult to track down. The heterogeneity of the soil, the vegetation, elevation, and the spatial distribution of the rainfall are all important for the peak flow response of the hydrologic system, which could lead to higher uncertainties and therefore different responses than the models used in this research. The uncertainties for the temporal and spatial distribution of the precipitation increase in the case of convective precipitation fields that can generate extreme flooding events in high-elevation, complex terrain configurations (Anquetin et al. 2005) (Emmanuel et al. 2017) (Sokol et al. 2021). This is then reflected in the models being better in simulating the low flows and therefore having higher NSlog-efficiencies than the NS-efficiencies.

Adding the GLDAS21-transpiration data to the calibration, causes the NS-efficiencies and the NSlog-efficiencies to be slightly lower for the streamflow simulation, especially the streamflow modeled by FLEX^{C}. This is to be expected when the amount of Monte Carlo simulations stays the same during the calibration, while there are two extra objective functions.

When the models are calibrated on streamflow the streamflow simulated by FLEX^{FAO} has very similar or sometimes even lower NS-efficiencies and NSlog-efficiencies than FLEX^{C}, which means that the FAO adjustment does not necessarily have a positive effect on the general capability of the FLEX model to simulate the streamflow, when only streamflow observations are used to calibrate the model. With the two extra parameters $K_{v,ini}$ and $K_{v,mid}$, there are 11 parameters to be calibrated instead of 9 in FLEX^{C}. This makes the calibration more extensive, which means the Monte Carlo needs more simulations to find the best performing parameter sets. In this research the ranges within which the parameters are calibrated is set relatively wide to make them suitable for a wide range of different catchments. This has the disadvantage that more simulations are needed to find the optimal set of parameters. Whether more Monte Carlo simulations would improve the simulations of FLEX^{C} and FLEX^{FAO} in this study is not investigated, because in this study the focus is more on the comparison than on optimizing the model. However, this is a purely mathematical explanation for the lack of impact of the FAO adjustment, which has nothing to do with the hydrological processes in the catchments. This means that the slightly more complex calibration process of FLEX^{FAO} weighs up to the impact of a more sophisticated way of including vegetation dynamics in the process of partitioning the evaporation in the model. When both the streamflow observations and the GLDAS21-transpiration data are included in the calibration process the FAO adjustment seems to cause a positive impact on the NS-efficiencies and the NSlog-efficiencies. The efficiencies of FLEX^{C} are slightly lower than when the models are calibrated on streamflow only due to the two extra objective functions, but the impact of the FAO adjustment on the model is higher, with the extra benefit that the transpiration is also better simulated. This positive ΔNS and ΔNSlog is consistent for the catchments with deciduous broadleaf forest and mixed forest. For catchments with evergreen needleleaf forest and grassland, it is more variable with more outliers, with especially the ΔNSlog for catchments with grassland being positive. These catchments are catchments with a relatively long period with little to no streamflow. It could be that for this reason, the model is compensating in the calibration process, which causes the underestimation during the peak flow. It can also be explained partly by the higher overall performance of the catchments with deciduous broadleaf forest and mixed forest, which indicate a higher quality of the input data. Besides that, there is no clear relationship visible between how large the impact is of the FAO adjustment and the vegetation type in the catchment. In the catchments with deciduous broadleaf forest and mixed forest, the LAI difference is the largest, which means that the vegetation dynamics throughout the seasons are the largest for these catchments. However, FLEX^{FAO}, which accounts for these vegetation dynamics, does not necessarily simulate the streamflow better.

For the FLEX^{C} calibrated on streamflow only, the mean seasonal sum of the streamflow is underestimated in spring for most catchments. This underestimation is logical, when combined with the overestimation of the mean seasonal sum of the transpiration in spring by FLEX^{C}. This underestimation of the mean seasonal sum of the streamflow in spring is solved mostly by the FLEX^{FAO}. This means that even though the overall performance of the FLEX^{FAO} is similar or worse than FLEX^{C}, the underestimation of the streamflow in spring is fixed by the FAO-adjustment.

The NS-efficiencies and the NSlog-efficiencies are for most catchments very similar, so the predictive capacity of the models is good. The differences between FLEX^{C} and FLEX^{FAO} are

small to nonexistent when looking at the efficiencies in the calibration period compared to the efficiencies in the evaluation period. The LAI data that is included in FLEX^{FAO} as extra information on vegetation dynamics could make it more difficult for the model to predict any of the water fluxes in the model since an extra model input needs to be predicted. However the model does not use the LAI data with a daily resolution, but only the timestamps of when the growing stages change. This in combination with the observation that the seasonal patterns in the LAI data were quite consistent through the years for the catchments in this research, would minimize the negative impact on the predictive capacity of the model.

4.2. Transpiration simulation

The performance of both FLEX^{C} and FLEX^{FAO} in simulating the transpiration varies more. When the GLDAS21-transpiration data is included in the calibration process, FLEX^{C} ability to simulate the peak flows is reasonable, which is reflected in the NS-efficiencies being above 0.5 for most catchments. However, the NSlog-efficiencies of the FLEX^{C} are worse. Most catchments have NSlog-efficiencies below 0.4, even when only GLDAS21-transpiration data is used in the calibration process. This can be explained by the absence of information on the vegetation dynamics in FLEX^{C}. The transpiration in FLEX^{C} is determined by the potential evaporation in combination with the amount of available water in the unsaturated soil reservoir(S_u) and the calibrated parameter S_{uMax}. This leads to an overestimation of the transpiration in winter and the start of spring, which was visible for all three calibration strategies in the mean seasonal sums of the transpiration. This overestimation of the transpiration in winter and the start of spring is also reflected in the fact that the NS-efficiencies of the transpiration simulation of FLEX^{C} were higher than the NSlog-efficiencies. The NSlog-efficiencies emphasize the low flows during the winter and early spring, which are the months that are simulated less accurately due to the absence of information on vegetation dynamics and phenology in the model.

The spread of the boxplots showing the efficiencies for the transpiration simulation of the models calibrated on GLDAS21-transpiration data only was also very low. This is due to the sensitivity of the models being very high for the shape coefficient(β), which is used in determining the runoff coefficient(C_r) in the model. Because the models are used in a wide range of catchments with different properties, the ranges which are used for the parameters in the calibration process are kept relatively wide, while for getting high efficiencies for the transpiration simulation a smaller range for β would be sufficient. The calibration process of selecting the 100 best performing parameter sets and using their median streamflow and transpiration simulation as results for the models leads to a narrow range of efficiencies. This does not occur as much when the models are also using streamflow observations in the calibration process, because for getting well-performing streamflow simulations, the models are much less sensitive for β.

Where a positive impact of the FAO adjustment on the performance of the streamflow was only situational, it is for the transpiration simulation much larger and consistent regardless of the calibration strategy. Even when the models were only calibrated on streamflow, the NS-efficiencies and NSlog-efficiencies of the transpiration simulation of FLEX^{FAO} were for most catchments well above 0.5. In the improvement is visible that ΔNSlog is higher for most catchments than ΔNS, so the model improves more in the low flows than the peak flows. This is to be expected, because in winter and early spring when the transpiration is lowest, FLEX^{C} is overestimating the transpiration due to the absence of information on vegetation dynamics in the model. The improvement in the performance of the model is present regardless of the dominant vegetation type of the catchment. However, the improvement seems largest

for catchments with grassland, even though these are not the catchments with the highest difference in LAI throughout the seasons or the catchment with trees that shed their leaves during autumn and winter. What can be noticed in the results is that these high values for $\triangle$NS and $\triangle$NSlog are more due to bad performances of $FLEX^{C}$ than to high performances of $FLEX^{FAO}$. The efficiencies of $FLEX^{FAO}$ are relatively close to each other, when the catchments are compared, while the efficiencies of $FLEX^{C}$ vary a lot.

What is important to mention as a critical sidenote to the calibration strategy used in this research, is that there is no actual observational data used to calibrate the transpiration simulation. The GLDAS21-transpiration data used to calibrate the model on transpiration is modeled data itself. The GLDAS21-transpiration data is a product of the Noah Land Surface Model (Rui and Beaudoing 2021). The Noah Land Surface Model uses a Jarvis scheme(Jarvis and Mcnaughton 1986) or Ball-Berry scheme(Ball, Woodrow, and Berry 1987) to calculate transpiration. This method also uses LAI data as input (Niu et al. 2011). Because there is no observational data on transpiration available, and the transpiration product of for example the Noah Land Surface Model is the best option around, this choice to use this data in the calibration of the model is justified.

5

Conclusion

This study has as its objective to test and investigate a model adjustment that includes information on vegetation dynamics and plant phenology to partition the evaporation in conceptual hydrological models. We compared the performance of a conventional conceptual hydrological model and the same model with the model adjustment, where the performance is separated into the general capacity of the model to simulate streamflow and transpiration and the predictive capacity of the model. On top of this, we investigated the relationship between the performance of the models and the catchment's characteristics on vegetation and the catchment's climate conditions. We did this by testing the FLEX model, which is a relatively inelaborate conceptual hydrological model (Gao et al. 2014), in its conventional setup and its adjusted setup on 28 catchments with different climate and land cover characteristics. This adjustment is based on the crop evaporation method of the Food and Agriculture Organization and partitions the evapotranspiration using LAI data (Allen et al. 1998). Because we investigated the general capacity to simulate both streamflow and transpiration, streamflow observations and transpiration data are used in the calibration process, both separately and together.

$FLEX^C$ and $FLEX^{FAO}$ are able to model the observed streamflow quite well, with higher performances for the low flows than for the peak flows. NS-efficiencies and the NSlog-efficiencies of $FLEX^C$ are in most cases higher than the NS-efficiencies and the NSlog-efficiencies of $FLEX^{FAO}$ when the model is calibrated only on streamflow observations. This could be due to the addition of two extra parameters that need to be calibrated for $FLEX^{FAO}$. However, it does suggest that the FAO adjustment does not necessarily improve the general ability of the model to streamflow when only streamflow observations are used for calibration. The negative $\triangle$NS and $\triangle$NSlog are consistent for catchments with deciduous broadleaf forest and mixed forest. For catchments with evergreen needleleaf forest and grassland, it is more variable with more outliers, with especially the $\triangle$NS for catchments with grassland being positive. This is due to an underestimation of $FLEX^C$ during the peak flow. These catchments are catchments with a relatively long period with little to no streamflow. It could be that for this reason, the model is compensating in the calibration process, which causes the underestimation during the peak flow. Also, it could be due to the data quality, which is relatively low for some catchments with evergreen needleleaf forest and grassland. When the model is calibrated on both streamflow and transpiration, the NS-efficiencies and the NSlog-efficiencies are in general lower than when only calibrated on streamflow. However The $\triangle$NS and $\triangle$NSlog after the FAO adjustment are positive for most catchments, and this is in combination with a large improvement in the ability of the model to simulate the transpiration.

In the transpiration simulation of FLEX^{C} and FLEX^{FAO}, there is a larger difference between both models. For all catchments FLEX^{FAO} produces NS-efficiencies and NSlog-efficiencies that are much higher than FLEX^{C}, regardless of the vegetation type of the catchment. The transpiration in FLEX^{C} is determined by the potential evaporation in combination with the amount of available water in the unsaturated soil reservoir(S_u) and the calibrated parameter S_{uMax}. This leads to an overestimation of the transpiration in winter and the start of spring, which was visible for all three calibration strategies in the mean seasonal sums of the transpiration. This overestimation of the transpiration in winter and the start of spring is also reflected in the fact that the NS-efficiencies of the transpiration simulation of FLEX^{C} were higher than the NSlog-efficiencies. Where the expectation was that the improvement would be highest for catchments with a large difference in LAI throughout the year, this is not reflected in the results. The improvement seems largest for catchments with grassland, even though these are not the catchments with the highest difference in LAI throughout the seasons or the catchment with trees that shed their leaves during autumn and winter. What can be noticed in the results is that these high values for $\triangle$NS and $\triangle$NSlog are more due to bad performances of FLEX^{C} than to high performances of FLEX^{FAO}.

The predictive capacity of both FLEX^{C} and FLEX^{FAO} are very similar. Regardless of the dominant vegetation type of the catchments or the climate indices of the catchments or the calibration strategy used, the NS-efficiencies and the NSlog-efficiencies of FLEX^{C} and FLEX^{FAO} in the calibration period are very similar to the evaluation period.

This study concludes that conceptual hydrological models that include information from plant phenology as in the FAO adjustment have a similar ability to simulate the streamflow as conceptual hydrological models that do not include this information. However, the impact on the ability to simulate transpiration can be very positive even for vegetation types that transpire throughout all seasons. To optimize the model and to be able to simulate both the streamflow and the transpiration simulation best, both streamflow and transpiration data could be used in the calibration process. The predictive capacity of a conceptual hydrological model is not impacted by the inclusion of information on vegetation dynamics if it is done in the structure of this research.

5.1. Recommendations

It would be valuable if the research on this promising combination of the FLEX model and the FAO adjustment would be extended. This research has its focus on homogeneous and consistent land cover conditions in which rapid changes in land cover by for example human activities are not included. Also, more research could be done on an even more generally applicable way of including vegetation dynamics could be investigated. For example, using more predictable variables like temperature to predict the growing stages in the future for which no direct LAI is available.

References

Adler, R. F. et al. (2003). "The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present)". In: *Journal of Hydrometeorology* 4.6, pp. 1147–1167. DOI: 10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/hydr/4/6/1525-7541_2003_004_1147_tvgpcp_2_0_co_2.xml (cit. on p. 8).

Aghakouchak, A. and E. Habib (2010). "Application of a conceptual hydrologic model in teaching hydrologic processes". In: *International Journal of Engineering Education* 26.4, pp. 963–973 (cit. on p. 1).

Allen, R. et al. (Jan. 1998). "Crop evapotranspiration guidelines for computing crop requirements. FAO Irrig. Drain. Report modeling and application". In: *J. Hydrol.* 285, pp. 19–40 (cit. on pp. 7–9, 12, 13, 53).

Anderson, E. A. (1973). "National Weather Service River Forecast System – Snow accumulation and ablation model". In: *NOAA Technical Memorandum, NWS, HYDRO-17, US Department of Commerce, Silver Spring, MD*, p. 217 (cit. on p. 4).

Anquetin, S. et al. (2005). "The 8 and 9 September 2002 flash flood event in France: a model intercomparison". In: *Natural Hazards and Earth System Sciences* 5.5, pp. 741–754. DOI: 10.5194/nhess-5-741-2005. URL: https://nhess.copernicus.org/articles/5/741/2005/ (cit. on p. 49).

Ball, J. T., I. E. Woodrow, and J. A. Berry (1987). "A Model Predicting Stomatal Conductance and its Contribution to the Control of Photosynthesis under Different Environmental Conditions". In: *Progress in Photosynthesis Research: Volume 4 Proceedings of the VIIth International Congress on Photosynthesis Providence, Rhode Island, USA, August 10–15, 1986*. Ed. by J. Biggins. Dordrecht: Springer Netherlands, pp. 221–224. ISBN: 978-94-017-0519-6. DOI: 10.1007/978-94-017-0519-6_48. URL: https://doi.org/10.1007/978-94-017-0519-6_48 (cit. on pp. 8, 52).

Bergström, S (1992). "The hbv model – its structure and applications". In: *SMHI Reports Hydrology* (cit. on p. 9).

Budyko, M. I., D. H. Miller, and D. H. Miller (1974). "Climate and life". In: *Academic press New York* (cit. on p. 5).

Derber, J. C., D. F. Parrish, and S. J. Lord (1991). *The new global operational analysis system at the National Meteorological Center.* (Cit. on p. 8).

Emmanuel, I. et al. (2017). "A method for assessing the influence of rainfall spatial variability on hydrograph modeling. First case study in the Cevennes Region, southern France". In: *Journal of Hydrology* 555, pp. 314–322. ISSN: 0022-1694. DOI: https://doi.org/10.1016/j.jhydrol.2017.10.011. URL: https://www.sciencedirect.com/science/article/pii/S0022169417306753 (cit. on p. 49).

Fenicia, F. et al. (2006). "Is the groundwater reservoir linear? Learning from data in hydrological modelling". In: *Hydrology and Earth System Sciences* 10.1, pp. 139–150. DOI: 10.5194/hess-10-139-2006. URL: https://hess.copernicus.org/articles/10/139/2006/ (cit. on p. 9).

Fenicia, F. et al. (2016). "From spatially variable streamflow to distributed hydrological models: Analysis of key modeling decisions". In: *Water Resour. Res.* 52, pp. 954–989. URL: https://doi.org/10.1002/2015WR017398 (cit. on p. 1).

Gao, H. et al. (2014). “Testing the realism of a topography-driven model (FLEX-Topo) in the nested catchments of the Upper Heihe, China”. In: *Hydrology and Earth System Sciences* 18.5, pp. 1895–1915. DOI: 10.5194/hess-18-1895-2014. URL: https://hess.copernicus.org/articles/18/1895/2014/ (cit. on pp. 1, 9, 10, 49, 53).

Good, S. P., D. Noone, and G. Bowen (2015). “Hydrologic connectivity constrains partitioning of global terrestrial water fluxes”. In: *Science* 349.6244, pp. 175–177. DOI: 10.1126/science.aaa5931. eprint: https://www.science.org/doi/pdf/10.1126/science.aaa5931. URL: https://www.science.org/doi/abs/10.1126/science.aaa5931 (cit. on p. 2).

Hrachowitz, M. and M. Clark (2017). “Hess opinions: The complementary merits of competing modelling philosophies in hydrology”. In: *Hydrol. Earth Syst. Sci.* 21, pp. 3953–3973. URL: https://doi.org/10.5194/hess-21-3953-2017 (cit. on p. 1).

Jarvis, P. and K. Mcnaughton (1986). “Stomatal control of transpiration: Scaling up from leaf to region. Advances in Ecological Research”. In: *Advances in Ecological Research* 15, pp. 1–49 (cit. on pp. 2, 8, 52).

Kool, D. et al. (2014). “A review of approaches for evapotranspiration partitioning”. In: *Agricultural and Forest Meteorology* 184, pp. 56–70. ISSN: 0168-1923. DOI: https://doi.org/10.1016/j.agrformet.2013.09.003. URL: https://www.sciencedirect.com/science/article/pii/S016819231300230X (cit. on p. 2).

Kumar, R., L. Samaniego, and S. Emad (2010). “The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics”. In: *J. Hydrol.* 392, pp. 54–69 (cit. on p. 1).

Liu, Z. et al. (2017). “Conceptual Hydrological Models”. In: *Handbook of Hydrometeorological Ensemble Forecasting*. Ed. by Q. Duan et al. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–23. ISBN: 978-3-642-40457-3. DOI: 10.1007/978-3-642-40457-3_22-1. URL: https://doi.org/10.1007/978-3-642-40457-3_22-1 (cit. on p. 12).

Livneh, B. et al. (2013). “A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions”. In: *Journal of Climate* 26.23, pp. 9384 –9392. DOI: 10.1175/JCLI-D-12-00508.1. URL: https://journals.ametsoc.org/view/journals/clim/26/23/jcli-d-12-00508.1.xml (cit. on p. 7).

Mert, O. (2021). “The Impact of Vegetation on the Partitioning of Evaporation in Conceptual Hydrological Models”. In: (cit. on p. 2).

Nash, J. and J. Sutcliffe (1970). “River flow forecasting through conceptual models part I — A discussion of principles”. In: *Journal of Hydrology* 10.3, pp. 282–290. ISSN: 0022-1694. DOI: https://doi.org/10.1016/0022-1694(70)90255-6. URL: https://www.sciencedirect.com/science/article/pii/0022169470902556 (cit. on p. 14).

Newman, A. et al. (2014). “A large-sample watershed-scale hydrometeorological dataset for the contiguous USA.” In: *Boulder, CO: UCAR/NCAR*. URL: https://dx.doi.org/10.5065/D6MW2F4D (cit. on pp. 4, 5).

Newman, A. et al. (Jan. 2015). “Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance”. In: *Hydrology and Earth System Sciences* 19, pp. 209–223. DOI: 10.5194/hess-19-209-2015 (cit. on p. 8).

Niu, G. et al. (2011). “The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements”. English (US). In: *Journal of Geophysical Research: Atmospheres* 116.12. ISSN: 2169-897X. DOI: 10.1029/2010JD015139 (cit. on pp. 8, 52).

Pan, S. et al. (2020). “Evaluation of global terrestrial evapotranspiration using state-of-the-art approaches in remote sensing, machine learning and land surface modeling”. In: *Hydrology*

and Earth System Sciences 24.3, pp. 1485–1509. DOI: 10.5194/hess-24-1485-2020. URL: https://hess.copernicus.org/articles/24/1485/2020/ (cit. on p. 2).

Ranga Myneni Knyazikhin, Y. (2015). "MOD15A3H MODIS/Combined Terra+Aqua Leaf Area Index/FPAR Daily L4 Global 500m SIN Grid". In: *NASA LP DAAC*. URL: http://doi.org/10.5067/MODIS/MOD15A3H.006 (cit. on pp. 8, 9).

Rodell, M. et al. (Mar. 2004). "The Global Land Data Assimilation System". In: *bams* 85, pp. 381–394. DOI: 10.1175/BAMS-85-3-381 (cit. on p. 8).

Rui, H. and H. Beaudoing (2021). "Readme document for nasa gldas version 2 data products". In: *NASA/GSFC, Greenbelt, MD, USA, NASA Goddard Earth Sciences Data and Information Services Center (GES DISC)* (cit. on pp. 8, 52).

Service, S. T. (2020). *What is Transpiration in Trees? How Trees Use Transpiration*. URL: https://www.steintree.com/what-is-transpiration-in-trees-how-trees-use-transpiration/ (visited on 09/29/2022) (cit. on p. i).

Sokol, Z. et al. (2021). "The Role of Weather Radar in Rainfall Estimation and Its Application in Meteorological and Hydrological Modelling—A Review". In: *Remote Sensing* 13.3. ISSN: 2072-4292. DOI: 10.3390/rs13030351. URL: https://www.mdpi.com/2072-4292/13/3/351 (cit. on p. 49).

Taikan, O. and K. Shinjiro (2006). "Global Hydrological Cycles and World Water Resources". In: *Science* 313.5790, pp. 1068–1072. DOI: 10.1126/science.1128845. eprint: https://www.science.org/doi/pdf/10.1126/science.1128845. URL: https://www.science.org/doi/abs/10.1126/science.1128845 (cit. on p. 2).

Thornton, P. et al. (2014). *Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2*. en. DOI: 10.3334/ORNLDAAC/1219. URL: http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1219 (cit. on p. 7).

Woods, R. A. (2009). "Analytical model of seasonal climate impacts on snow hydrology: Continuous snowpacks". In: *Advances in Water Resources* 32.10, pp. 1465–1481. ISSN: 0309-1708. DOI: https://doi.org/10.1016/j.advwatres.2009.06.011. URL: https://www.sciencedirect.com/science/article/pii/S030917080900102X (cit. on p. 5).

Xia, Y. et al. (2012). "Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products". English (US). In: *Journal of Geophysical Research: Oceans* 117.3. Copyright: Copyright 2018 Elsevier B.V., All rights reserved. ISSN: 0148-0227. DOI: 10.1029/2011JD016048 (cit. on p. 7).

Appendix A: Leaf area index and determination of growing stages

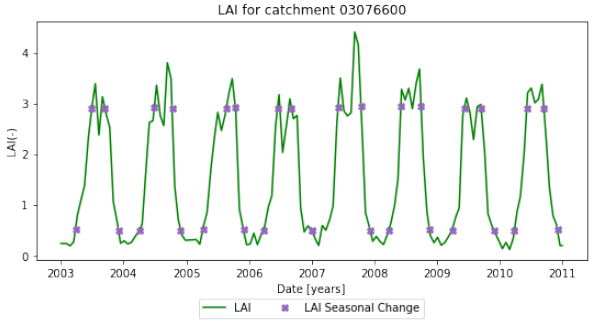


Figure A.1: LAI and growing stages catchment C1

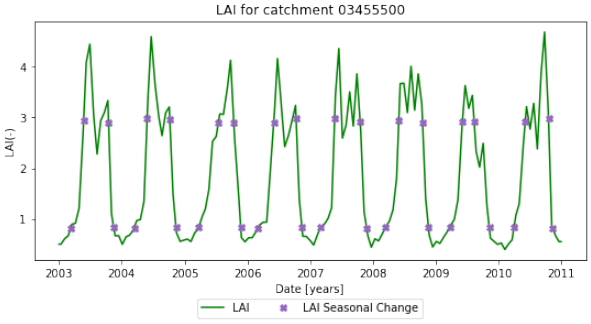


Figure A.2: LAI and growing stages catchment C2

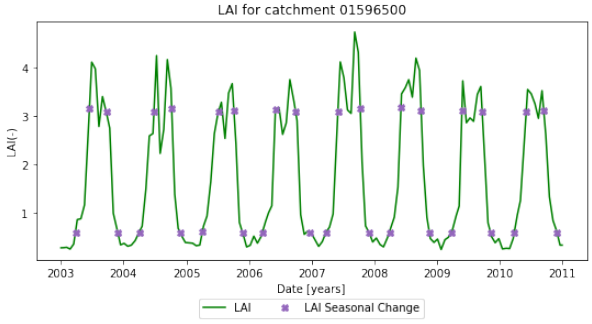


Figure A.3: LAI and growing stages catchment C3

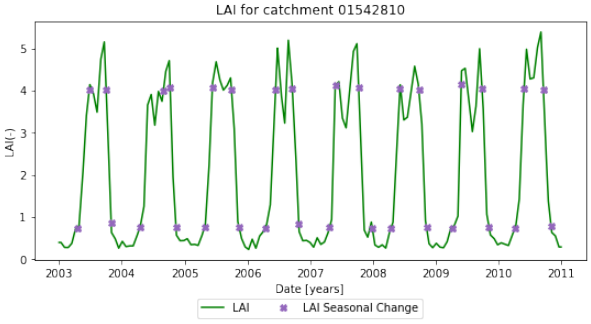


Figure A.4: LAI and growing stages catchment C4

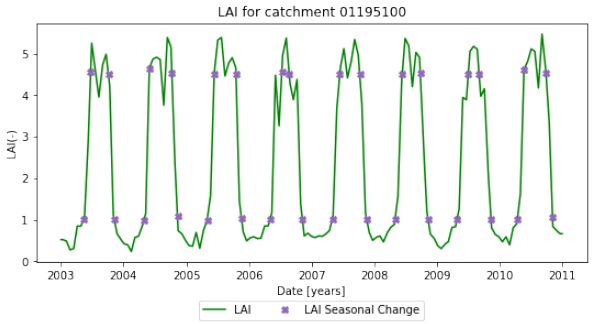


Figure A.5: LAI and growing stages catchment C5

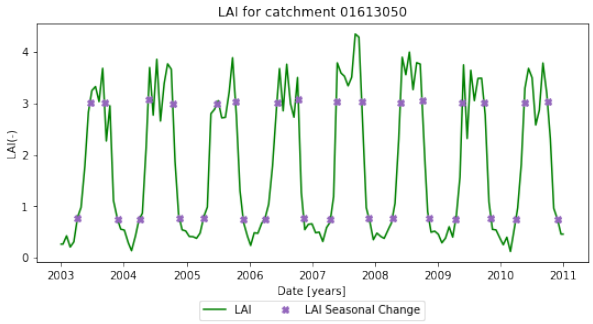


Figure A.6: LAI and growing stages catchment C6

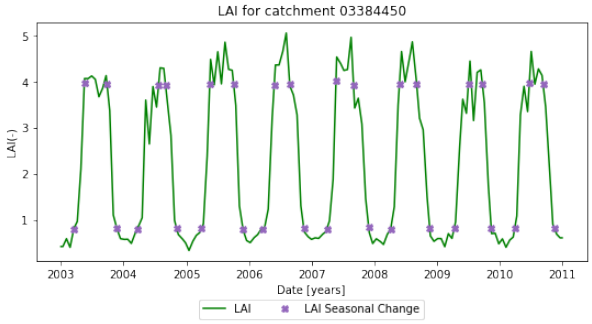


Figure A.7: LAI and growing stages catchment C7

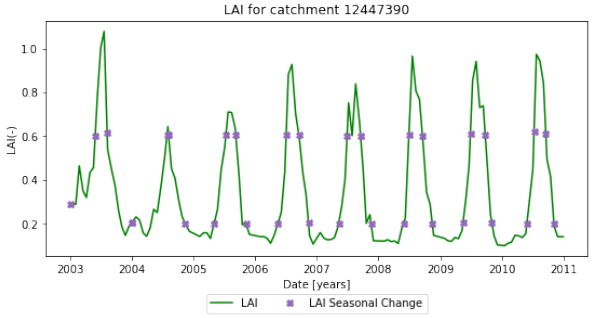


Figure A.8: LAI and growing stages catchment C8

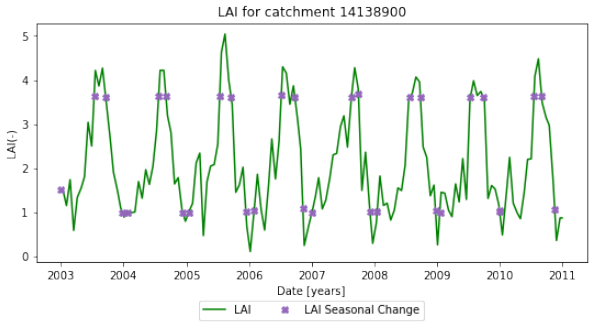


Figure A.9: LAI and growing stages catchment C9

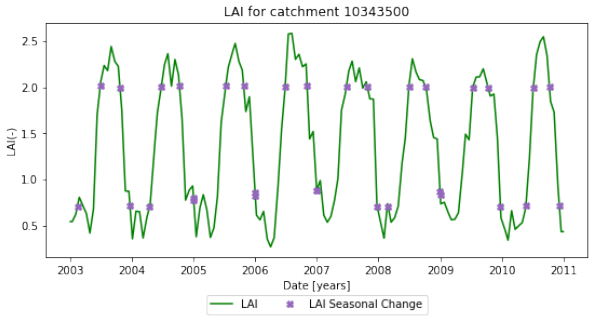


Figure A.10: LAI and growing stages catchment C10

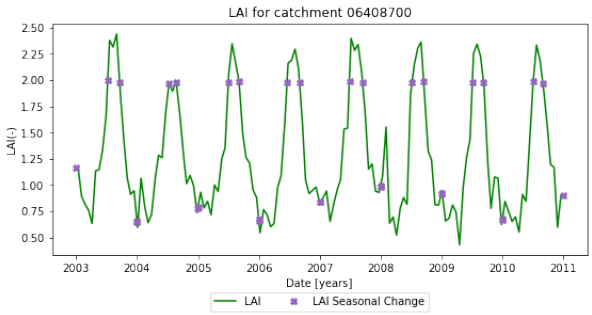


Figure A.11: LAI and growing stages catchment C11

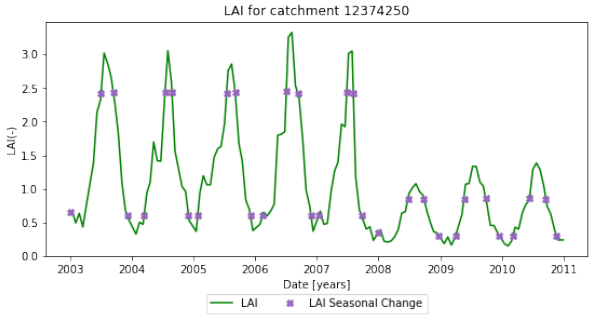


Figure A.12: LAI and growing stages catchment C12

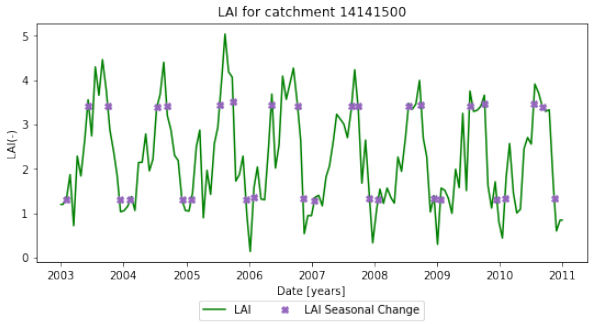


Figure A.13: LAI and growing stages catchment C13

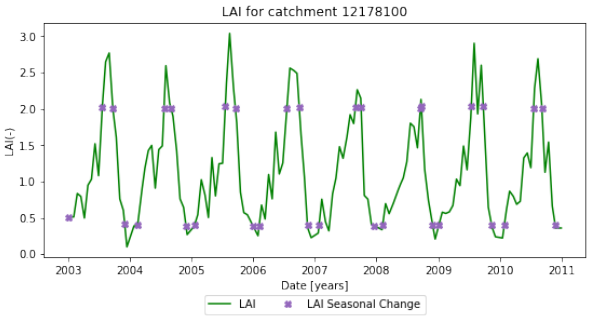


Figure A.14: LAI and growing stages catchment C14

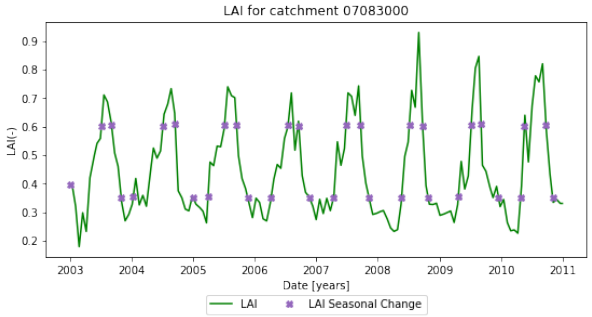


Figure A.15: LAI and growing stages catchment C15

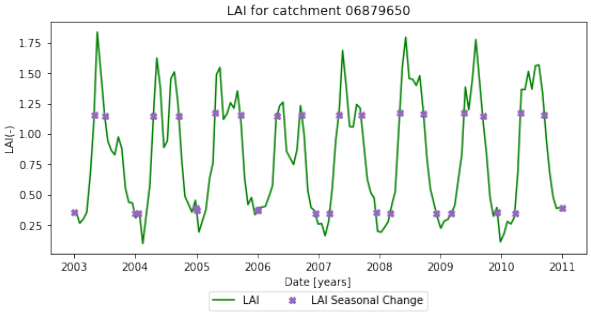


Figure A.16: LAI and growing stages catchment C16

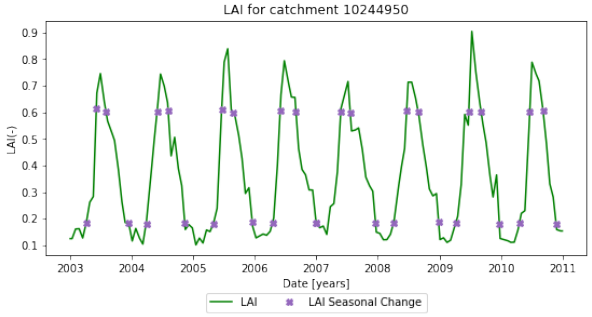


Figure A.17: LAI and growing stages catchment C17

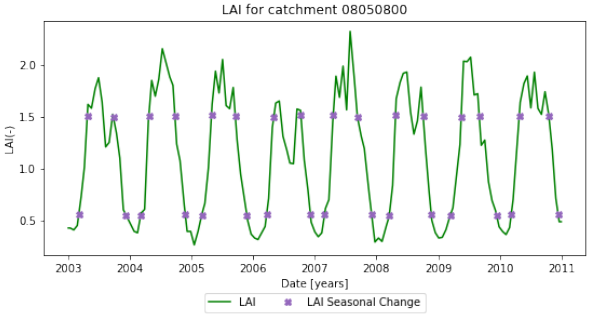


Figure A.18: LAI and growing stages catchment C18

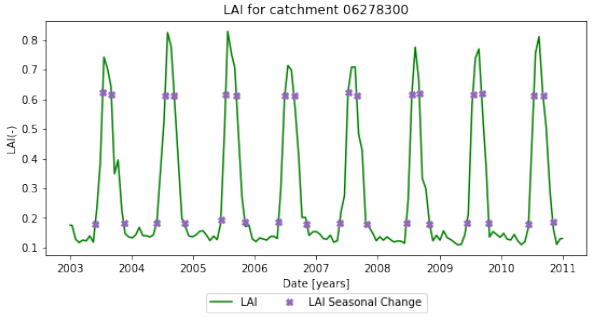


Figure A.19: LAI and growing stages catchment C19

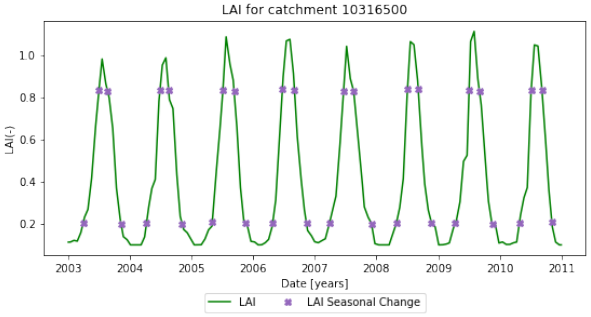


Figure A.20: LAI and growing stages catchment C20

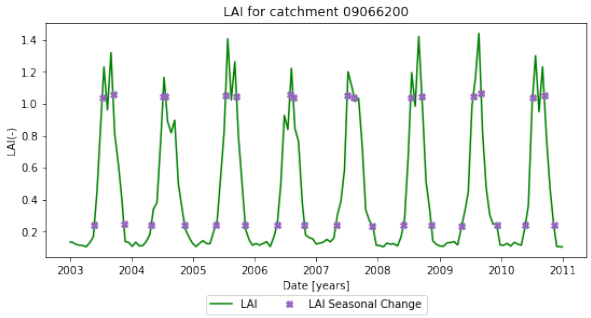


Figure A.21: LAI and growing stages catchment C21

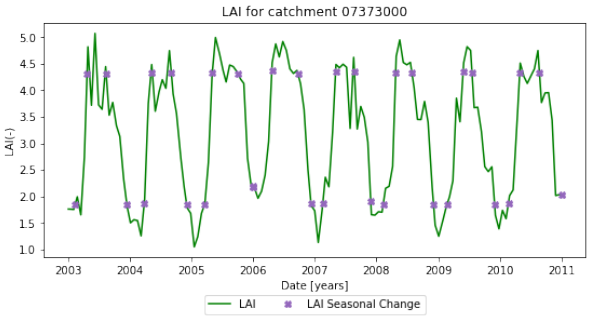


Figure A.22: LAI and growing stages catchment C22

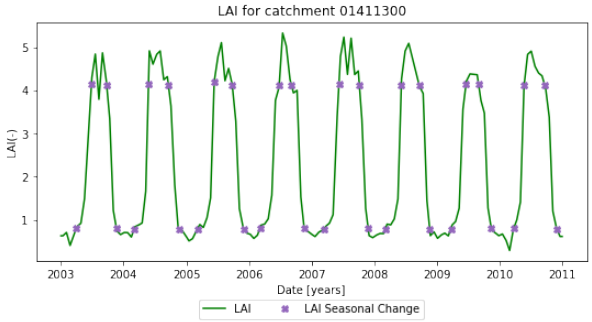


Figure A.23: LAI and growing stages catchment C23

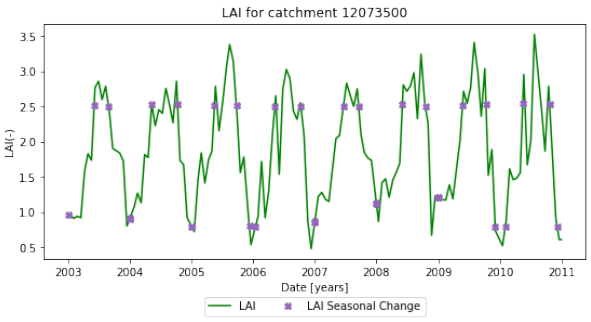


Figure A.24: LAI and growing stages catchment C24

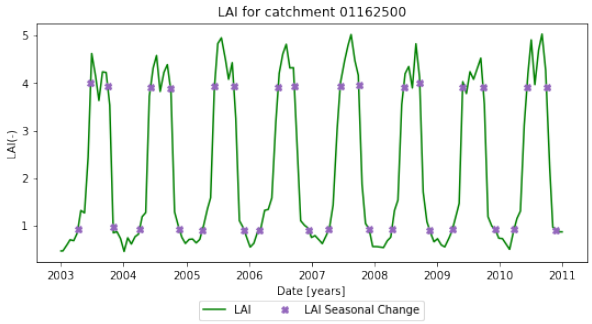


Figure A.25: LAI and growing stages catchment C25

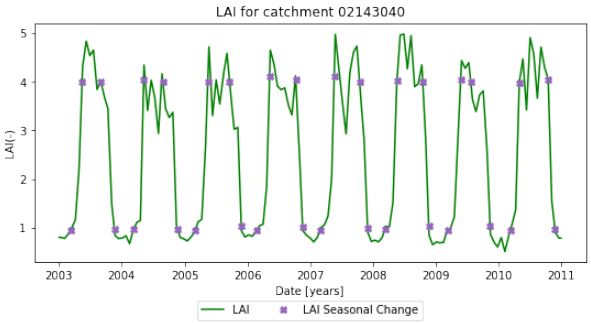


Figure A.26: LAI and growing stages catchment C26

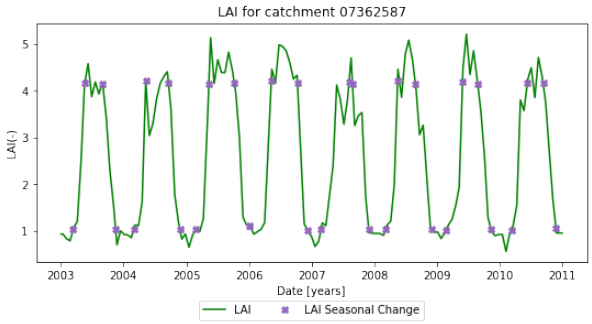


Figure A.27: LAI and growing stages catchment C27

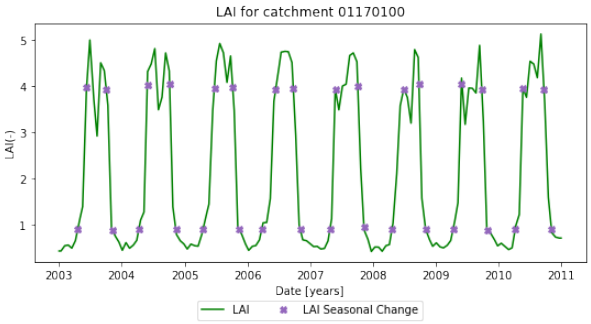


Figure A.28: LAI and growing stages catchment C28

B

Appendix B: Streamflow simulations

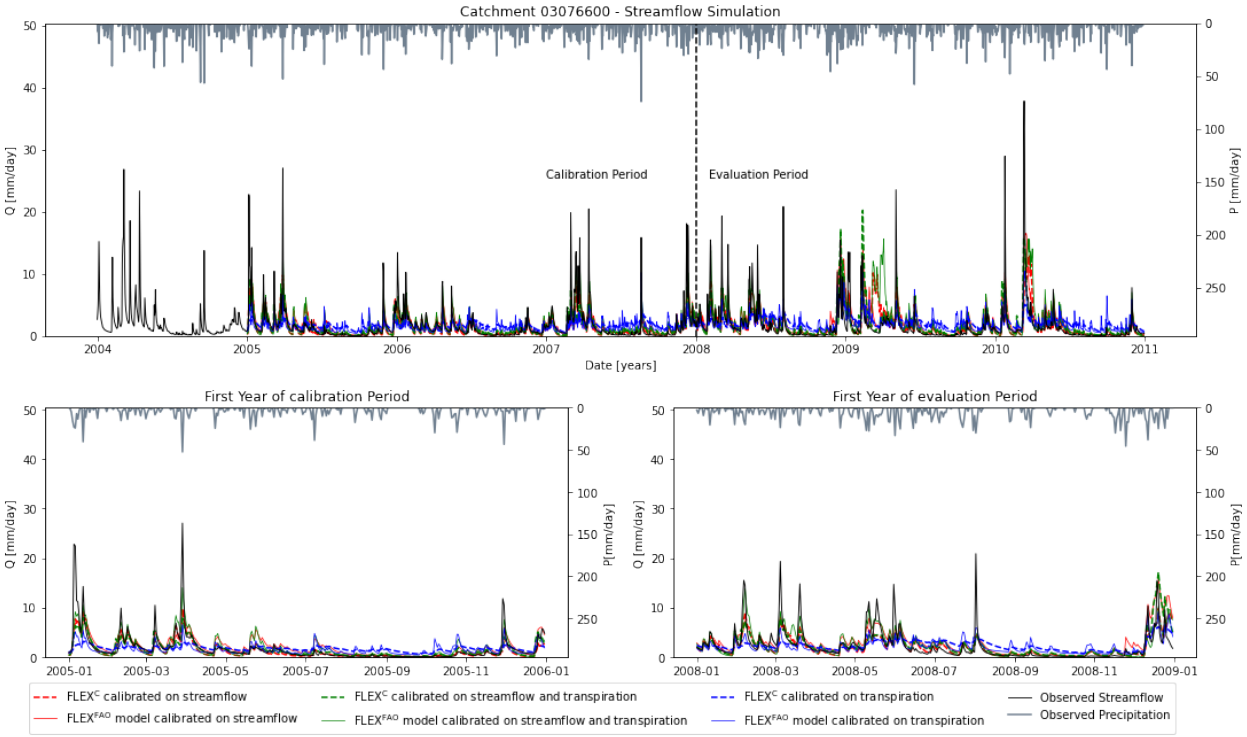


Figure B.1: Catchment C1 streamflow simulation

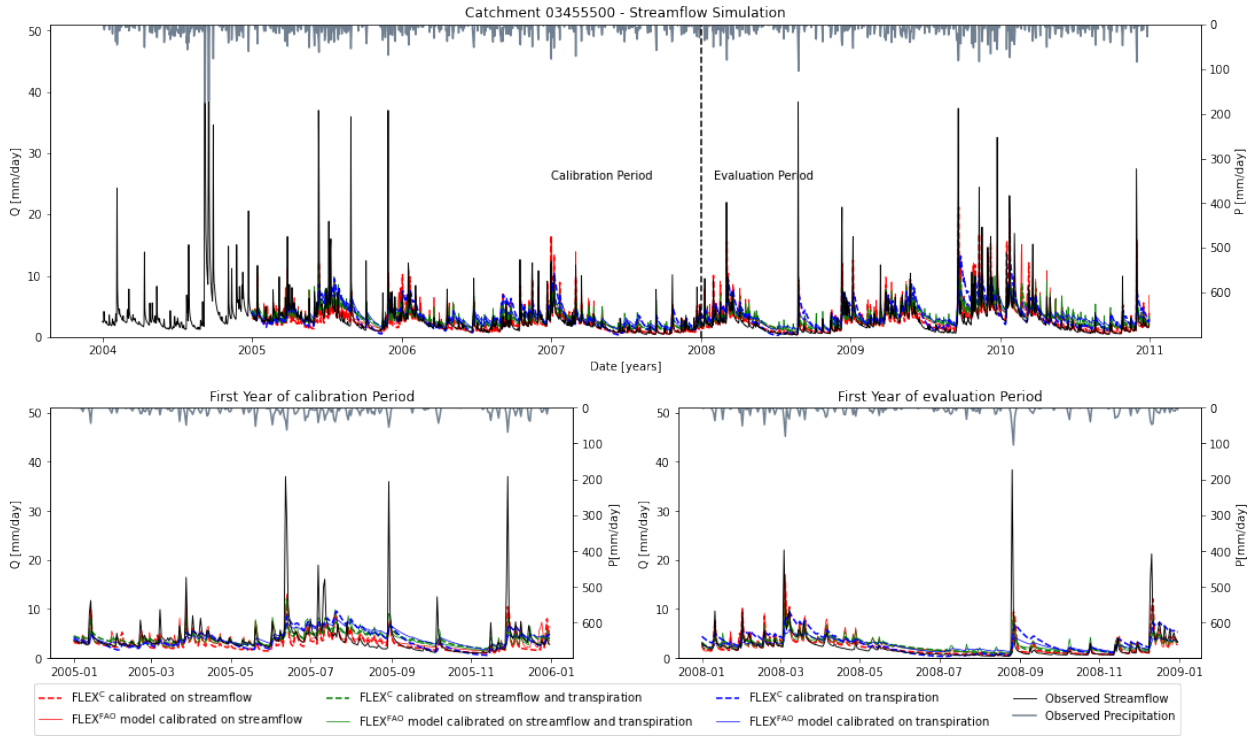


Figure B.2: Catchment C2 streamflow simulation

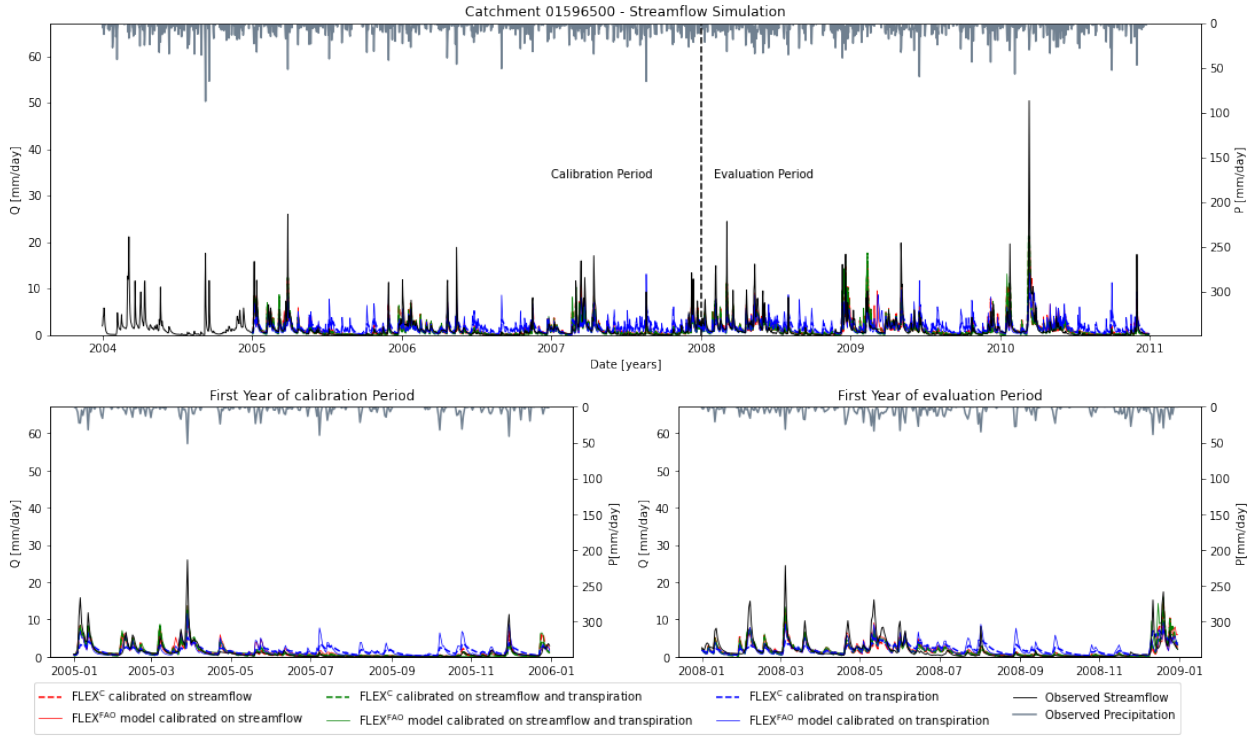


Figure B.3: Catchment C3 streamflow simulation

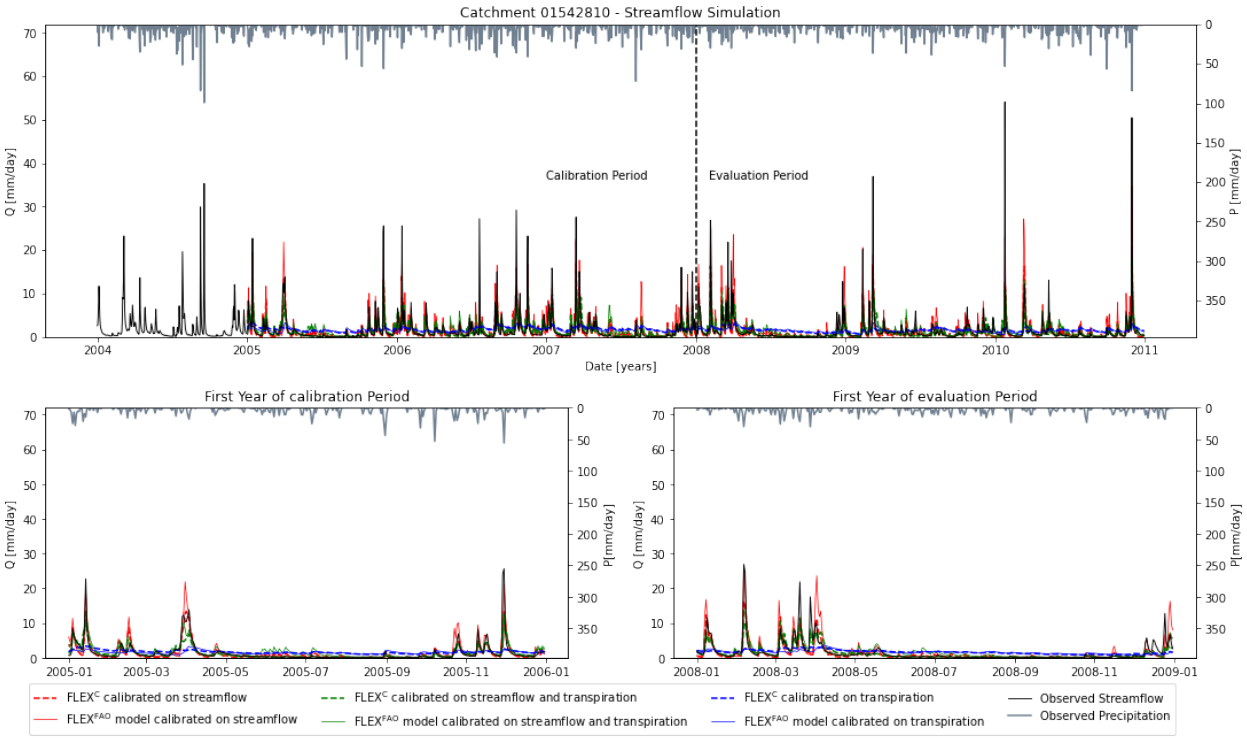


Figure B.4: Catchment C4 streamflow simulation

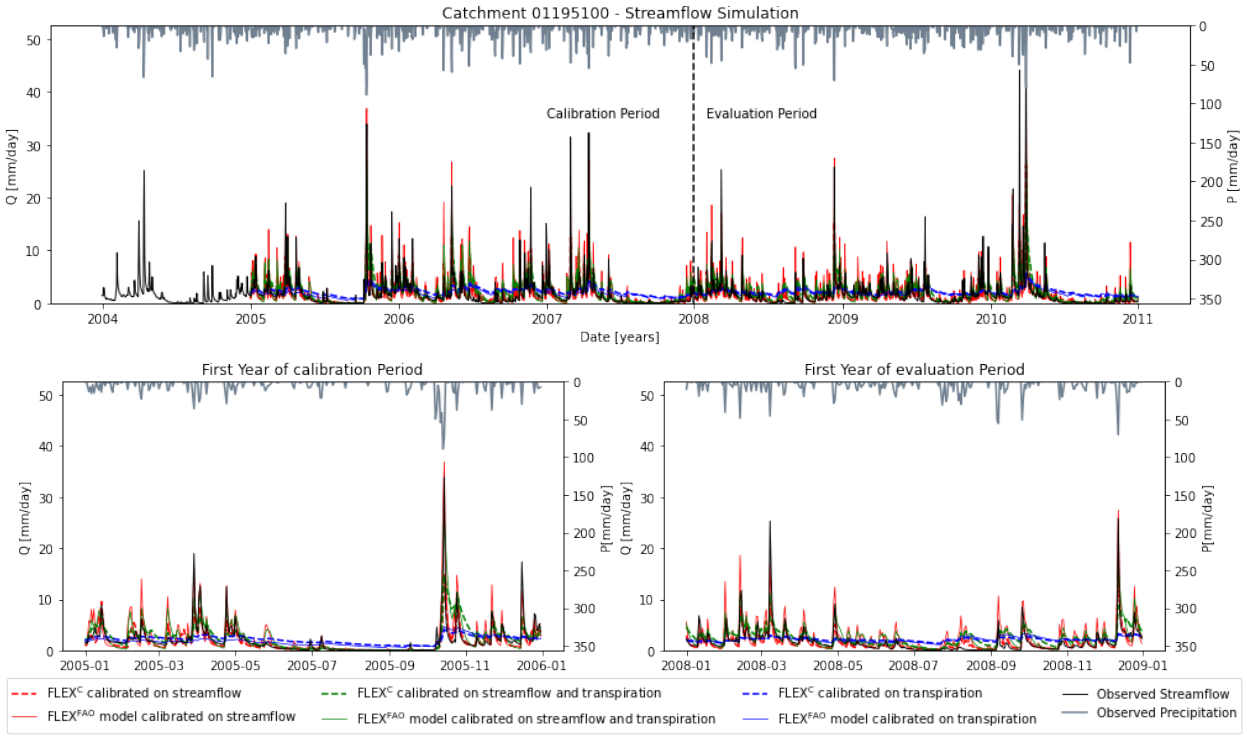


Figure B.5: Catchment C5 streamflow simulation

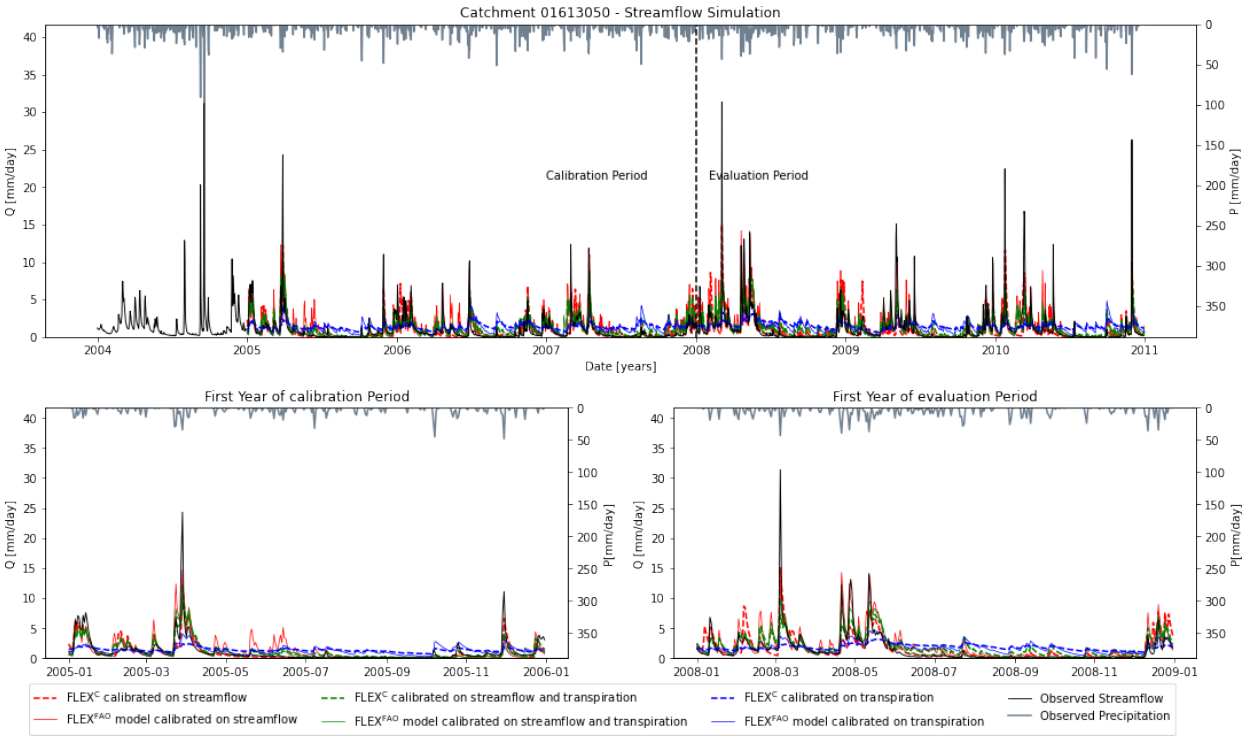


Figure B.6: Catchment C6 streamflow simulation

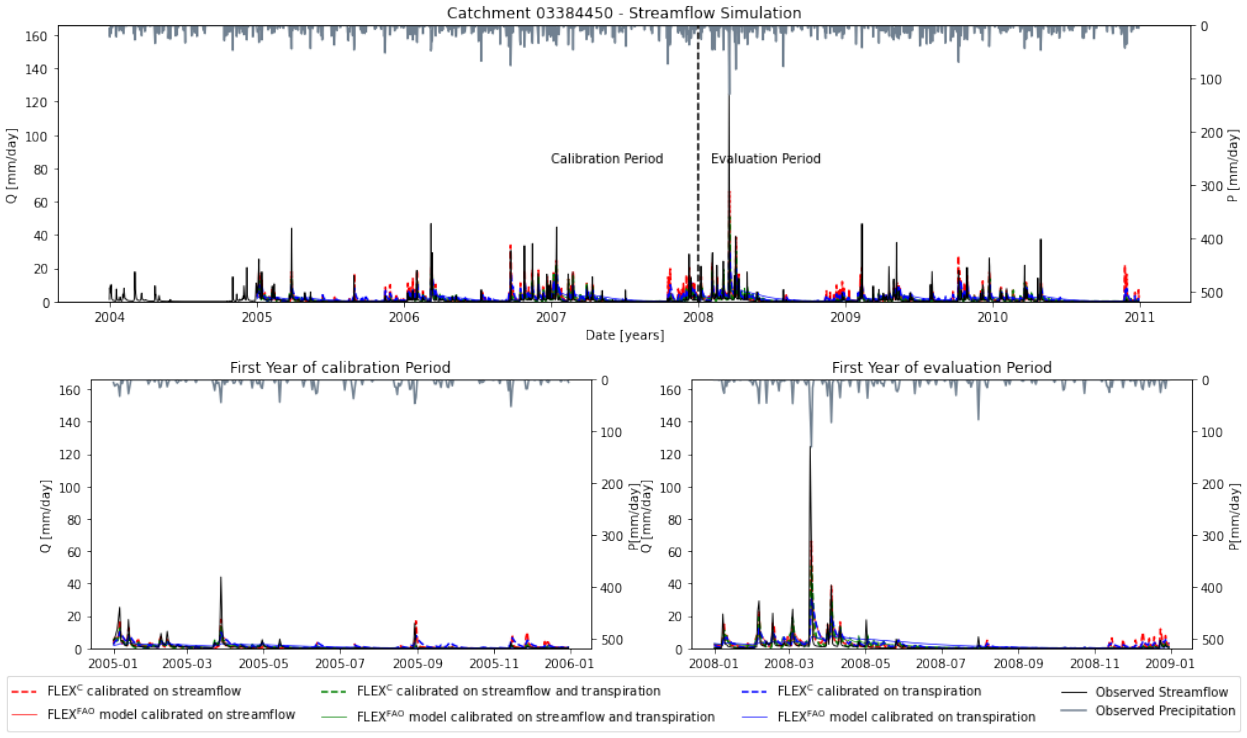


Figure B.7: Catchment C7 streamflow simulation

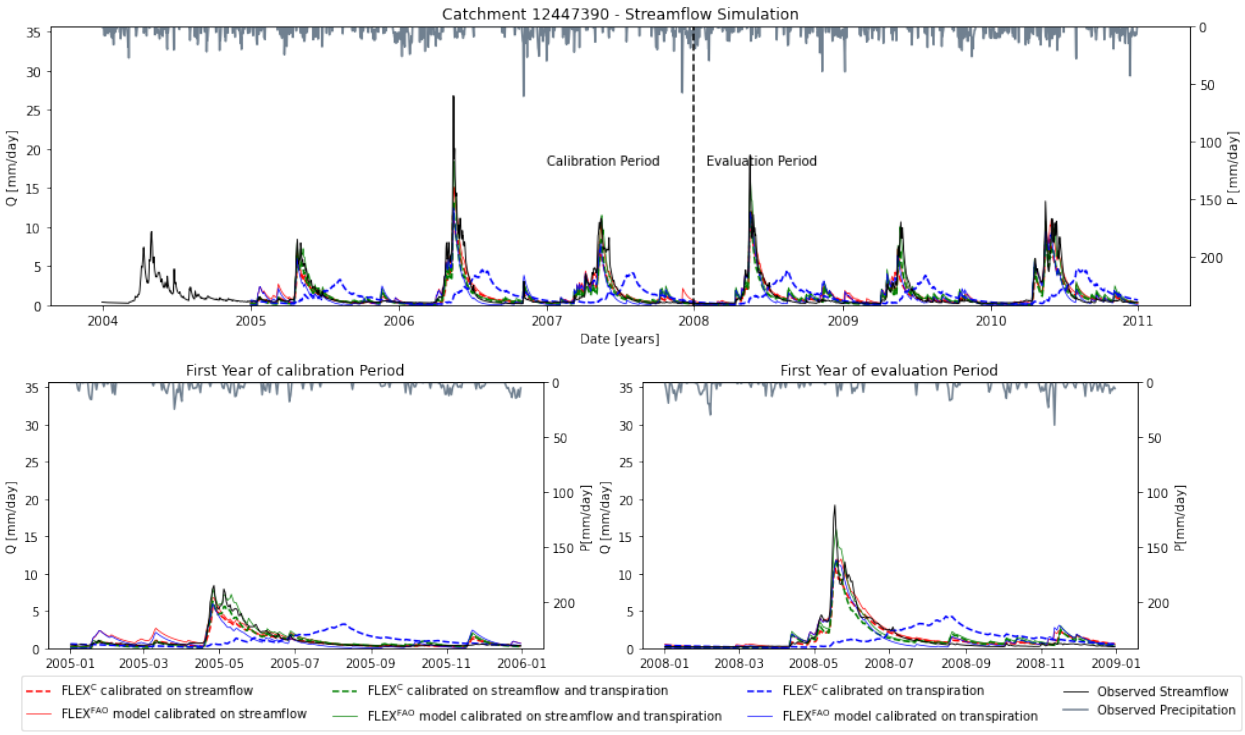


Figure B.8: Catchment C8 streamflow simulation

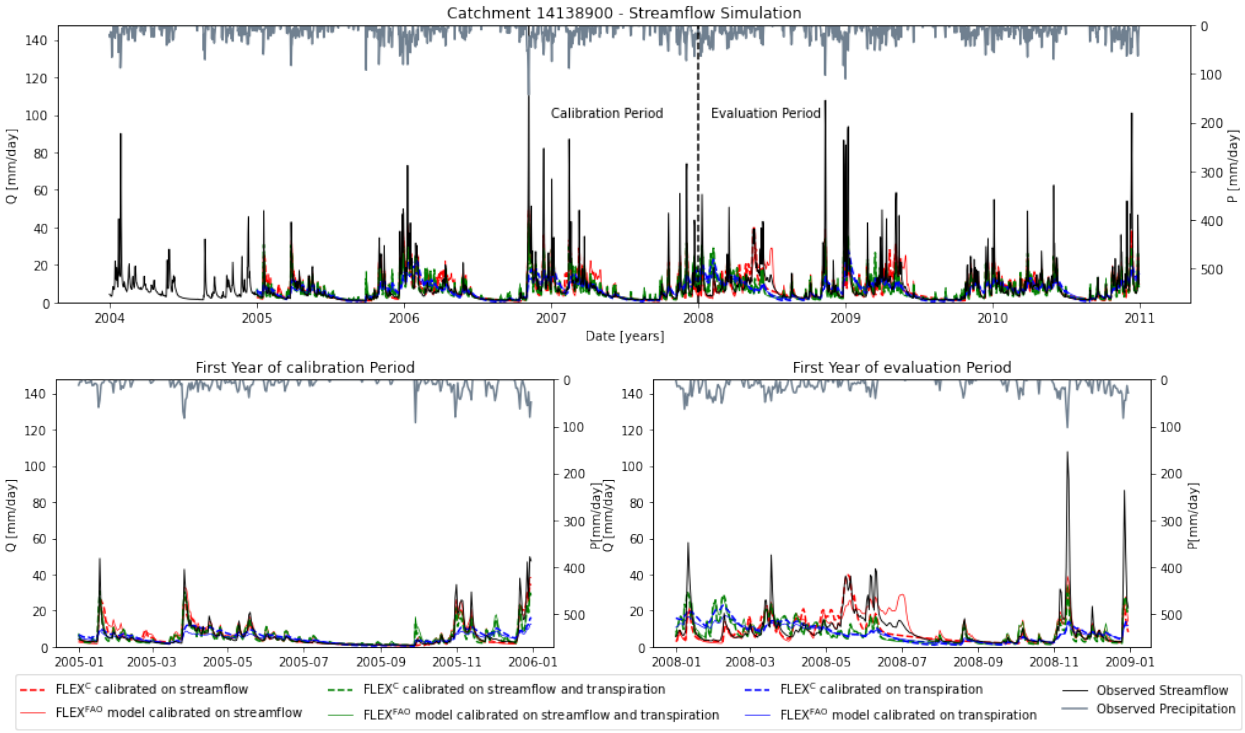


Figure B.9: Catchment C9 streamflow simulation

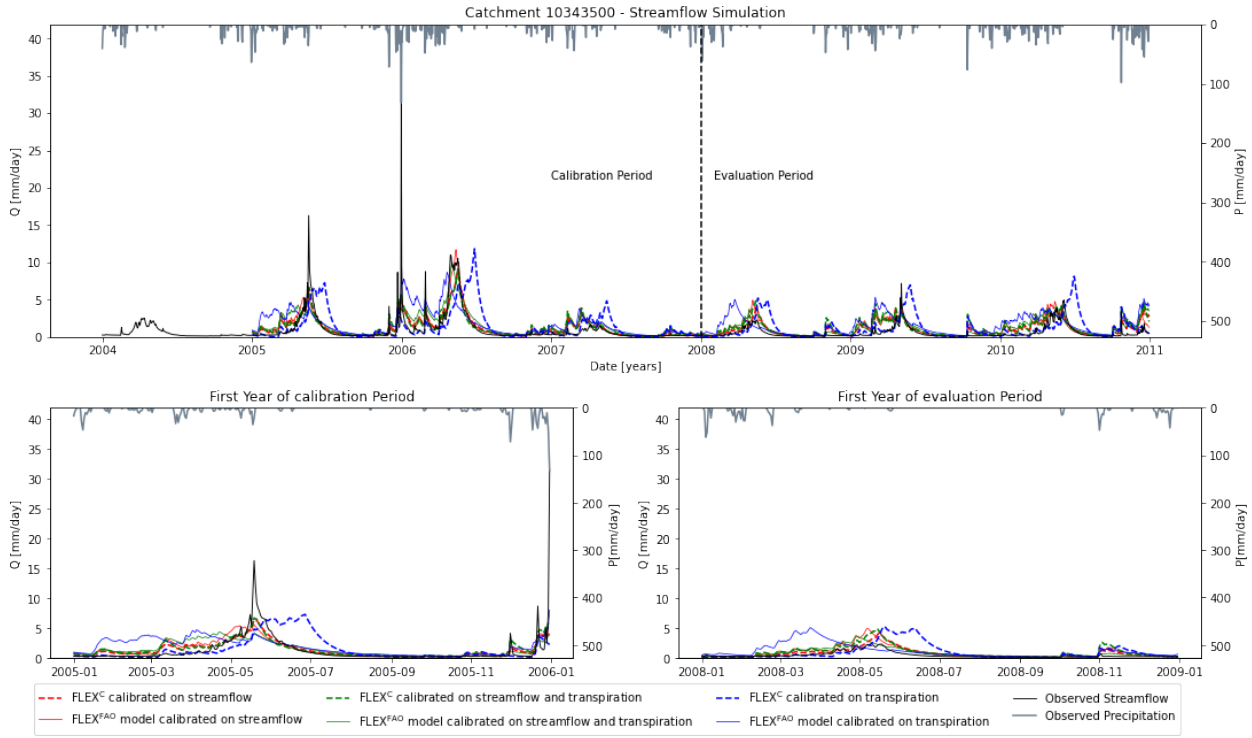


Figure B.10: Catchment C10 streamflow simulation

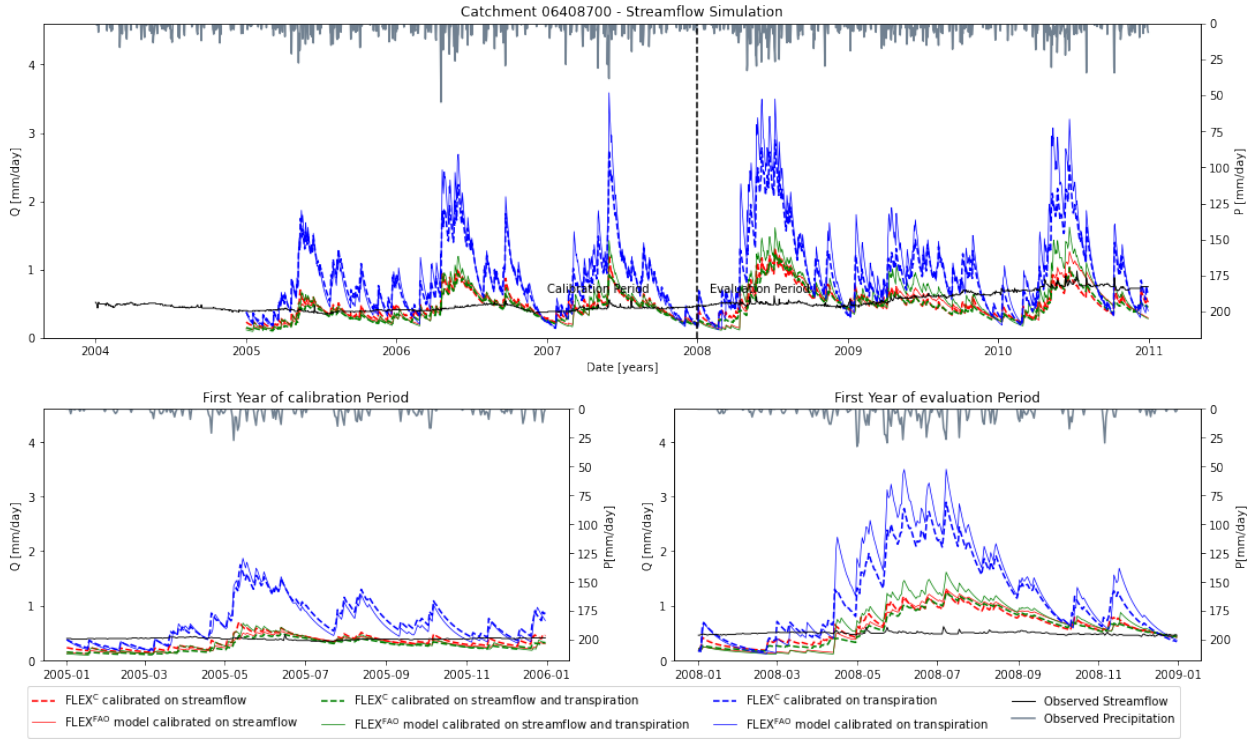


Figure B.11: Catchment C11 streamflow simulation

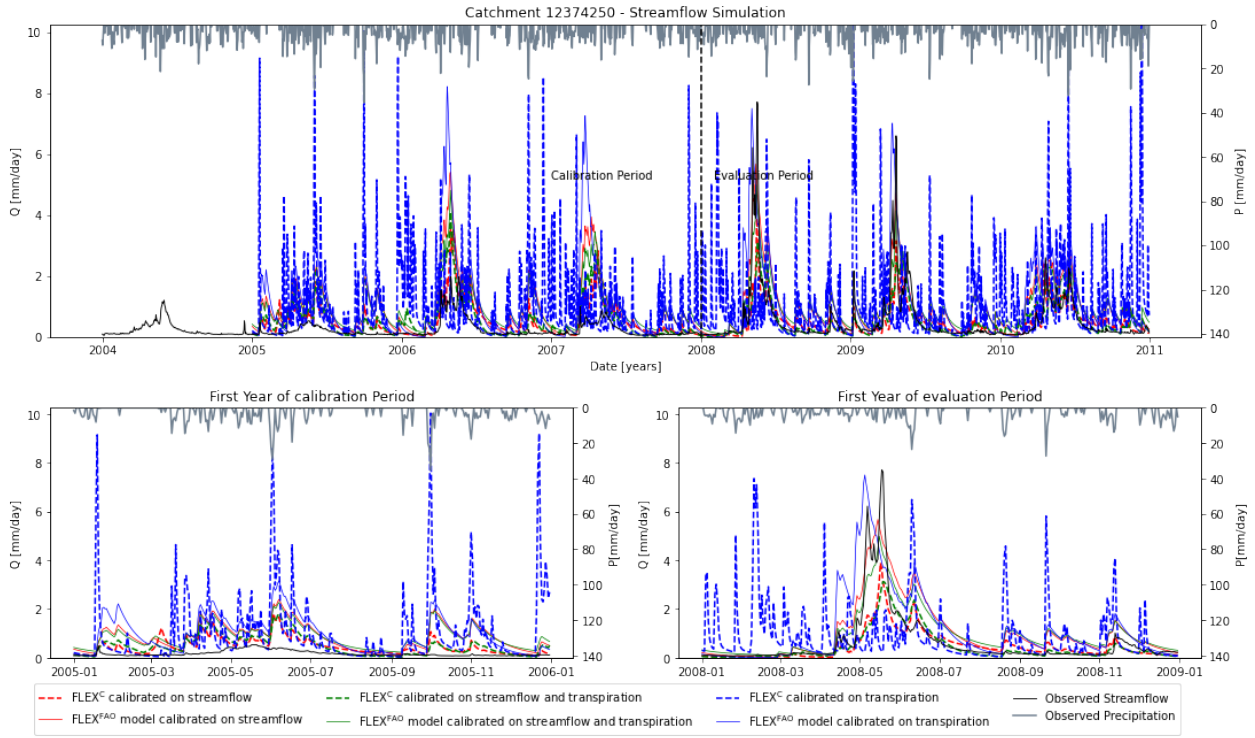


Figure B.12: Catchment C12 streamflow simulation

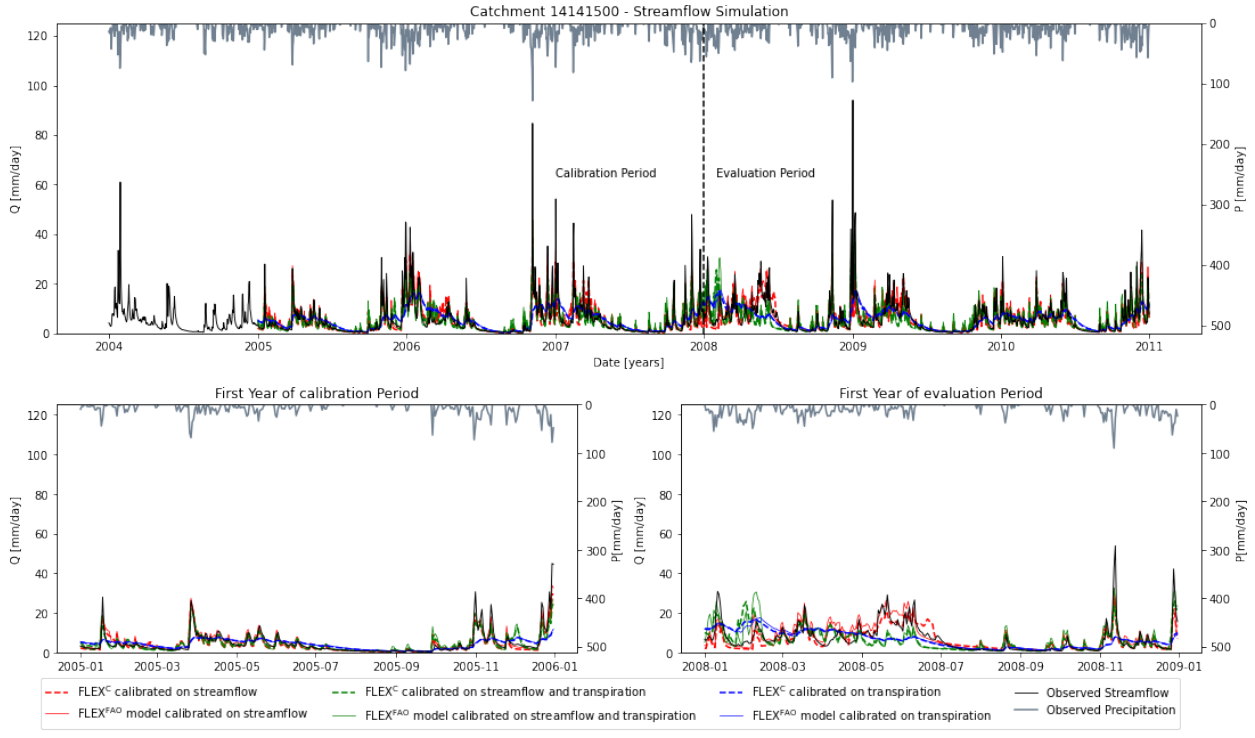


Figure B.13: Catchment C13 streamflow simulation

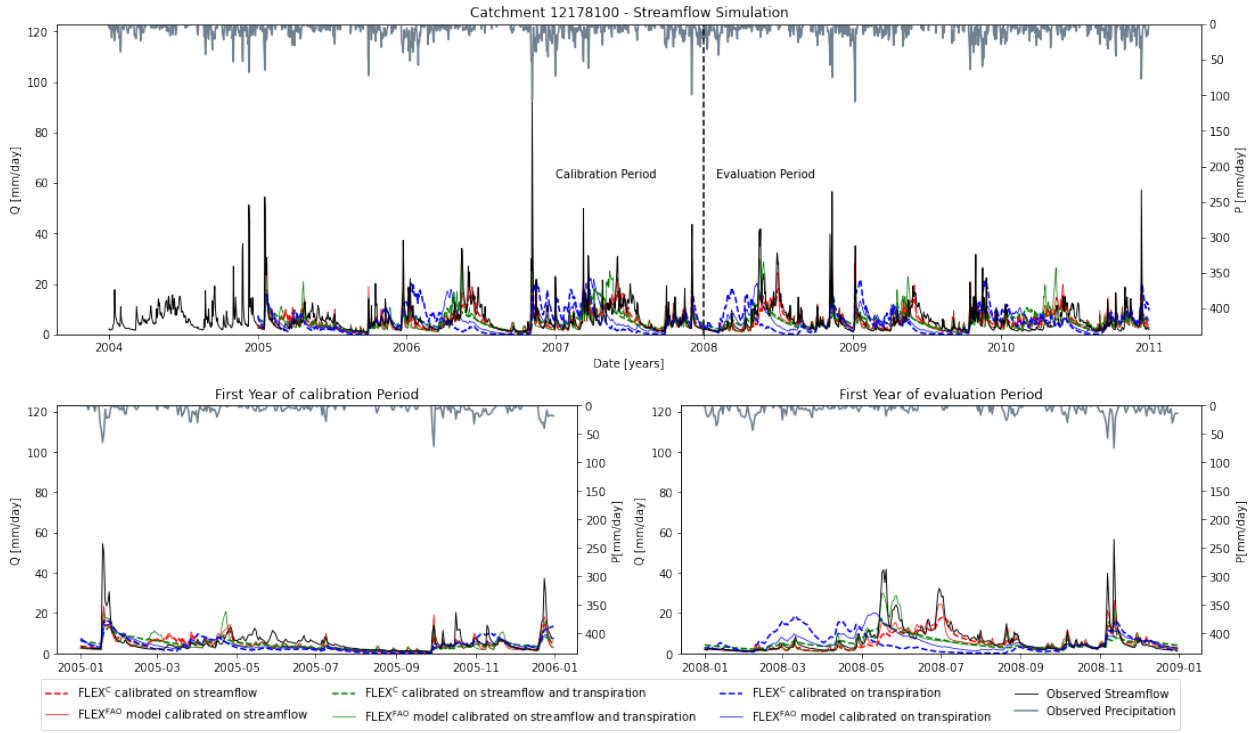


Figure B.14: Catchment C14 streamflow simulation

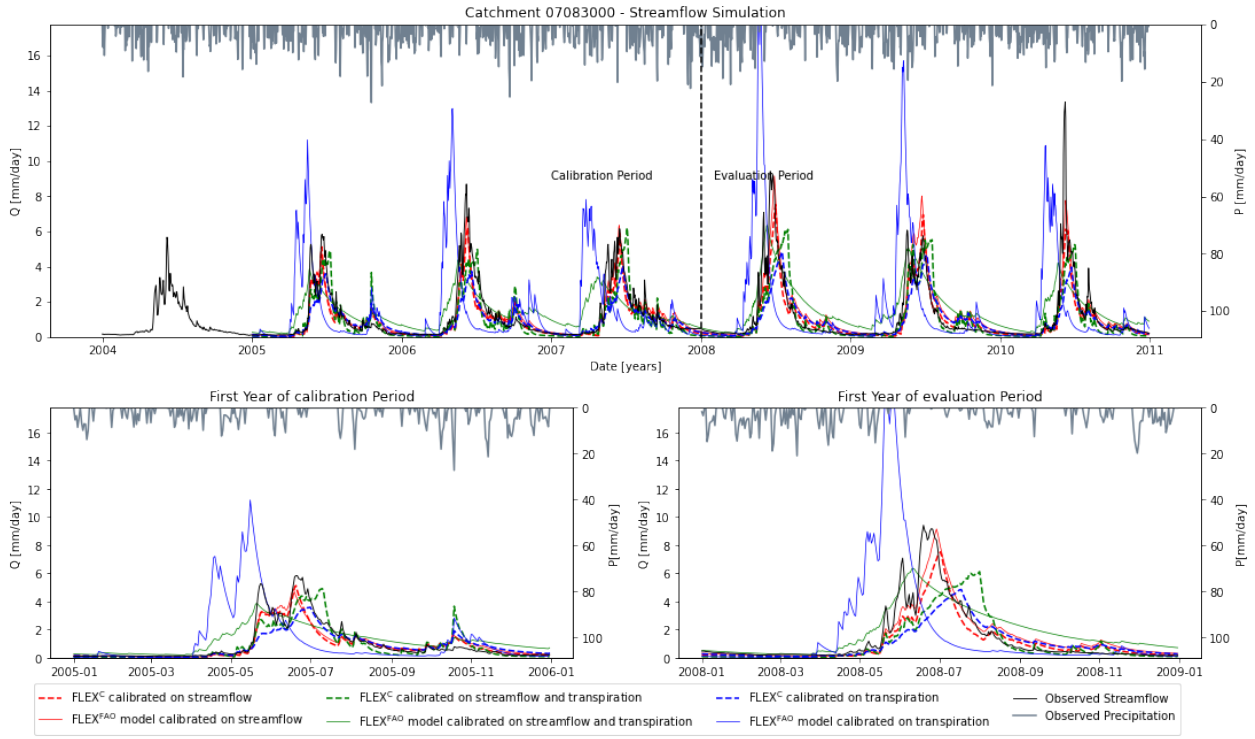


Figure B.15: Catchment C15 streamflow simulation

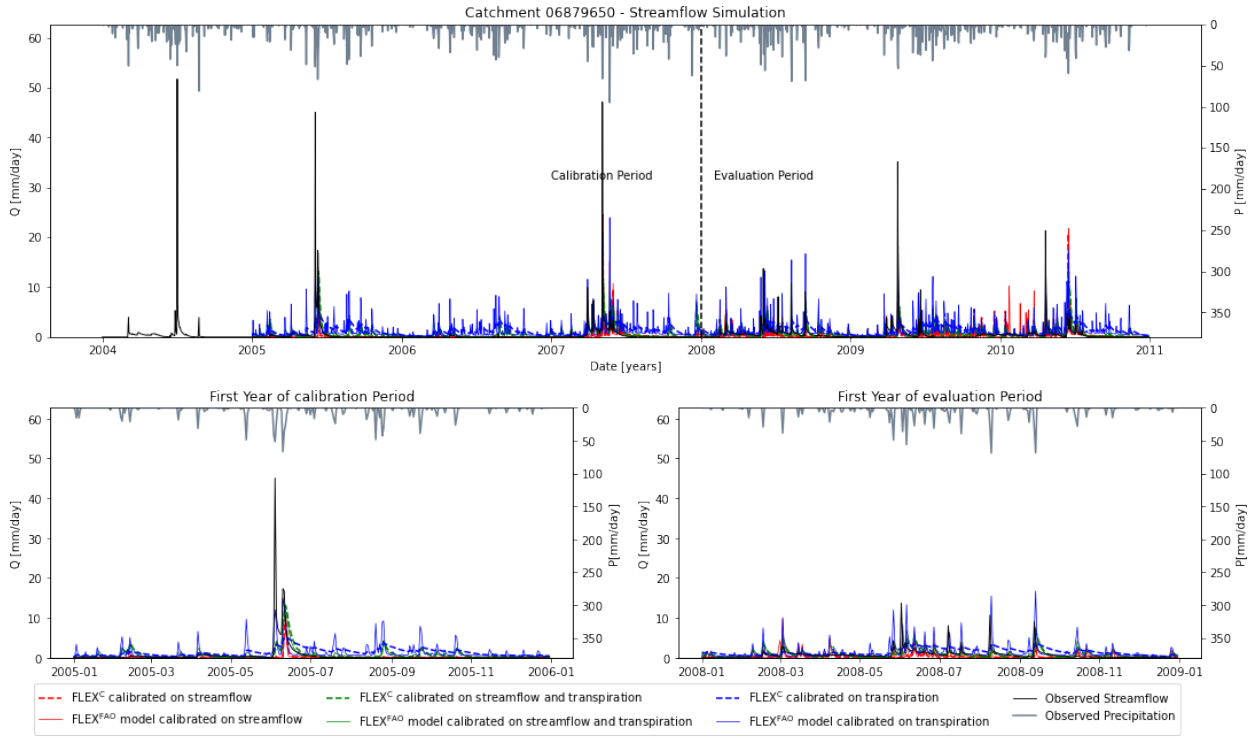


Figure B.16: Catchment C16 streamflow simulation

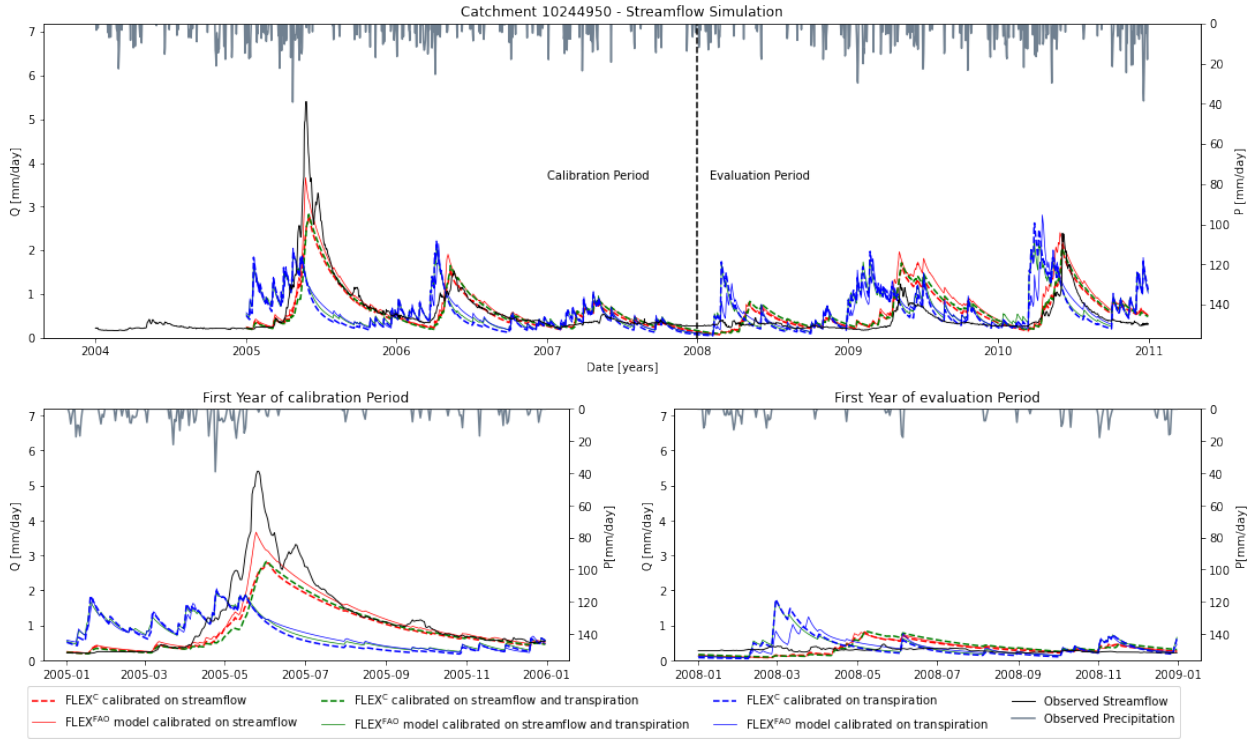


Figure B.17: Catchment C17 streamflow simulation

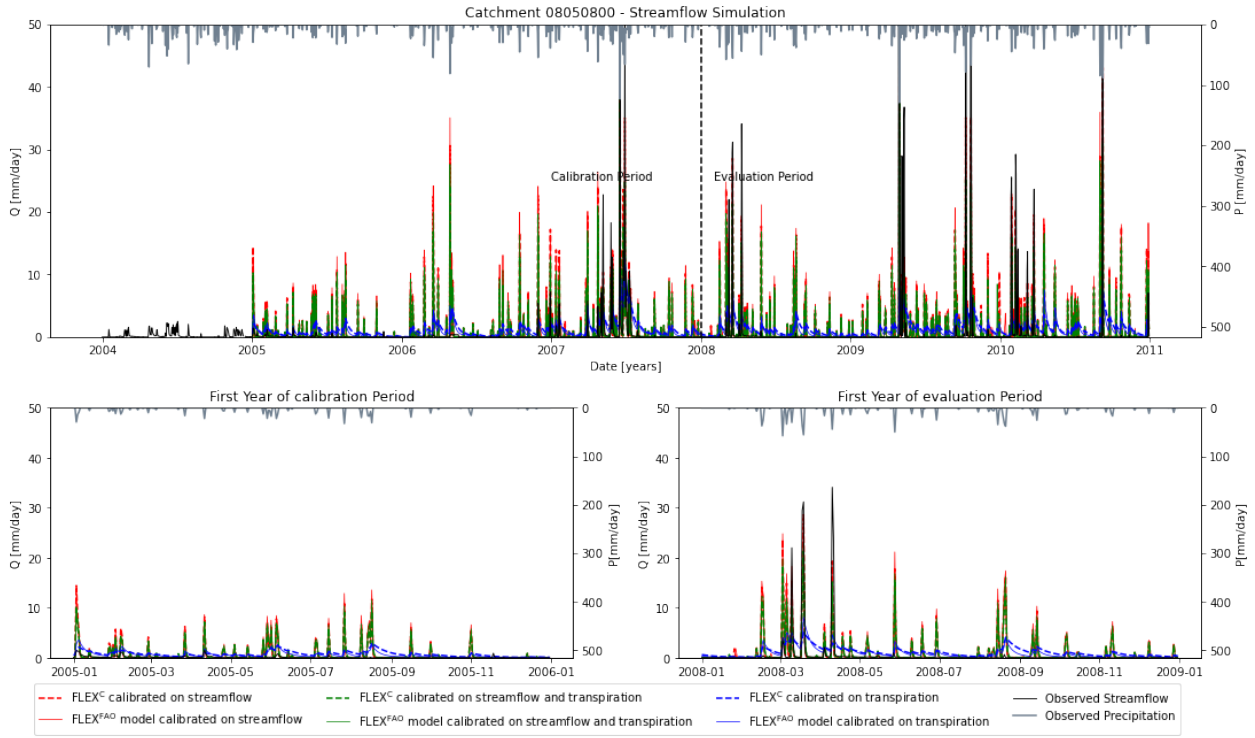


Figure B.18: Catchment C18 streamflow simulation

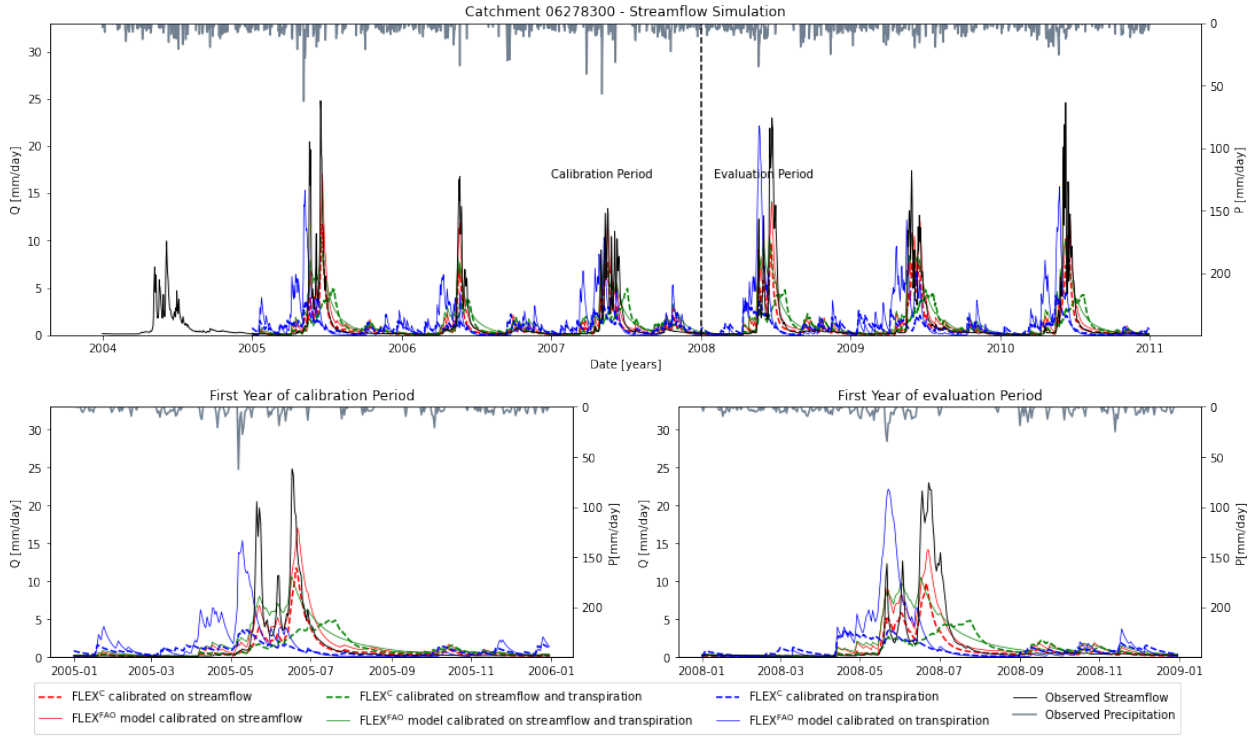


Figure B.19: Catchment C19 streamflow simulation

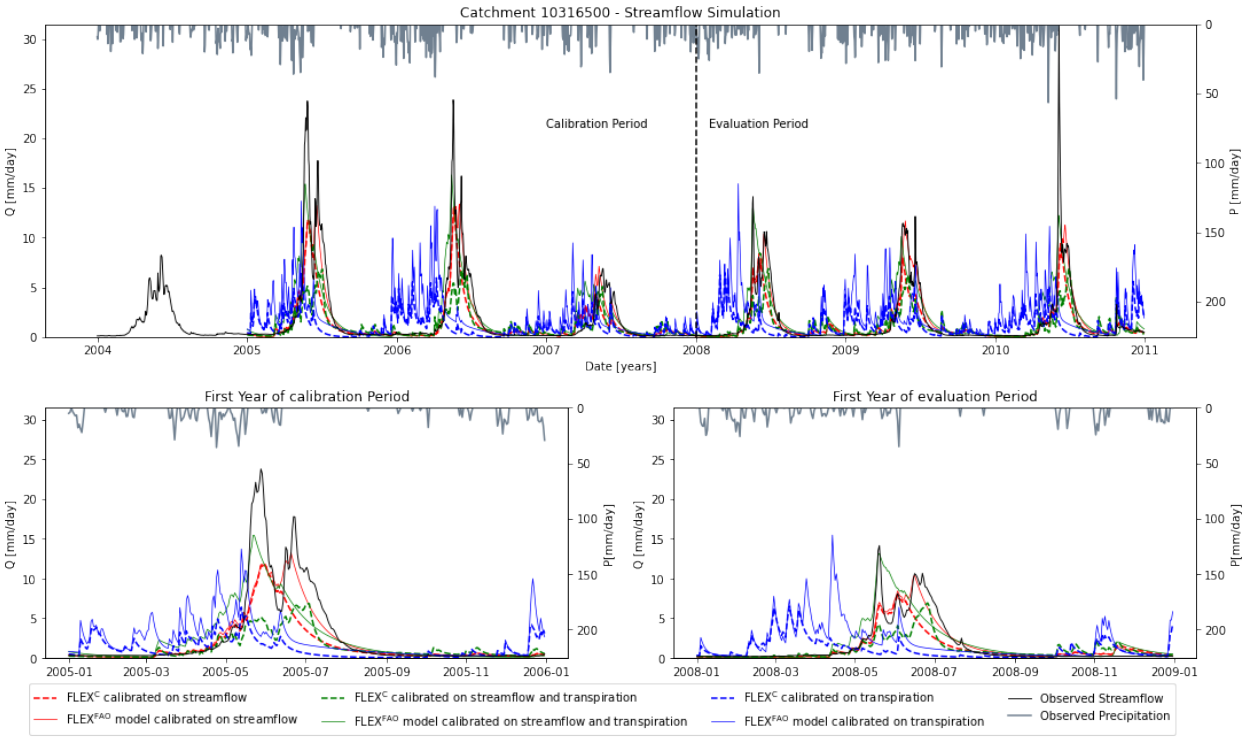


Figure B.20: Catchment C20 streamflow simulation

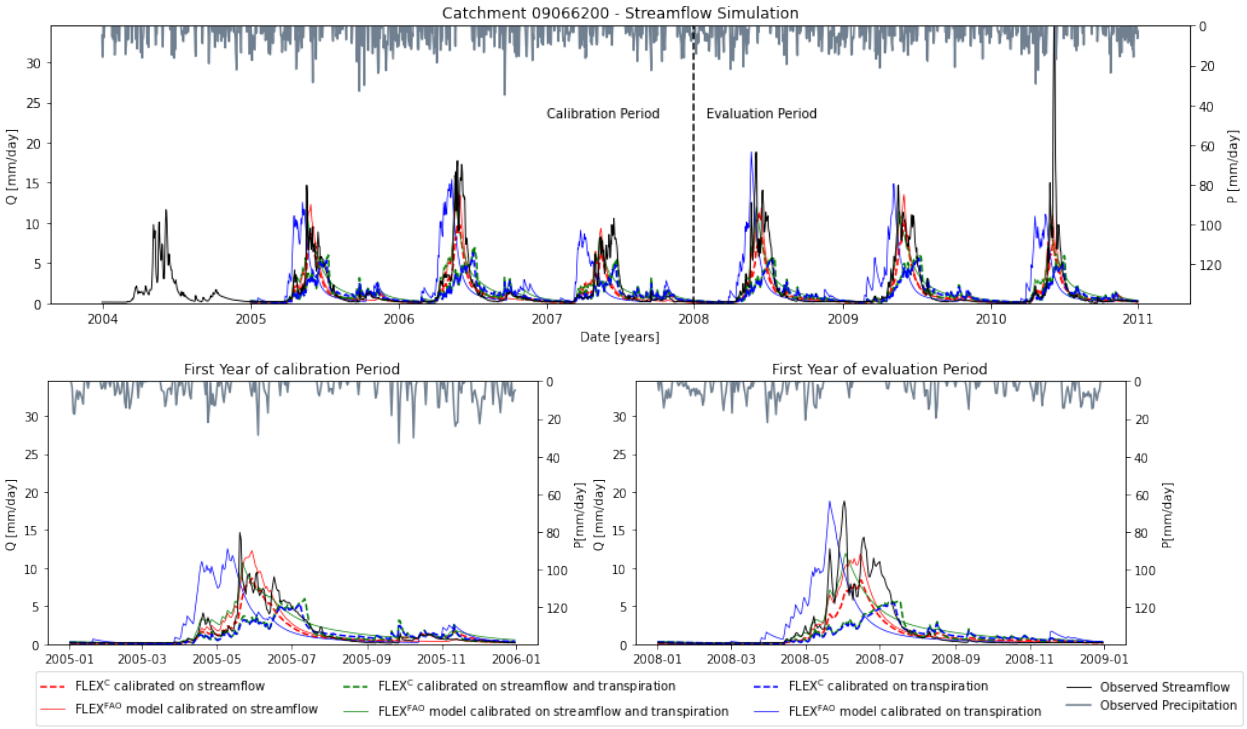


Figure B.21: Catchment C21 streamflow simulation

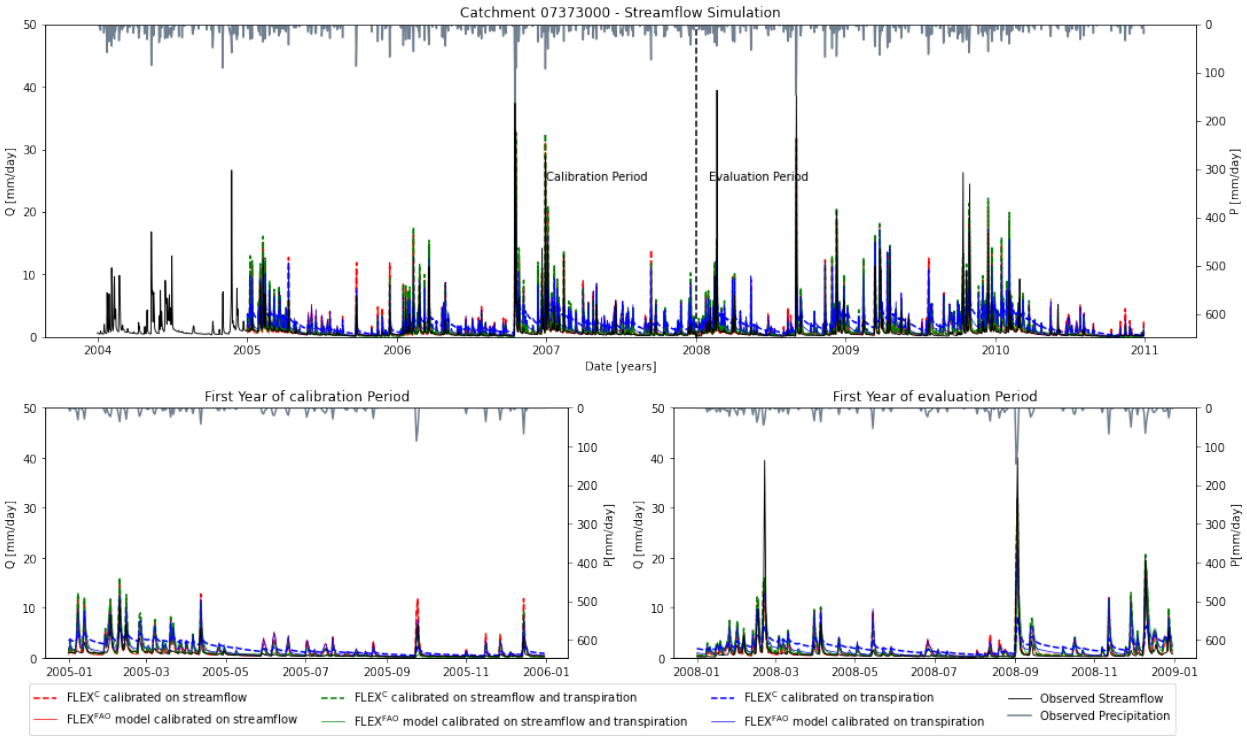


Figure B.22: Catchment C22 streamflow simulation

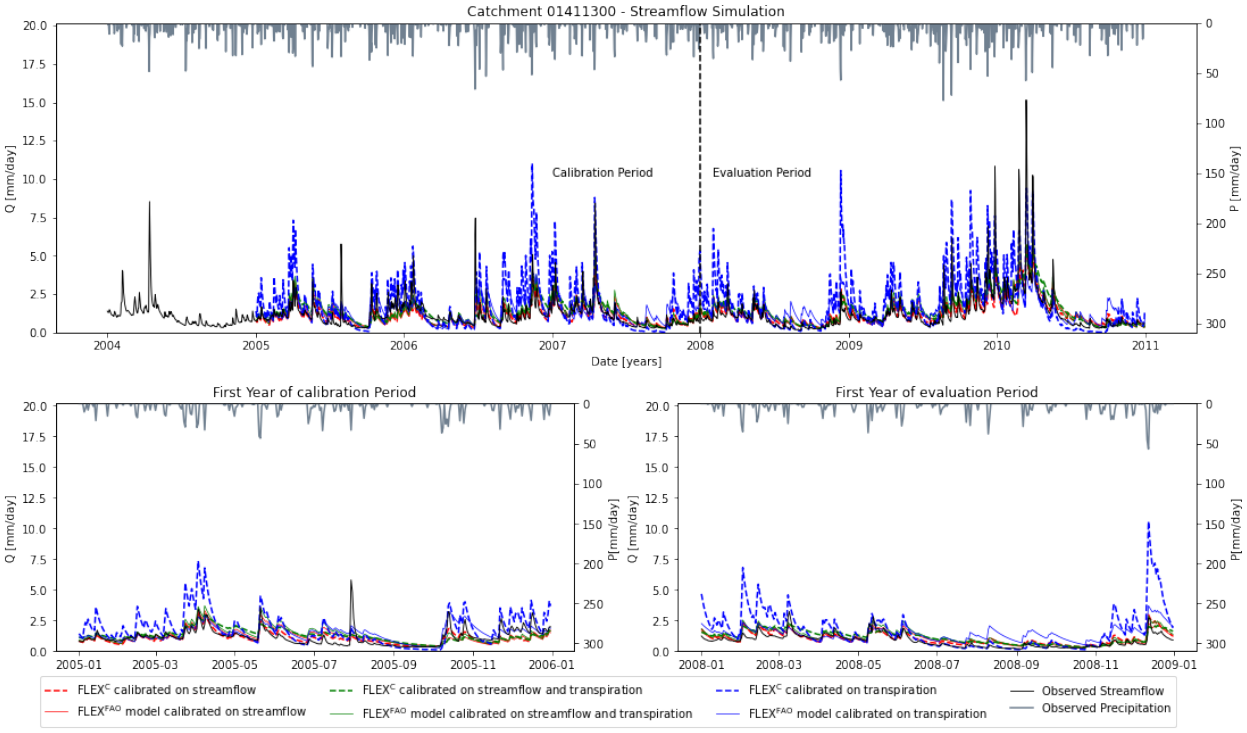


Figure B.23: Catchment C23 streamflow simulation

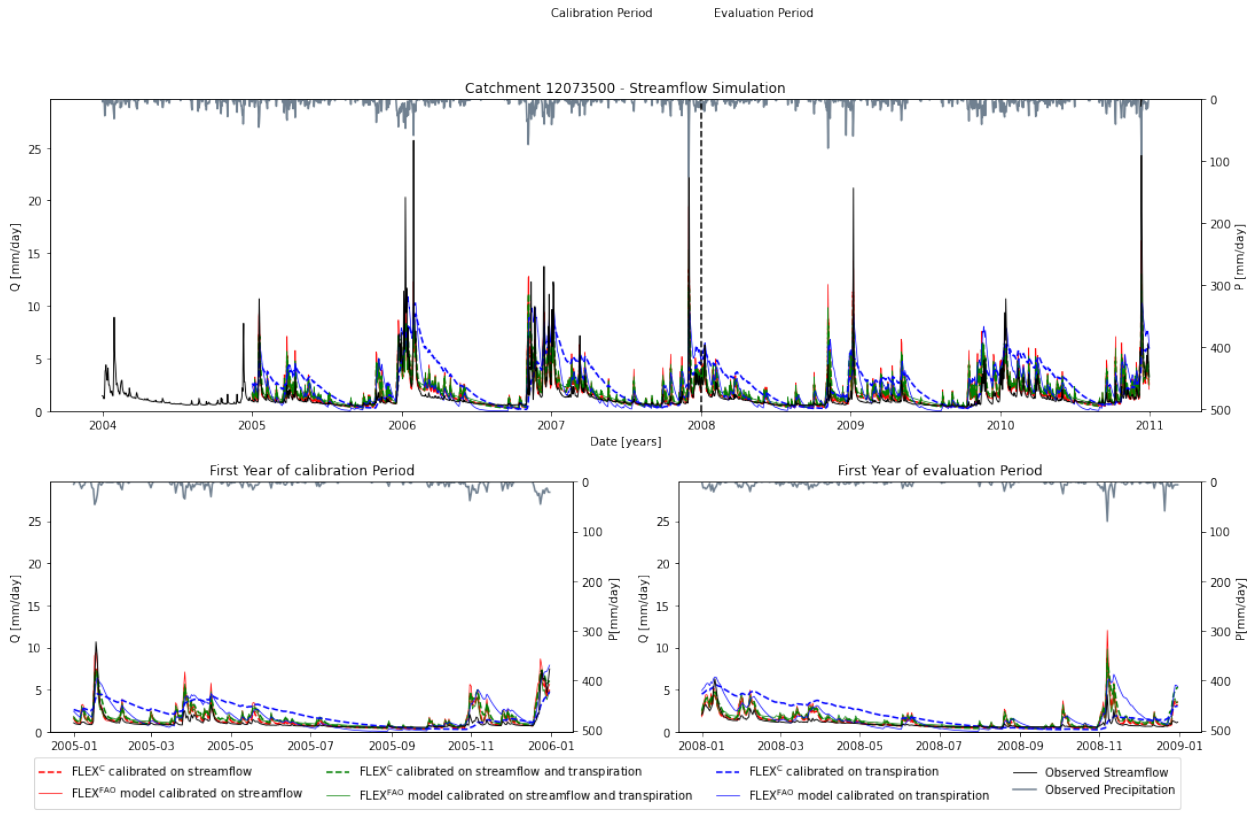


Figure B.24: Catchment C24 streamflow simulation

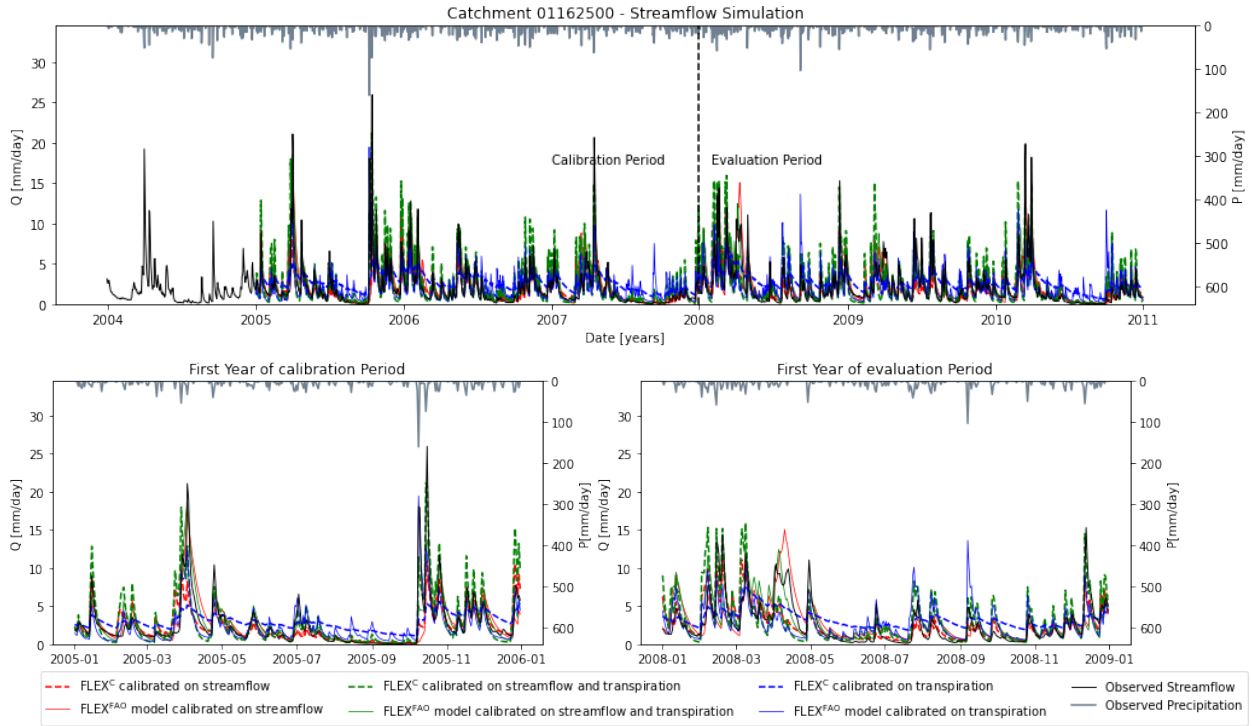


Figure B.25: Catchment C25 streamflow simulation

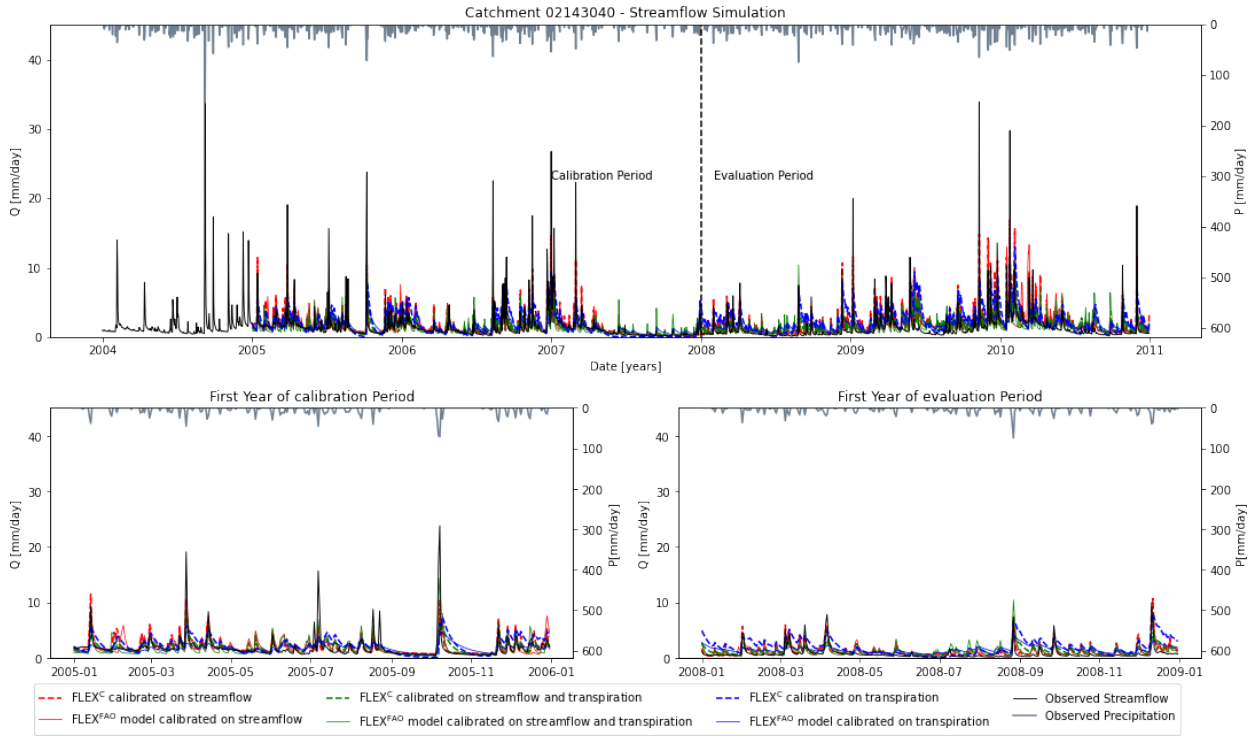


Figure B.26: Catchment C26 streamflow simulation

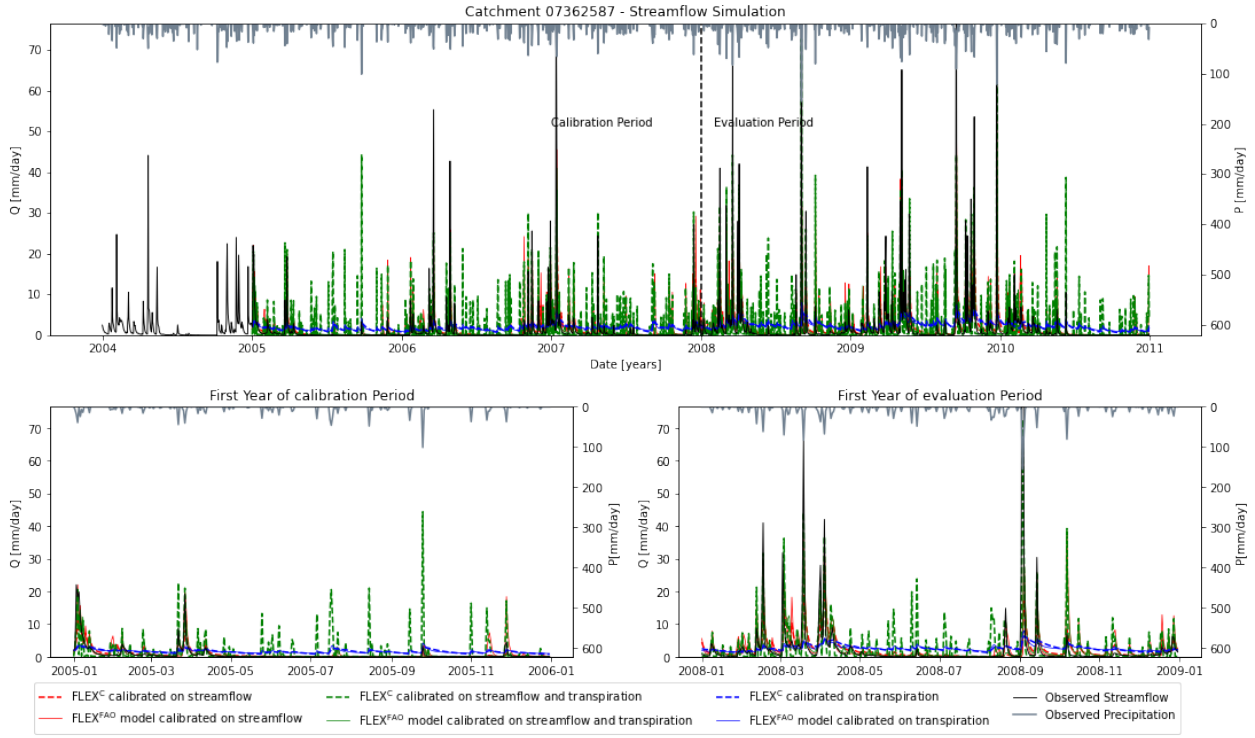


Figure B.27: Catchment C27 streamflow simulation

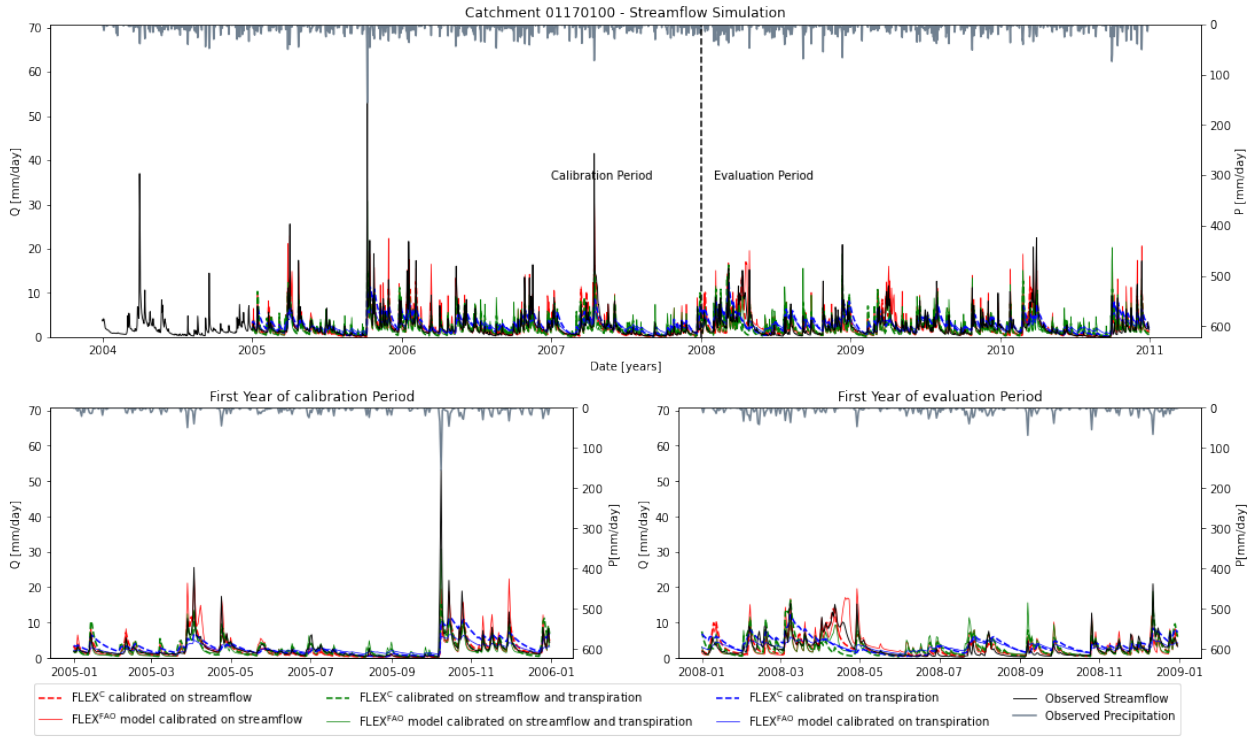


Figure B.28: Catchment C28 streamflow simulation

C

Appendix C: Transpiration simulations

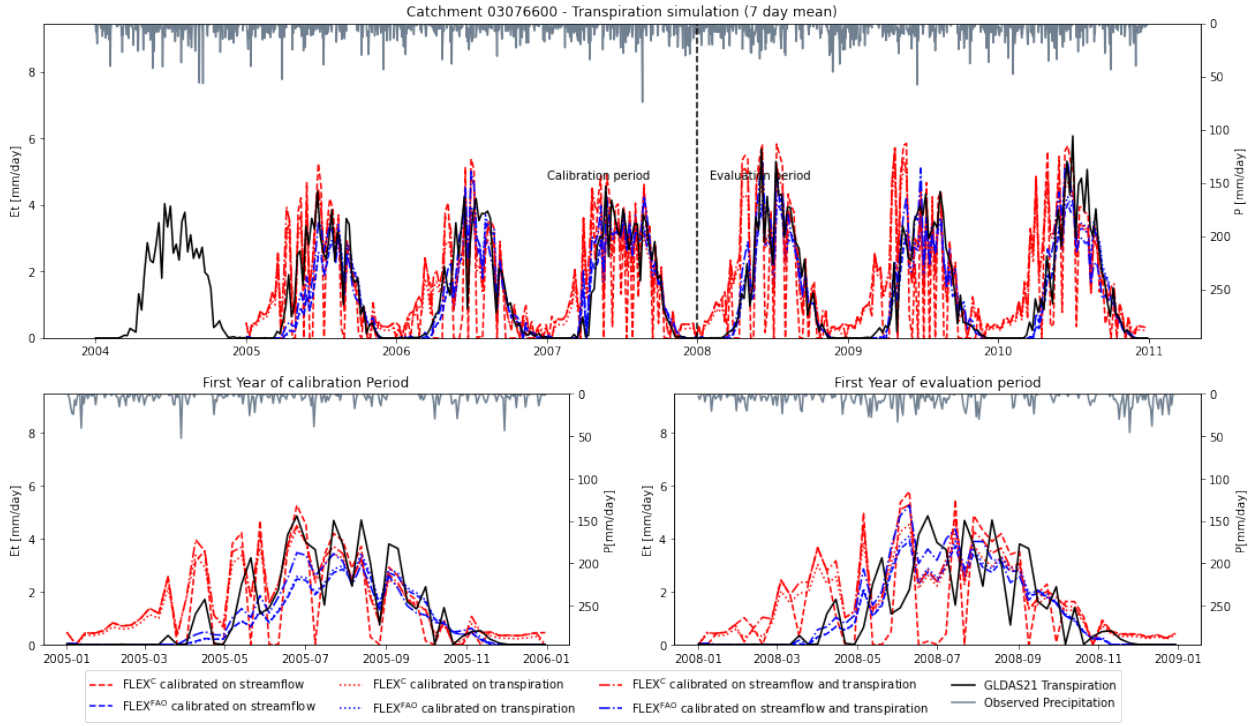


Figure C.1: Catchment C1 transpiration simulation

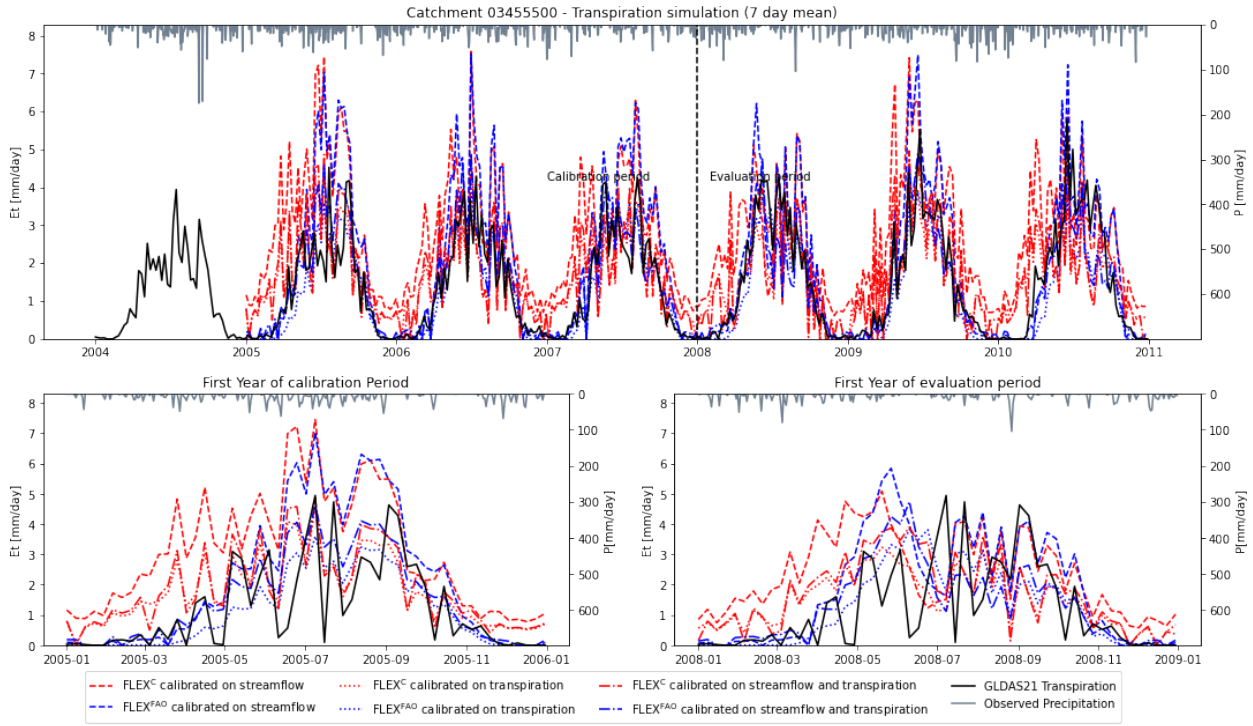


Figure C.2: Catchment C2 transpiration simulation

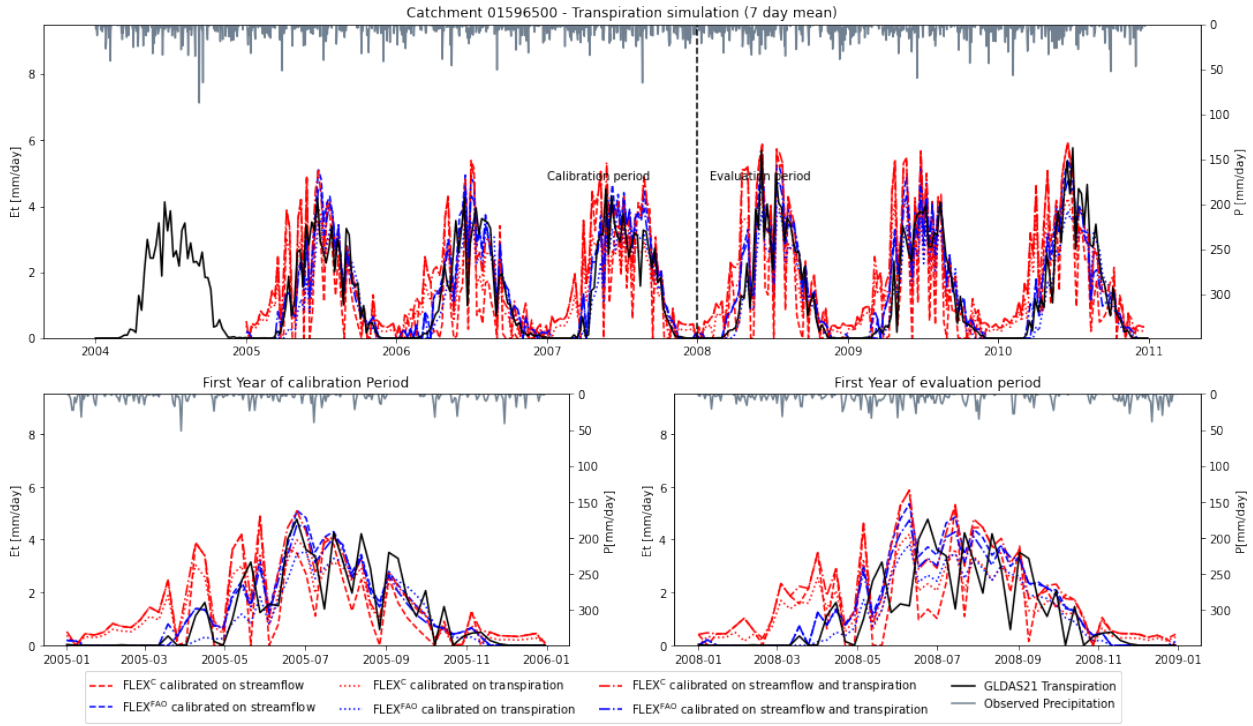


Figure C.3: Catchment C3 transpiration simulation

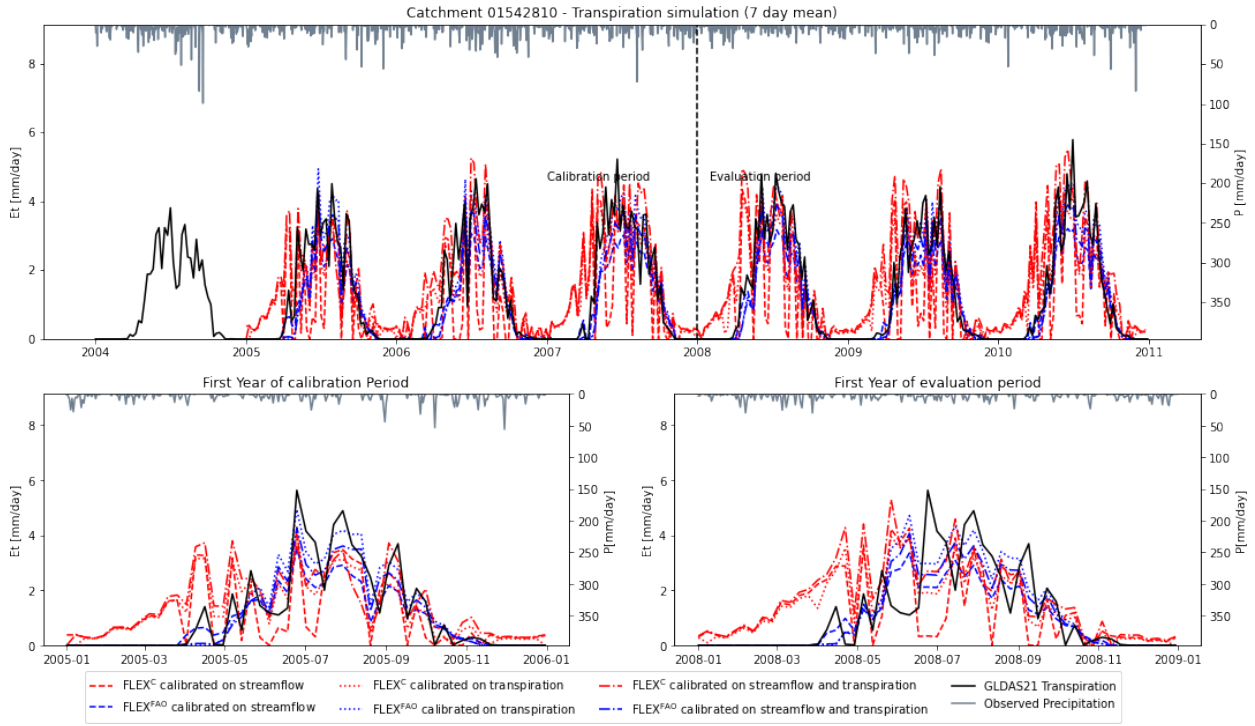


Figure C.4: Catchment C4 transpiration simulation

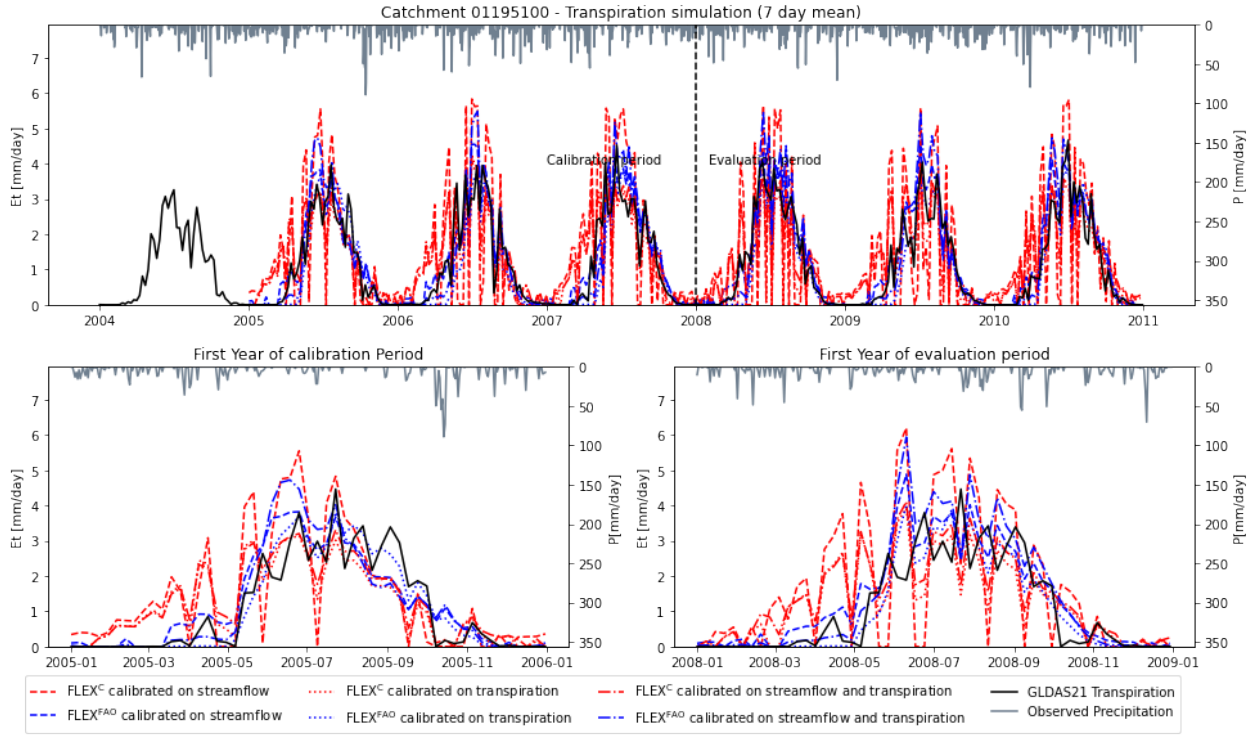


Figure C.5: Catchment C5 transpiration simulation

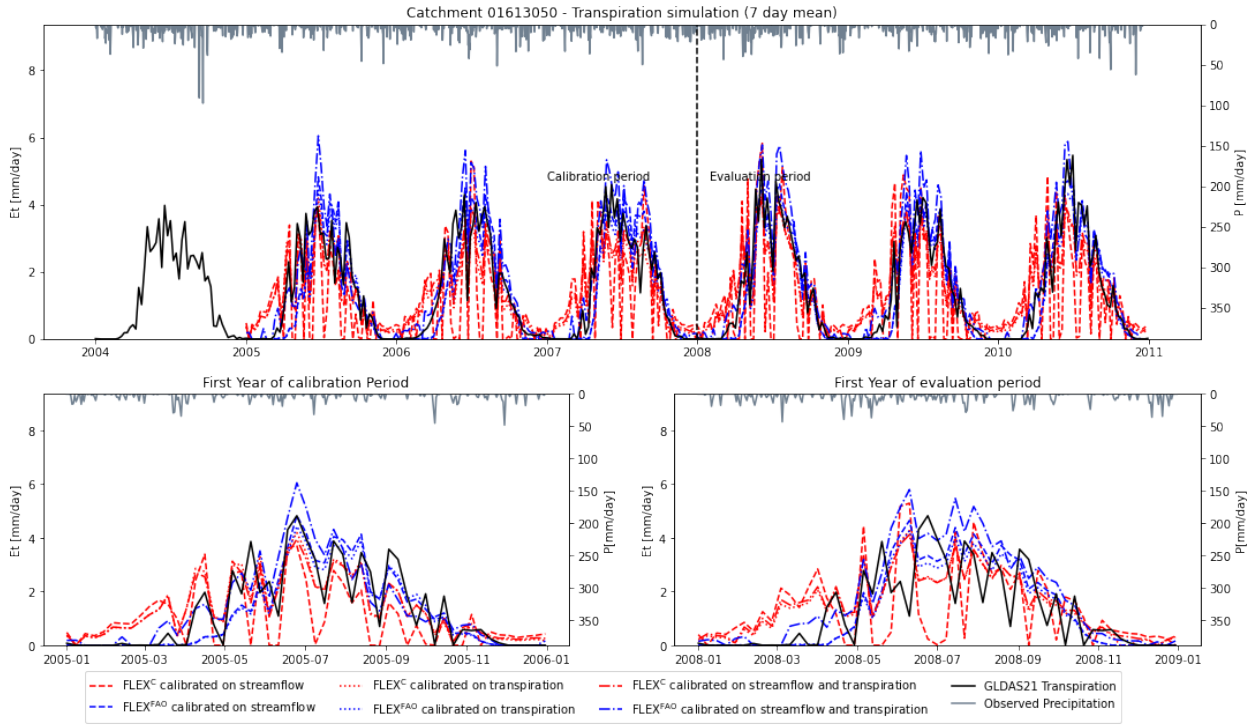


Figure C.6: Catchment C6 transpiration simulation

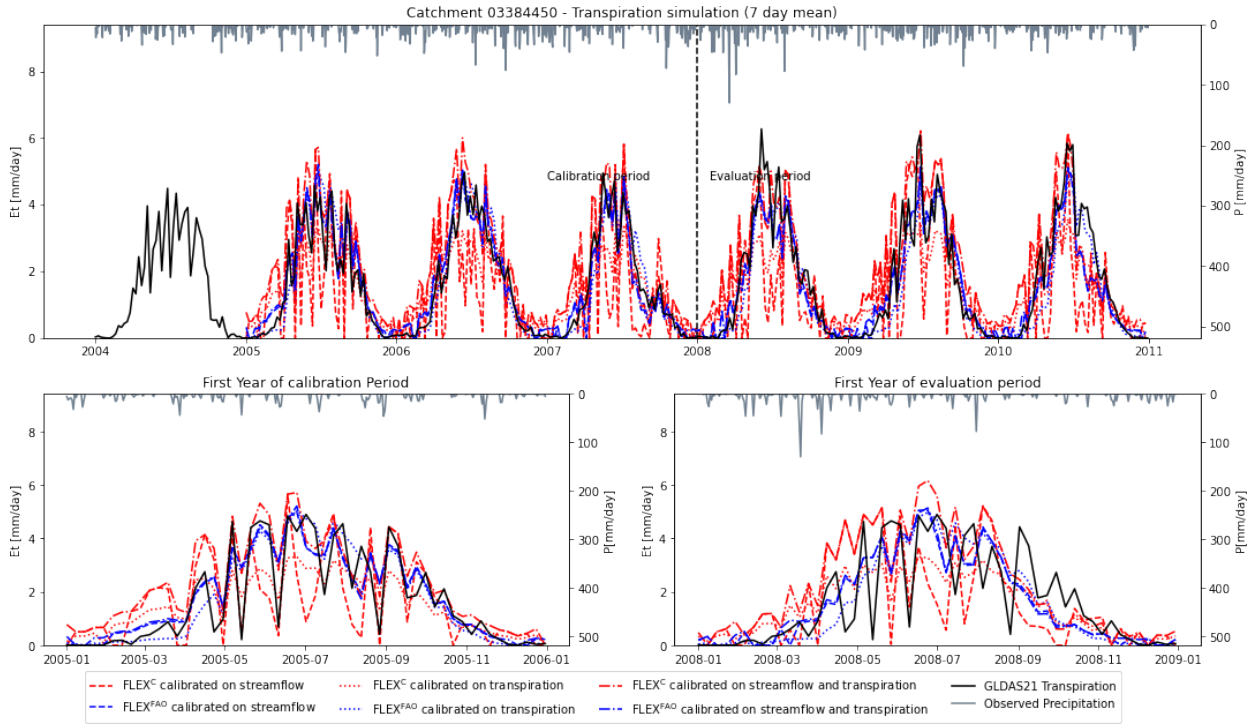


Figure C.7: Catchment C7 transpiration simulation

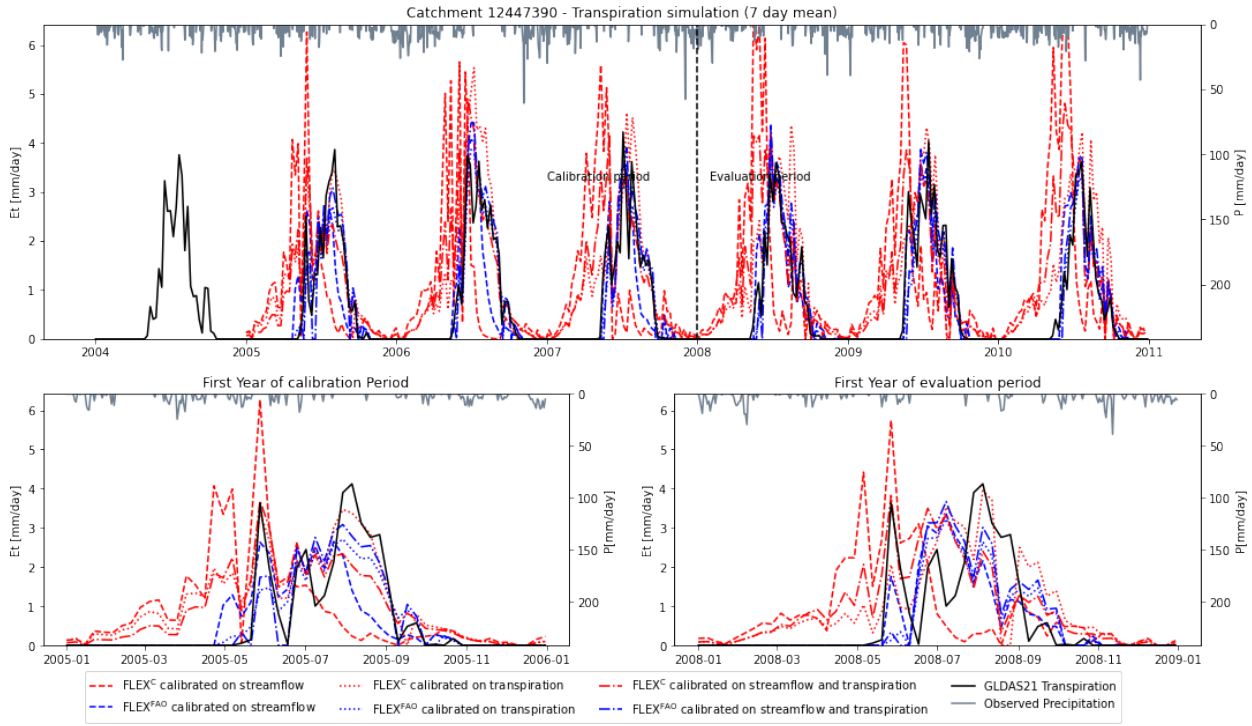


Figure C.8: Catchment C8 transpiration simulation

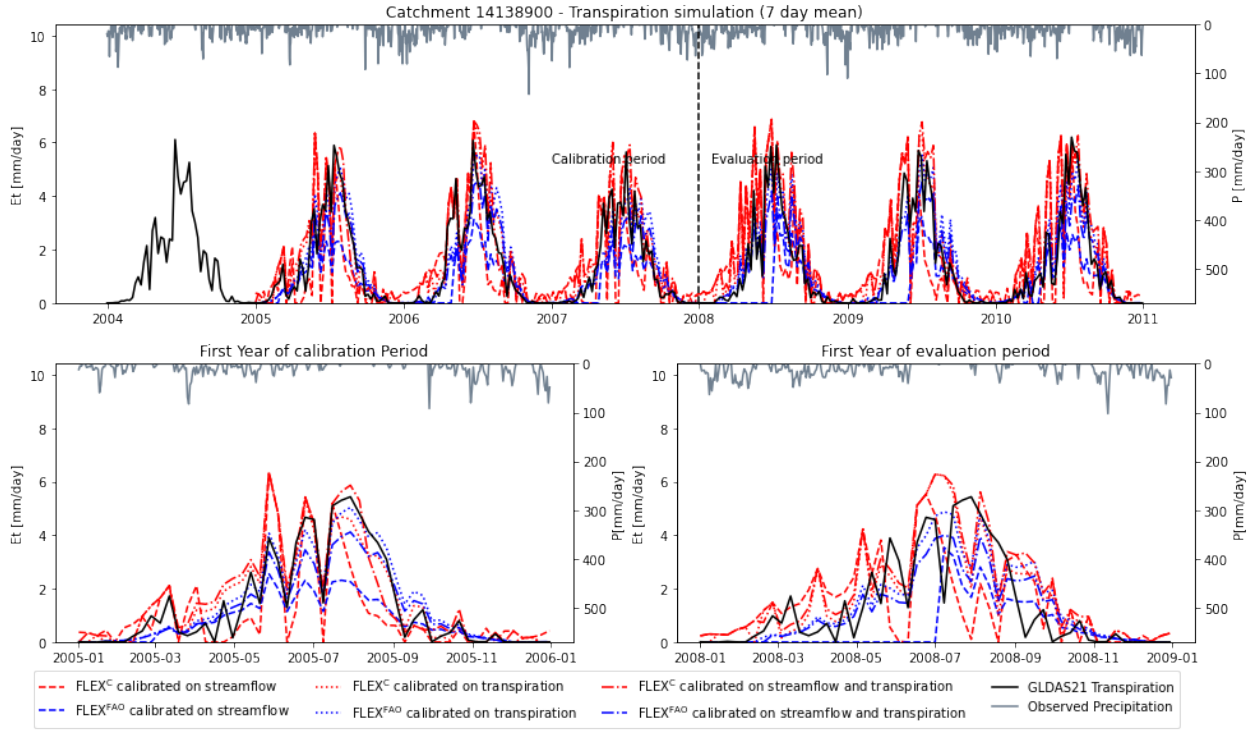


Figure C.9: Catchment C9 transpiration simulation

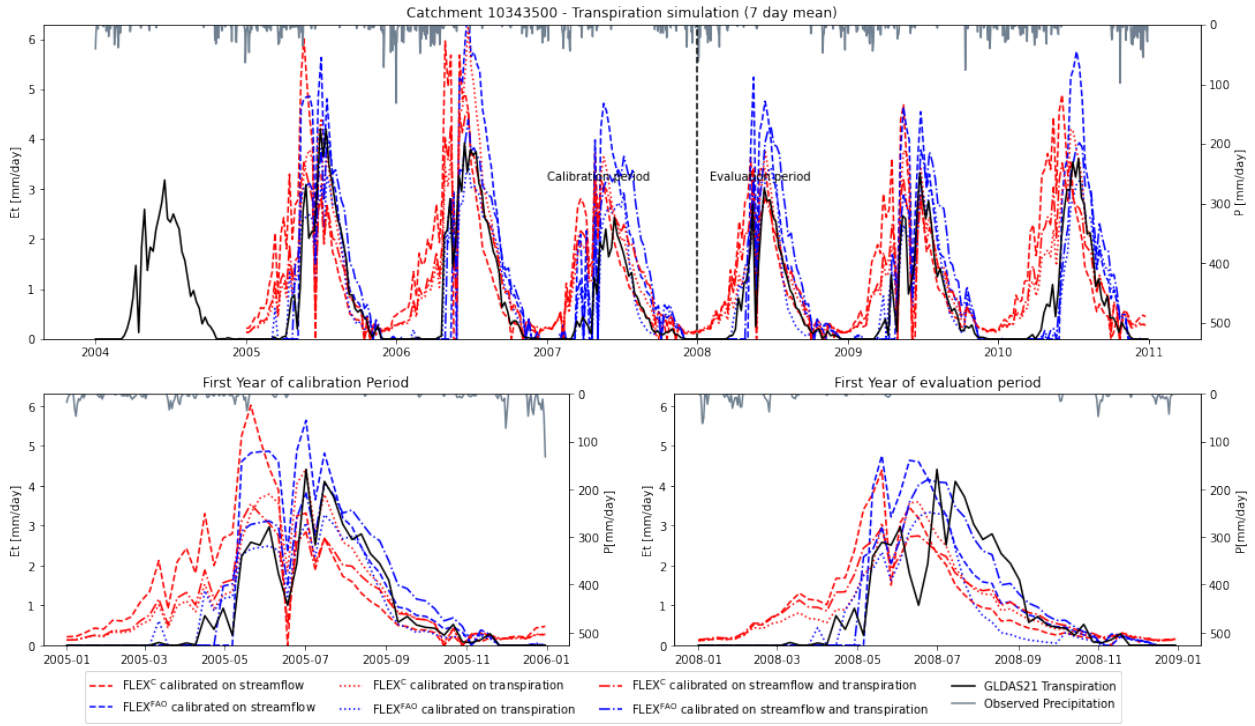


Figure C.10: Catchment C10 transpiration simulation

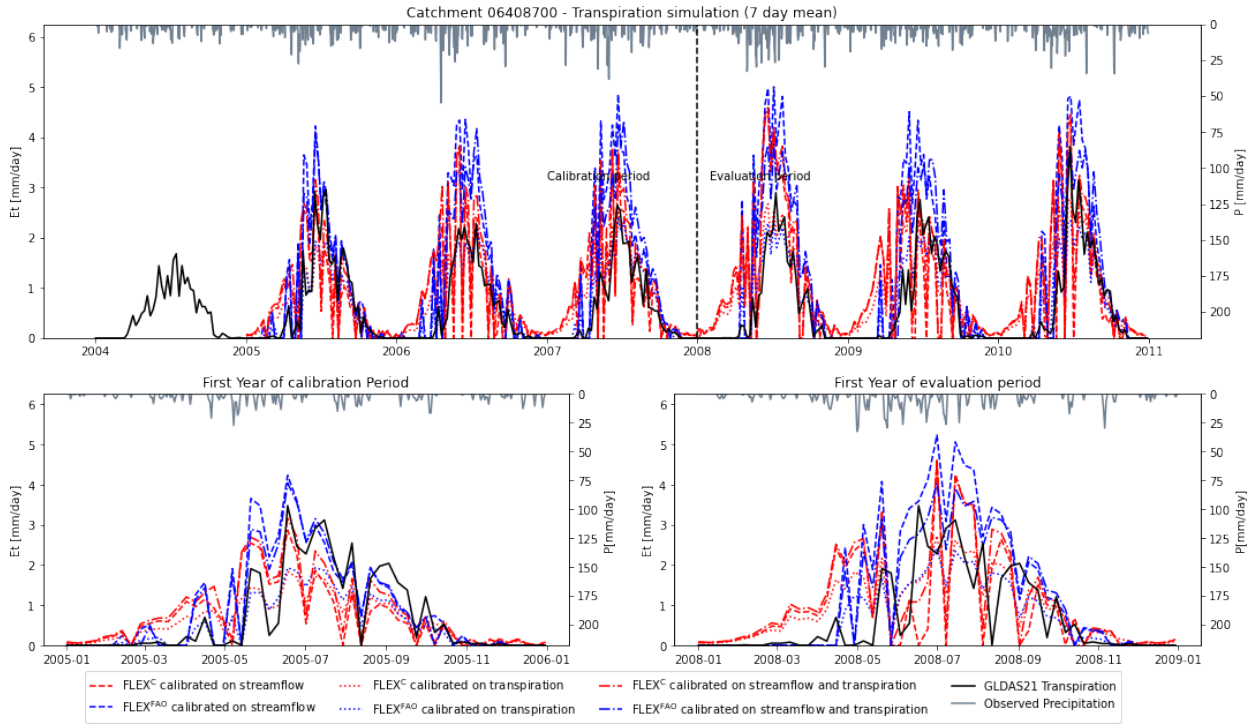


Figure C.11: Catchment C11 transpiration simulation

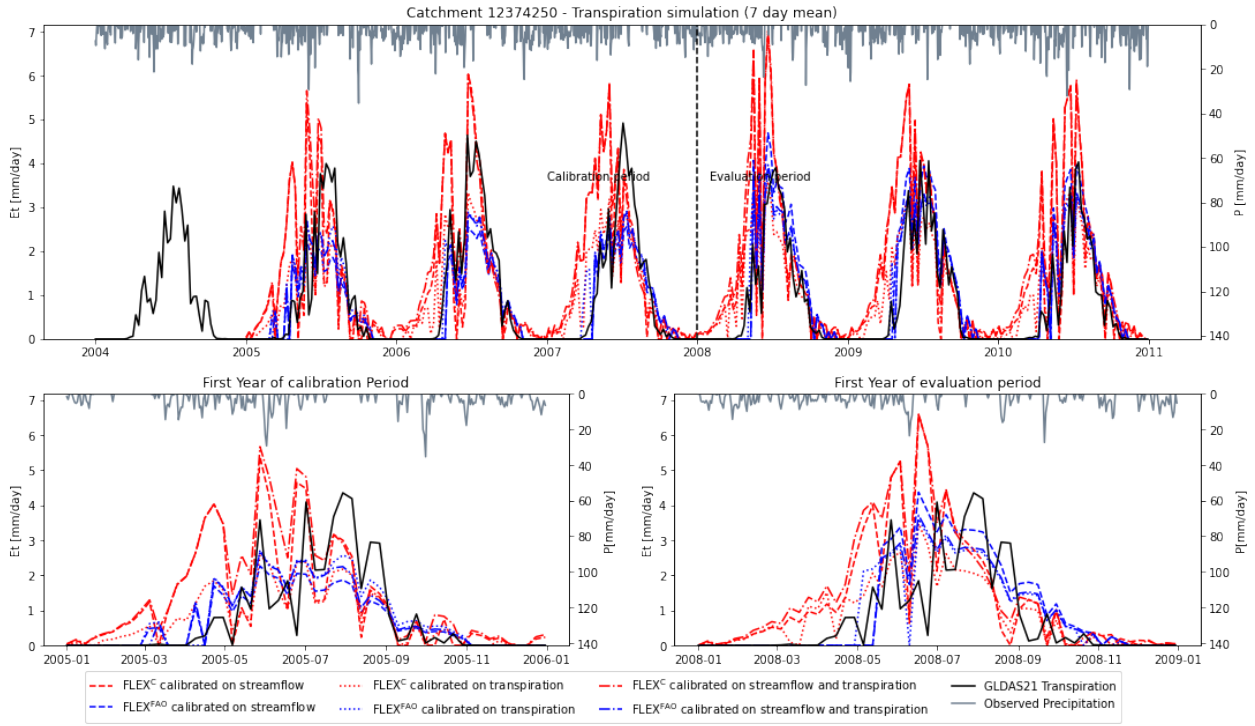


Figure C.12: Catchment C12 transpiration simulation

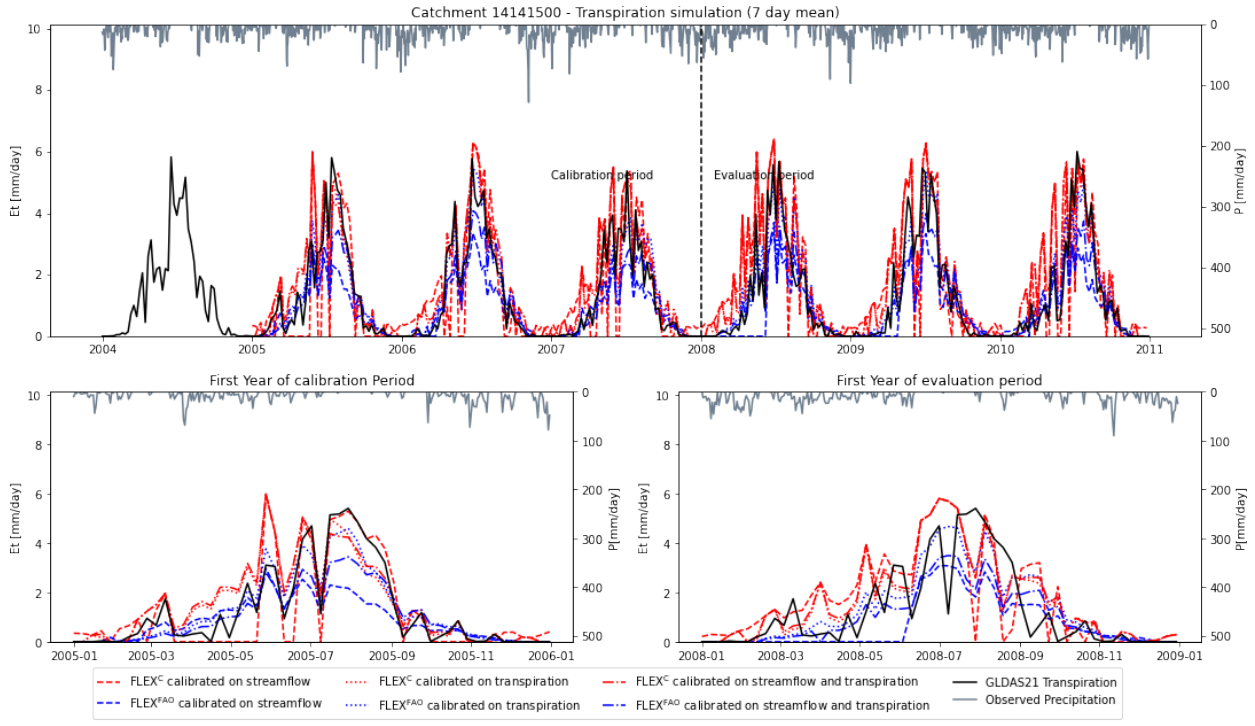


Figure C.13: Catchment C13 transpiration simulation

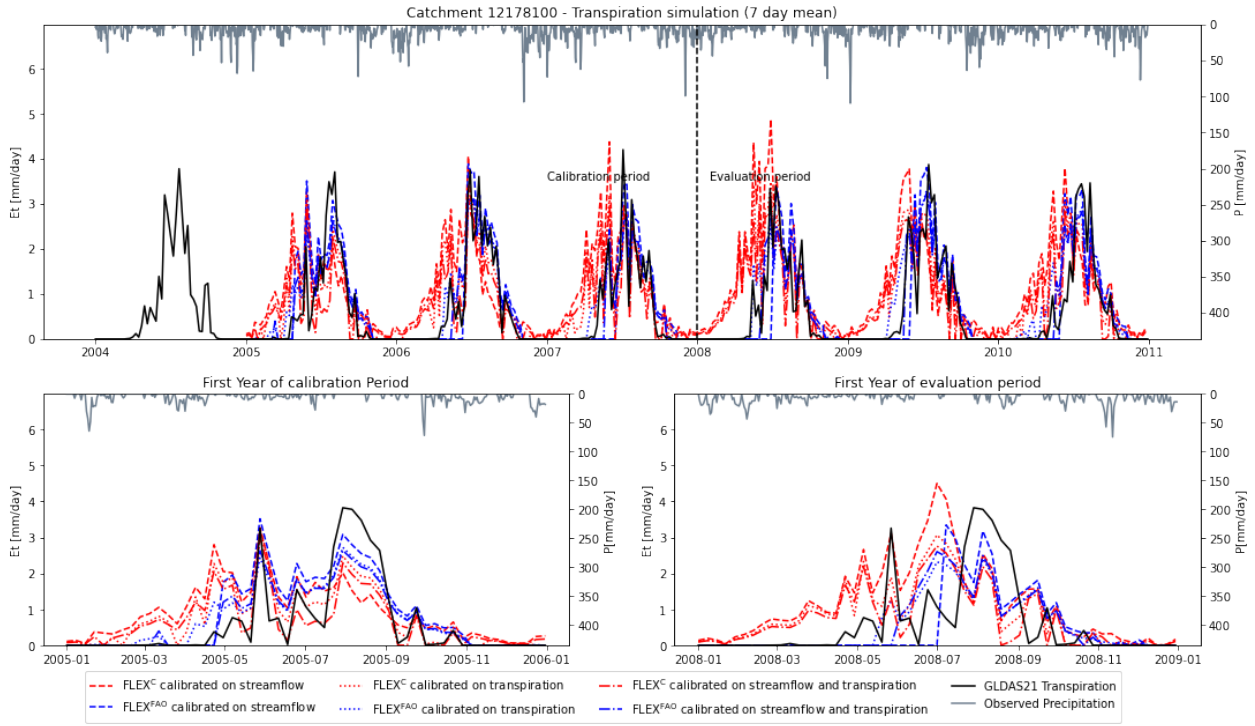


Figure C.14: Catchment C14 transpiration simulation

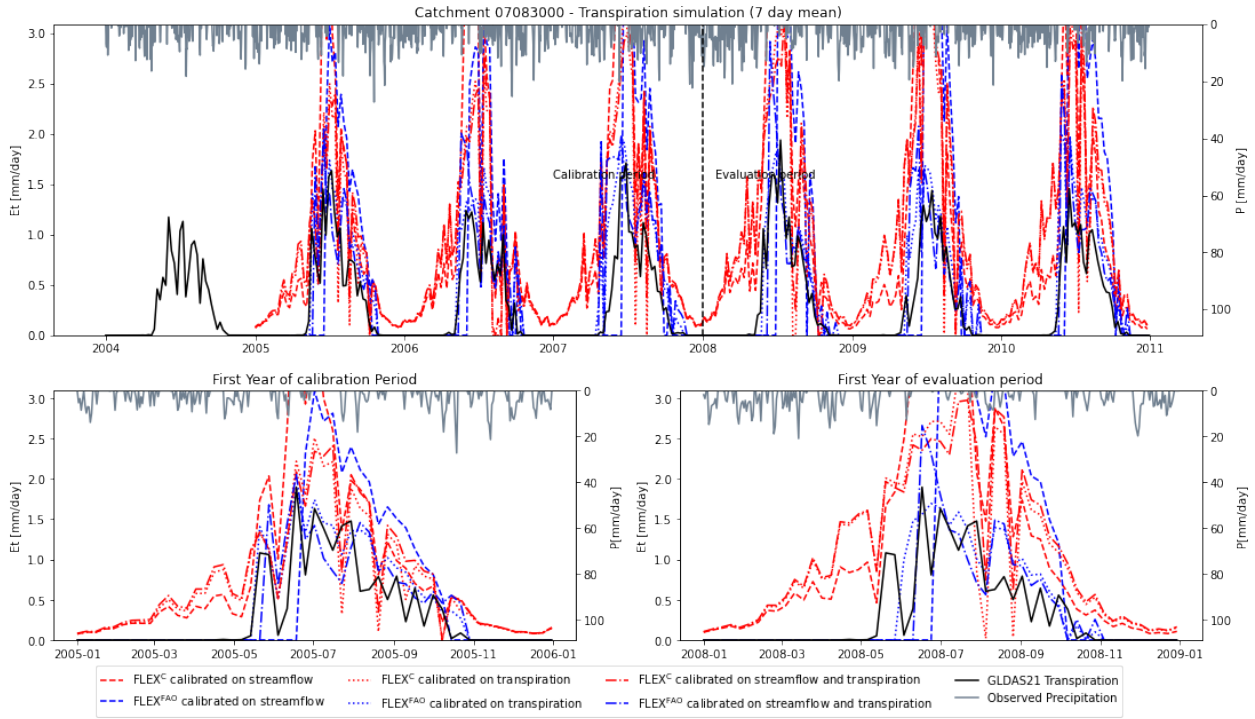


Figure C.15: Catchment C15 transpiration simulation

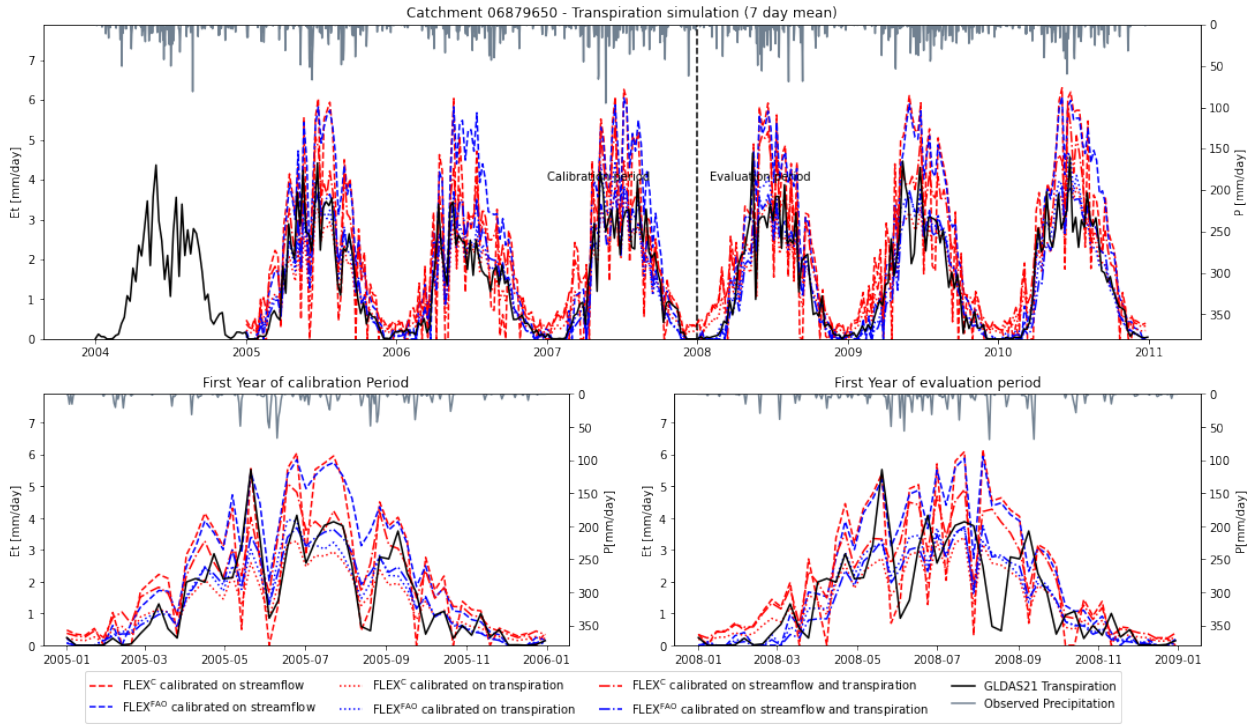


Figure C.16: Catchment C16 transpiration simulation

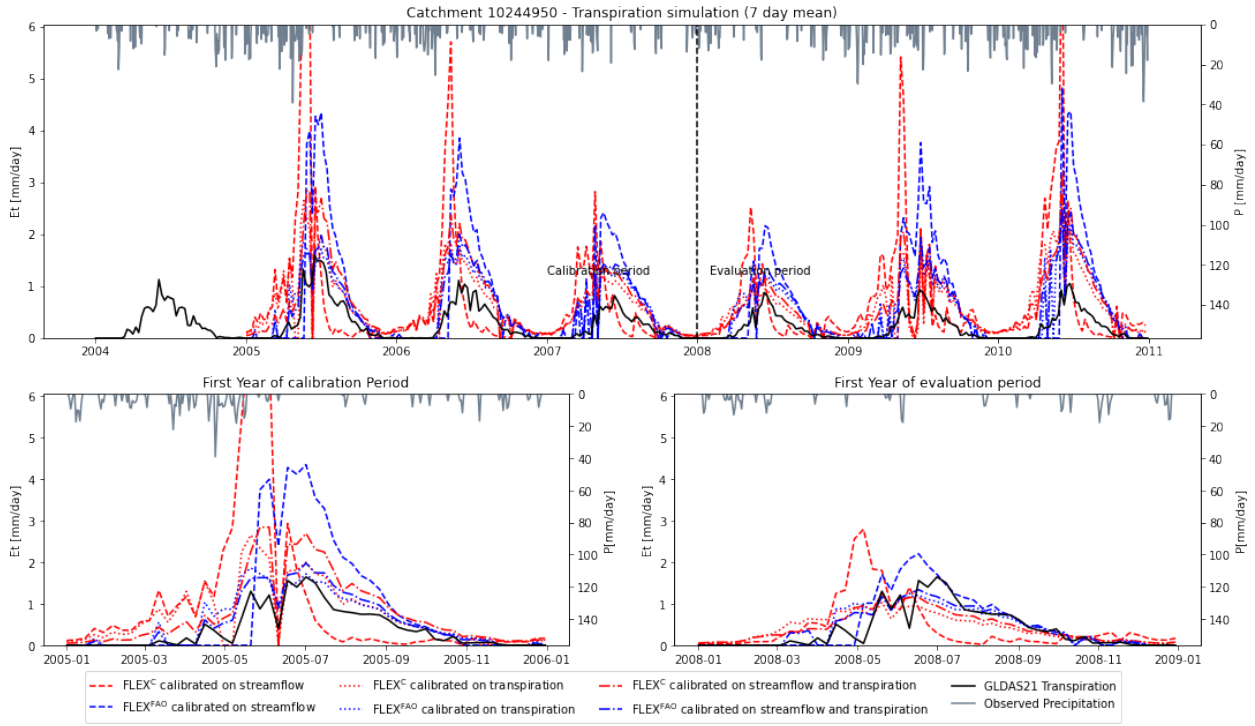


Figure C.17: Catchment C17 transpiration simulation

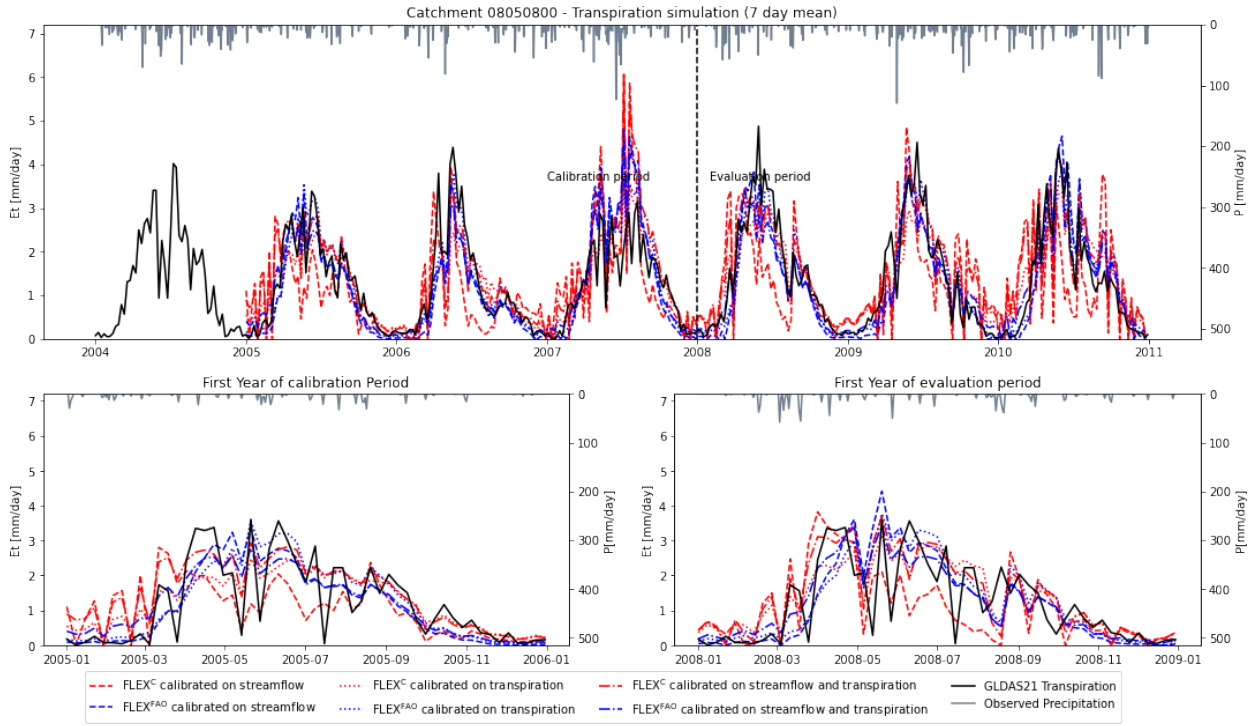


Figure C.18: Catchment C18 transpiration simulation

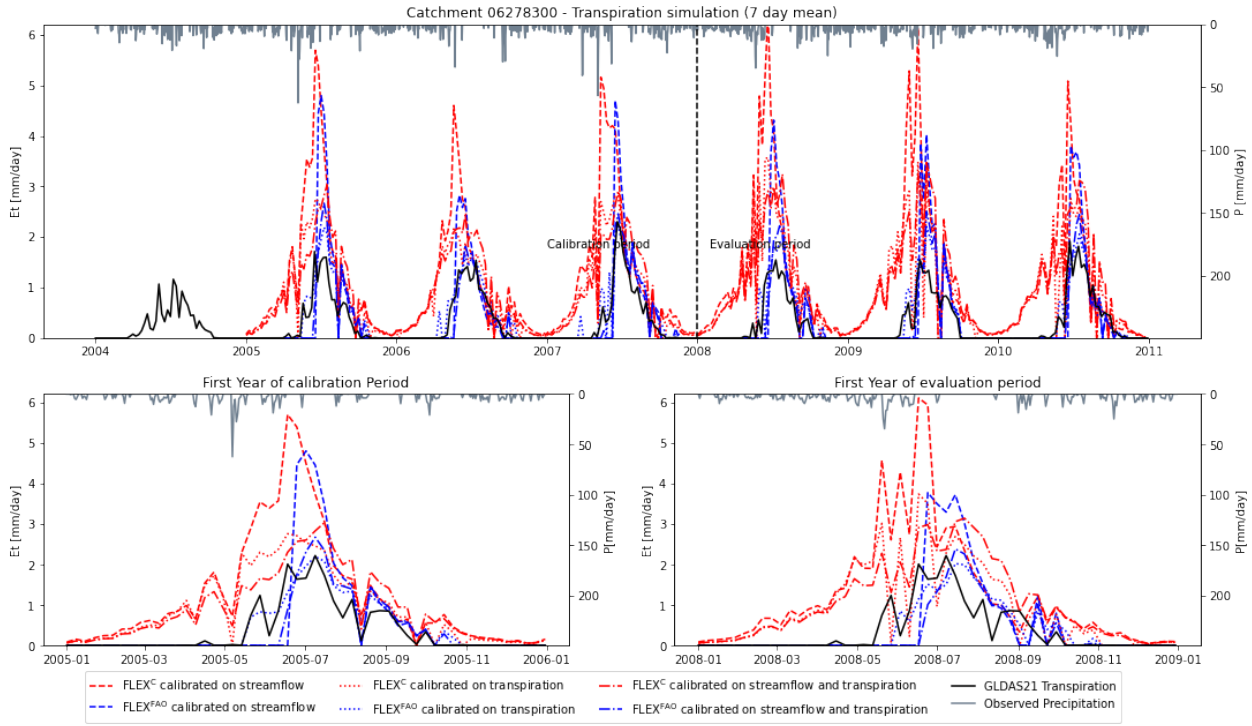


Figure C.19: Catchment C19 transpiration simulation

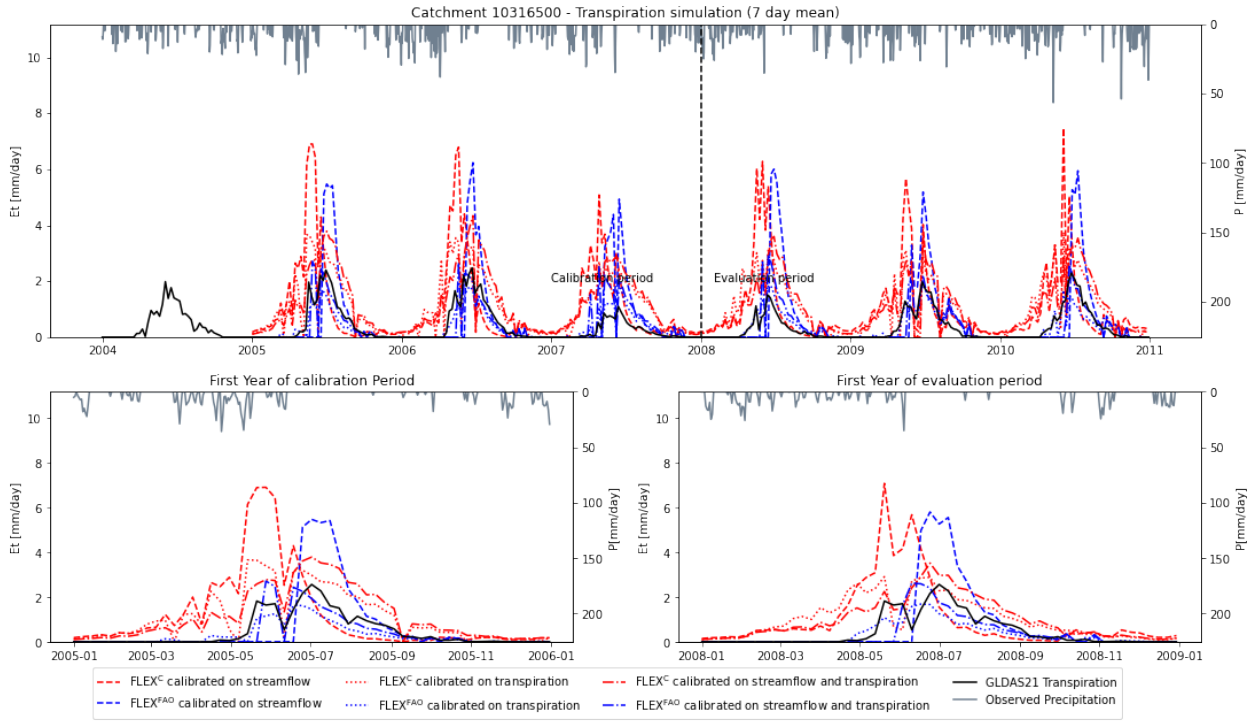


Figure C.20: Catchment C20 transpiration simulation

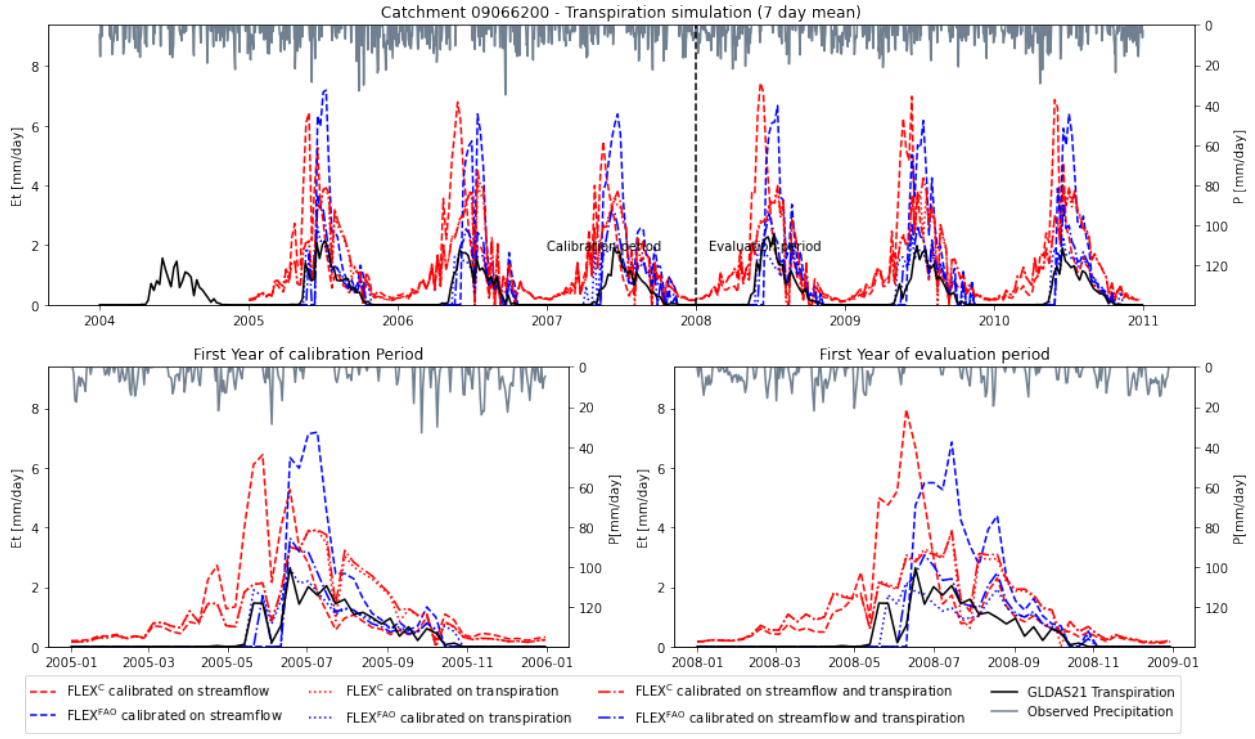


Figure C.21: Catchment C21 transpiration simulation

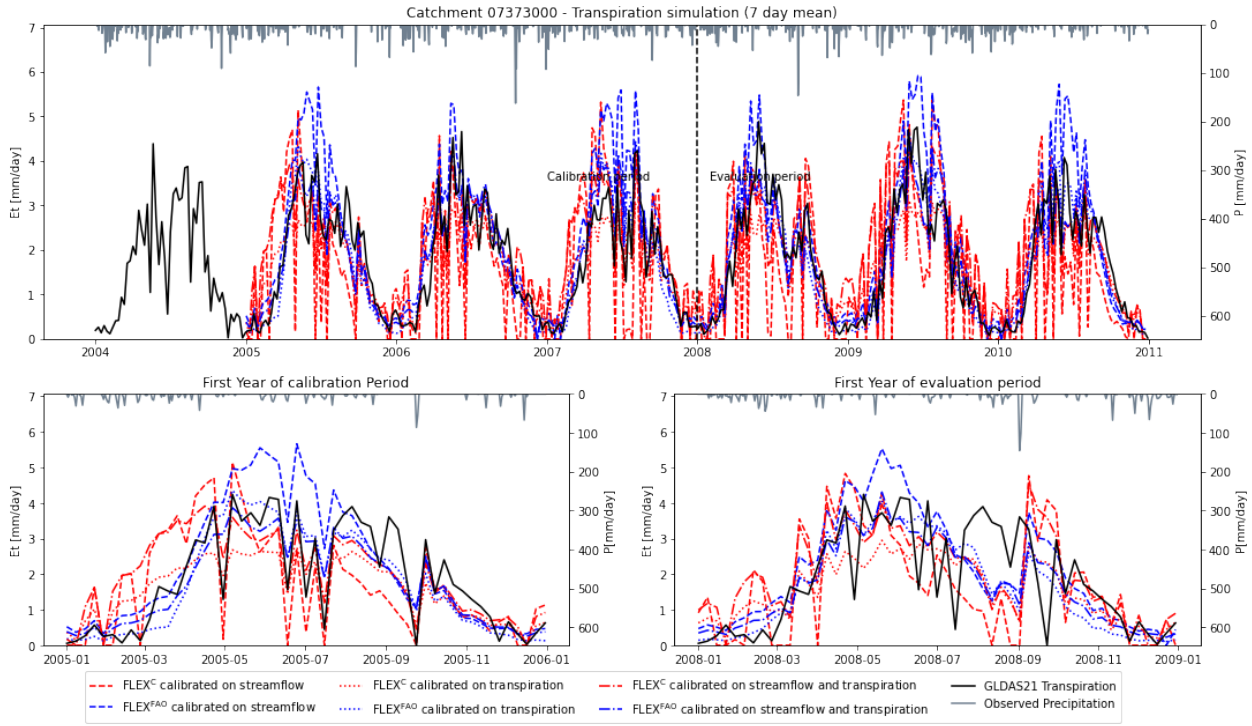


Figure C.22: Catchment C22 transpiration simulation

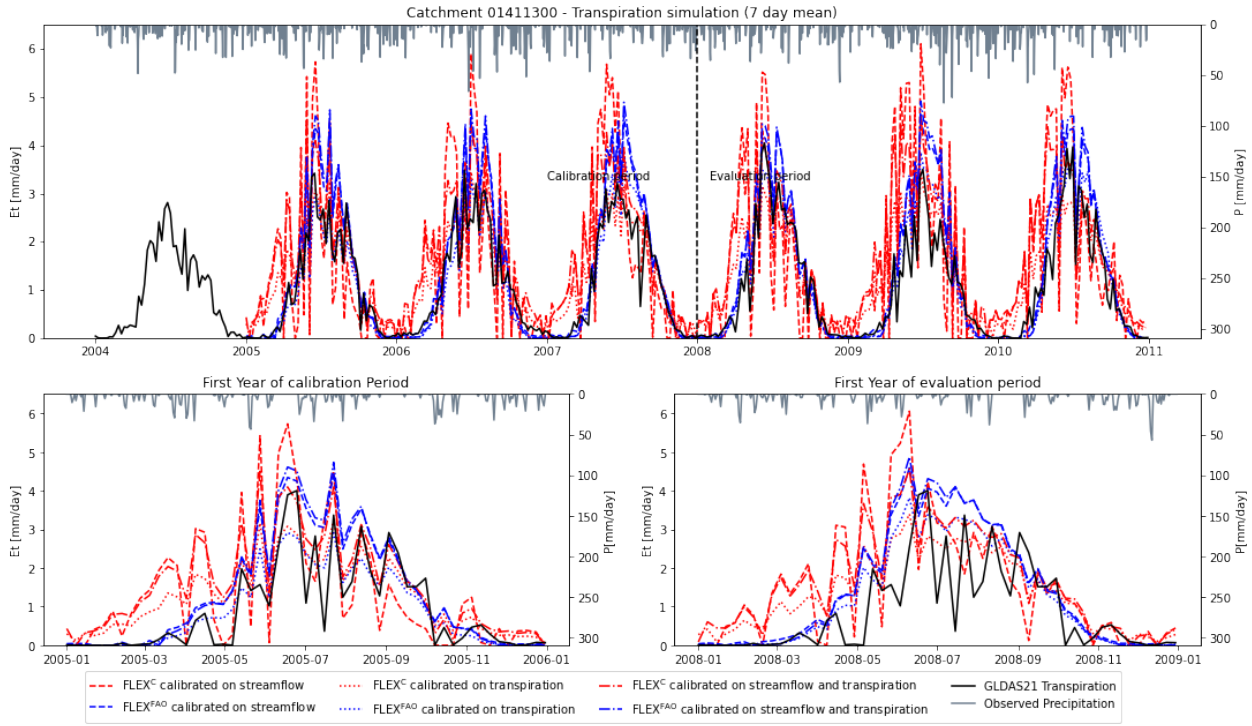


Figure C.23: Catchment C23 transpiration simulation

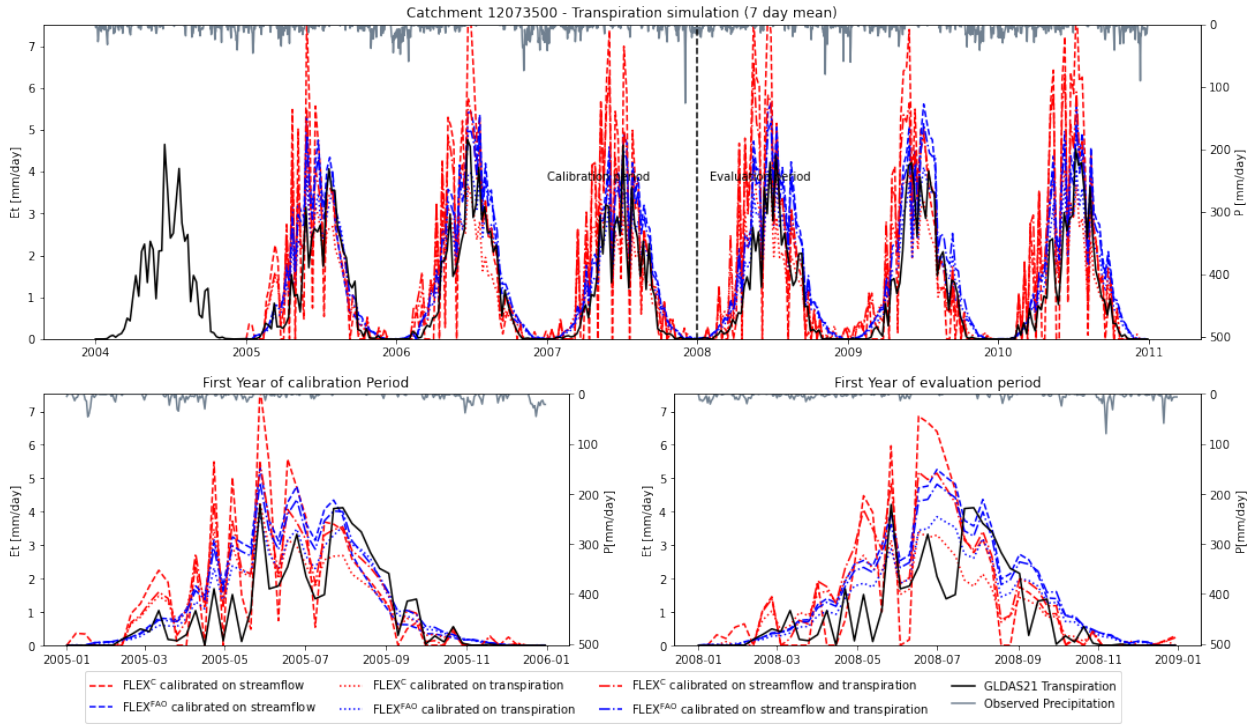


Figure C.24: Catchment C24 transpiration simulation

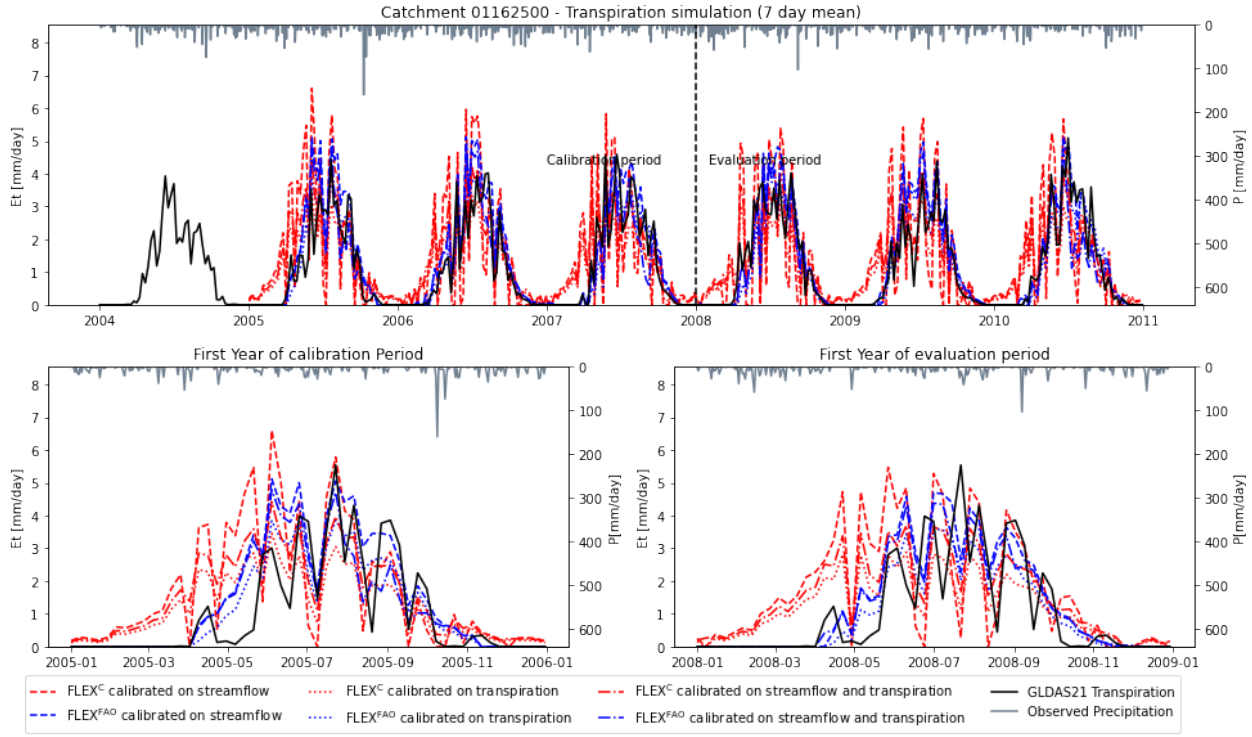


Figure C.25: Catchment C25 transpiration simulation

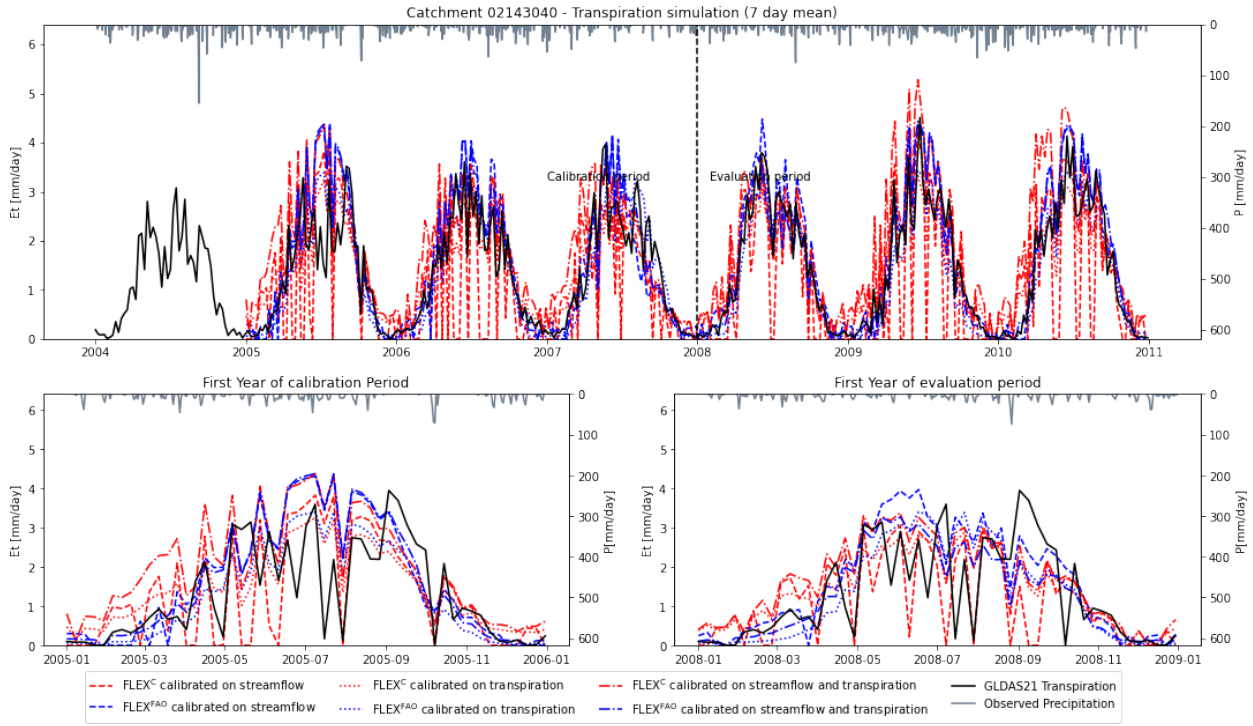


Figure C.26: Catchment C26 transpiration simulation

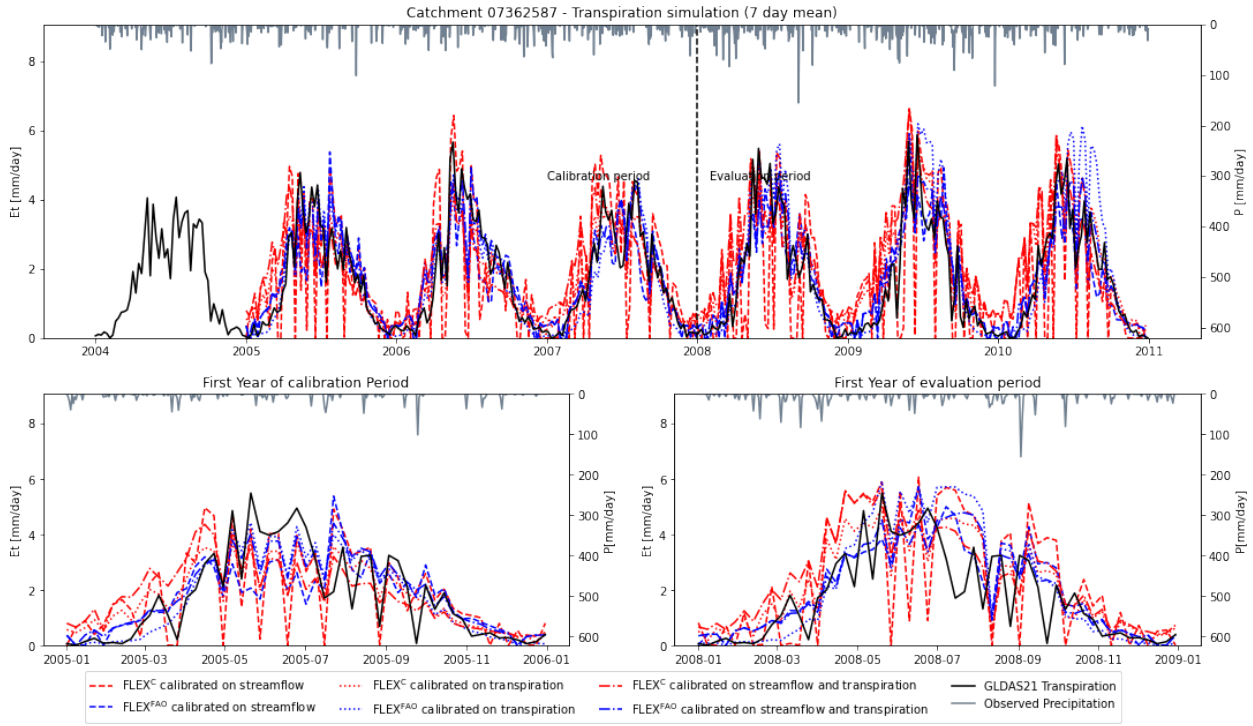


Figure C.27: Catchment C27 transpiration simulation

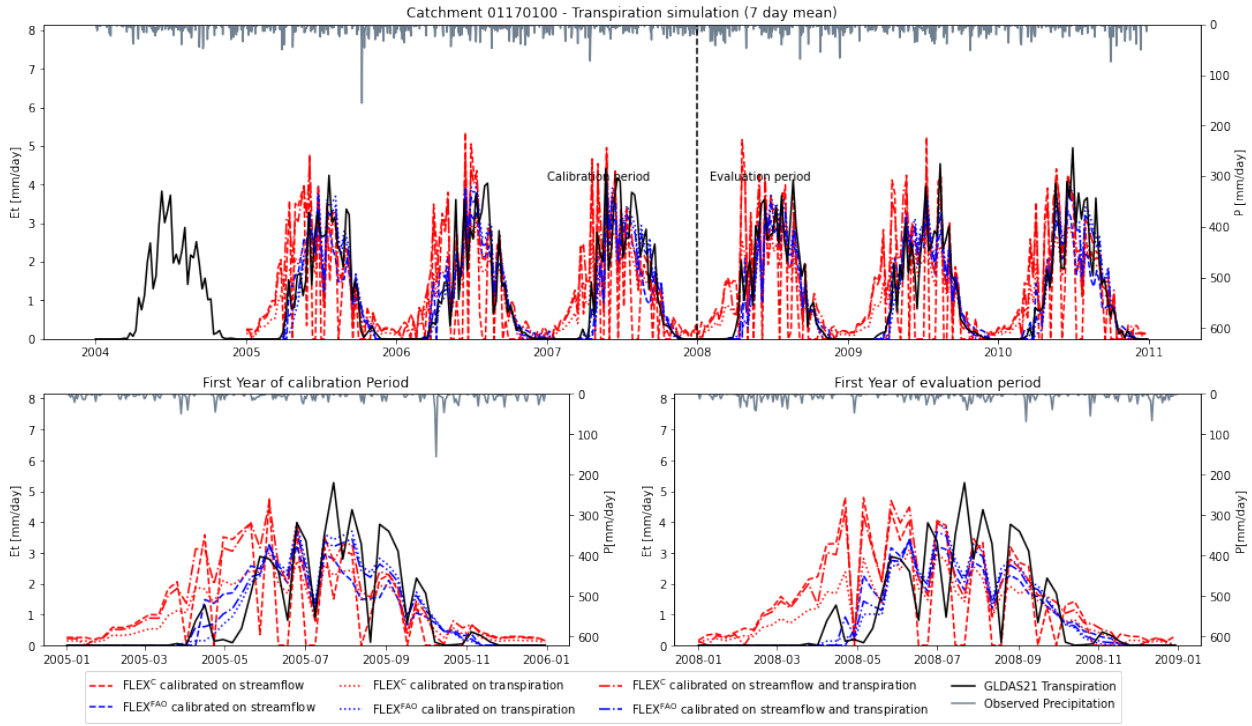


Figure C.28: Catchment C28 transpiration simulation

D

Appendix D: Performance results of streamflow simulation when models are calibrated on transpiration

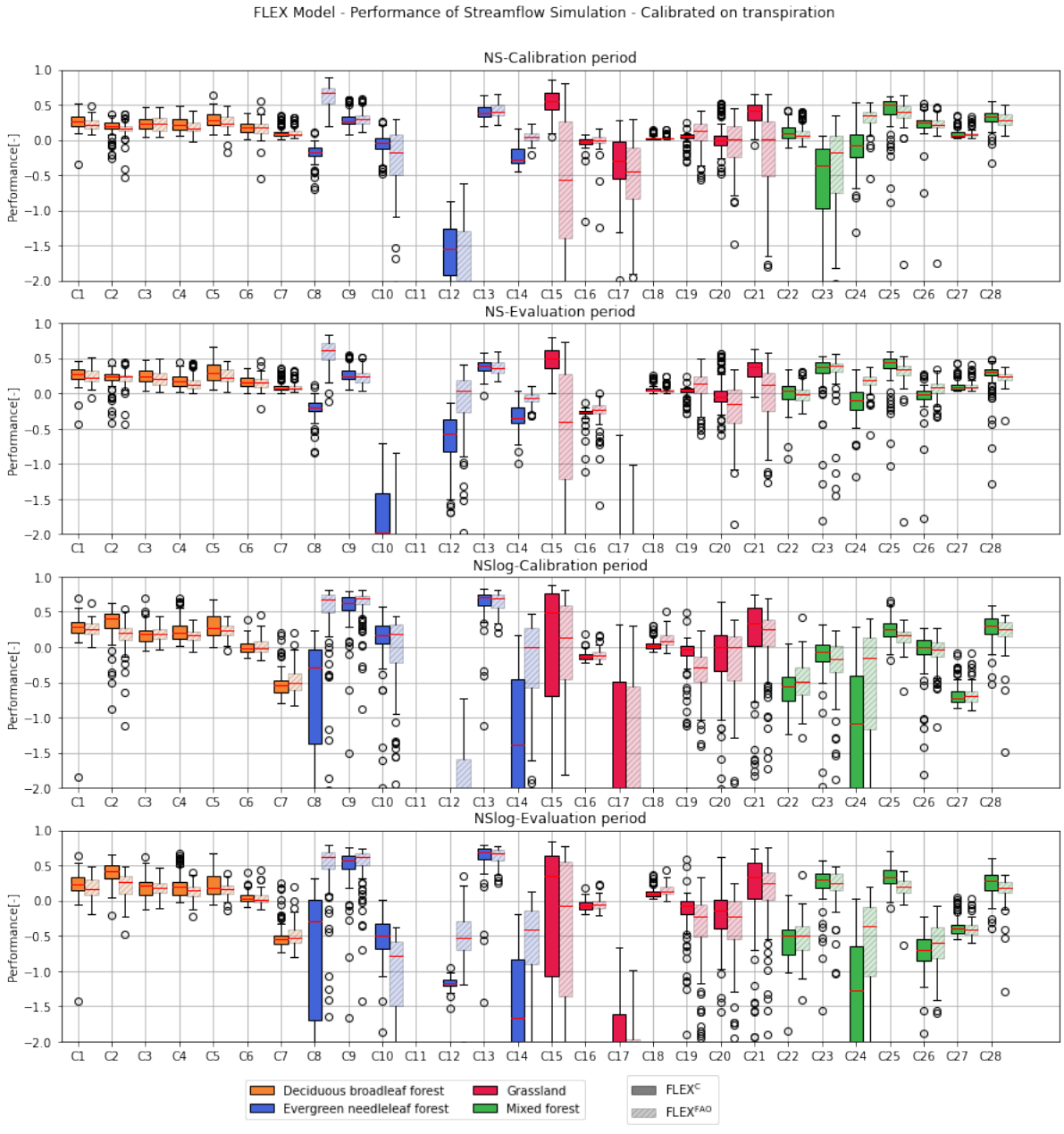


Figure D.1: Performance of the streamflow simulation of FLEX^{C} and FLEX^{FAO} when calibrated on transpiration, presented next to each other for every catchment. The four graphs separate the performance into the NS and NSlog objective function and the performance in the calibration period and the evaluation period. The colors show the dominant vegetation type in all the catchments. The x-axes show all the catchments that have been simulated.

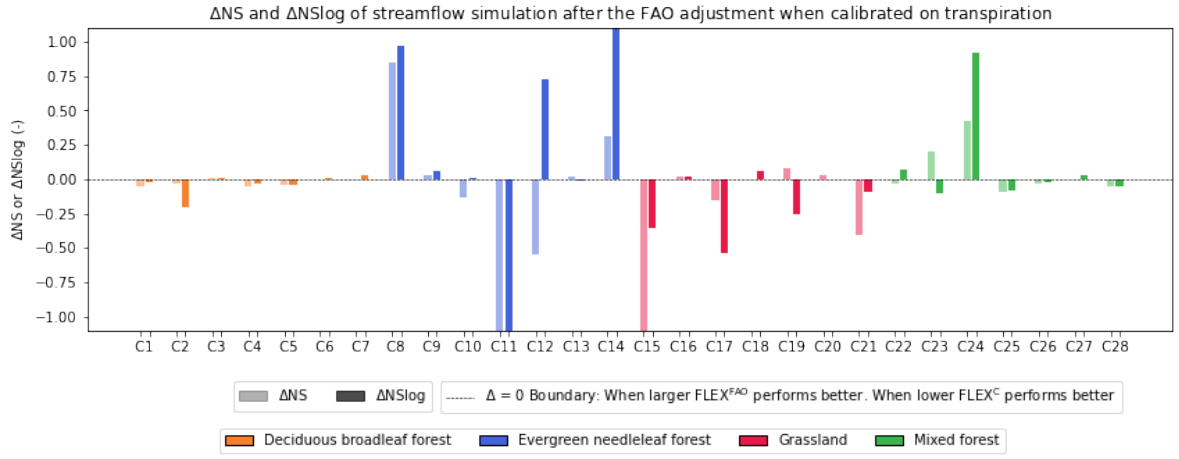


Figure D.2: Performance comparison of the streamflow simulation between FLEX^{C} and FLEX^{FAO} calibrated on transpiration. The figure shows the increase or decrease of the NS-efficiency and the NSlog-efficiency after the FAO adjustment for every catchment respectively. The colors indicate the dominant vegetation type.

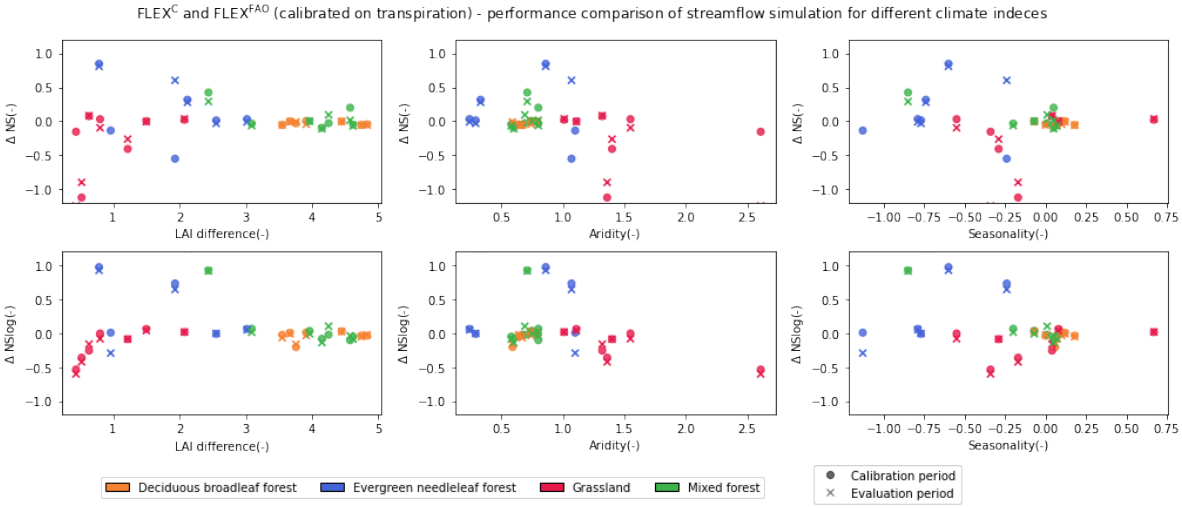


Figure D.3: Performance comparison between FLEX^{C} and FLEX^{FAO}. The top figures show the comparison of the NS-efficiency and the bottom figures show the comparison of the NSlog-efficiency. The y-axes show how much the NS/NSlog-efficiency is increased or decreased after the FAO adjustment. The dot marker shows the performance in the calibration period and the cross marker the performance in the evaluation period. The colors indicate the dominant vegetation type in the catchment. The x-axes show different climate indices.

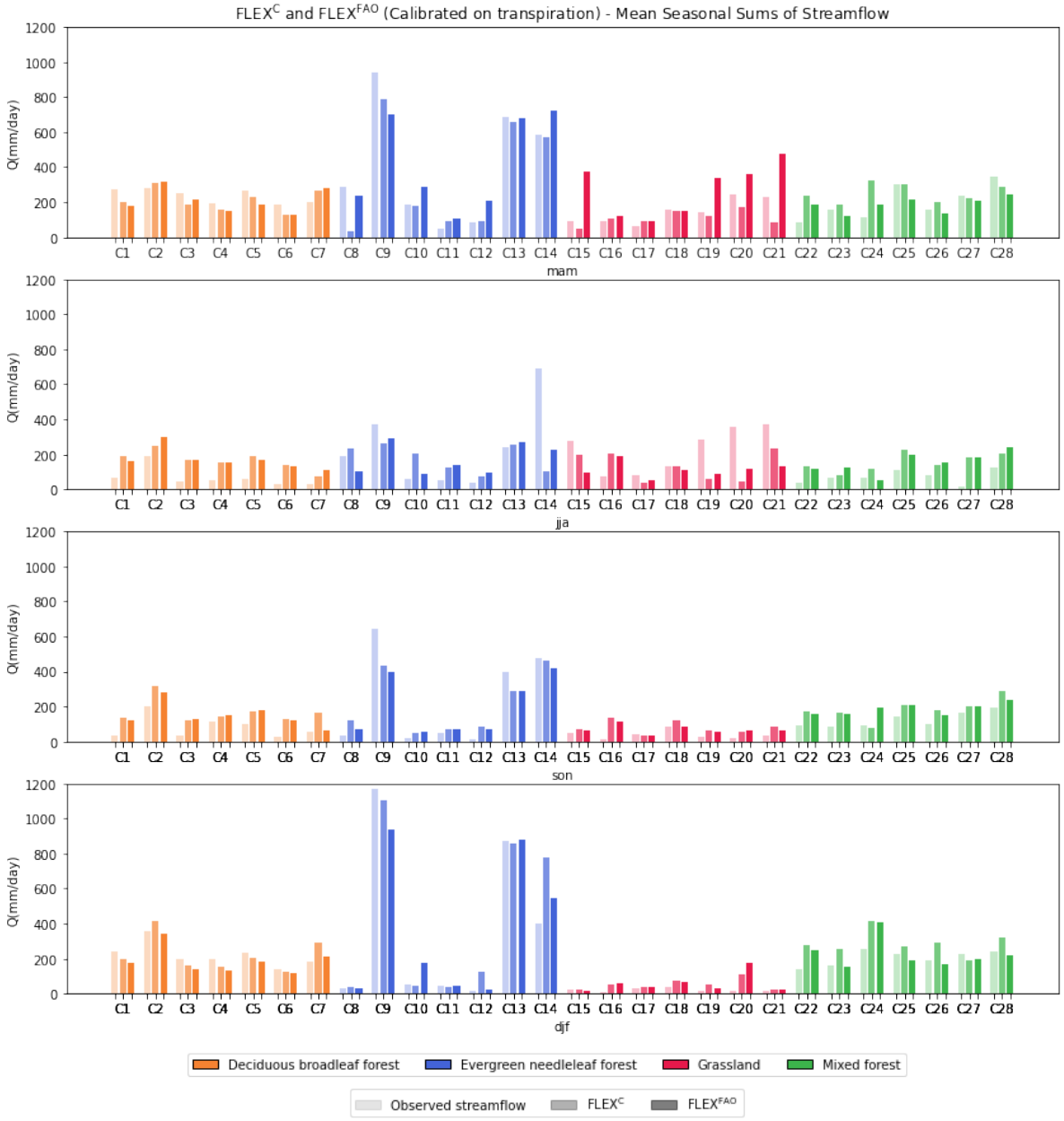


Figure D.4: Mean seasonal sums of the streamflow for the observed streamflow, FLEX^{C} and FLEX^{FAO}. The seasons are indicated by mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. The colors show the dominant vegetation type in the catchments, which are indicated on the x-axes. For every catchment, the observed streamflow, FLEX^{C} and FLEX^{FAO} are presented in that order.