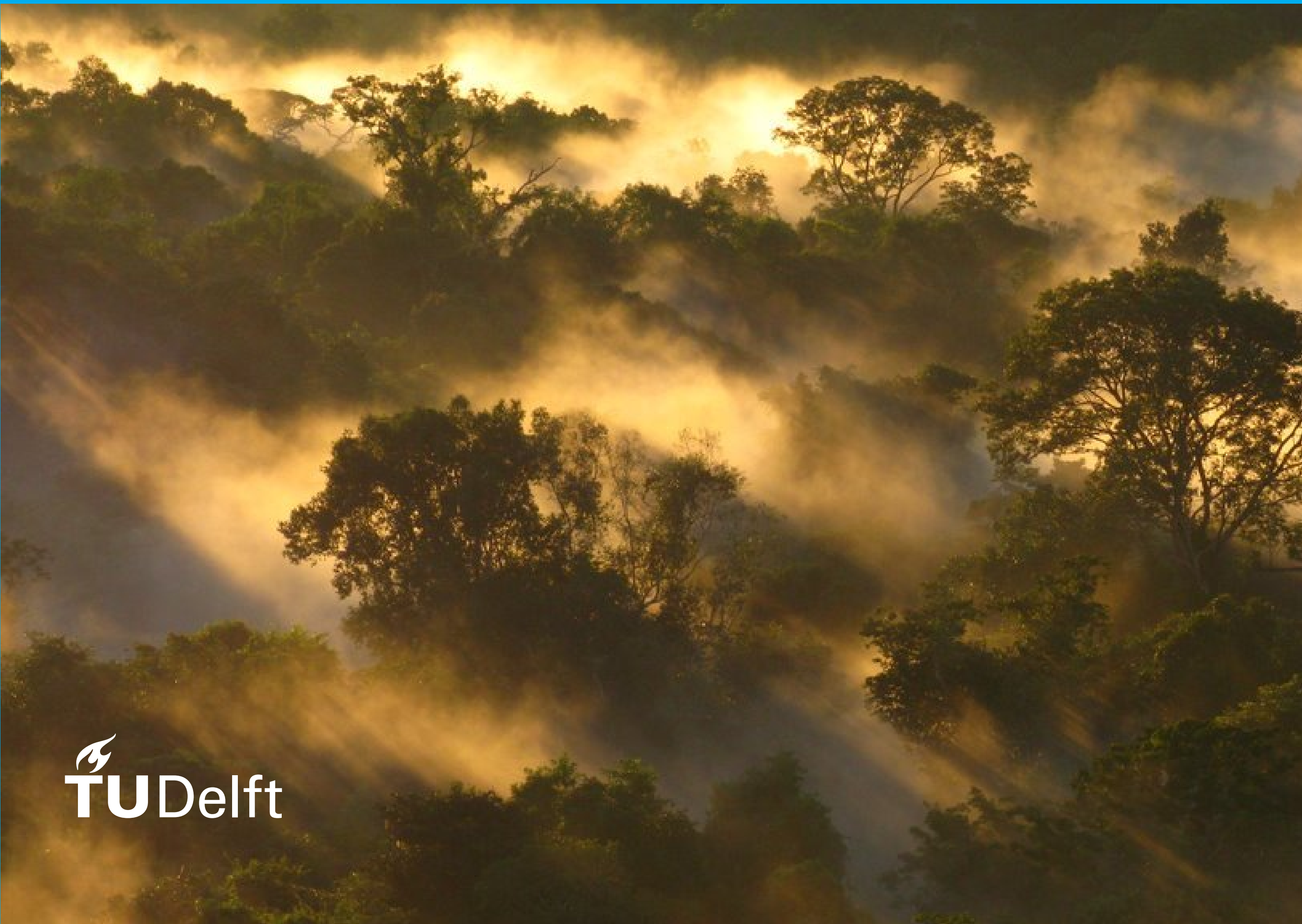
# The Impact of Vegetation on the Partitioning of Evaporation in Conceptual Hydrological Models

## O. Mert



TU Delft

# The Impact of Vegetation on the Partitioning of Evaporation in Conceptual Hydrological Models

by

## O. Mert

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday July 9, 2021 at 09:00 AM.

| | | |
|---|---|---|
| Student number: | 4750195 | |
| Thesis committee: | Dr. Markus Hrachowitz, | TU Delft, supervisor |
| | Prof. Dr. ir. Susan Steele-Dunne, | TU Delft |
| | Dr. ir. Miriam Coenders, | TU Delft |

Image on cover page: (van der Sleen, 2021)

TU Delft

# Abstract

Many conceptual hydrological models do not include any information on plant phenology, which represents the dynamic behavior of vegetation, to partition evaporation in a catchment. Such models often use potential evaporation to quantify evaporation from the interception reservoir and what remains afterward to quantify transpiration, or the other way around depending on the modeler's decisions. Such defined orders are the opposite of what is perceived in reality. Vegetation is alive, continuously developing, and impacts the partitioning of evaporation through its ability to transpire. Hence, vegetation is able to constantly change this defined order and thus the ratio of transpiration over interception evaporation over time.

This study aimed to include information from plant phenology, representing the dynamic behavior of vegetation, to partition evaporation in a catchment in conceptual hydrological models. This study hypothesized that the conceptual hydrological models that do include this information provide more reliable hydrological responses. To test this hypothesis, this study used two conceptual hydrological models, which are the lumped FLEX and GR4J models, to simulate daily values streamflow and evaporation for two distinct catchments in the United States.

To include information from plant phenology, this study applied three separate modifications to the conventional structures of the two conceptual hydrological models. The first modified structure uses a part of the Jarvis model, in which stomata respond to temperature, to partition evaporation. The second modified structure uses a method similar to the crop evaporation method of the Food and Agriculture Organization (Allen, Pereira, Raes, & Smith, 1998). Lastly, the third modified structure uses a combination of the other two structures.

This study concludes that conceptual hydrological models that include information from plant phenology, as in the three modified model structures, are able to provide streamflow simulations as good as models that do not include plant phenology. The streamflow simulations of conceptual hydrological models, such as the three modified model structures, are therefore not necessarily more reliable. However, conceptual hydrological models that do include information from plant phenology might still benefit hydrological studies that take into account the effects of long and short-term changes to the vegetation in a catchment, such as the effects of climate change, land-use change, and forest fires. Before plant phenology should be used in the models of such studies, challenges such as the uncertainties in the simulations of the evaporation components need to be addressed.

# Preface

Before you is the thesis ”The Impact of Vegetation on the Partitioning of Evaporation in Conceptual Hydrological Model”, which is the result of my graduation project to obtain the degree of Master of Science in Civil Engineering, track of Water Management, at the Delft University of Technology. I would like to thank my supervisor Markus Hrachowitz for his support and guidance during this study and my committee members, Susan Steele-Dunne and Miriam Coenders, for their valuable feedback.

Furthermore, I would like to thank my parents and brothers for their support and understanding during the past year, in which everyone had to adapt to different circumstances. Lastly, I would also like to thank my friends for their encouragement and advice.

*O. Mert*
*Delft, July 2021*

# Contents

# List of Figures

# List of Tables

# 1

# Introduction

A Roman scientist in the first century BC, named Marcus Vitruvius, developed one of the first theories about the hydrological cycle that is similar to our current understanding. With this theory, Vitruvius described the formation of springs and streams at the base of mountains as the result of infiltration and percolation of precipitation at the top of these mountains (Karterakis, Karney, Singh, & Guergachi, 2007). Despite Vitruvius' theory, up until the Middle Ages, many Europeans still believed that streamflow originated from subsurface caverns or lakes. It was only in the 16th century that a French engineer, named Bernard Palissy, was able to provide a science-based theory for the hydrological cycle. Palissy described the role of many hydrological processes and recognized the temporal discontinuity between precipitation and streamflow. Since then, numerous scientifically based experiments took place, which provided key insights for hydrological phenomena and stimulated the development of hydrology as a modern science (Rosbjerg & Rodda, 2019).

In the second half of the 20th century another major development for the hydrological science took place due to shifting societal needs (Sivapalan & Blöschl, 2017). During this period, countries that were recovering from the Second World War experienced economic growth through the investment of massive infrastructural projects such as the construction of dams and irrigation systems. The constructions of such systems needed economic optimality, which previous hydrological methods could not take into account. On top of this, the economic growth created another issue for many countries in the Western World. This issue was the contamination of water, for which the solution needed hydrological models that could predict the water quality in the absence of observations. The rise of computers and the advancement of technology provided many opportunities to satisfy the societal needs at that time and resulted in the creation of numerous hydrological models (Sivapalan & Blöschl, 2017; Sitterson et al., 2017).

Despite the digital revolution and the knowledge that has been built over multiple generations of scientists and engineers, many hydrological models still lack certainty and the ability to simulate reliable hydrological responses. This is due to several issues as described by Beven (2001a) and due to the uniqueness of any place around the globe. The uniqueness of a place complicates the extrapolation of data beyond their observed location and therefore complicates to take advantage of the available computational power for hydrological models that include micro-scale physics (Beven, 2000; Blöschl, 2001). One hydrological model that does not require micro-scale physics is the spatially lumped conceptual hydrological model. Such a model attempts to describe the hydrological system at the macro scale and therefore integrates and simplifies the micro-scale heterogeneities that are present in a catchment (Hrachowitz & Clark, 2017). As a result, conceptual hydrological models require relatively less complex structures and parameterizations. As a downside, however, conceptual hydrological models generally represent relatively less hydrological and thermodynamic processes in a physical way, which is often the case for the partitioning of evaporation.

In most conceptual hydrological models, the demand of the atmosphere for water is quantified by potential evaporation ($E_p$). Often in such models, the atmospheric water demand is first supplied by interception evaporation ($E_i$) and afterward, the remaining demand is supplied by transpiration ($E_t$), or the other way around depending on the modeler's decisions. The evaporation fluxes are not only limited by the atmospheric water demand, but also by the amount of water available for evaporation. Conceptual hydrological models generally use functions with certain parameters that determine the amount of energy and water available for evaporation (Zhao, Xia, Xu, & et al., 2013). These parameters require model calibration to obtain their values and are subsequently kept as constants throughout the entire simulation period. This method is easy to implement and hydrological models that use this method are often able to provide relatively accurate streamflow simulations. However, the defined order in which the atmospheric water demand is satisfied by the evaporation components is not always the same in reality. Even when water is not a limiting factor, vegetation is able to change this defined order as it is also affected by the number of leaves present on vegetation (Donohue, Roderick, & McVicar, 2007). The dynamics of vegetation are therefore significant for the partitioning of evaporation and thus also the generation of streamflow in a catchment.

Hydrological models that can accurately and precisely represent hydrological systems and predict reliable hydrological responses, are important for understanding and efficiently managing water fluxes and resources. According to Sivapalan and Blöschl (2017), in the near future, water management will be more complicated and failure to overcome challenges such as water security will likely result in problems beyond the borders of a country and cause political instability around the globe. Water security is already a huge challenge for 36 countries dealing with extreme water scarcity (Reig, Maddocks, & Gassert, 2013) and it is expected that half of the world's population will have to deal with lasting water scarcity by the year 2025 according to Organization (2019). To overcome challenges such as water security and other societal needs related to water, hydrological models that can predict reliable hydrological responses should be necessary.

This study attempted to improve the reliability of hydrological models by including the dynamics of vegetation on the partitioning of evaporation. The dynamics of vegetation are in this case described by phenology, a study that is concerned with the timing of recurring life cycle stages and their relation to the climate and environment. Furthermore, this study only focused on conceptual hydrological models due to their relatively simple structures and ease of modifications. The intention behind this was to start simple and gradually increase complexity when necessary.

## 1.1. Hydrological Models

The construction of a hydrological model consists of 5 stages, which are respectively the perceptual model, the conceptual model, the procedural model, model calibration and model validation (Beven, 2001b). Starting with the perceptual model, the modeler perceives the real world and correspondingly decides on the dominant processes to include in the model. In the conceptual model, the modeler decides on the necessary equations, the relations of the dominant processes, and anything else that is necessary for the model to run. The construction of the conceptual model is considered as the creative stage, in which the modeler builds the entire structure of the model. This stage should not be mistaken as the hydrological model type sharing the same name. To prevent any confusion, the latter is in this study described as conceptual hydrological model. During the third stage, the procedural model, the modeler translates the conceptual model into computer code. Furthermore, the fourth and the fifth stages are respectively the calibration of the model parameters and the validation of the model responses.

Hydrological models at the catchment scale can be categorized within a so-called spatial resolution-process complexity continuum (Hrachowitz & Clark, 2017). This continuum describes the models placed in one end as spatially lumped, conceptual, bucket-based and top-down and the models in the other end as spatially distributed, physically based, continuum-based and bottom-down.

### 1.1.1. Conceptual Hydrological Models

Looking at the models in the spatially distributed and high process complexity end of the spatial resolution-process complexity continuum (Hrachowitz & Clark, 2017), physically-based models take up the space. This model type represents the hydrological system by detailed processes of fluid movements, which are based on the Navier-Stokes and conservation equations. At the spatially lumped and low process complexity end of the continuum, conceptual hydrological models take up the space. The models in this very end represent the hydrological system by a top-down approach and use ''buckets'' to describe the water reservoirs at the necessary scales and relatively simple processes to describe the fluxes between these ''buckets''.

It could be argued that physically based models should always be able to produce reliable hydrological responses due to their dependence on micro-scale physics and their objective to represent the reality as detailed as possible. This might be true at the point scale, however, is likely not the case at larger spatial scales such as at the catchment scale. At larger spatial scales, multiple issues arise that hinder physically-based spatially distributed models to produce reliable responses (Beven, 2001a). A few of these issues are the issue of scale, heterogeneity, and equifinality. The issue of scale implies that the micro-scale physical properties and processes observed from the human perspective do not necessarily relate to those occurring at the larger spatial scales such as at the catchment scale. Heterogeneity refers to the issue that physical properties and processes are highly variable at any scale and that at larger scales this issue only becomes worse and more apparent (Beven, 2000; Blöschl, 2001; Sivapalan, 2003a). The issue of equifinality is associated with overparameterization and is frequently the case for complex physically-based models. Equifinality is the result of a model that performs equally well for different combinations of calibrated parameters (Beven, 1993). This makes it difficult to select the combination of calibrated parameters that can provide reliable model responses. The issues of scale, heterogeneity, and equifinality are not only limited to physically-based models. Conceptual hydrological models also suffer from the same issues, however, at the same time present some opportunities.

Physically-based models tend to zoom in to the micro-scale, whereas conceptual hydrological models tend to zoom out to the scale of the catchment. The latter presents opportunities to capture the self-organizing capacity of a catchment. The result of self-organization is from the emergence of spatial patterns through certain structures. The presence of these spatial patterns makes it possible to represent the micro-scale heterogeneities by only a few simple processes at the catchment scale (Dooge, 1986; Sivapalan, 2003a, 2006). In other words, conceptual hydrological models provide the opportunity to only include simple processes relevant at the catchment scale and to take advantage of the observations made at the appropriate scale. Remote sensing is in this case an useful tool for observations at the catchment scale. Furthermore, conceptual hydrological models require fewer parameters for model calibration and therefore have a reduced issue of equifinality compared to the physically-based models.

## 1.2. Evaporation

Evaporation is a combination of two major processes, in which water is transported from the earth's surface into the atmosphere. These two major processes consist of evaporation from the combined surfaces of open water ($E_o$), soil ($E_s$) and interception ($E_i$) and of transpiration ($E_t$) through vegetation. Evaporation ensures the phase change of liquid water into water vapor through an endothermic process, in which energy is taken up from solar radiation and the surrounding air temperature. The energy used by evaporation is defined as the latent heat of vaporization ($\lambda$) and is assumed to be constant at 2.45 MJ/kg. The latent heat of vaporization describes the amount of Joule needed to vaporize 1 kg of water. The rate of evaporation depends on a variety of factors of which a few are solar radiation, air temperature, relative humidity, and wind speed. Evaporation is a continuous process under the condition that water is available and the saturated air above the evaporating surface gets replaced by drier air (Allen et al., 1998).

Unlike the other group of evaporation components, $E_t$ is considered a three-stage process (Penman, 1956). The first stage can be considered as one in which the roots of plants take up water from the soil. During the second stage, plants transport this water throughout their bodies into the intercellular spaces of their leaves, where eventually vaporization takes place. In the final stage, plants use their stomata to remove the water vapor from the inter-cellular spaces of their leaves into the atmosphere. Stomata are small openings on the outer layers of leaves and are responsible for regulating plants' gas exchange with the atmosphere. Apart from the meteorological factors, $E_t$ also depends on the ability of plants to obtain water and transpire. This ability relies among others on factors such as the hydraulic properties of the soils around the roots, the chemical properties of the soils, the water saturation of the soils, the development stage of the plant, and the species of the plant (Allen et al., 1998).

### 1.2.1. Potential Evaporation

The rate for the combined evaporation processes can be estimated or measured by a variety of methods and devices. A commonly used method for the calculation of potential evaporation ($E_p$) is the Penman-Monteith equation (Monteith, 1965):

$$E_p = \frac{\Delta(R_n - G) + \frac{\rho_a * c_p}{r_a}(e_s - e_a)}{\rho_w * \lambda * (\Delta + \gamma(1 + \frac{r_s}{r_a}))} \tag{1.1}$$

Like the term ''potential'' suggests, $E_p$ refers to the amount of evaporation in case of the availability of unlimited water. In other words, $E_p$ is the demand of the atmosphere for water. This atmospheric water demand depends on the available energy and on the ability of the atmosphere to replace the air above the evaporation surface. The addition of the parameters surface resistance ($r_s$) and aerodynamic resistance ($r_a$) make the Penman-Monteith equation suitable to also calculate $E_p$ from soil, interception and transpiration from vegetation. A detailed description of each parameter and how to calculate the values can be retrieved from Monteith (1965) or Allen et al. (1998). A short description of each parameter is listed below:

$E_p$ = potential evaporation [$\frac{\mathrm{mm}}{\mathrm{day}}$];

$\Delta$ = slope of relationship between saturated vapor pressure and air temperature [$\frac{\mathrm{kPa}}{^\circ\mathrm{C}}$];

$R_n$ = net radiation [$\frac{\mathrm{MJ}}{\mathrm{m^2 day}}$];

$G$ = ground heat flux [$\frac{\mathrm{MJ}}{\mathrm{m^2 day}}$];

$\rho_a$ = density of air at constant pressure [$\frac{\mathrm{kg}}{\mathrm{m^3}}$];

$c_p$ = specific heat capacity of air [$\frac{\mathrm{MJ}}{\mathrm{kg}\ ^\circ\mathrm{C}}$];

$r_a$ = aerodynamic resistance [$\frac{\mathrm{s}}{\mathrm{m}}$];

$e_s$ = saturated vapor pressure [kPa];

$e_a$ = actual vapor pressure [kPa];

$\rho_w$ = density of water [$\frac{\mathrm{kg}}{\mathrm{m^3}}$];

$\lambda$ = latent heat of vaporization [$\frac{\mathrm{MJ}}{\mathrm{kg}}$];

$\gamma$ = psychrometric constant [$\frac{\mathrm{kPa}}{^\circ\mathrm{C}}$];

$r_s$ = bulk surface resistance [$\frac{\mathrm{s}}{\mathrm{m}}$].

Another commonly used method for the calculation of $E_p$ is the Priestley-Taylor equation (Priestly & Taylor, 1972) (Eq. 1.2). This equation is often used as a replacement for the Penman-Monteith equation in case of lack of observations. The Priestley-Taylor equation is therefore almost identical to the Penman-Monteith equation and mostly uses the same parameters. The Priestley-Taylor coefficient, $\alpha$, is the only parameter that is present in the Priestley-Taylor equation and absent in the Penman-Monteith equation. This additional coefficient makes up for the eliminated resistance values and vapor pressure deficit. Generally, the coefficient $\alpha$ has a value of 1.26 for wet areas (land and water) and higher values for arid areas (Priestly & Taylor, 1972; Morton, 1983). Further detailed information on each parameter in the Priestley-Taylor equation can be retrieved from (Priestly & Taylor, 1972).

$$E_p = \alpha \frac{\Delta(R_n - G)}{\rho_w * \lambda * (\Delta + \gamma)} \tag{1.2}$$

### 1.2.2. Actual Evaporation

Actual evaporation is the rate of evaporation that takes place in reality and is defined as the net result of the atmospheric water demand. A range of methods and devices exist to determine actual evaporation as a whole and each of its components individually. The suitability of a method and device changes according to the spatial scale of interest. At the plot scale, which is around 1 $\mathrm{m}^2$, a Lysimeter is usually suitable to measure actual evaporation above surfaces combined with soil, interception, and vegetation. At a spatial scale of around 10 ha, the methods Bowen ratio, which is combined with the energy balance, and eddy covariance become more useful to measure actual evaporation. For spatial scales above 10,000 ha, water balance methods become effective to estimate the actual evaporation. Considering these methods and devices, the lysimeter is the only practical tool to also measure each evaporation component separately.

In general, the methods and devices that determine actual evaporation can be categorized into direct and indirect methods. Direct methods include methods such as pan evaporation, lysimeter, sap flow, soil heat pulse, and chamber, and depending on their setup only measure one or a combination of actual evaporation components. These methods only provide reliable evaporation rates for the scale they are measured, which is usually at the plot scale, and therefore complicates to upscale the measurements to the catchment scale.

Indirect Methods include the methods Bowen ratio, eddy covariance, scintillometer, and water balance. These indirect methods can determine evaporation at larger spatial scales, such as at the catchment scale. However, they are usually not able to determine the components of evaporation individually. To obtain an estimation of the latter, partitioning techniques are often used. The integration of land surface models and remote sensing data provide opportunities to partition actual evaporation at the catchment scale and anywhere around the world.

#### Actual Evaporation (Open Water, Soil and Interception)

The Penman equation (Penman, 1948) can be used to calculate actual evaporation from open water. This equation is almost identical to the Penman-Monteith equation. The only difference is that the Penman equation does not include the additional parameters for the surface and aerodynamic resistances. Actual evaporation from open water can also be measured by the use of the pan evaporation method or a lysimeter filled with water. Additionally, the lysimeter has the advantage to measure actual evaporation above different surfaces, such as soil and interception, separately. This is because the lysimeter only measures actual evaporation above the surfaces of the samples it uses. Furthermore for the measurement of soil-only surfaces, the combination of soil heat pulses and energy balances can be used.

### Actual Transpiration

At the plot scale, actual $E_t$ can be measured by a lysimeter filled with vegetation or by the sap flow method for individual trees. For larger spatial scales, chambers can be used. Additionally, many other methods and devices exist that can partition actual evaporation and estimate the rates of each evaporation component separately (Kool et al., 2014; Stoy et al., 2019). The selection of such a method then mainly depends on the spatial scale of interest and the objective of the user. The combination of such methods might also be complementary and provide more accurate results.

## 1.3. Phenology

Phenology is a study that is concerned with the timing of recurring life cycle stages and their relation to the climate and environment. Phenology related to plants in specific describes the timing of phenological events such as leafing, flowering, senescence, and abscission. A phenological event, in this case, is thus considered as a recurring moment in the life cycle of plants and indicates the start or the end of a phenophase (Thomas, Denny, Miller-Rushing, Crimmins, & F., 2010). A phenophase is considered as a certain development stage of plants such as the period in which leaves or flowers are present. The timing of phenological events is not always the same. Plants can shift the timing of their phenological events according to any change in their environment and climate. Phenology can be used to include the dynamical behavior of vegetation in hydrological models. Vegetation and its dynamics should be considered important in hydrological models as approximately 71 % of land on earth is covered by vegetation (Latham, Cumani, Rosati, & Bloise, 2014). On top of that, vegetation takes up a major part in the allocation of energy and water over the earth's surface by impacting the rates of $E_i$ and $E_t$ and inversely of $E_s$ (Arora, 2002; Donohue et al., 2007; L. Zhang, Dawes, & Walker, 2001).

The development of plants depends on a variety of factors. Some major factors that drive this development are temperature, precipitation, and light (Manske, n.d.; Nord & Lynch, 2009). Temperature influences the length and timing of the growth seasons. Too high temperatures prevent the ability of plants to run their biological processes and too low temperatures freeze water which then becomes inaccessible to plants. The inaccessibility of water prevents again the ability of vegetation to run its biological processes and its ability to obtain other natural resources. Precipitation provides fresh water to the roots of the plants, which use this water for among others the growth and maintenance of their tissue and for the absorption of natural resources, such as nutrients and minerals. Light ensures the ability of plants to perform photosynthesis. Through photosynthesis, plants split water into oxygen and hydrogen and produce chemical energy (Arora, 2002). The hydrogen and chemical energy are then combined with carbon dioxide from the atmosphere to form carbohydrates. These carbohydrates are responsible for the production of biomass and for the operations of other cellular processes within plants.

Other factors that can impact the development of plants are related to their direct environment. Such factors include soil strength, soil temperature and soil moisture (McMichael & Quisenberry, 1993; Pregitzer, King, Burton, & Brown, 2000; Nord & Lynch, 2009). Soils with a high bulk density for example can significantly decrease the root growth of plants and therefore also decrease the ability of plants to obtain natural resources. Another significant factor to influence the development of plants is topography. Topography can for example affect the water table depths and therefore also affect the rooting depths of plants (Fan, Miguez-Macho, Jobbágy, Jackson, & Otero-Casal, 2017).

### 1.3.1. Leaf Area Index

Plant phenology attempts to investigate the timing and lengths of the growth seasons and the development of plants over multiple years. For its development, plants interact with soil water through $E_t$ and as a result impact the generation of streamflow at the catchment scale (Donohue et al., 2007). This interaction can be indicated by physiological properties of plants that appear at the catchment scale such as the Leaf Area Index (LAI) and the stomatal conductance (Arora, 2002). The development of leaves for example drives the ability of plants to obtain natural resources such as water, nutrients, and minerals through the process of $E_t$ (Nord & Lynch, 2009).

The Leaf Area Index indicates the density of vegetation covers and the capacity of vegetation to use light. The Leaf Area Index is a dynamic quantity that changes over time and space according to the development of the associated vegetation, the environment, and the climate. Additionally, LAI can significantly change over a short period due to human-induced forces such as deforestation and natural disasters such as wildfires. The Leaf Area Index indicates the capacity of foliage to absorb a fraction of the Photosynthetically Active Radiation (fPAR) and hence affects transpiration and evaporation rates (Donohue et al., 2007). The Leaf Area Index is defined as the ratio of the leaf area over the ground area:

$$LAI = \frac{Leaf\ Area\ [\mathrm{m}^2]}{Ground\ Area\ [\mathrm{m}^2]} \tag{1.3}$$

A range of methods exists to determine LAI and generally can be categorized into direct and indirect methods. The direct methods include ground-based measurement techniques such as destructive sampling and litterfall collection and are generally considered as labor-intensive and time-consuming. The indirect methods include remote sensing data and are usually considered to overcome the disadvantages of the direct methods. However, as a drawback, the indirect methods tend to underestimate LAI in dense vegetation covers due to overlying leaves.

## 1.4. Problem Definition

At the global scale, $E_t$ is considered one of the dominant fluxes within the sum of terrestrial evaporation processes and is estimated to account for roughly 80 % at most of this sum

(Coenders-Gerrits, van der Ent, Bogaard, & et al., 2014). Transpiration is a complex process that is interconnected with the development of vegetation and is influenced by a range of factors. Hence the reason a variety of methods and devices exist to estimate $E_t$ at different spatial and temporal scales (Stoy et al., 2019). Despite the significance of $E_t$, many conceptual hydrological models simplify this process and neglect its interconnection with the dynamical behavior of vegetation.

In most conceptual hydrological models, the amount of energy available to evaporation is defined by $E_p$, or in other words the atmospheric water demand. When there is no limit to the water availability, the atmospheric water demand in such models is often first satisfied by $E_i$ and the demand that remains afterward by $E_t$, or depending on the modeler's decisions the other way around. Usually, the availability of water is, however, limited and therefore also puts a constraint on both evaporation fluxes. Many conceptual hydrological models use relatively simple functions to describe the amount of energy and water available for both evaporation components. Such functions use parameters, which need model calibration for their values to be determined (Zhao et al., 2013). These calibrated parameter values subsequently remain constant for the entire simulation period. The advantage of this method is that it is easy to implement and hydrological models that use this method might provide accurate streamflow simulations. However, such models might not always provide accurate $E_i$ and $E_t$ simulations due to the defined order in which the atmospheric water demand is satisfied by both evaporation fluxes. In other words, hydrological models that use this method might provide accurate streamflow simulations, but not necessarily for the right reasons and could therefore be considered as unreliable.

The assumption of a defined order, in which the atmospheric water demand is satisfied by the evaporation fluxes, is in contrast with the perception of the real world. In reality, vegetation is able to change this defined order over time by impacting the partitioning of evaporation as it is able to change the rates of $E_t$ and $E_i$. and inversely change the rate of $E_s$, according to the number of leaves and litter present (Donohue et al., 2007). This dynamic behavior of vegetation is not only able to affect the partitioning of evaporation, but also the generation of streamflow. Phenology can be used to describe the dynamics of vegetation, as it relates the timing of phenological events such as leafing and abscission to the changes in seasons and climates. The development of leaves is for instance a major driver that determines the ability of vegetation to transpire and additionally influence the rates of $E_i$ and $E_s$ (Arora, 2002; Donohue et al., 2007; Nord & Lynch, 2009; L. Zhang et al., 2001).

The application of physically-based spatially distributed models to include the dynamic behavior of vegetation does not necessarily provide a better solution as it usually results in overly complex models that are unable to simulate reliably responses due to equifinality and several other issues (Beven, 2001a). It might therefore be more effective to explore holistic and simple methods to represent the relation between vegetation and evaporation at the catchment scale.

## 1.5. Research Objective, Hypothesis and Questions

### Research objective

The objective of this study is to include information from plant phenology on the partitioning of evaporation in conceptual hydrological models.

### Hypothesis

Conceptual hydrological models that include information from plant phenology on the partitioning of evaporation will simulate more reliable hydrological responses compared to those that do not.

### Research questions

1. How can conceptual hydrological models include information from vegetation phenology to partition evaporation at the catchment scale?
2. How do conceptual hydrological models that use the conventional approach perform compared to the conceptual hydrological models that include information from vegetation phenology?
3. How do conceptual hydrological models that include information from vegetation phenology perform for catchments with different dominant land covers?

## 1.6. Thesis Outline

Chapter 2 includes information about the two catchments for which the conceptual hydrological models are applied and includes information about the data that is used for the operation and evaluation of the models

Chapter 3 includes the methodology of this study. This chapter describes the conceptual hydrological models that are used, the three additional model structures for the inclusion of plant phenology, and the methods for the calibration and validation of the models

Chapter 4 describes the results of the simulations of each model for both catchments. Chapter 5 includes the discussion in which the results of the simulations are interpreted. Lastly, Chapter 6 describes the conclusion of this study and the recommendations for future studies.

# 2

# Study Areas and Data

Every vegetation type exerts a different dynamic behavior towards the development of leaves and therefore influences the partitioning of evaporation differently. Vegetation types can be categorized by biomes, which are communities of animals and plants that have formed in response to their environment and climate. Walter and Breckle (2002) classifies nine separate biomes on Earth and relates the boundary of each biome to the differences in the distribution of vegetation types. To account for the different behavior of vegetation types, this study focused on two catchments in a separate biome with a distinct vegetation type. The two selected catchments are located in the United States and are selected due to characteristics of their dominant vegetation types and the availability of data.

According to the classification of Walter and Breckle (2002), six of the nine biomes are present within the contiguous United States. These biomes are Woodland/Shrubland, Subtropical Desert, Temperate Rainforest, Temperate Seasonal Forest, Temperate Grassland/Desert, and Boreal Forest. One of the selected catchments is located within a Temperate Seasonal Forest in the eastern part of the United States and the other in a Temperate Rainforest in the western part of the United States. The biome Temperate Seasonal Forest includes deciduous broadleaf forests in a temperate climate with cold winters and the biome Temperate Rainforest includes evergreen needleleaf forests in a temperate climate with heavy rainfall and high moist conditions.

## 2.1. Catchments

The United States Forest Service (USFS) is tasked with the administration of national forests and grasslands within the United States of America and manages approximately 780,000 $\mathrm{km}^2$ of land. The United States Forest Service includes a network of long-term experimental lands across the United States that are dedicated to research regarding land management and ecological change. The network started in 1908 and since then 84 experimental forests and ranges have been founded. The two selected catchments are located in such experimental forests and are Hubbard Brook Experimental Forest's watershed 3 and HJ Andrews Experimental Forest's watershed 2. The reasons for the selection of these two catchments are due to their dominant vegetation types, sizes and availability of long-term observational data. The land cover in Hubbard Brook's watershed 3 is dominated by deciduous broadleaf forests, whereas the land cover in HJ Andrews' watershed 2 is dominated by evergreen needleleaf forests. This way the hypothesis of this study could be tested for two catchments with distinct dominant vegetation types, as different vegetation types differently impact the partitioning of evaporation. One catchment with a deciduous broadleaf forest for a highly variable pattern in the LAI and the other catchment with an evergreen needleleaf forest for a less variable pattern in the LAI. Both of the selected catchments have approximately the same size such that they include heterogeneity at the same spatial scale. The long-term observational data is required to include the dynamical behavior of the vegetation in the catchments over multiple seasons.

### 2.1.1. Hubbard Brook

One of USFS's experimental forests near the east coast of the United States is the Hubbard Brook Experimental Forest (Study, 2021). Hubbard Brook is located within the White Mountain National Forest in the state of New Hampshire and consists of an area of approximately 32 $\mathrm{km}^2$. The experimental forest was established in 1955 and has since collected data and knowledge about the responses of animals, plants, soils, water, and air to the management of land.

The Hubbard Brook Experimental Forest is characterized by a continental climate, in which long, cold winters take place and short, mild summers. The average annual precipitation is approximately 1400 mm and for evaporation around 500 mm, which is estimated by taking the difference between the measured precipitation and streamflow. The average temperature in January is -9 °C and in July is 18 °C. Around a third of the precipitation falls as snow and on average 145 days a year are frost-free.

The geological structure of the eastern part of the Hubbard Brook Experimental Forest, which includes watershed 3, consists of igneous and metasedimentary rocks. The soils in the experimental forest consist of mostly Spodosols and sandy loam. On top of the soil is usually a layer between 20 and 200 mm of forest floor, which mostly consists of leaf litter and other organic material. Many streams in the experimental forest frequently dry up in the summers and quickly fill after a precipitation event or snowmelt due to the high soil porosity.

The major landcover in the Hubbard Brook Experimental Forest consists of a northern hardwood forest, which is dominated by three deciduous broadleaf tree species and another mixture of deciduous and evergreen conifer species. The deciduous broadleaf species are comprised of American beech, yellow birch, and sugar maple. The leaves of these trees fully develop around 15 May and start falling around 15 September.

Hubbard Brook contains ten experimental watersheds in which long-term monitoring has taken place of streamflow and water chemistry. Each of the watersheds drains into the Hubbard Brook stream, which eventually flows into the Pemigewasset River. Some of the watersheds have been subject to experimental treatments for scientific purposes. Other watersheds have not experienced any treatment and therefore serve as reference watersheds. Among these none treated watersheds is watershed 3 (WS3), which acts as a hydrological reference in the

Hubbard Brook Experimental Forest.

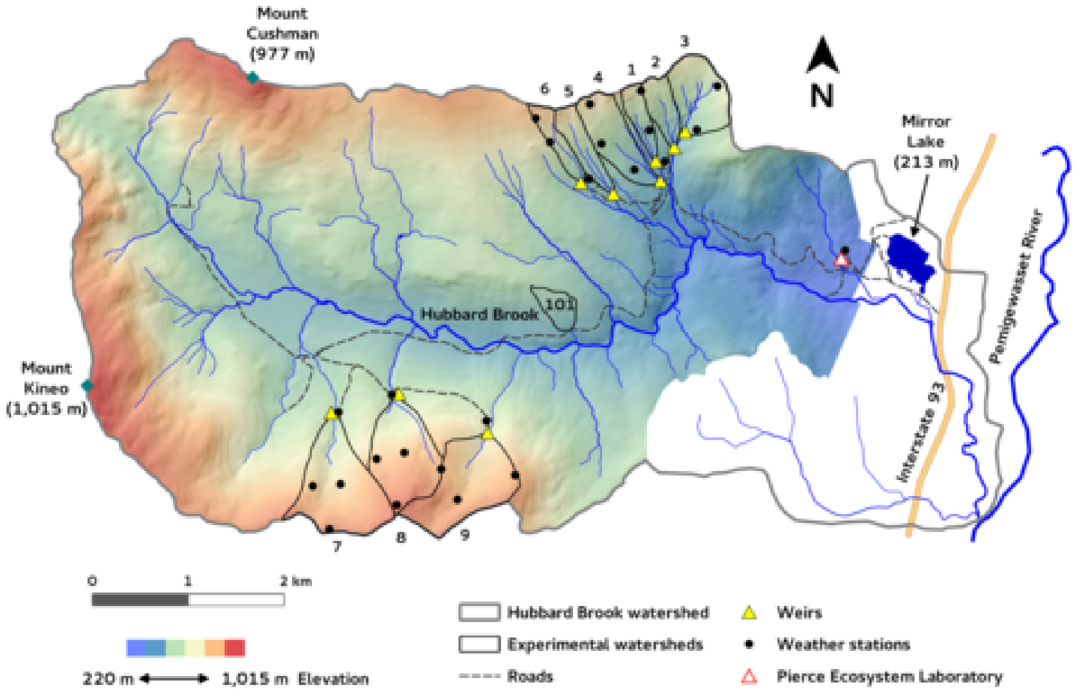


Figure 2.1: Map of Hubbard Brook Experimental Forest with the boundaries of the watersheds, the locations of the weather stations, and the locations of the gauging stations (Study, 2021). Selected catchment is watershed 3.

Watershed 3 is located on the northeast of the Hubbard Brook Experimental Forest and has a size of approximately 42.4 ha. The watershed has a slope of 12.1° and an elevation that ranges from 527 to 732 m above sea level. Watershed 3 is gauged by a V-notch weir for the daily observation of streamflow and contains two weather stations for the daily observation of precipitation. The weather station near the headquarters (Pierce Ecosystem Laboratory) provides daily meteorological data such as air temperature, solar radiation, vapor pressure, and wind speed.

### 2.1.2. HJ Andrews

One of USFS's experimental forests near the west coast of the United States is the HJ Andrews Experimental Forest (Station, 2003). HJ Andrews is located within the Western Cascades in the state of Oregon and has an area of approximately 64 $\mathrm{km}^2$. The experimental forest was established in 1948 and similarly to Hubbard Brook has collected data and conducted research about forest management, ecology, and hydrology.

The HJ Andrews Experimental Forest has a maritime climate, that is characterized by wet, mild winters and dry, cool summers. Precipitation mostly takes place in the months from November to March and on average is approximately 2300 mm per year near lower elevations. At higher elevations, precipitation increases to an average of approximately 3550 mm per year. Snowfall is common at the higher elevations and usually starts in November. Snowfall at the lower elevations makes up around a quarter of the precipitation. The average temperature is 1 °C in January and 18 °C in July.

Volcanic rocks make up the geological structure at the lower elevations of HJ Andrews, whereas bedrock formed by andesite lava flows dominates at the higher elevations. The landscape of HJ Andrews is characterized by steep slopes due to erosion and glaciation and the soils consist mostly of Inceptisols with less widespread Alfisols and Spodosols.

Dense evergreen conifer forests cover the landscape of HJ Andrews. The dominant tree species at the lower elevations are western red cedar, western hemlock, and Douglas-fir. At the higher elevations of HJ Andrews, the dominant tree species are western hemlock, Douglas-fir, Pacific silver fir, and noble fir.

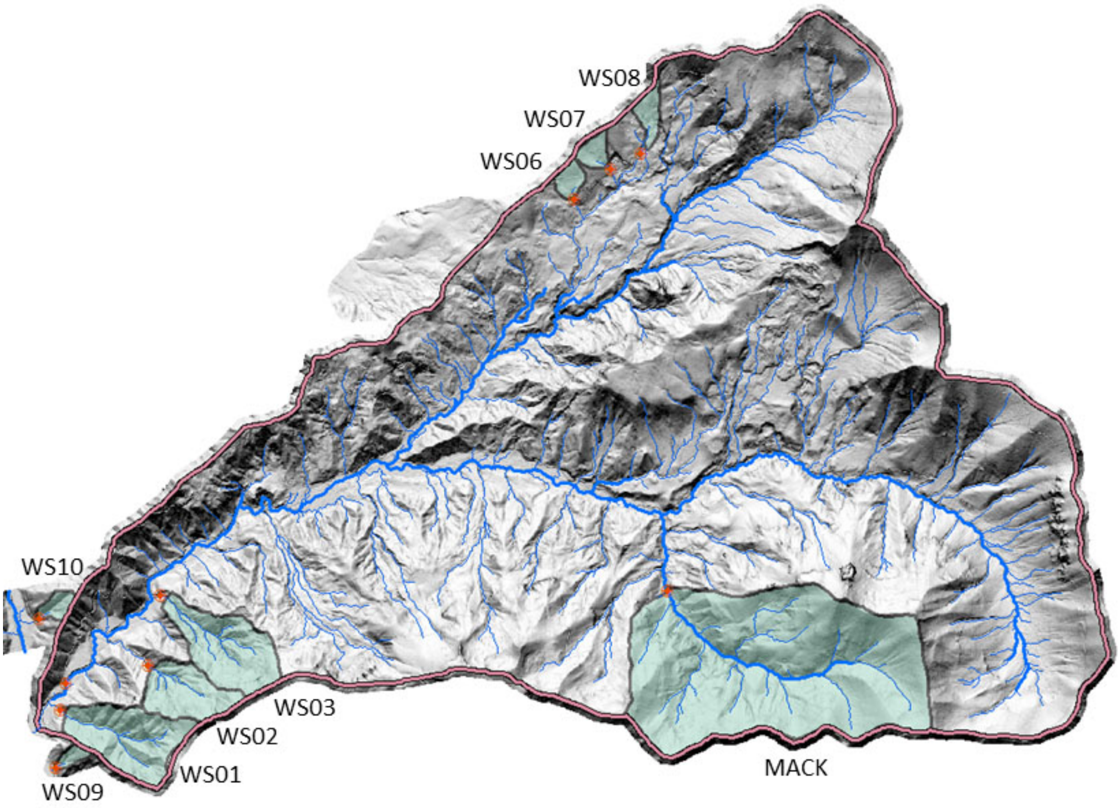


Figure 2.2: Map of HJ Andrews Experimental Forest with the boundaries of the watersheds and the locations of the gauging stations (Forest, 2021). The selected catchment is watershed 2.

Long-term monitoring of streamflow and water chemistry has taken place for nine experimental watersheds in HJ Andrews. These watersheds drain into the Lookout Creek, which eventually flows into the Blue River. Similar to Hubbard Brook, some of the experimental watersheds in HJ Andrews have been deliberately disturbed for hydrological and ecological studies. The remaining watersheds, such as watershed 2 (WS2), are undisturbed and act as references.

Watershed 2 is situated on the southwest of the HJ Andrews Experimental Forest and has an area of approximately 60 ha. Watershed 2 has a slope of 28.01°, a minimum elevation of 545 m, and a maximum elevation of 1079 m above sea level. Watershed 2 is gauged by a trapezoidal flume with an additional V-notch plate during the summer, low flow periods and provides daily streamflow observations. The meteorological station in WS2 provides daily observations of precipitation and air temperature and the primary meteorological station near the headquarters of HJ Andrews additionally provides daily observations of solar radiation. The headquarters is located near the mouth of the experimental forest.

## 2.2. Evaporation

The integration of land surface models with remote sensing data provide opportunities to partition evaporation at the catchment scale and anywhere around the globe. Due to this reason, a collection of evaporation products, which used this integration of land surface models and remote sensing data, are used to obtain evaporation estimates for the selected catchments.

### 2.2.1. GLDAS Version 2

The NASA Global Land Data Assimilation System (GLDAS) produces a range of land surface states and fluxes around the globe on a 0.25° to 1.0° spatial resolution and 3 hourly to monthly temporal resolution (Rodell et al., 2004; Rui & H., 2021). To produce the relevant data products, GLDAS combines data that is observed from space and ground with advanced techniques such as land surface modeling. For this, GLDAS makes use of four different land surface models, which are the Variable Infiltration Capacity (VIC), the Community Land Model (CLM), the Catchment (CLSM), and the National Centers for Environmental Prediction/Oregon State University/Air Force/Hydrological Research lab (Noah). Land surface models can simulate the movement of energy and water between the Earth's surface and atmosphere.

Version 2 of GLDAS (GLDAS-2) contains two components, namely GLDAS-2.0 and GLDAS-2.1, that provide new and reprocessed data products. The first component, GLDAS-2.0, uses the updated Princeton Global Meteorological Forcing Dataset as the input data and provides the relevant data products for the period between 1948 and 2014. The second component, GLDAS-2.1, uses a combination of observed and modeled data as input data and provides the relevant data products since the year 2000. Both components provide evaporation estimation data products, which are described as canopy water evaporation, direct evaporation from bare soil, and transpiration and are quantified in W/$\mathrm{m}^2$. This study regards these evaporation data products respectively as estimates for interception evaporation ($E_i$), soil evaporation ($E_s$) and transpiration ($E_t$).

This study used the GLDAS-2.0 and GLDAS-2.1 0.25° and 3 hourly products from the offline Noah land surface model (Chen et al., 1996; Chen & Dudhia, 2001; Ek et al., 2003). The products are provided in rates and therefore up-scaled to daily values through temporal averaging. The Noah model is an one-dimensional column model and numerically integrates the relevant equations representing the physical processes over soil, vegetation and snowpack. The Noah model is able to simulate soil moisture, soil temperature, skin temperature, snowpack water equivalent, snowpack depth, canopy water contents, and the energy and water fluxes of the surface energy and water balances.

The offline Noah land surface model, in which the simulations are carried out by driving the model with forcing datasets, estimates total evaporation as the sum of evaporation from canopy interception (interception evaporation), evaporation from the uppermost bare soil layer (soil evaporation), and transpiration. Interception evaporation in the Noah land surface model is determined by equation 2.1, in which $W_c$ is the amount of intercepted canopy water and $S$ is the maximum interception capacity of $W_c$. Furthermore, $\sigma_f$ is the fraction of green vegetation, which is the photosynthesis active green canopy that intercepts midday downward solar insolation, of a model grid cell and $E_p$ is the potential evaporation, which is calculated by the stability-dependent Penman equation (Mahrt & Ek, 1984). When $W_c$ increases above $S$, excess precipitation enters the soil as throughfall.

$$E_i = \sigma_f * E_p * \left(\frac{W_c}{S}\right)^n \tag{2.1}$$

Soil evaporation in the Noah land surface model refers to the water vapor flux from the bare soil, non-vegetated, fraction of a model grid cell and is defined by the following formulations:

$$FX = \frac{\theta_1 - \theta_{dry}}{\theta_{sat} - \theta_{dry}} \tag{2.2}$$

$$E_s = (1 - \sigma_f) * E_p * (FX)^{fx} \tag{2.3}$$

In this formulation, $FX$ is the soil moisture saturation fraction in the upper soil layer and $\theta_1$, $\theta_{dry}$, and $\theta_{sat}$ are respectively the soil moisture in the upper soil layer, the dry, and the saturation values. Finally, $fx$ is and empirical coefficient.

Transpiration in the Noah land surface model is determined by equation 2.4, in which $B_c$ indicates the canopy resistance with soil moisture stress and $(W_c/S)^n$ acts as a coefficient restraining $E_t$ in favor of $E_i$ when the surface of a canopy becomes more wet.

$$E_t = \sigma_f * E_p * B_c * \left[1 - \left(\frac{W_c}{S}\right)^n\right] \tag{2.4}$$

### 2.2.2. PML Version 2

Version 2 of the Penman-Monteith-Leuning Evaporation (PML-2) model produces estimated data for the evaporation and gross primary product around the globe on a 500 m spatial and 8-day temporal resolution for the period between 2002 and 2017 (Y. Zhang et al., 2019, 2016; Gan et al., 2018). PML-2 includes evaporation estimation data products, which are described as interception from vegetation canopy, soil evaporation, and vegetation transpiration and quantified in mm/d. This study regards these evaporation data products respectively as $E_i$, $E_s$ and $E_t$.

The PML model estimates the total evaporation at each model grid cell as the sum of $E_i$, $E_s$ and $E_t$. Transpiration and soil evaporation are respectively estimated by equations 2.5 and 2.6. The available energy for absorption by the surface ($A$) is partitioned into energy absorption by canopy ($A_c$) and soil ($A_s$) using LAI. The Leaf Area Index is also used for the estimation of $G_c$. Interception evaporation is estimated by a modified version of the Gash rainfall interception model (van Dijk & Bruijnzeel, 2001). The PML model takes as input data the meteorological forcings from GLDAS-2.1, the remote sensing data from the MODIS collection 6 products, and the land cover types data from the NASA dataset center.

$$E_t = \frac{\varepsilon * A_c + (\rho * c_p/\gamma) * D_a * G_a}{\lambda(\varepsilon + 1 + (G_a/G_c))} \tag{2.5}$$

$$E_s = \frac{f * \varepsilon * A_s}{\lambda(\varepsilon + 1)} \tag{2.6}$$

$\varepsilon$ = $s$ / $\gamma$, $s$ = slope of relationship between saturated vapor pressure and air temperature $[\frac{\mathrm{kPa}}{{}^\circ\mathrm{C}}]$, $\gamma$ = psychrometric constant $[\frac{\mathrm{kPa}}{{}^\circ\mathrm{C}}]$;

$\rho$ = density of air at constant pressure $[\frac{\mathrm{kg}}{\mathrm{m}^3}]$;

$\lambda$ = latent heat of vaporization $[\frac{\mathrm{MJ}}{\mathrm{kg}}]$;

$c_p$ = specific heat capacity of air $[\frac{\mathrm{MJ}}{\mathrm{kg}\ {}^\circ\mathrm{C}}]$;

$D_a$ = water vapor pressure deficit of air [kPa];

$G_a$ = aerodynamic conductance $[\frac{\mathrm{MJ}}{\mathrm{m}^2\mathrm{d}}]$;

$G_c$ = canopy conductance $[\mathrm{m/s}]$;

$f$ = fraction of precipitation to equilibrium soil evaporation $[\frac{\mathrm{MJ}}{\mathrm{m}^2\mathrm{d}}]$;

$A_c$ = energy absorbed by canopy $[\frac{\mathrm{MJ}}{\mathrm{m}^2\mathrm{d}}]$;

$A_s$ = energy absorved by soil $[\frac{\mathrm{MJ}}{\mathrm{m}^2\mathrm{d}}]$;

## 2.3. Leaf Area Index

This study used a remote sensing product to obtain the Leaf Area Index (LAI) for the catchments due to the advantage of remote sensing to obtain data around the globe for long periods. The selected LAI dataset is a product of the MODIS sensors that are situated on the Terra and Aqua satellites of NASA (Myneni et al., 2015). This product is referred to as the MCD15A3H V6 level 4 and since July 2002 provides data for LAI and the Combined Fraction of Photosynthetically Active Radiation (FPAR) around the globe on a 500 m spatial and 4-day temporal resolution. The Leaf Area Index of this MODIS product is for broadleaf canopies defined as the one-sided green leaf area over the unit ground area and for coniferous canopies as half of the total needle surface area over the unit ground area.

The algorithm of the MODIS LAI product consists of the main process that is based on a Look-up-Table produced by the 3D radiative transfer equation and a backup process that utilizes empirical relationships between LAI and Normalized Difference Vegetation Index (NDVI). The main process makes use of the spectral data from the MODIS near-infrared (NIR) surface reflectance at 858 nm (NIR) and the red at 648 nm. The algorithm takes as input the structural type of the vegetation, the geometry of the sun-sensor, the Bidirectional Reflectance Factors (BRF) at a specified NIR, and red spectral band and the uncertainties.

Figure 2.3 contains two graphs that schematically show the implementation of the main algorithm process. Graph A consists of vegetated pixels that are distributed over their corresponding surface reflectances at the NIR and red spectral bands. A point on this graph with an area around it, such as the red dot and yellow ellipse, is regarded as the BRF, measured at a certain sun-sensor geometry, and its uncertainty. Any combination of the canopy/soil parameters and the corresponding FPAR values is considered an acceptable solution if the resulting modeled BRF falls within the area such as the yellow ellipse. Graph B shows the density distribution function of the acceptable solutions. The Leaf Area Index with its uncertainty is retrieved from the mean LAI and its standard deviation.

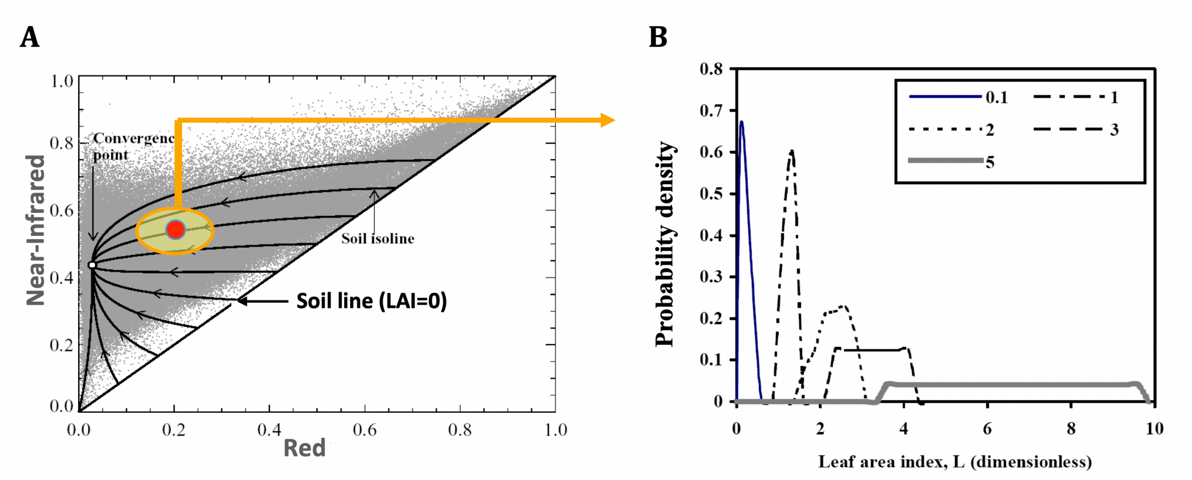


Figure 2.3: Illustration of MODIS LAI main algorithm (Myneni et al., 2015).

# 3

# Methodology

This study aimed to increase the reliability of conceptual hydrological models by including information from plant phenology on the partitioning of evaporation. To do so, this study focused on modifying two existing conceptual hydrological models instead of creating a new one to the already extensive list of hydrological models. The modification of the two conceptual hydrological models occurred in three different structures. The first structure used a function of the Jarvis model, in which stomata only respond to temperature. The second structure used the crop evaporation method as defined by the Food and Agriculture Organization (FAO) Allen et al. (1998) and the third structure used a combination of the first two structures.

## 3.1. Hydrological Models

The two selected conceptual hydrological models are the FLEX model (Fenicia et al., 2006) and the GR4J model (Perrin et al., 2003). Both models are considered spatially lumped, conceptual, and bucket-based. The selection of these two models was based on their relatively simple structures, which make it convenient to add or remove processes to the catchment as a whole. The intention behind this was to start simple and only add more complexity whenever necessary. The reason for the application of two models instead of one was to test the hypothesis of this study for different conceptual hydrological models and assess any potential bias towards a model's structure. Hence, the selection of the FLEX model and GR4J model, which are similar in terms of spatially lumped, conceptual, and bucket-based, however, have completely different structures. The main differences between the structures of the two models are that the FLEX model uses significantly more parameters than the GR4J model and does not include inter-catchment groundwater flow, whereas the GR4J model aims to use minimal parameters and does include inter-catchment groundwater flow.

### 3.1.1. FLEX Model

The FLEX model (Fenicia et al., 2006) is usually considered as a spatially lumped conceptual hydrological model with relatively low process complexity. The FLEX model is based on the HBV model, which has been accepted as a robust and successful rainfall-runoff model for numerous catchments around the world (Bergström, 1992). The FLEX model only includes hydrological processes that are relevant at the catchment scale and usually includes four reservoirs that indicate the various water storages at the appropriate scale. The four reservoirs consist of the interception reservoir ($S_i$), the unsaturated soil reservoir ($S_u$), the fast-reacting reservoir ($S_f$), and the slow reacting reservoir ($S_s$). This study used a slightly different structure for the FLEX model due to the notable number of snowy days and the fast rainfall-runoff responses in the selected catchments. This slightly different structure is based on the lumped FLEX model in Gao, Hrachowitz, Fenicia, et al.(2014). Hence, the FLEX model in this study included an additional snow reservoir ($S_w$).

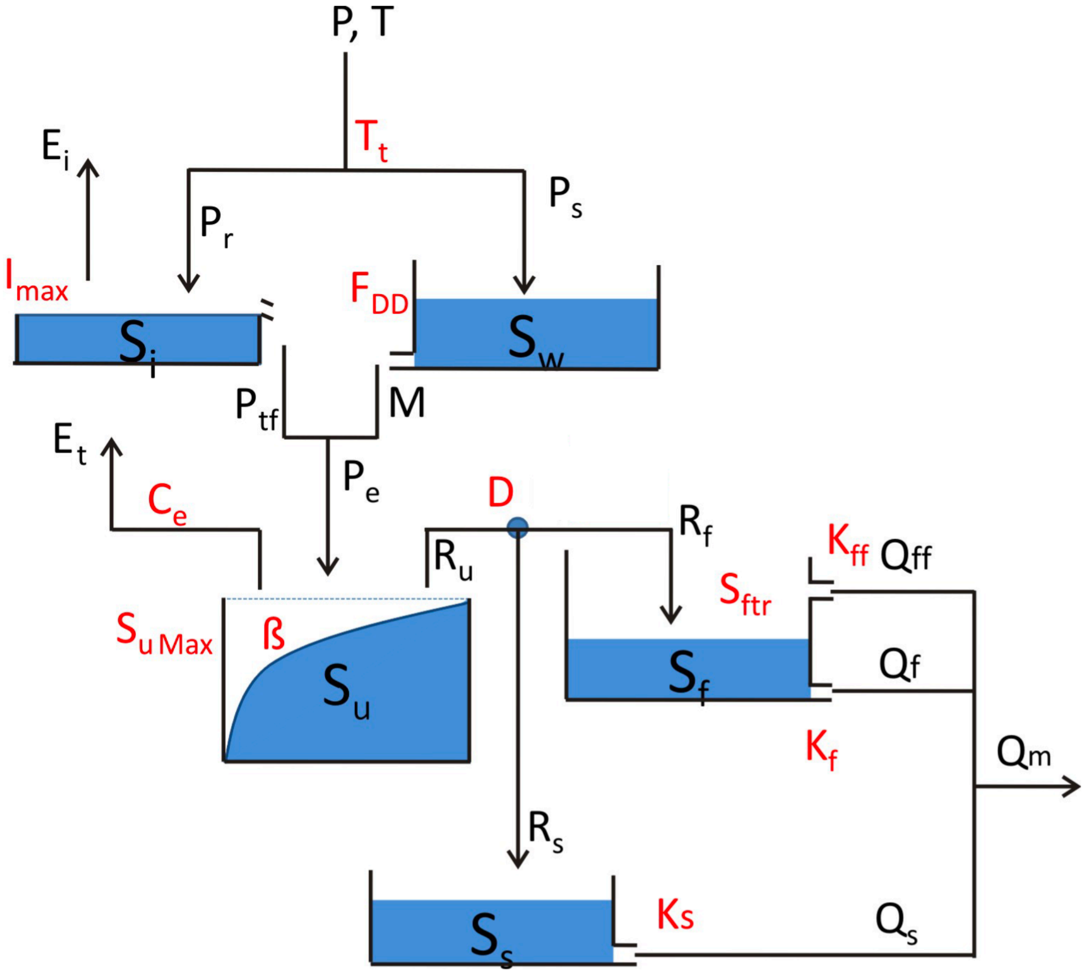


Figure 3.1: Structure of the lumped FLEX model (Gao, Hrachowitz, Fenicia, et al., 2014).

### Snow reservoir

Precipitation ($P$) enters $S_w$ and accumulates as snow ($P_s$) when the temperature ($T$) is below the threshold temperature ($T_t$). Water only leaves $S_w$ as effective precipitation ($P_e$) when $T$ is above $T_t$ and snow is still present. Effective precipitation from $S_w$ is then a combination of $P$, if present, and snowmelt ($M$). The calculation of $M$ consists of a simple degree-day model, which is the multiplication between the degree-day factor ($F_{DD}$) and the difference between $T$ and $T_t$. The degree-day factor is defined as the amount of water that is melted per day and Celsius degree.

$$M = \min(S_w, F_{DD}(T - T_t)) \quad \text{if} \quad T > T_t \tag{3.1}$$

### Interception reservoir

Precipitation enters $S_i$ as rain $P_r$ and accumulates as intercepted water when $T$ is above $T_t$ and no snow is present. Rain can only be intercepted up until a certain threshold value ($I_{max}$). Above this threshold value, excess rain ($P_{tf}$) leaves $S_i$ as $P_e$. Intercepted water is reduced through interception evaporation ($E_i$), which is quantified by potential evaporation ($E_p$) and the amount of intercepted water.

$$P_{tf} = \max(0, S_i - I_{max}) \tag{3.2}$$

$$E_i = \min(S_i, E_p) \tag{3.3}$$

### Unsaturated soil reservoir

A part of $P_e$ enters $S_u$ according to the runoff coefficient ($C_r$). This runoff coefficient depends on another coefficient ($\beta$), the amount of water stored in the reservoir, and the maximum storage capacity of the reservoir ($S_{u,Max}$). The runoff coefficient is determined by an S-shaped function, which is heavily influenced by $\beta$ as seen in figure 3.2.

$$C_r = \frac{1}{1 + \exp\left(\frac{-S_u/S_{u,max}+1/2}{\beta}\right)} \tag{3.4}$$

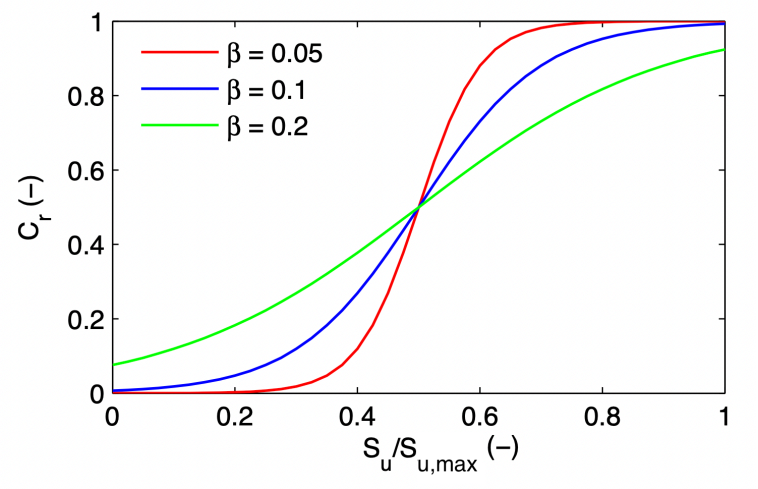


Figure 3.2: Example of $C_r$ as a function of $\frac{S_u}{S_{u,max}}$ for multiple values of $\beta$ (Fenicia et al., 2006).

Water stored in $S_u$ is reduced by transpiration ($E_t$). Transpiration is quantified by $E_p$ that remains after $E_i$ and by the amount of water available to $E_t$. The amount of water available to $E_t$ is determined by a function of the amount of stored water, $S_{u,max}$ and a certain coefficient ($C_e$). The coefficient, $C_e$, indicates the constraint of the amount of stored water on the remaining $E_p$.

$$E_t = (E_p - E_i)\min\left(1, \frac{S_u}{S_{u,max} * C_e}\right) \tag{3.5}$$

### Fast reacting reservoir

The other part of $P_e$ is considered as excess water leaving $S_u$ as runoff ($R_u$). Part of this $R_u$ enters $S_f$ according to a certain coefficient ($D$). The fast reacting reservoir subsequently generates the surface and fast runoffs, $Q_{ff}$ and $Q_f$ respectively.

$$R_f = C_r * D * P_e \tag{3.6}$$

$$Q_{ff} = \frac{\max(0, S_f - S_{ftr})}{K_{ff}} \tag{3.7}$$

$$Q_f = \frac{S_f}{K_f} \tag{3.8}$$

### Slow reacting reservoir

The remaining part of $R_u$ enters $S_s$ as preferential flow ($R_s$). The slow reacting reservoir subsequently generates the slow runoff ($Q_s$). The total simulated streamflow ($Q_m$) then consists of the sum of $Q_{ff}$, $Q_f$ and $Q_s$.

$$R_s = C_r * (1 - D) * P_e \tag{3.9}$$

$$Q_s = \frac{S_s}{K_s} \tag{3.10}$$

$$Q_m = Q_{ff} + Q_f + Q_s \tag{3.11}$$

### Parameters

The FLEX model contains 11 parameters that require model calibration for obtaining their values. The parameters are summed in table 3.1.

| Parameter | Unit | Description |
|---|---|---|
| $F_{DD}$ | L/(T*TEMP) | Degree day factor |
| $T_t$ | TEMP | Threshold temperature for snow |
| $I_{max}$ | L | Maximum storage capacity of $S_i$ |
| $S_{u,max}$ | L | Maximum storage capacity of $S_u$ |
| $C_e$ | - | Coefficient for constraint on $E_p$ |
| $\beta$ | - | Coefficient for shape of $C_r$ |
| $D$ | - | Coefficient for partitioning of runoff |
| $S_{ftr}$ | L | Maximum storage capacity of $S_f$ |
| $K_{ff}$ | T | Time scale for $Q_{ff}$ |
| $K_f$ | T | Time scale for $Q_f$ |
| $K_s$ | T | Time scale for $Q_s$ |

Table 3.1: Parameters of the Lumped FLEX model. L = Length, T = Time, and TEMP = Temperature.

### 3.1.2. GR4J Model

The GR4J model is, similarly to the FLEX model, considered as a spatially lumped conceptual hydrological model with relatively low process complexity. The GR4J model mostly distinguishes itself from other conceptual hydrological models due to its low number of parameters and the inclusion of inter-catchment groundwater flow. The GR4J model contains four parameters and two reservoirs. This study, however, used a slightly different structure of the GR4J model due to the significant number of snowy days and the fast rainfall-runoff responses in the selected catchments.

The GR4J model includes an interception reservoir ($S_i$) with zero storage capacity and therefore immediately takes the difference between precipitation ($P$) and potential evaporation ($E_p$) to determine the net precipitation ($P_n$) or evaporation ($E_n$). If precipitation is above $E_p$, the difference between the two is considered the $P_n$, which is similarly to the effective precipitation ($P_e$) as in the FLEX model. In this case all $E_p$ is used to vaporize precipitation and no $E_p$ is left for transpiration ($E_t$) from another reservoir. It $E_p$ is above $P$, the difference between the two is considered $E_n$ and some $E_p$ is left for $E_t$ from another reservoir.

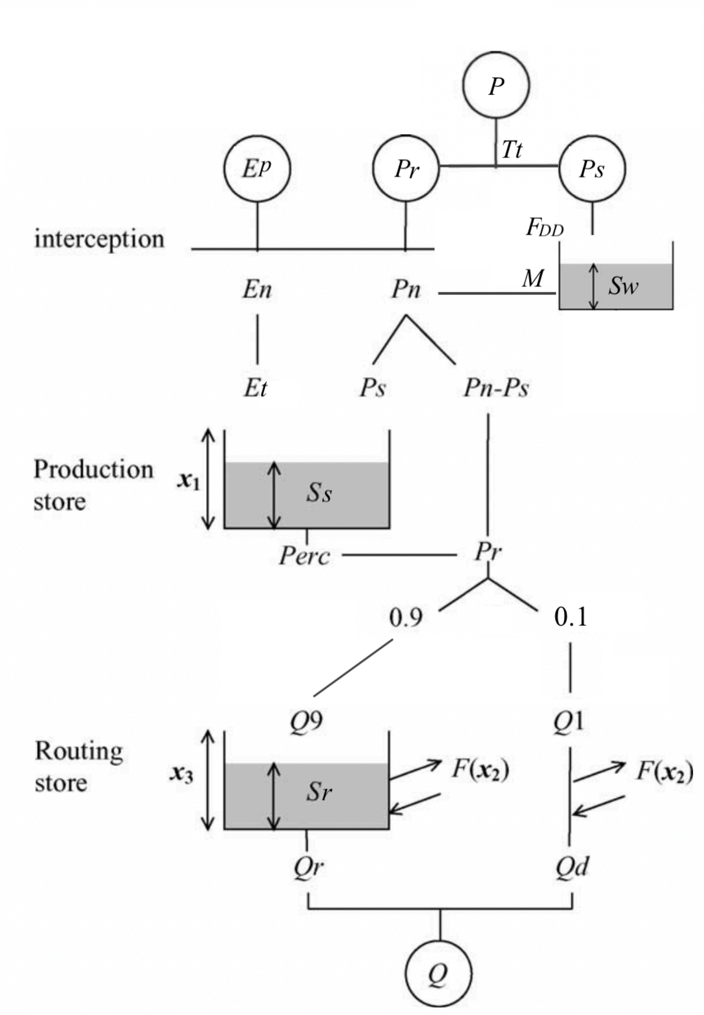


Figure 3.3: Structure of the lumped GR4J model (Perrin et al., 2003).

### Snow reservoir

The snow reservoir ($S_w$) is the same as defined in the description of the FLEX model. When ($T$) is below the threshold temperature ($T_t$), $P$ accumulates as snow in $S_w$. When $T$ is above $T_t$ and snow is present, the combination of $P$, if present, and snow melt ($M$), which is calculated by the degree-day model as in equation 3.1, is taken as $P_n$. When $T$ is above $T_t$ and no snow is present, $P_n$ or $E_n$ is determined by the difference between $P$ and $E_p$.

### Production reservoir

A part of $P_n$, if present, enters the production reservoir ($S_s$) according to equation 3.12. This equation is a function of the amount of water stored in the reservoir, the maximum storage capacity of the reservoir ($x_1$), and $P_n$.

$$P_s = \frac{x_1 * \left(1 - \left(\frac{S_s}{x_1}\right)^2\right) * \tanh\left(\frac{P_n}{x_1}\right)}{1 + \frac{S_s}{x_1} * \tanh\left(\frac{P_n}{x_1}\right)} \tag{3.12}$$

Water stored in $S_s$ is reduced by transpiration ($E_t$) and percolation ($Q_{perc}$). Transpiration is quantified by equation 3.13, which is a function of the amount of stored water, $x_1$ and $E_n$.

$$E_t = \frac{S_s * \left(2 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{E_n}{x_1}\right)}{1 + \left(1 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{E_n}{x_1}\right)} \tag{3.13}$$

$$Q_{perc} = S_s * \left\{1 - \left[1 + \left(\frac{4}{9}\frac{S_s}{x_1}\right)^4\right]^{-1/4}\right\} \tag{3.14}$$

## Routing reservoir

The other part of $P_n$, with the addition of $Q_{perc}$, is further divided according to a 10-90% split. The larger split of 90% ($Q_9$) first enters the routing reservoir ($S_r$), in which inter-catchment groundwater flow ($F$) is added or removed. Whereas the other 10% ($Q_1$), after the addition or removal of $F$, directly contributes to the total simulated streamflow. The inter-catchment groundwater flow is quantified by a function of the amount of stored water in $S_r$, the maximum storage capacity of $S_r$ ($x_3$), and the groundwater exchange coefficient ($x_2$). Depending on the value of $x_2$, water can either be added to or removed from $S_r$ and $Q_1$ as $F$.

$$F = x_2\left(\frac{S_r}{x_3}\right)^{7/2} \tag{3.15}$$

The outflow from $S_r$ is calculated by equation 3.16, which is a function of the amount of water stored in the reservoir and $x_3$. The total simulated streamflow ($Q$) then consists of the sum of the outflow from $S_r$ ($Q_r$) and $Q_1$ after the addition or removal of $F$ ($Q_d$).

$$Q_r = S_r * \left\{1 - \left[1 + \left(\frac{S_r}{x_3}\right)^4\right]^{-1/4}\right\} \tag{3.16}$$

$$Q = Q_r + Q_d \tag{3.17}$$

The GR4J model contains 3 parameters that require model calibration for obtaining their values. The parameters are summed in table 3.2.

| **Parameter** | **Unit** | **Description** |
|---|---|---|
| $F_{DD}$ | L/(T*TEMP) | Degree day factor |
| $T_t$ | TEMP | Threshold temperature for snow |
| $x_1$ | L | Maximum storage capacity of $S_s$ |
| $x_2$ | L | Coefficient for groundwater exchange |
| $x_3$ | L | Maximum storage capacity of $S_r$ |

Table 3.2: Parameters of the lumped GR4J model. L = Length, T = Time, and TEMP = Temperature.

## 3.2. Integration with Phenology

The conventional FLEX and GR4J models only partition evaporation into two separate processes. For the FLEX model, this is in evaporation from the interception reservoir and transpiration from the unsaturated soil reservoir. For the GR4J model, this is in evaporation from the interception reservoir, which has zero storage capacity, and transpiration from the production reservoir. To quantify evaporation fluxes, hydrological models generally use a method from two groups (Zhao et al., 2013). The first group consists of methods that estimate evaporation for the interception, soil, and transpiration separately according to the land cover distribution and integrates the estimated values to obtain evaporation at the catchment scale. The second group consists of methods that use a certain soil moisture extraction function to transform potential evaporation ($E_p$) into actual evaporation or transpiration ($E_t$) as in equation 3.18. The FLEX and GR4J models both use a method that fits within the second group, which does not include any information on the dynamic behavior of vegetation.

$$E_t = E_p\left(\frac{soil\,moisture}{field\,capacity\,soil\,moisture}\right) \tag{3.18}$$

The modification of the FLEX and GR4J models intended to include information from plant phenology, which indicates the dynamic behavior of vegetation, on the partitioning of evaporation at the catchment scale. This study realized this modification by three separate structures, which are the Jarvis-based, FAO-based, and the combined structure. In short, the modification consisted of an additional function that puts a dynamic constraint on the evaporation processes. Depending on the structure, the function requires different system inputs and indicates if the conditions in a catchment are favorable for vegetation to transpire.

### 3.2.1. Jarvis-Based Structure

According to Jarvis and Mcnaughton (1986), two groups of scientists exist with contrasting beliefs in the dominant factor controlling $E_t$. One group consists of physiologists who conclude that stomata are the main controller of $E_t$ and the other group consists of meteorologists who conclude that weather is the main controller of $E_t$. Jarvis and Mcnaughton (1986), however, explains that both groups are only correct at the spatial scales of their associated observations and suggests that the dominant factor controlling $E_t$ changes according to the following order:

- Stoma: Transpiration at the spatial scale of an individual stoma is determined by the amount of water vapor leaving through an individual stomatal pore. This is the result of water evaporating within the substomatal cavity and afterward vapor diffusing into the air surrounding the stomatal pore. At this spatial scale, it is assumed that the vapor pressure deficit near the surface of an entire leaf is not affected by $E_t$ through an individual stomatal pore. Hence, it is assumed that a change in the conductance of an individual stomatal can proportionally change $E_t$ through the same pore.
- Leaf to canopy: For larger spatial scales such as at an entire leaf, plant, and canopy, a change in the conductance of all stomatal pores can significantly affect the vapor pressure deficit at the surfaces of the associated scales. However, it is assumed that the vapor pressure deficits of the air surrounding the boundary layer of the surfaces of the

associated scales, such as at the leaf, plant, and canopy, are not affected by $E_t$ through their stomatal pores. Therefore, $E_t$ at such spatial scales is mainly controlled by stomata or weather according to the relationships between the vapor pressure deficits at the surfaces and the air surrounding the boundary layers of the associated scales. Generally, Jarvis and Mcnaughton (1986) assumes that the dominant factor controlling $E_t$ shifts from stomata to weather as the spatial scale of interest increases.

- Region: When the extent of a canopy is large enough, a change in canopy conductance can affect the vapor pressure deficit in the mixed layer of the planetary boundary layer. For a region of this scale, Jarvis and Mcnaughton (1986) considers weather, concerning the atmospheric water demand, as the dominant controller of $E_t$ and the influence of the canopy as minimal. For this spatial scale, several potential evaporation equations exist such as the Penman and the Priestley and Taylor equations.

The planetary boundary layer is considered as the lowest part within the atmosphere and is directly impacted by the surface of the Earth (De Bruin, 1989). This layer is characterized by turbulence and strong vertical mixing of physical properties such as water vapor and heat. The atmosphere above the planetary boundary layer is mostly characterized by laminar and stable stratified flows (Jarvis & Mcnaughton, 1986). The planetary boundary layer can be divided into three layers, which are the surface layer, the mixed layer and the interfacial layer (Deardorff, 1979). The surface layer is considered as the bottom layer in which the turbulent fluxes do not deviate much from their surface values (Wyngaard, 1985). The mixed layer of the planetary boundary layer is located above the surface layer and is considered to contain the convective-driven turbulent flows. These convective-driven turbulent flows cause sufficient mixing such that potential temperature and specific humidity become almost constant within the mixed layer (De Bruin, 1989).

A change in canopy conductance over an area of a few kilometers or more can impact the vapor pressure deficit throughout the entire mixed layer of the planetary boundary layer (Jarvis & Mcnaughton, 1986). At such spatial scales, the importance of stomatal or canopy conductance on the control of $E_t$ increases only again for certain temperatures when almost stomatal closure occurs (McNaughton & Jarvis, 1991). Additionally, $E_t$ often reveals a peaked response to increasing values of vapor pressure deficit, which usually co-occurs with increasing values of air temperature (Duursma et al., 2014). For increasing vapor pressure deficits, $E_t$ increases until an optimum value and decreases afterward as the result of an apparent feed-forward response. Duursma et al. (2014) describes this feed-forward response in which photosynthesis changes stomatal conductance and reveals a peaked response to increasing values of air temperature.

The Jarvis model is able to estimate stomatal conductance ($K_s$) as a function of photon flux density ($Q_p$), leaf temperature ($T$), vapor pressure deficit ($\delta e$), leaf water potential ($\Psi_l$) and ambient CO2 concentration ($C_a$). This study only considered part of the function, in which the response of stomata depends on temperature due to the availability of data and the important relationship between plant phenology and temperature. For this, it is assumed that leaf temperature is equal to air temperature. Stomatal conductance as a function of temperature ($K_s(T)$) is expressed by a parabola opening down and varies between the values of 0, which occurs at the least optimum temperatures, and 1, which occurs at the optimum temperature. This function is used to partition the mean daily potential evaporation according to the mean daily air temperature.

$$K_s(Q_p, T, \delta e, \Psi_l, C_a) = K_s(Q_p) * K_s(T) * K_s(\delta e) * K_s(\Psi_l) * K_s(C_a) \tag{3.19}$$

$$K_s(T) = b_3 * (T - T_l) * (T_h - T)^{b_4} \tag{3.20}$$

$$b_3 = \frac{1}{(T_o - T_l) * (T_h - T_o)^{b4}} \tag{3.21}$$

$$b_4 = \frac{T_h - T_o}{T_h - T_l} \tag{3.22}$$

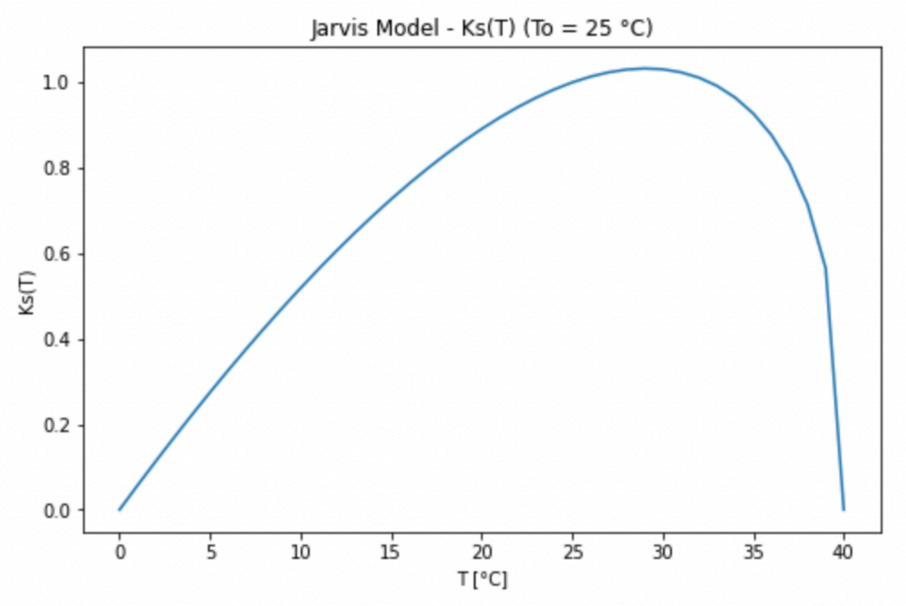


Figure 3.4: Example of Jarvis model $K_s(T)$ for $T_o$ = 25 °C.

Within the Jarvis-based modification of the models, $K_s(T)$ is calculated by using the average daily temperature as $T$ and assuming the least optimum temperatures at 0 °C and 40 °C, as $T_l$ and $T_h$ respectively. The optimal temperature, $T_o$, is regarded as a parameter, which value is obtained after model calibration. The function of $K_s(T)$ generates a value that changes with temperature and is used to put an additional dynamic constraint for the evaporation components in the FLEX model as following:

$$E_i = \min(S_i, (1 - K_s(T)) * E_p) \tag{3.23}$$

$$E_t = (K_s(T) * E_p)\min\left(1, \frac{S_u}{S_{u,max} * C_e}\right) \tag{3.24}$$

The Jarvis-based modification for the GR4J model is as following:

$$E_i = \min(P, (1 - K_s(T)) * E_p) \tag{3.25}$$

$$E_t = \frac{S_s * \left(2 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{K_s(T) * E_p}{x_1}\right)}{1 + \left(1 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{K_s(T) * E_p}{x_1}\right)} \tag{3.26}$$

### 3.2.2. FAO-Based Structure

The life cycle of vegetation and plants is among others dependent on the climate, environment, and plant species. In general, plants can be classified as annual, biennial, or perennial plants. As the term already suggests, annual plants have a life cycle that ends within one year and thus needs to be replanted after a single growth season. Biennial plants have a life cycle covering two years with a dormant season in between the two growth seasons. Lastly, perennial plants have a life cycle that covers more than two years and a dormant season in between the growth seasons. Transpiration varies along with the different development stages within the life cycle of vegetation.

The FAO crop evaporation can be used to estimate $E_t$ through crops (Allen et al., 1998). The FAO crop evaporation consists of three different levels, of which only the first two are relevant for the FAO-based modification of the models. The first level can be considered as the reference crop evaporation ($E_{t,0}$), which is defined as the atmospheric water demand for a reference surface and therefore similar to $E_p$. The second level can be considered as the crop evaporation under standard conditions ($E_{t,c}$), which is defined by the multiplication of $E_{t,0}$ and a crop-dependent coefficient ($K_c$). The crop evaporation under standard conditions, therefore, reflects the atmospheric water demand for different crop surfaces and over different development stages of the crop. The crop-dependent coefficient integrates the differences between the reference surface and the relevant crop surface and can be separated over multiple development stages of the crop.

According to Allen et al. (1998), the growth season of a crop can be divided into four development stages. These development stages are the initial stage, the crop development stage, the mid-season stage, and the late-season stage. The initial stage is defined as the period between the moment of the crop planting until the moment the crops have developed to cover around 10% of the soil. In case the life cycle of the crops takes longer than one year, with for example perennial crops, the initial stage starts at the moment of the development of new leaves. During the initial stage, the expansion of green vegetation is small and the exposure of bare soil high and therefore results in dominant soil evaporation rates. Hence during the initial stage, $K_c$ is small when the soil is dry and high when the soil is wet. The second stage, the crop development stage, is the period between the moment the crops have developed to cover around 10% of the soil until the moment the crops have developed to effectively cover the soil. The full effective coverage by the crops can be considered to take place at the start of flowering. During this stage, $K_c$ gradually increases with the development of the crops and eventually results in dominant $E_t$ rates. The third stage, the mid-season stage, takes place from the moment of full effective coverage until the start of maturity. The start of maturity indicates the start of senescence and abscission, falling, of the crop leaves. During the third stage, $K_c$ is constant and at its peak. Finally, the late-season stage is the period between the start of maturity until the end of senescence. In this stage, $K_c$ and the associated $E_t$ rates decrease. Allen et al. (1998) recommends the following four steps for $E_{t,c}$:

1. The first step consists of identifying the development stages of the crops, defining the length of each stage, and the corresponding $K_c$ values.
2. The second step consists of modifying the value of $K_c$ in the initial stage by using an indication of the soil wetness to include the impact of soil evaporation.
3. The third step is constructing the $K_c$ curve as in figure 3.5.

4. Finally, the fourth step is calculating $E_{t,c}$ by the multiplication of $E_{t,0}$ with $K_c$.

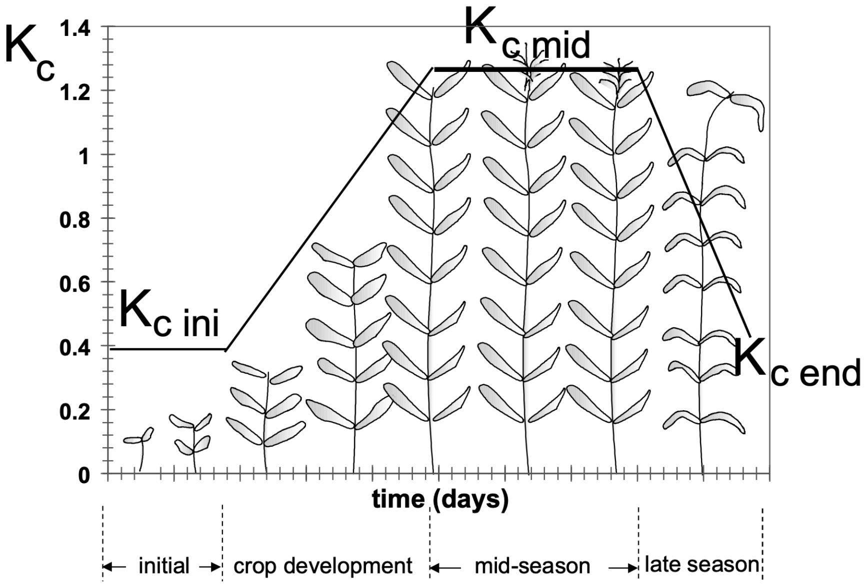


Figure 3.5: Example of FAO $K_c$ curve (Allen et al., 1998).

Within the FAO-based modification of the models, a concept similar to that of the FAO crop evaporation is used to partition the mean daily $E_p$ according to the development stage of the relevant vegetation. For this $E_{t,0}$ is replaced by the mean daily $E_p$ and $K_c$ is separated according to the appropriate development stages of the relevant vegetation. The crop-dependent coefficient for each development stage consists then of a dimensionless value between 0 and 1 and is obtained after model calibration. The recommended steps to calculate the FAO crop evaporation are used in a similar way to partition $E_p$ in the FAO-based modification of the models. For this, the leaf area index (LAI) is used due to remote sensing becoming an interesting tool at the catchment scale to use for observing the seasonal changes in vegetation covers (Xue & Su, 2017; Towers, Strever, & Poblete-Echeverria, 2019). This is performed in the following way:

1. The lengths of the development stages of the relevant vegetation are identified by investigating LAI over time. The dormant season is marked by a significant period of low LAI values and the peak season is marked by a significant period of high LAI values. The periods in which a sudden change takes place from low to high LAI are marked as the periods in which vegetation leaves its dormant season and starts to grow leaves. The period in which a sudden change takes place from high to low LAI is marked as periods in which vegetation enters its dormant season and starts to drop its leaves.
2. The values of $K_c$ for the dormant and peak seasons of the relevant vegetation are obtained through model calibration. The values of $(K_c)$ in between the two major seasons are determined by linear interpolation. Compared to the $K_c$ curve in figure 3.5, $K_c$ for the dormant seasons would cover the initial development stage and $K_c$ for the peak season would cover the mid-season development stage. The interpolated values of $K_c$ in between the two major seasons would cover the crop development and late-season development stages.

3. The timings of the dormant and peak seasons of the relevant vegetation and the calibrated values of $K_c$ for these seasons are used as additional dynamical constraints for the evaporation components in the FLEX Model as following:

$$E_i = \min(S_i, (1 - K_c(time)) * E_p) \tag{3.27}$$

$$E_t = (K_c(time) * E_p)\min\left(1, \frac{S_u}{S_{u,max} * C_e}\right) \tag{3.28}$$

The FAO-based modification for the GR4J model is as following:

$$E_i = \min(P, (1 - K_c(time)) * E_p) \tag{3.29}$$

$$E_t = \frac{S_s * \left(2 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{K_c(time) * E_p}{x_1}\right)}{1 + \left(1 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{K_c(time) * E_p}{x_1}\right)} \tag{3.30}$$

### 3.2.3. Combined Structure

The third structure of the modification of the models consisted of a combination of the other two structures. This combined structure considered both the influence of air temperature and the development stage of vegetation on the ability of vegetation to transpire. Similar to the other two structures, an additional coefficient is used to partition $E_p$ over time. This coefficient is determined by the average of the values from the coefficients of the other two structures as following for the FLEX model:

$$E_i = \min\left(S_i, \left(1 - \frac{K_s(T) + K_c(time)}{2}\right) * E_p\right) \tag{3.31}$$

$$E_t = \left(\frac{K_s(T) + K_c(time)}{2} * E_p\right)\min\left(1, \frac{S_u}{S_{u,max} * C_e}\right) \tag{3.32}$$

The combined modification for the GR4J model is as following:

$$E_i = \min\left(P, \left(1 - \frac{K_s(T) + K_c(time)}{2}\right) * E_p\right) \tag{3.33}$$

$$E_t = \frac{S_s * \left(2 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{\frac{K_s(T) + K_c(time)}{2} * E_p}{x_1}\right)}{1 + \left(1 - \frac{S_s}{x_1}\right) * \tanh\left(\frac{\frac{K_s(T) + K_c(time)}{2} * E_p}{x_1}\right)} \tag{3.34}$$

## 3.3. Model Calibration

Conceptual hydrological models $(M)$ basically relate the system input $(X)$, which are the meteorological forcing of $P$ and $E_p$, to the system state $(S)$, which are the water storages in the associated reservoirs, and the system output ($Y$), which are the hydrological responses of streamflow and evaporation, as in equation 3.35.

$$M(X, \theta, S_0) \rightarrow Y \tag{3.35}$$

The operation of conceptual hydrological models requires quantitative data for the system input ($X$), the model parameters ($\theta$), and the initial state values ($S_0$). Such models are with their operations able to relate the system input to the system states and outputs, which can be observed. For the operation of these models, the model parameters need to be quantified. The quantification of the models' parameters is highly important as this can easily result in unreliable simulations of hydrological models. Therefore, it is crucial to define the model parameters with identifiable values. Using direct measurements to define the model parameters is however extremely complex and sometimes even impossible with the current technology due to the issues of scale and uniqueness of place (Beven, 2000). Therefore, automatic calibration is a commonly used method to define the model parameters and hence is also used in this research.

Automatic calibration is in this study performed by an optimization method, in which objective functions are minimized or maximized according to some generated sets of parameter values. The Nash-Sutcliffe efficiency (eq. 3.36) is used as the objective function and Monte Carlo sampling is used for optimizing this function. Each model is calibrated separately according to the daily streamflow observed at the outlet of the two catchments and to the daily transpiration estimated by GLDAS-21 for each of the two catchments. The Monte Carlo method used 5000 iterations for the calibration of each model and used random samples from a uniform distribution with the minimum and maximum parameters values as defined by tables 3.3 and 3.4. All simulations with a Nash-Sutcliffe efficiency above 0.7 are considered as behavioral. The calibration period for the catchment in Hubbard Brook is between January 1, 2004, and December 31, 2006. The calibration period for the catchment in HJ Andrews is between January 1, 2006, and December 31, 2007. The initial state values are obtained for both catchments by training these in a one-year spin-up period before the calibration period. The potential evaporation is calculated by the Priestley-Taylor function. The parameter $\alpha$ of this function is quantified by the calibration of each model.

$$N_{NS} = 1 - \frac{\sum_{i=1}^{n}(Q_{s,i} - Q_{o,i})^2}{\sum_{i=1}^{n}(Q_{o,i} - \overline{Q}_{o,i})^2} \tag{3.36}$$

| FLEX Model | | | |
|---|---|---|---|
| **Parameter** | **Unit** | **Range WS3 Hubbard Brook** | **Range WS2 HJ Andrews** |
| $F_{DD}$ | mm/(d °C) | (1, 5) | (1, 5) |
| $T_t$ | °C | (-1, 0) | (-1, 0) |
| $I_{max}$ | mm | (1, 5) | (1, 5) |
| $S_{u,max}$ | mm | (10, 500) | (100, 1000) |
| $C_e$ | - | (0.5, 1) | (0.5, 1) |
| $\beta$ | - | (0.01, 1) | (0.01, 1) |
| $D$ | - | (0.5, 1) | (0.5, 1) |
| $S_{ftr}$ | mm | (10, 30) | (10, 30) |
| $K_{ff}$ | d | (1, 2) | (1, 2) |
| $K_f$ | d | (2, 10) | (2, 10) |
| $K_s$ | d | (10, 100) | (10, 100) |
| $\alpha$ | - | (1.26, 1.75) | (1.26, 1.75) |
| $T_o$ | °C | (10, 30) | (10, 30) |
| $K_{c,dormant}$ | - | (0, 0.1) | (0, 0.5) |
| $K_{c,peak}$ | - | (0.5, 1) | (0.5, 1) |

Table 3.3: Calibration ranges for the parameters of the lumped FLEX model.

| GR4J Model | | | |
|---|---|---|---|
| **Parameter** | **Unit** | **Range WS3 Hubbard Brook** | **Range WS2 HJ Andrews** |
| $F_{DD}$ | mm/(d °C) | (1, 5) | (1, 5) |
| $T_t$ | °C | (-1, 0) | (-1, 0) |
| $x_1$ | mm | (10, 500) | (100, 1000) |
| $x_2$ | mm | (-10, 10) | (-10, 10) |
| $x_3$ | mm | (10, 100) | (10, 100) |
| $\alpha$ | - | (1.26, 1.75) | (1.26, 1.75) |
| $T_o$ | °C | (10, 30) | (10, 30) |
| $K_{c,dormant}$ | - | (0, 0.1) | (0, 0.1) |
| $K_{c,peak}$ | - | (0.5, 1) | (0.5, 1) |

Table 3.4: Calibration ranges for the parameters of the lumped GR4J model.

## 3.4. Model Validation and Comparison

For the simulation of hydrological responses, this study used two conceptual hydrological models, the FLEX and GR4J model, in four different structures, which are the conventional, Jarvis-based, FAO-based, and combined structure. This study, therefore, included eight different hydrological models and calibrated each model to the observed values of streamflow and estimated values of transpiration separately for each of the two selected catchments. In other words, this study operated 32 models in total. Each model is validated by comparing the simulated streamflow values over the non calibrated period to the observed values over the same period. The validation period for the catchment in Hubbard Brook is between January 1, 2007, and December 31, 2010. The validation period for the catchment in HJ Andrews is between January 1, 2008, and December 31, 2010. Due to absence of other observational data for both catchments and to investigate additional differences between the four model structures, the simulated evaporation and maximum storage capacity of the unsaturated soil reservoir ($S_{u,max}$)/production reservoir ($x_3$), which are indicative of the amount of water available to transpiration, are compared to their values estimated by other methods. These methods are the integration of land surface models with remote sensing data as the evaporation estimations and the root zone storage capacity (Gao, Hrachowitz, Schymanski, et al., 2014; Nijzink et al., 2016) as the $S_{u,max}/x_3$ estimations. The GLDAS-2.0, GLDAS-2.1 and PML-2 datasets are used to obtain the evaporation estimations and the methods described in Gao, Hrachowitz, Schymanski, et al.(2014) and Nijzink et al.(2016) are used to determine the root zone storage capacities. These estimation methods are selected due to the available data. Only behavioral simulations are used for the validation and comparisons of the model results.

### 3.4.1. Streamflow

A selection of objective functions is used to evaluate the performance of the daily simulated values of streamflow according to their daily observed values. These objective functions are correspondingly the Nash-Sutcliffe efficiency (eq. 3.36), the log Nash-Sutcliffe efficiency (eq. 3.37), the coefficient of determination (eq. 3.38) and the root mean square error (RMSE) (eq. 3.39). Each of these functions evaluates different features of the simulated values. The Nash-Sutcliffe efficiency and the RMSE evaluate the high flows, the log Nash-Sutcliffe efficiency evaluates the low flows and the coefficient of determination evaluates the timing of the flows.

$$N_{logNS} = 1 - \frac{\sum_{i=1}^{n}(log(Q_{s,i}) - log(Q_{o,i}))^2}{\sum_{i=1}^{n}(log(Q_{o,i}) - \overline{log(Q_{o,i})})^2} \tag{3.37}$$

$$N_{R^2} = \frac{\sum_{i=1}^{n}(Q_{s,i} - \overline{Q}_{o,i})^2}{\sum_{i=1}^{n}(Q_{o,i} - \overline{Q}_{o,i})^2} \tag{3.38}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_{s,i} - Q_{o,i})^2} \tag{3.39}$$

For the objective functions of the Nash-Sutcliffe efficiency, the log Nash-Sutcliffe efficiency and the coefficient of determination values that are closer to one indicate better agreement between simulations and observations. The root mean square error values closer to zero indicate better agreement.

### 3.4.2. Evaporation

The objective function RMSE that is used to measure the performance of the streamflow simulations is also used to compare the transpiration simulations to the transpiration estimations.

The performances of the simulated values of transpiration are compared to the estimated values by GLDAS-2.0, GLDAS-2.1, and PML-2. The daily estimated values of interception and soil evaporation by GLDAS-2.0 and GLDAS-2.1 are used for the comparison of the daily simulated values of evaporation from the interception reservoir. Similarly, daily estimated values of transpiration by again GLDAS-2.0 and GLDAS-2.1 are used for the comparison of the daily simulated values of transpiration. The estimated values by PML-2 have an 8-day temporal resolution. The daily simulated values of evaporation are therefore upscaled to 8-day temporal resolution for their performance compared to the estimated values by PML-2.

### 3.4.3. Root Zone Storage Capacity

The root zone storage capacity ($S_R$) is used as an additional measure for the comparison of the model results. The root zone storage capacity represents the capacity of a catchment to provide vegetation constant access to water. The root zone storage capacity is quantified by the methods described in Gao, Hrachowitz, Schymanski, et al.(2014) and Nijzink et al.(2016). For this, a long term water balance is used to determine the effective precipitation ($P_e$), which is used for the estimation of the average annual transpiration $\overline{E}_t$. This long term water balance takes into account and interception reservoir $S_i$, precipitation $P$, interception evaporation $E_i$, potential evaporation $E_p$ and the maximum storage capacity of the interception reservoir $I_{max}$. The following steps are used for the calculation of $P_e$:

$$S_i = P - E_i - P_e \tag{3.40}$$

$$E_i = \min(S_i, E_p) \tag{3.41}$$

$$P_e = \max(0, S_i - I_{max}) \tag{3.42}$$

Subsequently the mean annual transpiration is calculated by taking into account the mean annual $P_e$ and streamflow ($Q$) as following:

$$\overline{E}_T = \overline{P}_e - \overline{Q} \tag{3.43}$$

To consider the effects of seasonality, the $\overline{E}_t$ is scaled by the ratio of the mean daily $E_p$ over mean annual $E_p$.

$$E_t(t) = \frac{E_p(t)}{\overline{E}_p} * \overline{E}_T \tag{3.44}$$

The root zone storage capacity is estimated by the cumulative sum of the difference between the daily $E_t$ and $P$ for the period between the timings when this cumulative sum is equal to zero as seen in the example in figure 3.6.

$$S_R = \max \int_{T_0}^{T_1} (E_t - P_e)dt \tag{3.45}$$

It is considered that ecosystems at the catchment scale develop their root zone to overcome droughts with return periods of around 10-20 years (Gao, Hrachowitz, Schymanski, et al., 2014; Nijzink et al., 2016). Hence, the root zone storage capacity with a return period of 20 years ($S_{R,20yr}$) is compared with the parameters for the maximum storage capacities of the unsaturated soil reservoir in the FLEX model and the production reservoir in the GR4J model. Gumbel extreme value distribution and the annual $S_R$ is used to estimate $S_{R,20yr}$ of each catchment.

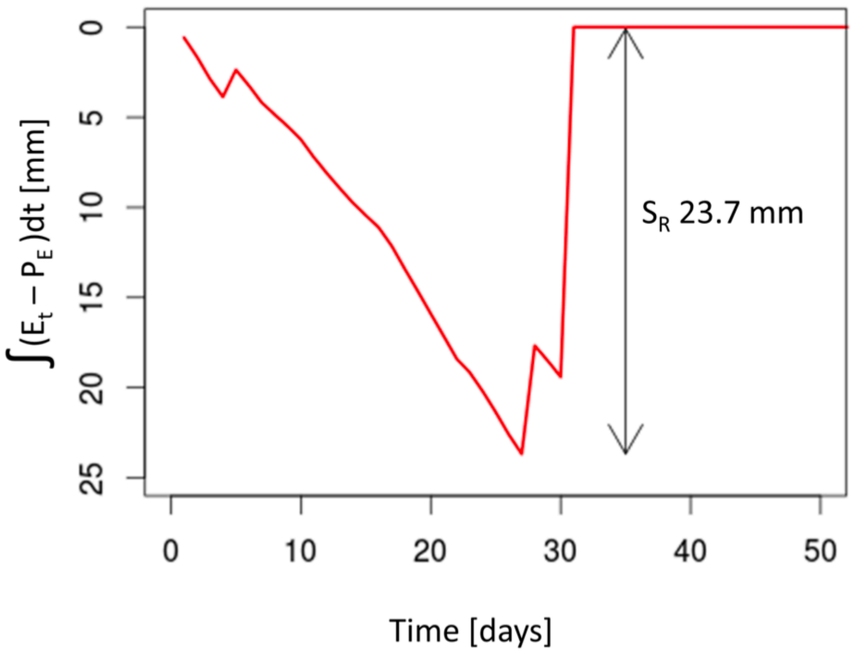


Figure 3.6: Example of derivation $S_R$ (Nijzink et al., 2016).

# 4

# Results

## 4.1. Hubbard Brook

Daily values of streamflow and evaporation are simulated for Hubbard Brook's watershed 3 (WS3) in the period between January 1, 2003, and December 31, 2010. The first year of this period, between January 1, 2003, and December 31, 2003, served as the spin-up period. The second period, between January 1, 2004, and December 31, 2006, served as the calibration period and the last period, between January 1, 2007, and December 31, 2010, served as the validation period.

### 4.1.1. Leaf Area Index

The MODIS-derived Leaf Area Index (LAI) for WS3 is visible in figure 4.1 for the calibration and validation periods. The Leaf Area Index for WS3 reveals a strong seasonal pattern, in which values are high during the summers and low during the winters. This pattern is typical for deciduous broadleaf forests, which is also WS3's dominant vegetation type. The pattern shows that when LAI exceeds the value of 1, it drastically increases to a value around 5. During the summer LAI stays almost the same, apart from the noise that can be observed, until around mid-fall. Around mid-fall, LAI drastically decreases to a value around 1 after which it does not change much again until around mid-spring. Around mid-spring, LAI continues the same cycle as described. Based on this pattern, the dormant season of the vegetation in WS3 is considered as the period between two LAI values of 1 and the peak season as the period between two LAI values of 5. The periods between the LAI values of 1 and 5 and between 5 and 1 are considered as the transitional seasons, in which vegetation respectively grows and drops leaves.

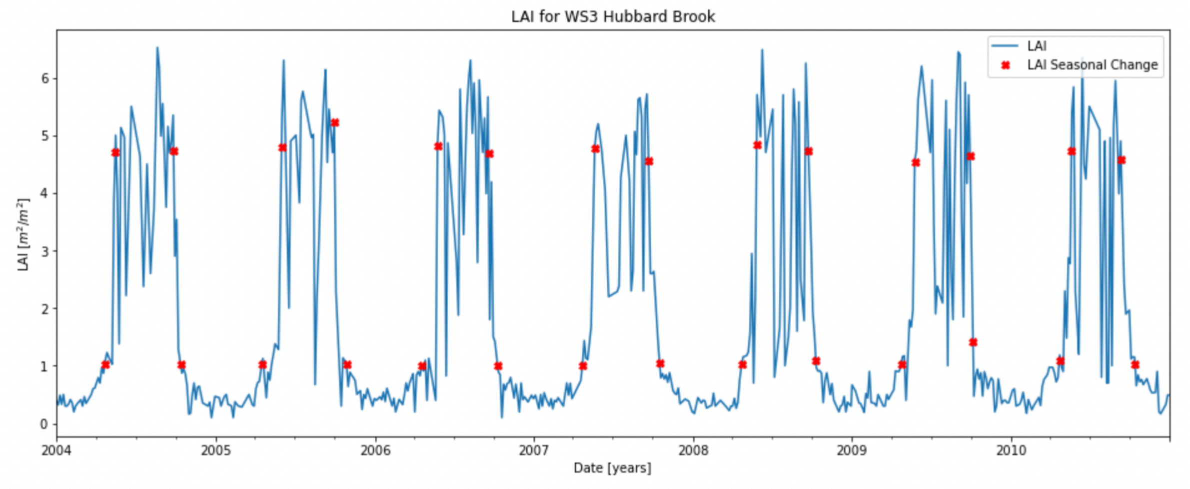


Figure 4.1: Leaf Area Index for WS3 in Hubbard Brook. The red dots indicate the transition of the vegetation's development stages. The periods between the red dots around LAI equal to 1 indicate the dormant seasons. The periods between the red dots around LAI equal to 5 indicate the peak seasons.

### 4.1.2. FLEX Model

The FLEX model is used with each of the four model structures, which are the conventional structure, Jarvis-based structure, FAO-based structure, and the combined structure, to simulate the daily values of streamflow and evaporation for WS3. The four FLEX model structures are separately calibrated by the uniform Monte Carlo method using 5000 iterations. The performances of the streamflow simulations are measured by the Nash-Sutcliffe efficiency (NS) using the streamflow observations as the reference data. The simulations that have a NS above 0.7 for the calibration period are considered behavioral. The number of behavioral simulations is 102 for the conventional model structure, 187 for the Jarvis-based model structure, 194 for the FAO-based model structure, and 194 for the combined model structure. The results in this section only consider the behavioral simulations that are calibrated to the observed streamflow. Each of the four FLEX model structures is additionally calibrated to the daily GLDAS-21 transpiration estimations in a similar way. The results of the transpiration calibrated models are similar to the streamflow calibrated models and are visible in appendix A.

#### Streamflow

The FLEX model streamflow simulations for WS3 in the calibration and validation periods are visible in figure 4.2. Figure 4.2 consists of three graphs, in which the upper graph contains the daily streamflow observations, the daily precipitation observations, and the median of the daily streamflow simulations by the conventional FLEX model. The lower graphs contain the daily streamflow observations and the medians of the daily streamflow simulations by all four structures of the FLEX model for the first year of the calibration period and the first year of the validation period.

Visual inspection of the graphs in figure 4.2 indicates that the daily streamflow simulations by each of the four structures of the FLEX model perform equally well according to the observations and barely deviate from each other. The streamflow simulations appear highly similar to the streamflow observations during both the calibration and the validation periods. This indicates that the calibrated parameter values of each of the four FLEX model structures can provide relatively accurate streamflow simulations outside the calibration period. Overall, the streamflow simulations only seem to lack accuracy during the extreme streamflow peaks.

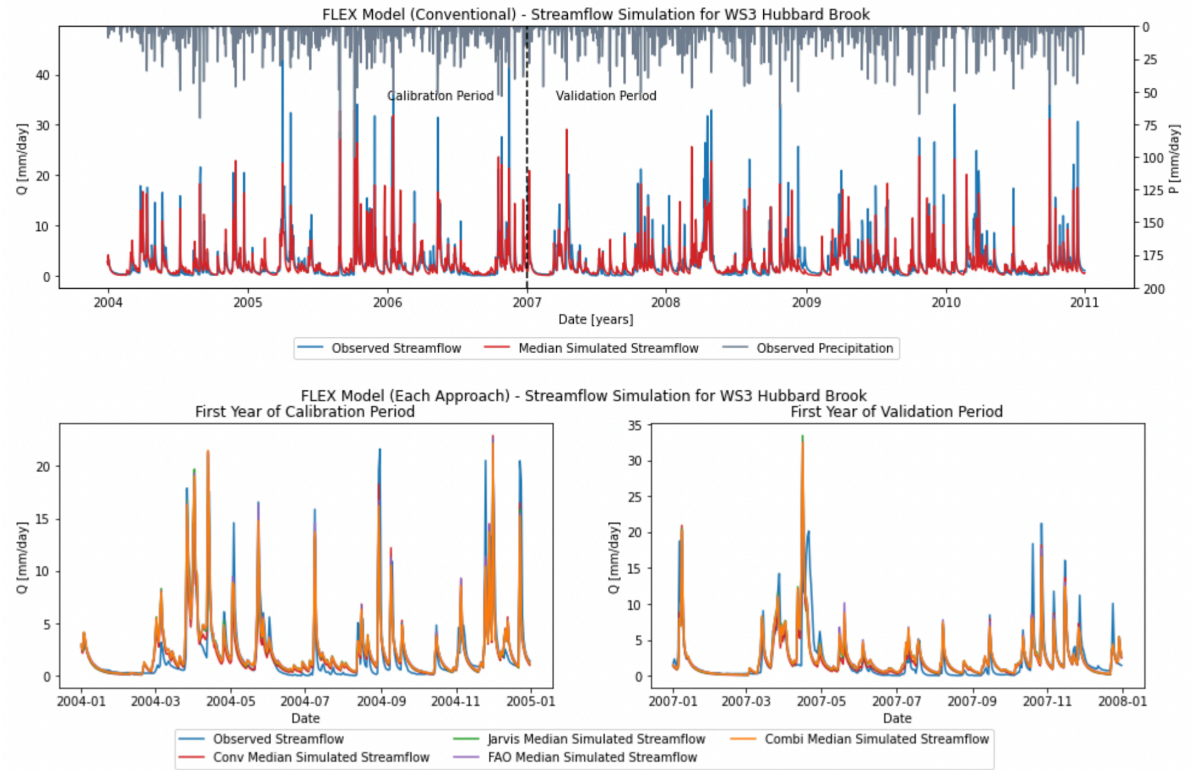


Figure 4.2: FLEX Model - Streamflow Simulations for WS3 in Hubbard Brook. The upper graph consists of the daily observations of streamflow and precipitation together with the median of the daily simulations of streamflow by the conventional FLEX model. Lower graphs consist of the daily observations of streamflow together with the medians of the daily simulations of streamflow by each of the four structures of the FLEX model. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

Figure 4.3 contains boxplots indicating the performances of the daily streamflow simulations by each of the four FLEX model structures, quantified by the four objective functions separately for the calibration and validation periods. The results of the four objective functions, which are the NS efficiency, the log NS efficiency ($logNS$), the coefficient of determination ($R^2$), and the root mean square error ($RMSE$), do not reveal many differences between the performances of the streamflow simulations of each of the four FLEX model structures. The daily streamflow simulations of each FLEX model structure thus perform equally well compared to each other. However, the performances of the streamflow simulations during the calibration period are slightly better than during the validation period, which indicates that the calibrated parameter values are not necessarily as suitable as during the calibration period as for any other period. Additionally, the results of $R^2$ show that the medians of the performances for the streamflow simulations by the conventional model are slightly lower than those by the three modified models, which are the Jarvis-based, FAO-based, and combined models. This indicates that there is a slightly better agreement between the timing of the streamflow observations and the streamflow simulations by the three modified FLEX models.

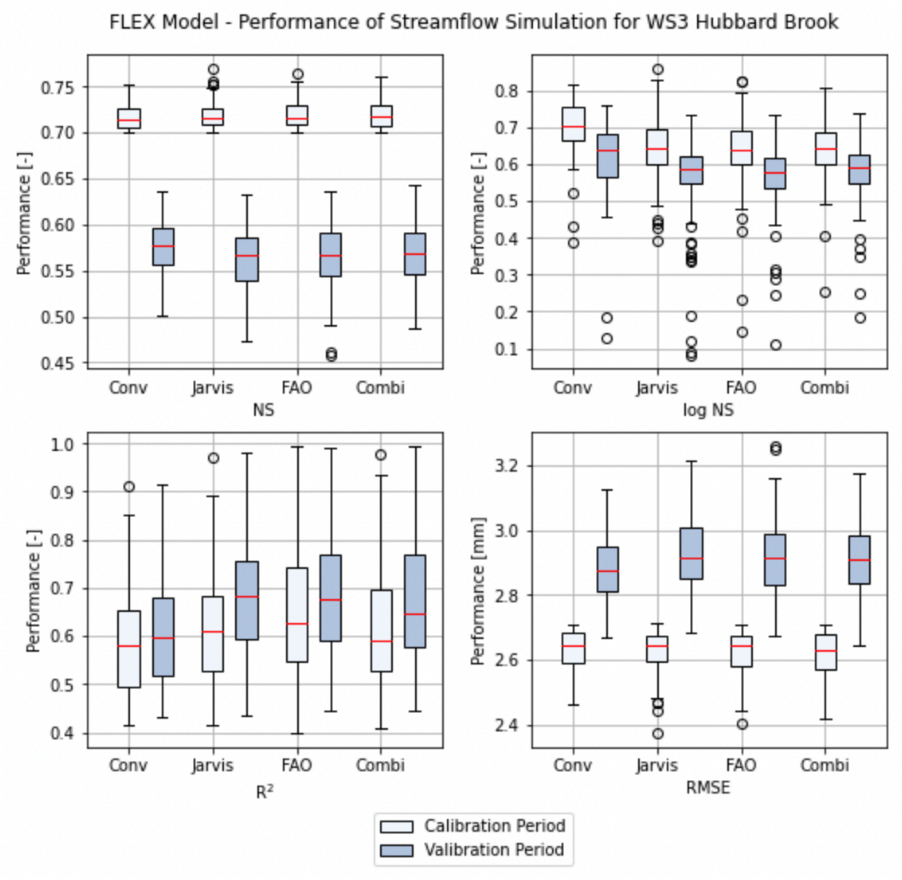


Figure 4.3: FLEX Model - Performance of Streamflow Simulations for WS3 in Hubbard Brook. Performances of all behavioral streamflow simulations by each of the structures of the FLEX model. The performances are quantified by the objective functions of the Nash-Sutcliffe efficiency (NS), the log Nash-Sutcliffe efficiency (log NS), the coefficient of determination ($R^2$), and the root mean square error (RMSE). Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

The mean seasonal sums of the daily streamflow observations and the daily streamflow simulations by the four FLEX model structures can be seen in figure 4.4 for the calibration and validation periods. The seasons are separated into four groups of three months, which correspondingly are March, April, and May for spring (mam), June, July, and August for summer (jja), September, October, and November for fall (son), and December, January, and February for winter (djf). For the seasons son and djf there is not much difference between the mean seasonal sums of the streamflow simulations. For the season mam, spring, the mean seasonal sums of the streamflow simulations by the modified model structures are slightly more similar to the sums of the observations compared to the sums of the conventional model structure. However, for the season jja, summer, the mean seasonal sums of the streamflow simulations by the conventional model structure are slightly more similar to the sums of the observations compared to the sums of the modified model structures. This indicates that during the spring the streamflow simulations by the modified FLEX models perform slightly better compared to the streamflow simulations by the conventional FLEX model and the other way around during the summer.

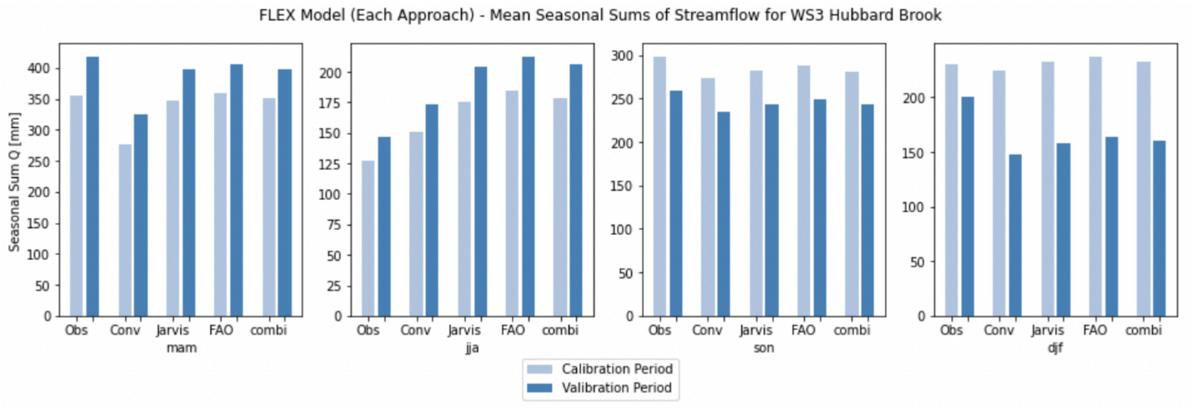


Figure 4.4: FLEX Model - Mean Seasonal Sums of Streamflow for WS3 in Hubbard Brook. Mean sums of the streamflow observations and simulations by each of the four structures of the FLEX model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. Obs = observations, Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

By plotting the flow duration curve, which describes the percent of time specified streamflow values are reached or exceeded, for the streamflow observations and simulations by each FLEX model structure (figure 4.5), again not much difference can be seen. The very high streamflow values, such as the peaks in the plots in figure 4.2, are less often reached or exceeded by the simulations compared to the observations and the highest observed streamflow values, above approximately 35 mm/day, are not reached at all by the simulations by each of the four FLEX model structures.

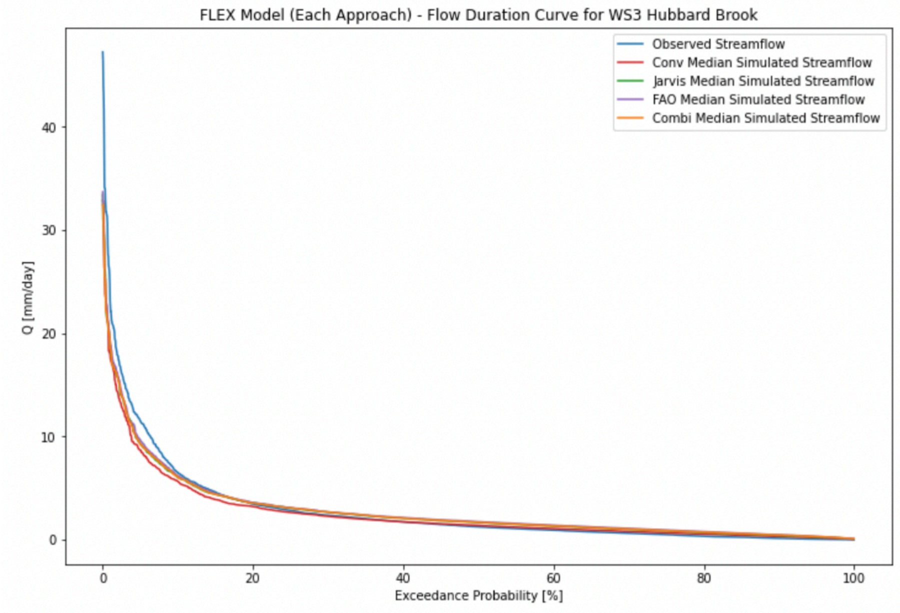


Figure 4.5: FLEX Model - Flow Duration Curve for WS3 in Hubbard Brook. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

## Evaporation

The FLEX model transpiration simulations for WS3 in the calibration and validation periods can be seen in figure 4.6. Figure 4.6 consists of three graphs, in which the upper graph contains the transpiration estimations, generated by the remote sensing data integrated land surface models GLDAS-20, GLDAS-21, and PML-V2, together with the median of the transpiration simulations by the conventional FLEX model. The lower graphs contain the same transpiration estimations together with the medians of the transpiration simulations of each four of the FLEX model structures for the first year of the calibration period and the first year of the validation period. The daily transpiration estimations of GLDAS-20 and GLDAS-21 and the daily transpiration simulations by each FLEX model structure are up-scaled in the graphs to an 8-day temporal resolution to be comparable with the 8-day temporal resolution of the PML-V2 estimations.

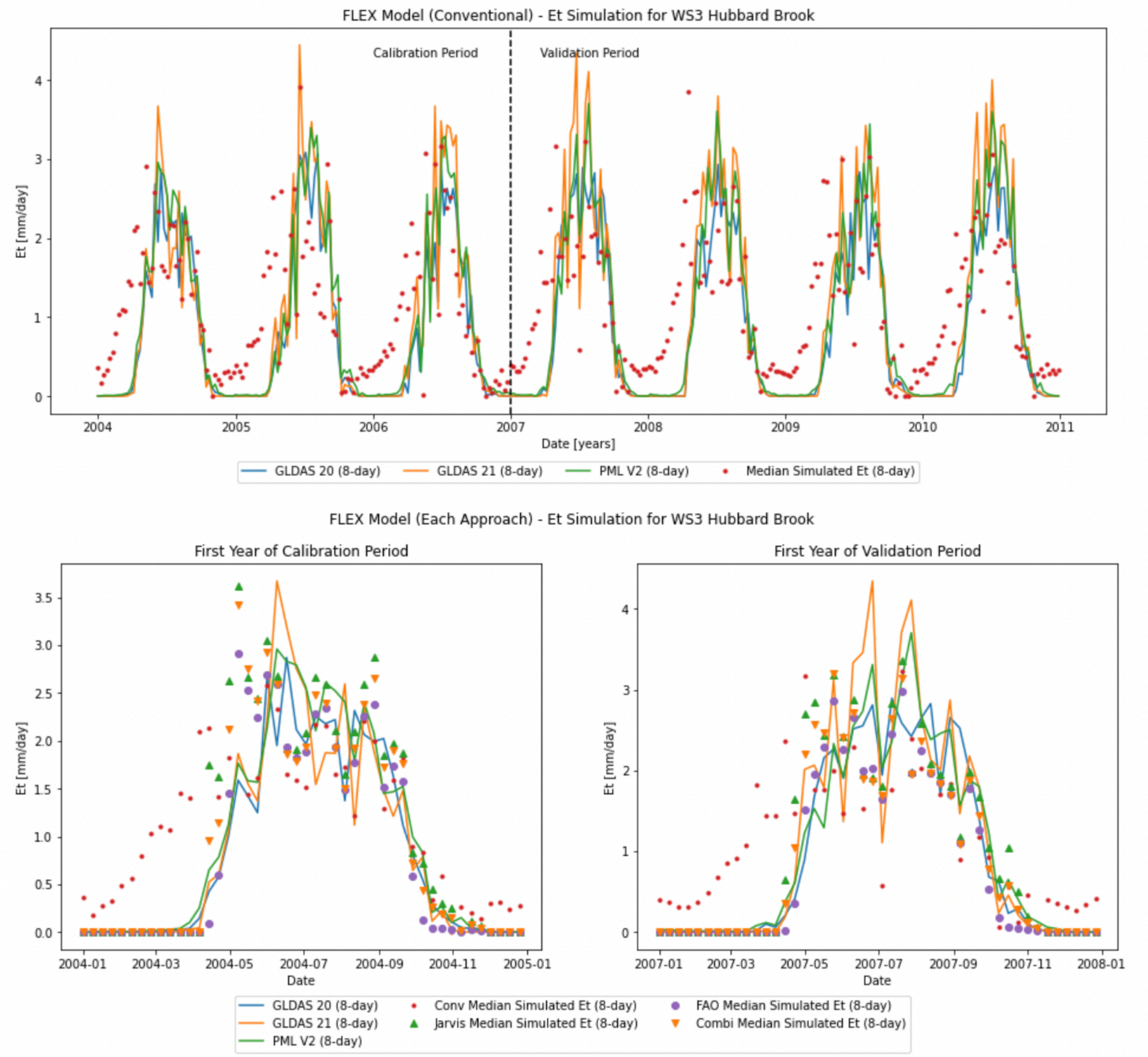


Figure 4.6: FLEX Model - Transpiration Simulations for WS3 in Hubbard Brook. The upper graph consists of the estimations of transpiration together with the median of the simulations of transpiration by the conventional FLEX model. Lower graphs consist of the estimations of transpiration together with the medians of the simulations of transpiration by each of the four structures of the FLEX model. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

The transpiration estimations of GLDAS-20, GLDAS-21, and PML-V2 reveal a pattern that agrees with the LAI of WS3. The transpiration estimations are zero, or close to zero, during the dormant seasons of the vegetation in WS3 and are highest during the peak seasons of the vegetation, similarly to the LAI of WS3. The transpiration simulations by each of the four FLEX model structures seem to follow the same pattern as the transpiration estimations. However, the transpiration simulations by the conventional model structure are slightly higher during the dormant seasons and lower during the peak seasons compared to the transpiration estimations. The transpiration simulations by the three modified model structures are more similar to the transpiration estimations compared to those by the conventional model. This is particularly the case during the dormant seasons. The transpiration simulations by the modified model structures do not deviate much from each other. However, it can be perceived that the transpiration simulations by the Jarvis-based model structure are slightly higher than the transpiration simulations by the other modified model structures and that the transpiration simulations by the FAO-based model structure approach the transpiration estimations the most.

Figure 4.7 includes boxplots for the results of the root mean square error ($RMSE$) that measured the performances of the daily transpiration simulations by each of the four FLEX model structures according to the daily transpiration estimations of GLDAS-20 and GLDAS-21 and the performances of the 8-day averaged transpiration simulations of the four model structures to the transpiration estimations of PML-V2. The results of $RMSE$ reveal similar performances to that of the visual inspection of the graphs in figure 4.6. According to figure 4.7, the transpiration simulations by the FAO-based model structure perform best according to the transpiration estimations of GLDAS-20, GLDAS-21, and PML-V2. The transpiration simulations by the conventional model structure perform the worst compared to the other structures.

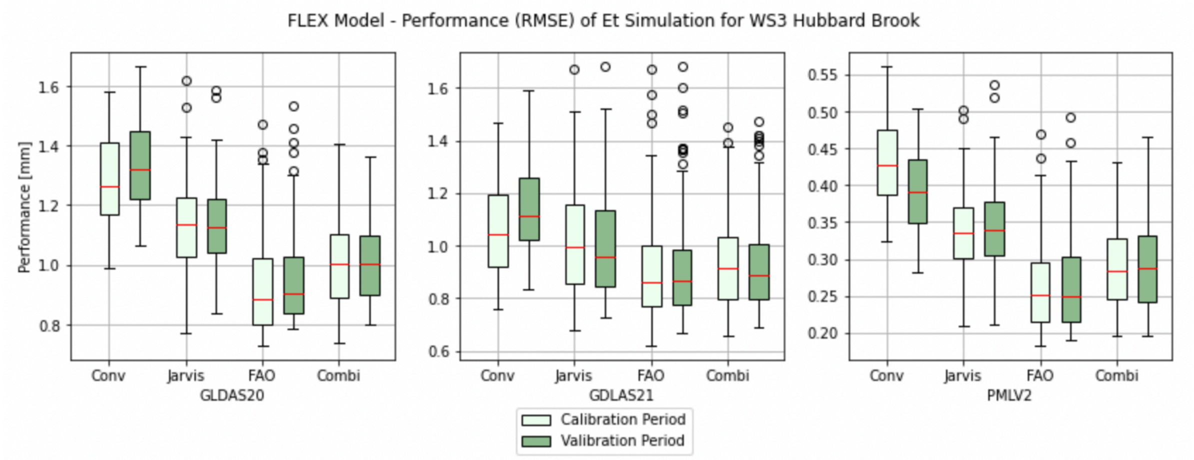


Figure 4.7: FLEX Model - Performance of Transpiration Simulations for WS3 in Hubbard Brook. Performances of all behavioral transpiration simulations by each of the structures of the FLEX model. The performances are quantified by the objective function of the root mean square error (RMSE). Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

The mean seasonal sums of the estimated and simulated values of transpiration and interception evaporation are visible in figure 4.8. The seasons are again divided over four groups of three months, spring for the months March, April, May (mam), summer for June, July, August (jja), fall for September, October, November (son), and winter for December, January, and February (djf). The mean seasonal sums of the transpiration simulations by the three modified model structures are most similar to the sums of the transpiration estimations. The mean seasonal sums of the transpiration simulations by the conventional model structure are least similar to the sums of the transpiration estimations, as these are too high during the winter and spring and too low during the summer. The mean seasonal sums of the interception evaporation simulations by each of the four model structures do now reveal many similarities to the sums of the interception evaporation estimations.

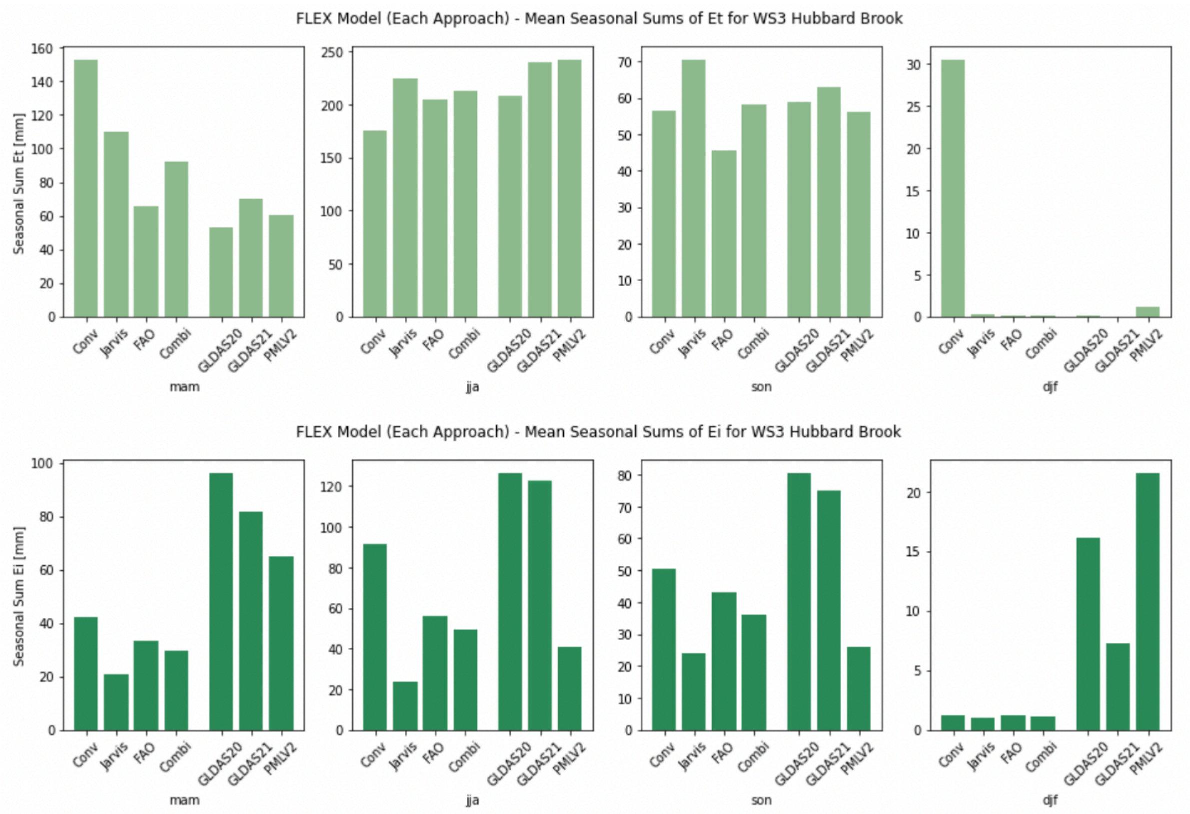


Figure 4.8: FLEX Model - Mean Seasonal Sums of Evaporation for WS3 in Hubbard Brook. Mean sums of the transpiration and interception evaporation estimations and simulations by each of the four structures of the FLEX model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December,January, and February. Obs = observations, Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

### Root Zone Storage Capacity

The Gumbel estimated root zone storage capacity ($S_R$) for WS3 is visible in figure 4.9 for a range of values for the maximum storage capacity of the interception reservoir ($I_{max}$) and a range of values for the return periods. The root zone storage capacity is similar to the parameter of the maximum storage capacity of the unsaturated soil reservoir $S_{u,max}$ in the FLEX model. Hence, the medians of the calibrated $S_{u,max}$ values by each of the four FLEX model structures are also plotted in figure 4.9 at the return period of 20 years. The root zone storage capacity at the return period of 20 years is considered for ecosystems to develop their root zone to overcome droughts (Gao, Hrachowitz, Schymanski, et al., 2014; Nijzink et al., 2016). None of the median calibrated $S_{u,max}$ fit between the upper ($I_{max}$ = 1 mm) and lower ($I_{max}$ = 5 mm) boundary of $S_R$ at the return period of 20 years. This means that the medians of the calibrated $S_{u,max}$ values by each of the four FLEX model structures are not consistent with $S_R$ estimated by another method (Gao, Hrachowitz, Schymanski, et al., 2014; Nijzink et al., 2016).

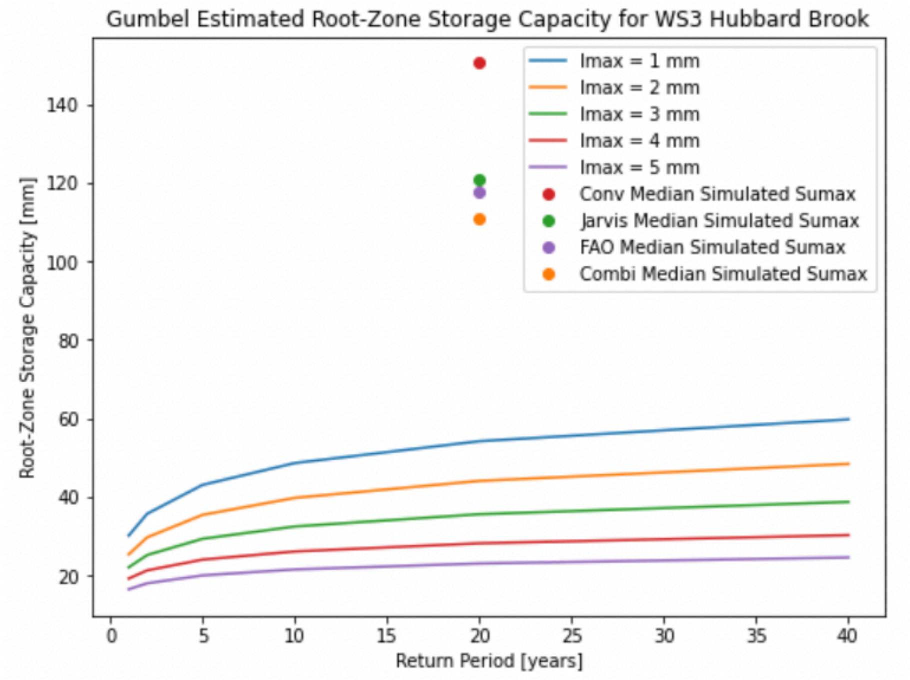


Figure 4.9: FLEX Model - Root Zone Storage Capacity Simulations for WS3 in Hubbard Brook. Gumbel estimated root zone storage capacities calculated for multiple values of $I_{max}$ and for multiple return periods. The colored dots are the median simulated values of the parameter $S_{u,max}$ of each of the four FLEX model structures. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

### 4.1.3. GR4J Model

The GR4J model is used with each of the four model structures, the conventional, Jarvis-based, FAO-based, and combined model structure, to simulate daily values of streamflow and evaporation for WS3. Each of the four GR4J model structures is calibrated by the uniform Monte Carlo method with the daily streamflow observations as the reference data. The Monte Carlo method used 5000 iterations for each GR4J model structure and resulted in the number of behavioral simulations, which have a Nash-Sutcliffe efficiency above 0.7 in the calibration period, of 34 for the conventional structure, 110 for the Jarvis-based structure, 141 for the FAO-based structure and 142 for the combined structure. The results below only included the behavioral simulations. Each of the four GR4J model structures is in a similar way calibrated to the daily GLDAS-21 transpiration estimation. The results of the transpiration calibrated models are similar to the streamflow calibrated models and are included in appendix A.

#### Streamflow

The medians of the daily streamflow simulations by each of the four GR4J model structures together with the streamflow and precipitation observations in the calibration and validation periods are visible in figure 4.10. By visual inspection of the graphs in figure 4.10, the streamflow simulations by the GR4J model structures seem to perform similar to the streamflow simulations by the FLEX model structures. The streamflow simulations seem similar to the streamflow observations except for the very high peaks. No significant difference can be perceived in the streamflow simulations between each of the four GR4J model structures.

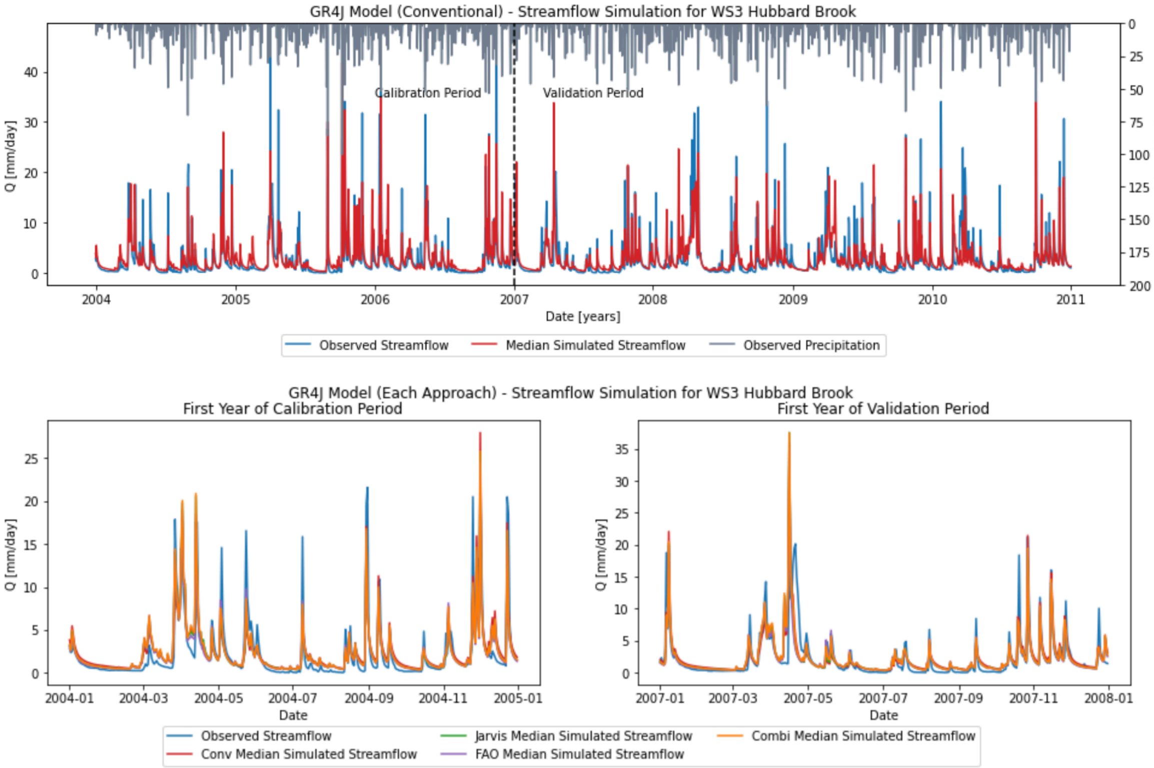


Figure 4.10: GR4J Model - Streamflow Simulations for WS3 in Hubbard Brook. The upper graph consists of the daily observations of streamflow and precipitation together with the median of the daily simulations of streamflow by the conventional GR4J model. Lower graphs consist of the daily observations of streamflow together with the medians of the daily simulations of streamflow by each of the four structures of the GR4J model. Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The performances of the daily streamflow simulations of each GR4J model structure are visible in figure 4.11. The performances in this figure are determined by the four objective functions, the Nash-Sutcliffe efficiency ($NS$), the log Nash-Sutcliffe efficiency ($logNS$), the coefficient of determination ($R^2$) and the root mean square error ($RMSE$), and mostly seem to indicate the same results as by the visual inspection of the graphs in 4.10. Overall, the performances of the streamflow simulations in the calibration period are better than the validation period and no significant difference can be perceived in the performances between the simulations by each of the four GR4J model structures. However, the streamflow simulations by the three modified model structures seem to slightly perform better than those by the conventional model structure according to $logNS$ and the other way around according $R^2$

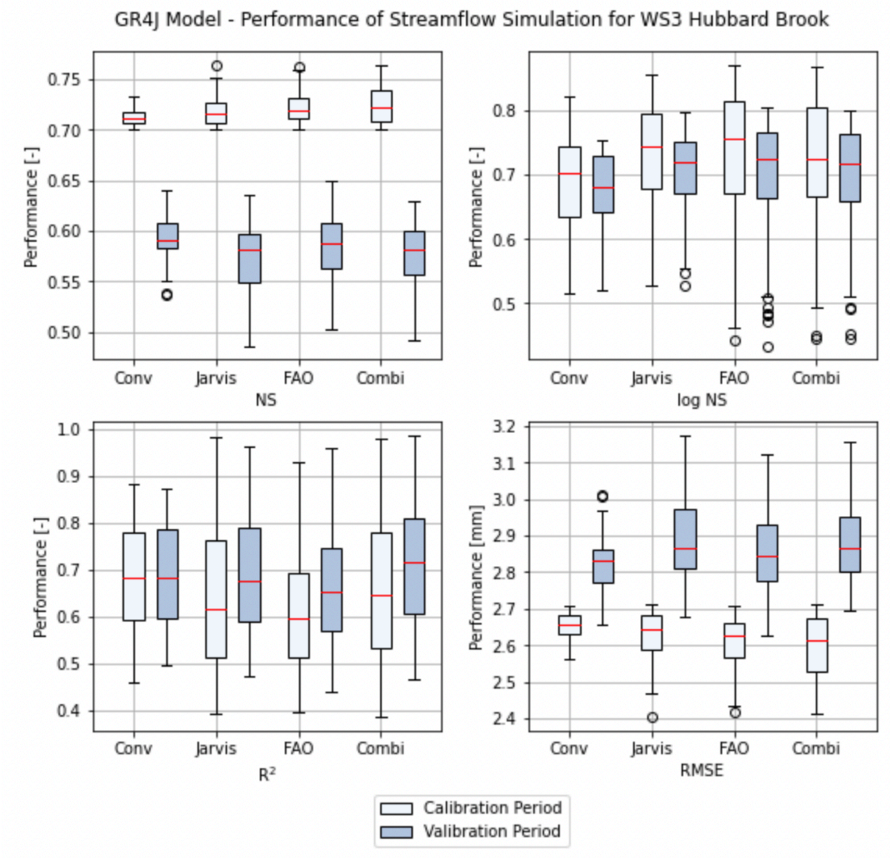


Figure 4.11: GR4J Model - Performance of Streamflow Simulations for WS3 in Hubbard Brook. Performances of all behavioral streamflow simulations by each of the structures of the GR4J model. The performances are quantified by the objective functions of the Nash-Sutcliffe efficiency (NS), the log Nash-Sutcliffe efficiency (logNS), the coefficient of determination (R2), and the root mean square error (RMSE). Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The mean seasonal sums of the daily streamflow observations and simulations by the four GR4J model structures can be seen in figure 4.12. There is not much difference between the seasonal sums of the streamflow observations and the streamflow simulations by any of the four model structures.

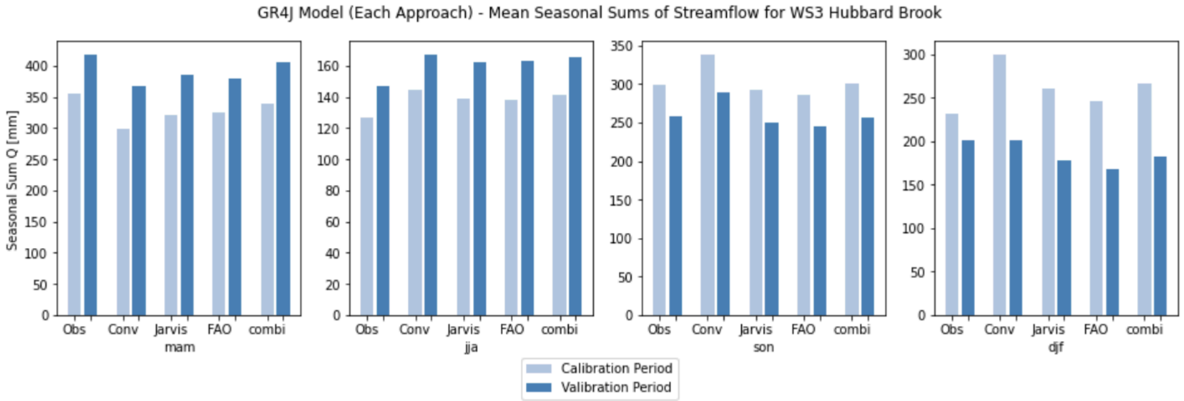


Figure 4.12: GR4J Model - Mean Seasonal Sums of Streamflow for WS3 in Hubbard Brook. Mean sums of the streamflow observations and simulations by each of the four structures of the GR4J model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. Obs = observations, Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The flow duration curve for the streamflow observations and the simulations of the four GR4J model structures is visible in figure 4.13. There is not much difference between the flow duration curve for the streamflow observations and the streamflow simulations by each of the four model structures. Similar to the flow duration curves generated by the FLEX model structures for WS3, the flow duration curves for the GR4J model simulations show that higher streamflow values are less often reached, or not reached at all, compared with the streamflow observations.

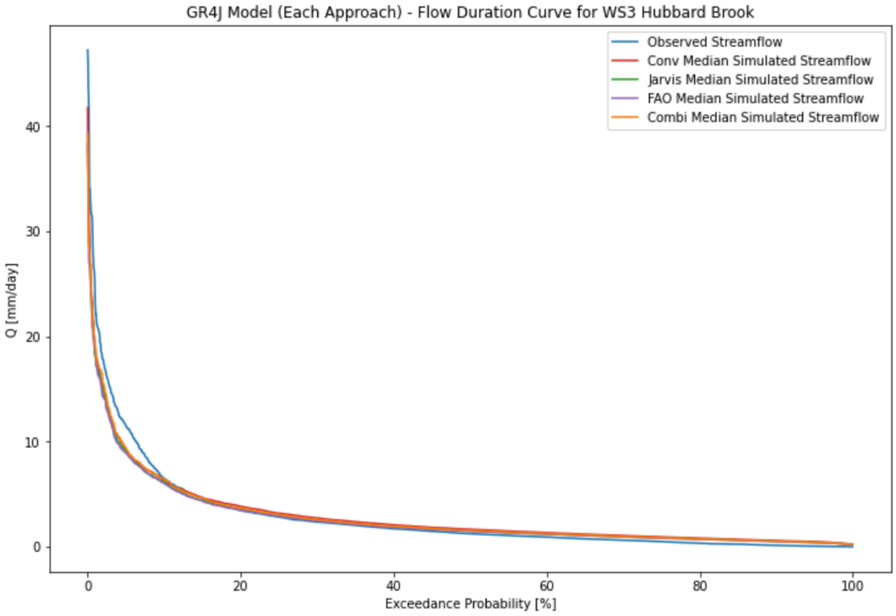


Figure 4.13: GR4J Model - Flow Duration Curve for WS3 in Hubbard Brook. Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

### Evaporation

The medians of the transpiration simulations of each GR4J model structure are visible in figure 4.14. Again, the transpiration simulations by the GR4J models are very similar to the transpiration simulations by the FLEX models. The transpiration simulations of the modified GR4J model structures approach the transpiration estimations of GLDAS-20, GLDAS-21, and PML-V2 much more than the conventional GR4J model. The transpiration simulations by FAO-based model structure seem to come closest to the transpiration estimations, followed by those by the combined structure and then the Jarvis-based structure. The transpiration simulations by the conventional model structure are too high during the winters and too low during the summers compared to the transpiration estimations.

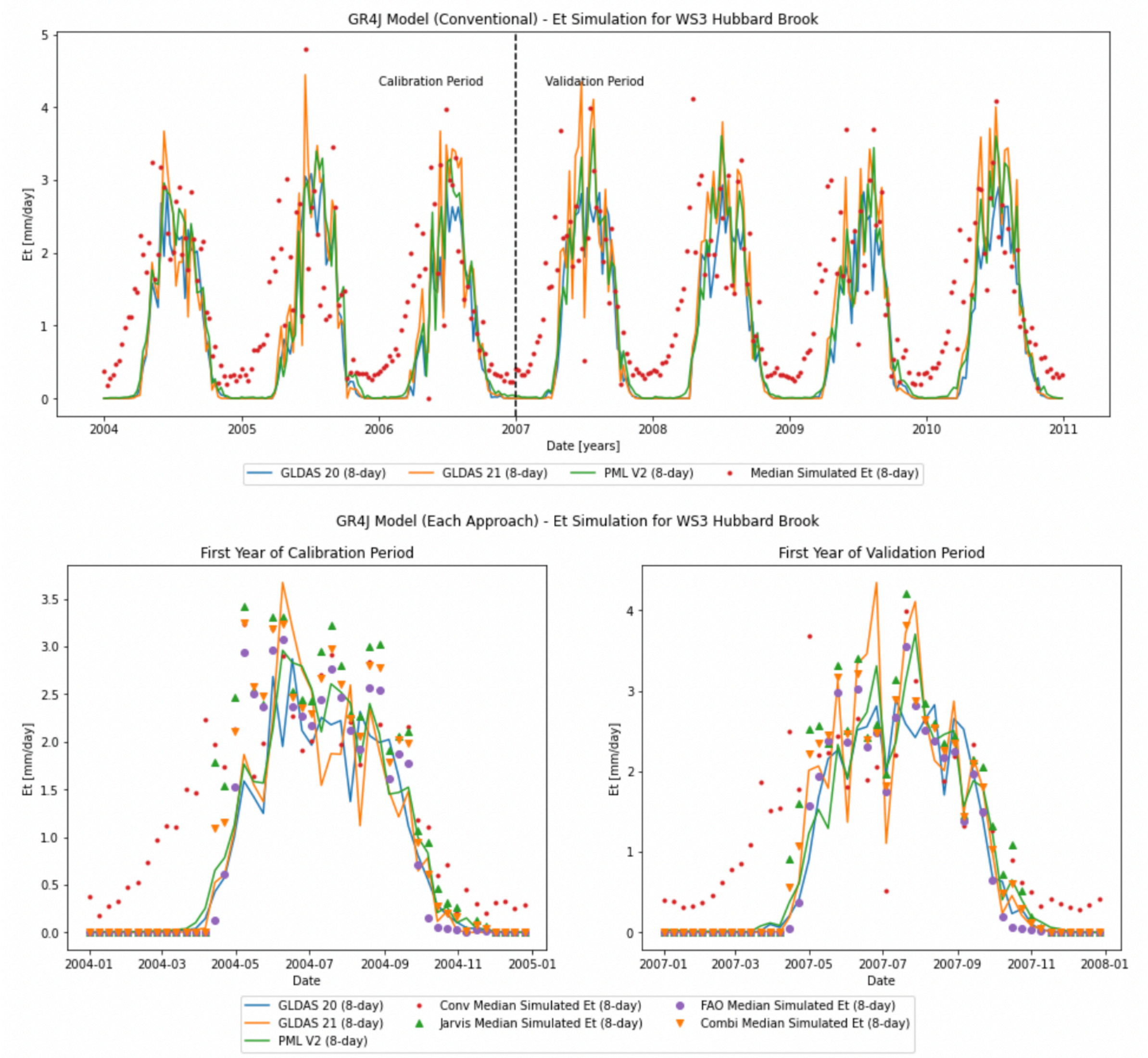


Figure 4.14: GR4J Model - Transpiration Simulations for WS3 in Hubbard Brook. The upper graph consists of the estimations of transpiration together with the median of the simulations of transpiration by the conventional GR4J model. Lower graphs consist of the estimations of transpiration together with the medians of the simulations of transpiration by each of the four structures of the GR4J model. Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The performances of the transpiration simulations according to the RMSE are visible in figure 4.15. It shows that the performances of the transpiration simulations, according to the transpiration estimations, by the modified model structures are much better than those by the conventional model structures. The transpiration simulations of the FAO-based model structure seem to perform best, according to the transpiration estimations, compared to those by the other model structures. The transpiration simulations mostly perform similarly for the calibration and the validation periods.

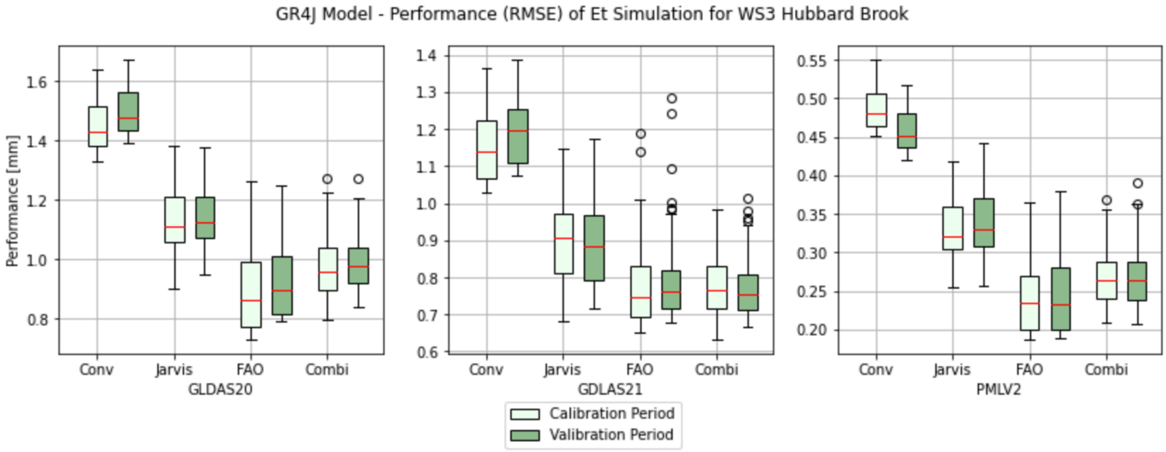


Figure 4.15: GR4J Model - Performance of Transpiration Simulations for WS3 in Hubbard Brook. Performances of all behavioral transpiration simulations by each of the structures of the GR4J model. The performances are quantified by the objective function of the root mean square error (RMSE). Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The mean seasonal sums of the estimated and simulated values of transpiration and interception evaporation are visible in figure 4.16. The seasonal transpiration sums of the FAO-based model structure are most similar to the estimated sums. Whereas, the seasonal transpiration sums of the conventional model structure are least similar to the estimated sums. None of the seasonal interception evaporation sums of the model structures are similar to the sums of the estimations.

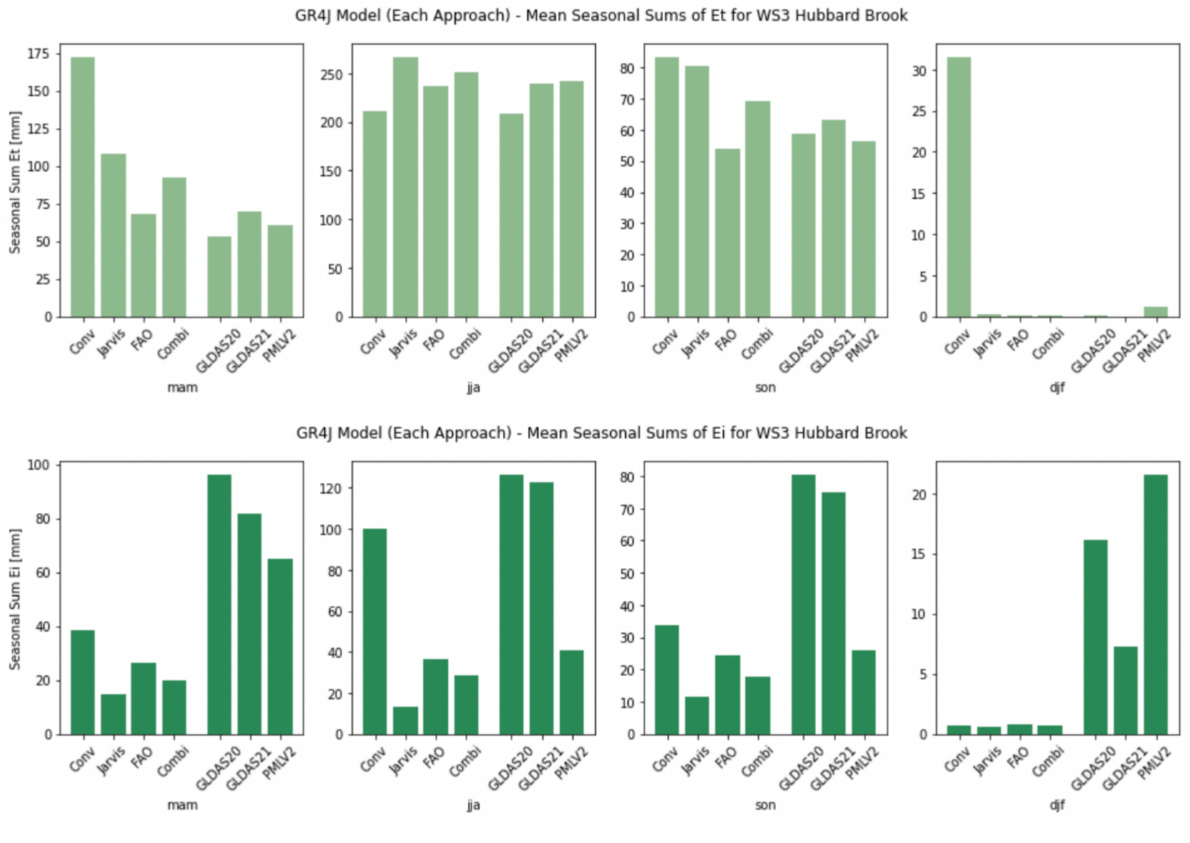


Figure 4.16: GR4J Model - Mean Seasonal Sums of Evaporation for WS3 in Hubbard Brook. Mean sums of the transpiration and interception evaporation estimations and simulations by each of the four structures of the GR4J model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December,January, and February. Obs = observations, Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

### Root Zone Storage Capacity

The medians of the calibrated values for the parameter maximum storage capacity of the production reservoir ($x_1$) are plotted beside the Gumbel estimated $S_R$ at the return period of 20 years. Similarly to the medians of the simulated $S_{u,max}$ of the FLEX model for WS3, the medians of the simulated $x_1$ do not fit within the upper ($I_{max}$ = 1 mm) and lower ($I_{max}$ = 5 mm) boundaries of the Gumbel estimated $S_R$ at the return periods of 20 years. Hence, the median $x_1$ simulations of the GR4J model structures are also not consistent with $S_R$ estimated by other methods (Gao, Hrachowitz, Schymanski, et al., 2014; Nijzink et al., 2016) for WS3.

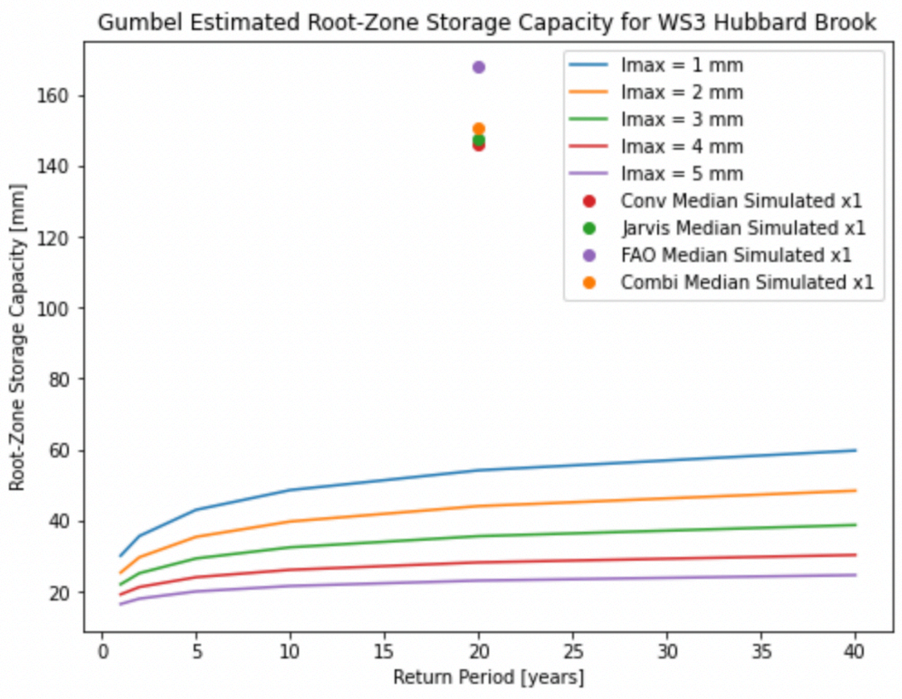


Figure 4.17: GR4J Model - Root Zone Storage Capacity Simulations for WS3 in Hubbard Brook. Gumbel estimated root zone storage capacities calculated for multiple values of $I_{max}$ and for multiple return periods. The colored dots are the median simulated values of the parameter $x_1$ of each of the four GR4J model structures. Conv= conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi =combined GR4J model.

## 4.2. HJ Andrews

The simulations of the daily streamflow and evaporation for watershed 2 (WS2) in the experimental forest of HJ Andrews started on January 1, 2005, and ended on December 31, 2010. The first year of the simulation period, between January 1, 2005, and December 31, 2005, served as the spin-up period. The second period, between January 1, 2006, and December 31, 2007, served as the calibration period and the last period, between January 1, 2008, and December 31, 2010, served as the validation period.

### 4.2.1. Leaf Area Index

The MODIS-derived LAI for HJ Andrews' WS2 can be seen in figure 4.18 for the calibration and validation periods.

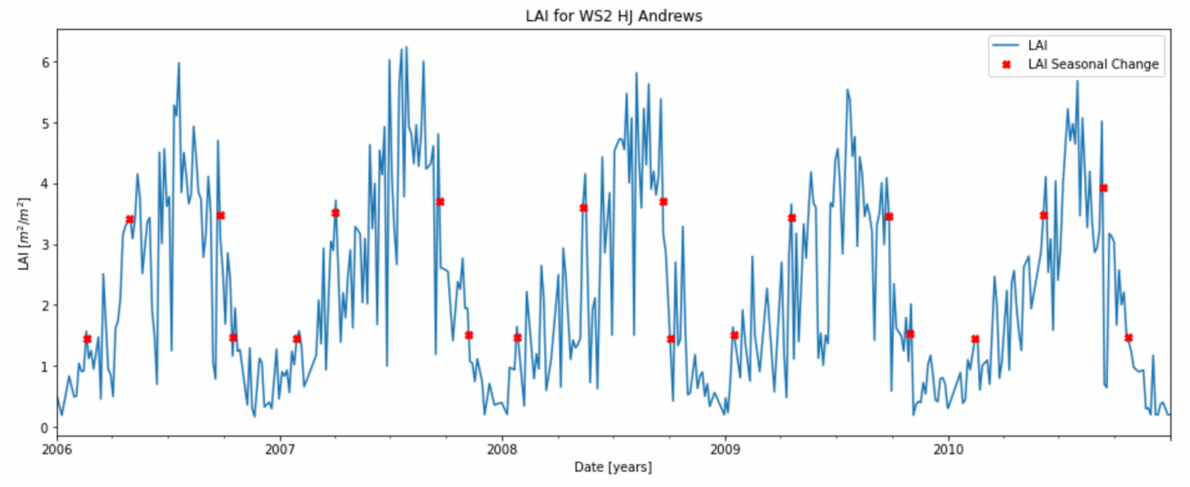


Figure 4.18: Leaf Area Index for WS2 in HJ Andrews. The red dots indicate the transition of the vegetation's development stages. The periods between the red dots around LAI equal to 1 indicate the dormant seasons. The periods between the red dots around LAI equal to 3.5 indicate the peak seasons.

The Leaf Area Index for the catchment in HJ Andrews reveals a pattern with moderate seasonal variation. The pattern consists of peaked values during the summers and minimum values during the winters. Compared to LAI for the catchment in Hubbard Brook, LAI for the catchment in HJ Andrews is for shorter periods minimal and takes longer to obtain peaked values. The pattern for the catchment in HJ Andrews shows that when LAI exceeds the value of 1, LAI gradually increases to a value around 3.5. Afterward, LAI does not increase much more and drops around mid-fall. Around mid-fall, LAI decreases to a value around 1 after which it does deviate much again until around mid-spring. Around mid-spring, LAI continues the same cycle as described. Based on this pattern, the dormant season of the vegetation in the catchment is defined as the period between two LAI values of 1 and the peak values between two LAI values of 3.5. The periods in between are defined as the transitional periods, in which vegetation grows leaves for increasing LAI values and drops leaves for decreasing LAI values.

### 4.2.2. FLEX Model

The FLEX model with each of the four described structures is used to simulate the daily values of streamflow and evaporation for the catchment in HJ Andrews. Each of the four FLEX model structures is calibrated using 5000 iterations in the GLUE method. The number of behavioral simulations, which have a Nash-Sutcliffe efficiency above 0.7 for the calibration period, is 427 for the conventional structure, 239 for the Jarvis-based structure, 157 for the FAO-based structure, and 212 for the combined structure. The results below only included the behavioral simulations. The results of the transpiration calibrated models are similar to the streamflow calibrated models and are visible in appendix B.

#### Streamflow

Figure 4.19 consists of three graphs, in which the upper graph contains the observed values of daily streamflow and precipitation and the median of the simulated values of daily streamflow by the conventional FLEX model structure for the catchment in HJ Andrews. The lower graphs contain the observed values and the medians of the simulated values of daily streamflow by each FLEX model structure for the first year of the calibration period and the first year of the validation period.

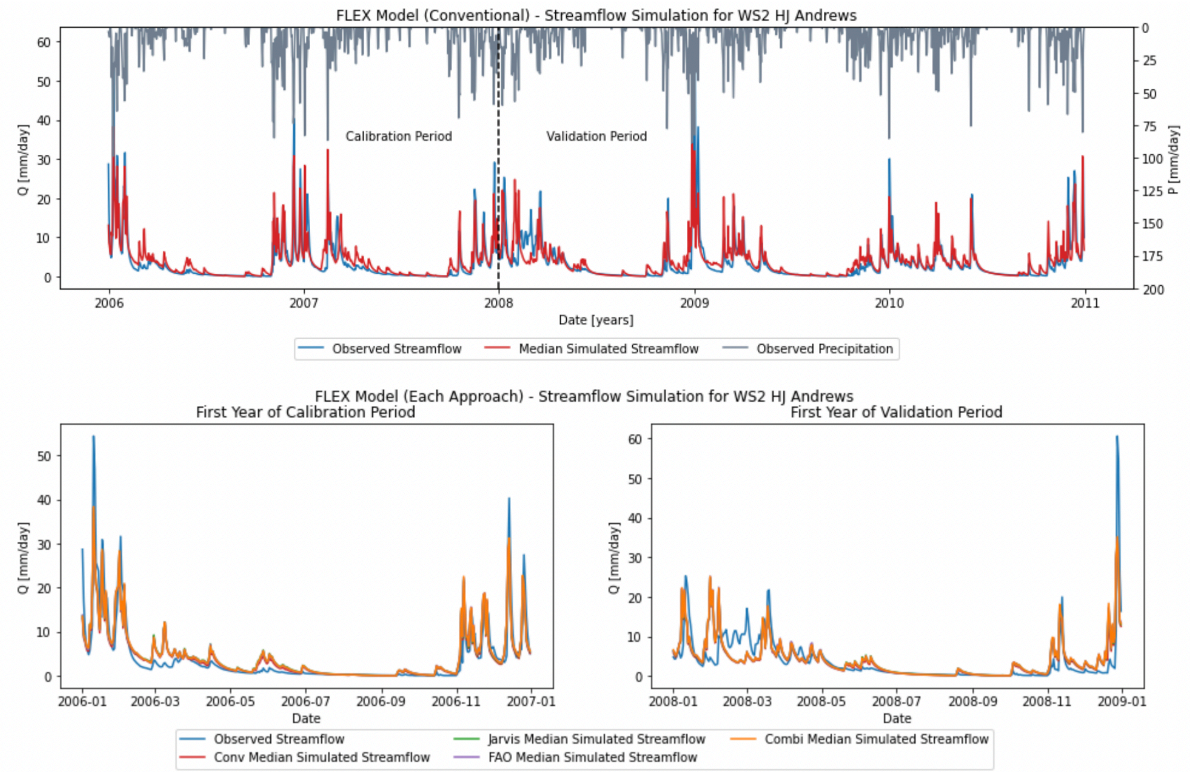


Figure 4.19: FLEX Model - Streamflow Simulations for WS2 in HJ Andrews. The upper graph consists of the daily observations of streamflow and precipitation together with the median of the daily simulations of streamflow by the conventional FLEX model. Lower graphs consist of the daily observations of streamflow together with the medians of the daily simulations of streamflow by each of the four structures of the FLEX model. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

Visual inspection of the graphs in figure 4.19 indicates that the daily streamflow simulations by the conventional model structure are highly similar to the daily streamflow observations, except for the period around March in 2008. During this period the streamflow observations lag behind the precipitation observations and the streamflow simulations. The daily streamflow simulations by the three modified structures perform equally well compared to the conventional model structure. In the graphs of figure 4.19, almost no difference can be perceived between the streamflow simulations of each model structure.

Figure 4.20 contains boxplots for the performances of the streamflow simulations of each FLEX model structure, which is determined by the four objective functions of NS, logNS, $R^2$, and RMSE with the streamflow observations as the reference data. Overall, the streamflow simulations during the calibration period perform slightly better than during the validation period. The performances of the streamflow simulations of each model structure are mostly equal, with exception of the performances quantified by $R^2$. For this, the median performances of the streamflow simulations by the FAO-based and the combined model structure are slightly better than the other two model structures.

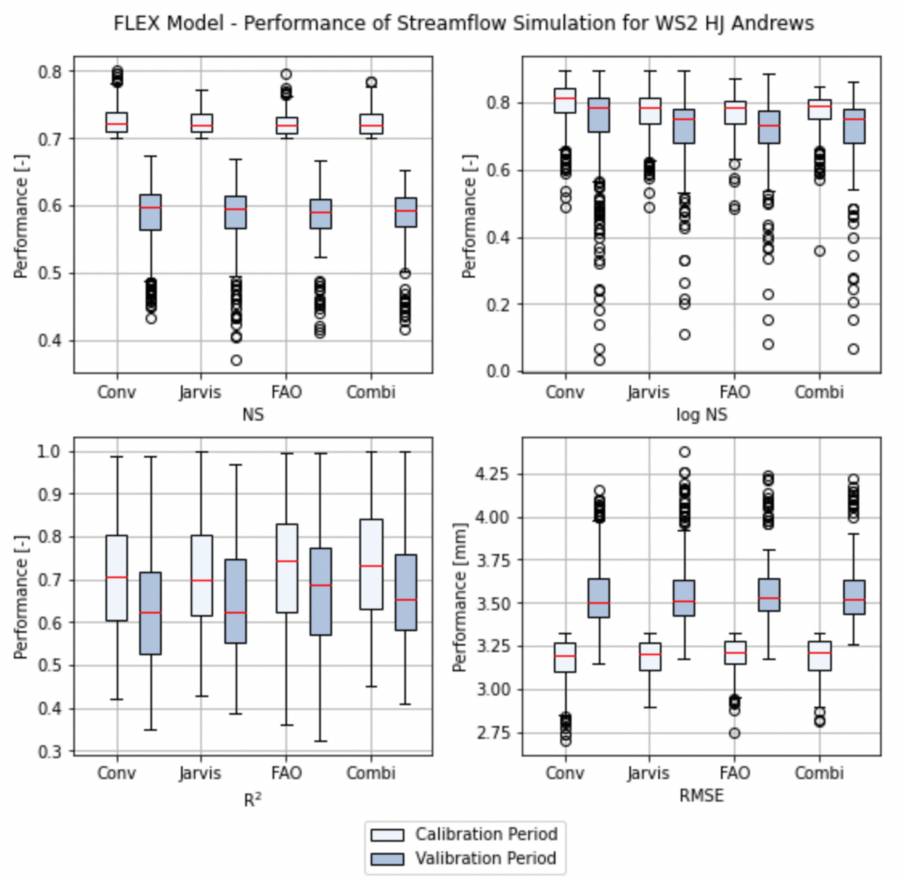


Figure 4.20: FLEX Model - Performance of Streamflow Simulations for WS2 in HJ Andrews. Performances of all behavioral streamflow simulations by each of the structures of the FLEX model. The performances are quantified by the objective functions of the Nash-Sutcliffe efficiency (NS), the log Nash-Sutcliffe efficiency (logNS), the coefficient of determination (R2), and the root mean square error (RMSE). Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

Figure 4.21 contains the mean seasonal sums of the daily streamflow observations and simulations of each FLEX model structure. No major difference is perceived between the mean seasonal sums of the streamflow simulations. Overall, the mean seasonal sums of the streamflow observations are slightly lower compared to those of the simulations for the seasons mam, spring, and jja, summer.

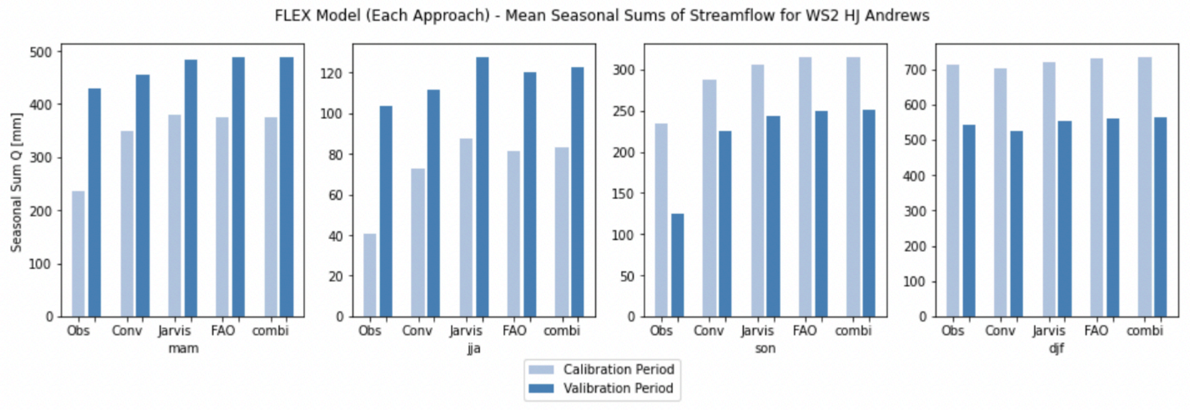


Figure 4.21: FLEX Model - Mean Seasonal Sums of Streamflow for WS2 in HJ Andrews. Mean sums of the streamflow observations and simulations by each of the four structures of the FLEX model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. Obs = observations, Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

Figure 4.22 consists of the flow duration curves generated by the daily streamflow observations and simulations of each FLEX model structure. It is perceived that the streamflow simulations are unable to reach or exceed the extremely high streamflow values as by the streamflow observations. Furthermore, for the lower streamflow values between 0 and 10 mm/day the streamflow simulations can slightly reach or exceed these values more often than the observations.

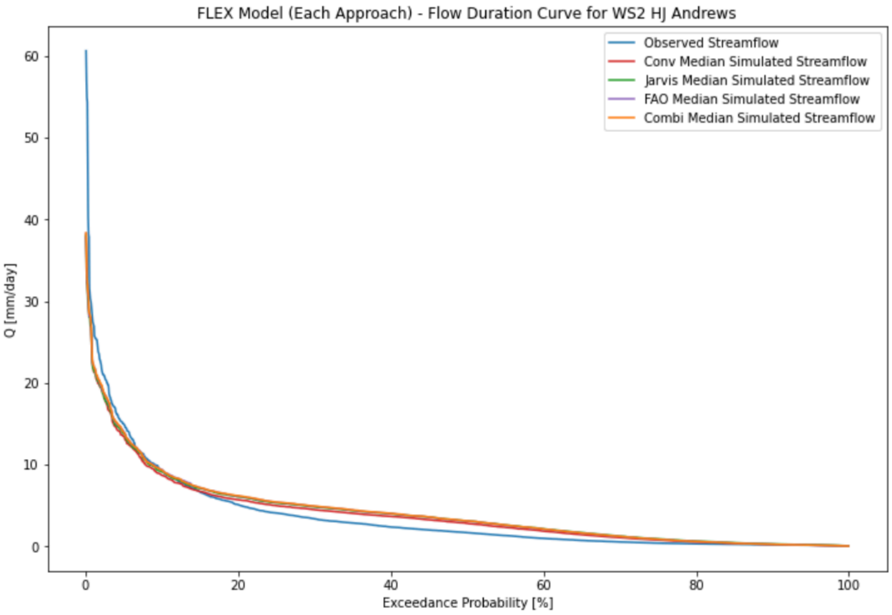


Figure 4.22: FLEX Model - Flow Duration Curve for WS2 in HJ Andrews. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

### Evaporation

Figure 4.23 consists of three graphs, in which the upper graph contains the transpiration estimations, by GLDAS-20, GLDAS-21, and PML-V2, and the median of the simulations of the conventional model structure on an 8-day temporal resolution for the catchment in HJ Andrews. The two lower graphs contain the transpiration estimations and the medians of the simulations of each FLEX model structure for the first year of the calibration period and the first year of the validation period.

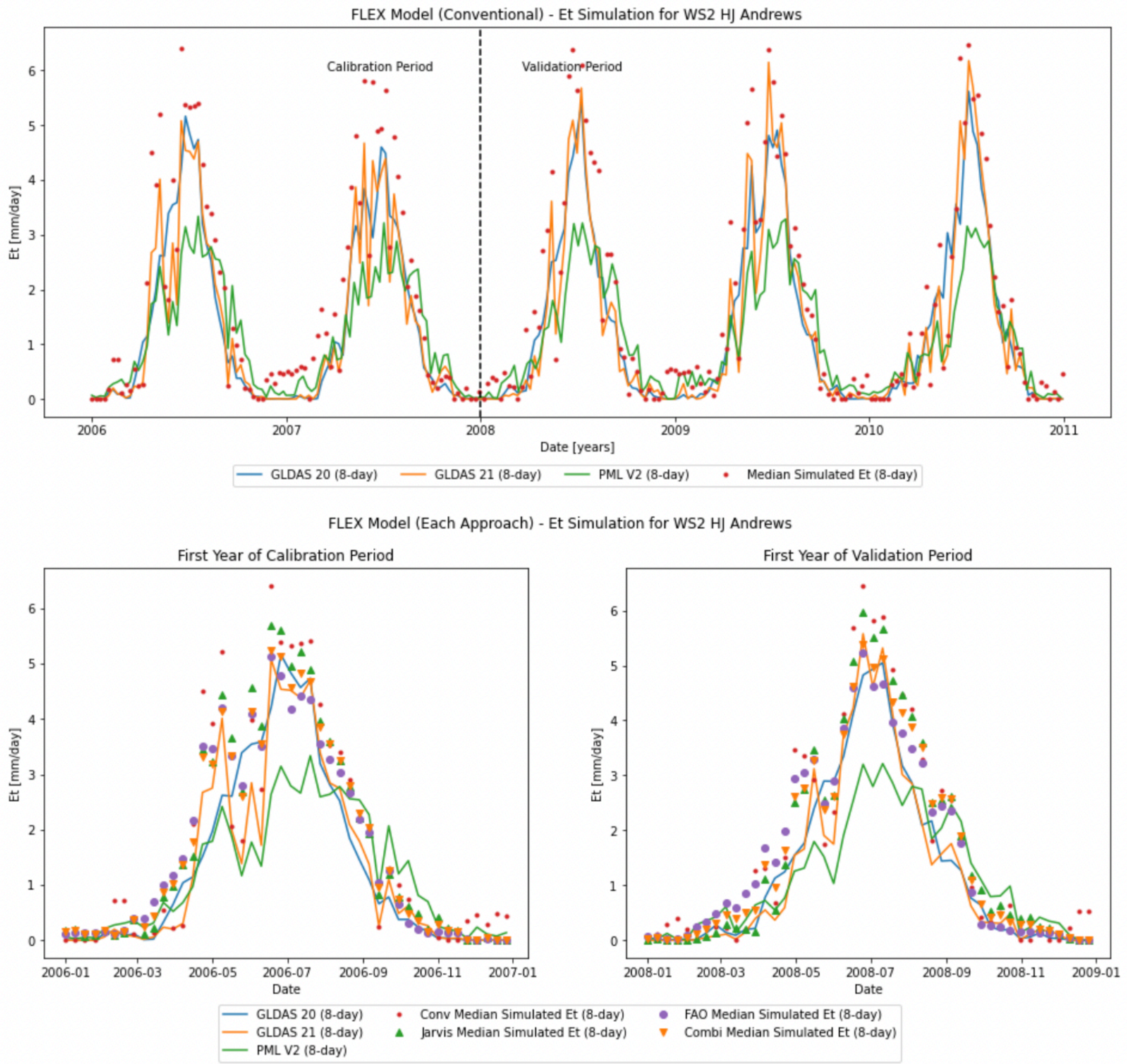


Figure 4.23: FLEX Model - Transpiration Simulations for WS2 in HJ Andrews. The upper graph consists of the estimations of transpiration together with the median of the simulations of transpiration by the conventional FLEX model. Lower graphs consist of the estimations of transpiration together with the medians of the simulations of transpiration by each of the four structures of the FLEX model. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

Visual inspection of the graphs in figure 4.23 shows that transpiration estimated by GLDAS-20 and GLDAS-21 are highly similar to one another and that transpiration estimated by PML-V2 is slightly lower during the summers. The transpiration simulated by each model structure shows many similarities to the values estimated by GLDAS-20 and GLDAS-21. Additionally, the transpiration simulated by each model structure shows many similarities to one another. Overall, however, the transpiration simulations of the conventional and the Jarvis-based model structures seem slightly higher compared to transpiration simulations of the other model structures. The transpiration estimations and simulations show a similar pattern to the LAI of the catchment in HJ Andrews, in that the values peak during the summers and minimize during the winters.

Figure 4.24 contains boxplots for the performances, determined by RMSE, of the transpiration simulations by each FLEX model structure according to the transpiration estimations. The median performances of the transpiration simulations by the FAO-based and combined model structure perform best according to the transpiration estimations. The performances of the transpiration simulations by the conventional model structure perform the least according to the estimations.

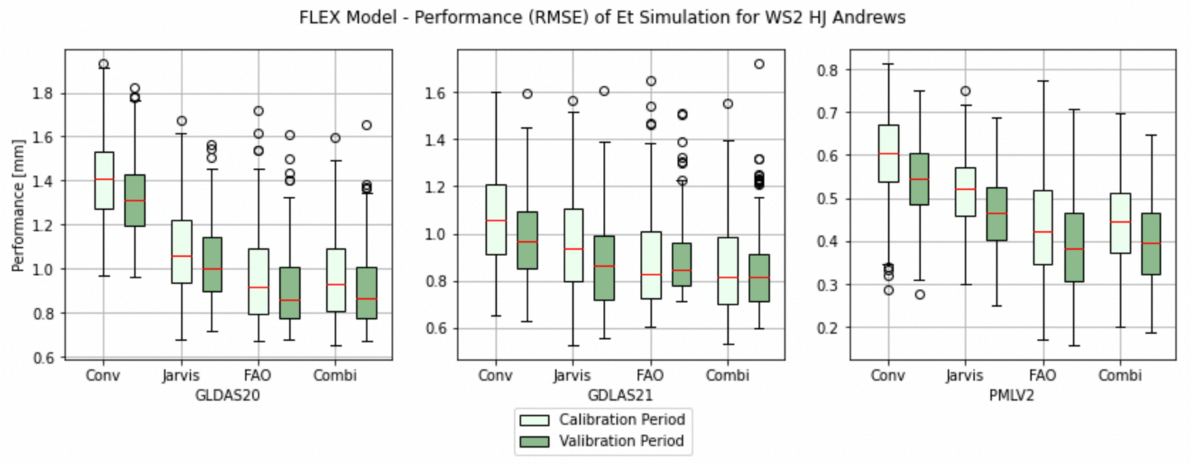


Figure 4.24: FLEX Model - Performance of Transpiration Simulations for WS2 in HJ Andrews. Performances of all behavioral transpiration simulations by each of the structures of the FLEX model. The performances are quantified by the objective functions of the root mean square error (RMSE). Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model

Figure 4.25 shows the mean seasonal sums of the transpiration and interception evaporation estimations and simulations. The seasonal sums of transpiration simulated by each model structure tend to be slightly higher than the estimated seasonal sums. For the winter season, which consists of the months December, January, and February (djb), the seasonal sum of transpiration simulated by the conventional model structure is significantly higher than the estimated seasonal sums and simulated sums of the other model structures. The seasonal sums of interception evaporation simulated by the conventional structure tend to be closer near the estimated seasonal sums of GLDAS-20 and GLDAS-21. Whereas, the seasonal sums of interception evaporation simulated by the modified structures tend to be closer to the estimated seasonal sums of PML-V2.

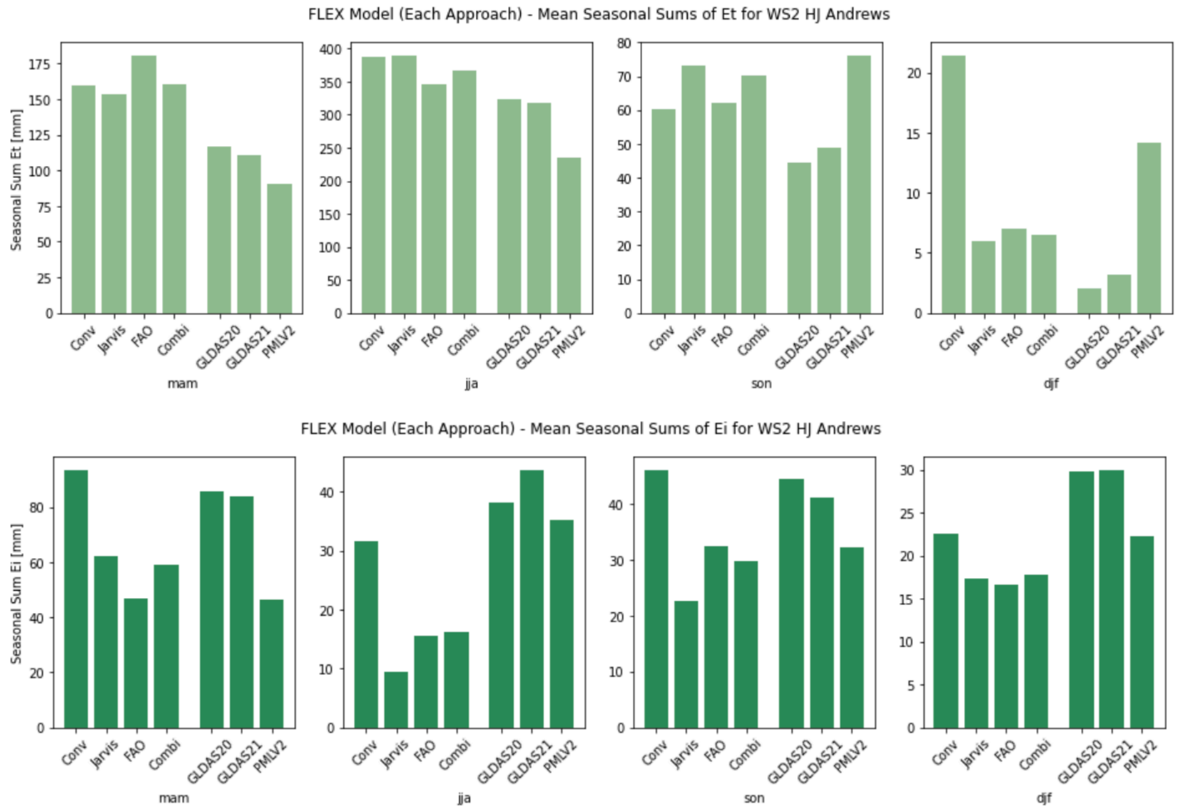


Figure 4.25: FLEX Model - Mean Seasonal Sums of Evaporation for WS2 in HJ Andrews. Mean sums of the transpiration and interception evaporation estimations and simulations by each of the four structures of the FLEX model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December,January, and February. Obs = observations, Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi = combined FLEX model.

## Root Zone Storage Capacity

Figure 4.26 shows the Gumbel estimated root zone storage capacity of the catchment in HJ Andrews including the median calibrated values for the parameter of the maximum storage capacity of the unsaturated soil reservoir ($S_{u,max}$). The median calibrated $S_{u,max}$ of the conventional and the FAO-based model structure do not fit within the upper ($I_{max}$ = 1 mm) and lower ($I_{max}$ = 5 mm) boundaries of the Gumbel estimated root zone storage capacity at the return period of 20 years.

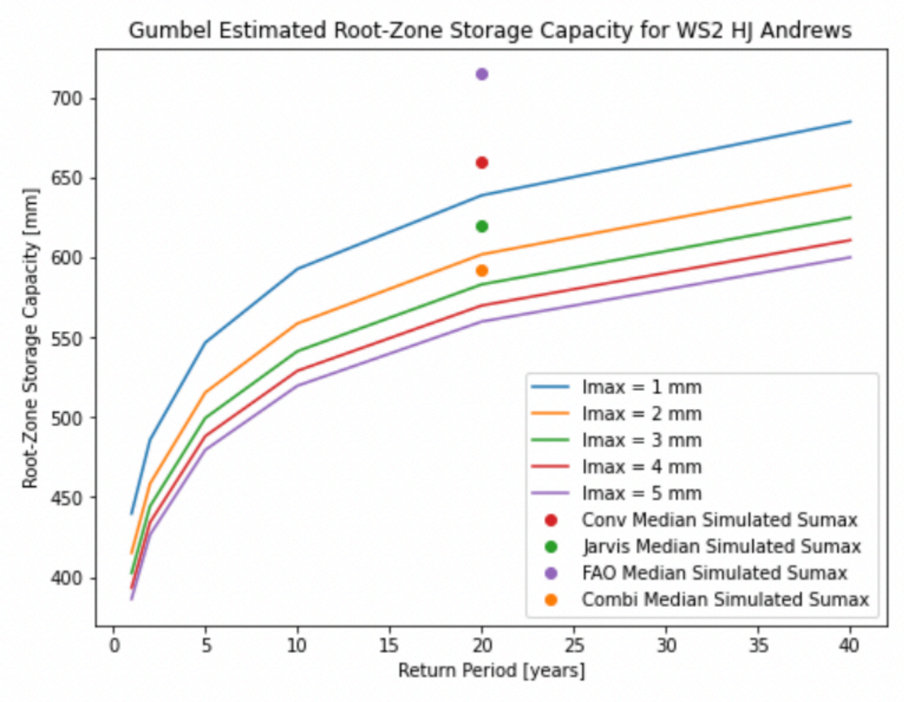


Figure 4.26: FLEX Model - Root Zone Storage Capacity Simulations for WS2 in HJ Andrews. Gumbel estimated root zone storage capacities calculated for multiple values of $I_{max}$ and for multiple return periods. The colored dots are the median simulated values of the parameter $S_{u,max}$ of each of the four FLEX model structures. Conv = conventional FLEX model, Jarvis = Jarvis-based FLEX model, FAO = FAO-based FLEX model and Combi =combined FLEX model.

### 4.2.3. GR4J Model

The GR4J model with each of the four described model structures is used to simulate the daily values of streamflow and evaporation for the catchment in HJ Andrews. Each of the four GR4J model structures is calibrated using 5000 iterations in the GLUE method. The number of behavioral simulations, which have a Nash-Sutcliffe efficiency above 0.7 for the calibration period, is 9 for the conventional model structure, 8 for the Jarvis-based model structure, 7 for the FAO-based model structure, and 10 for the combined model structure. The results below only included the behavioral simulations. The results of the transpiration calibrated models are similar to the streamflow calibrated models and are included in appendix B.

#### Streamflow

Figure 4.27 contains graphs of the medians of the daily streamflow simulations by each of the four GR4J model structures for the catchment in HJ Andrews. The simulations are highly similar to the simulations of the FLEX models and approach the observations for both the calibration and validation periods. The only period, in which the simulations are noticeable different than the observations is around March 2018, similarly to the simulations of the FLEX models.

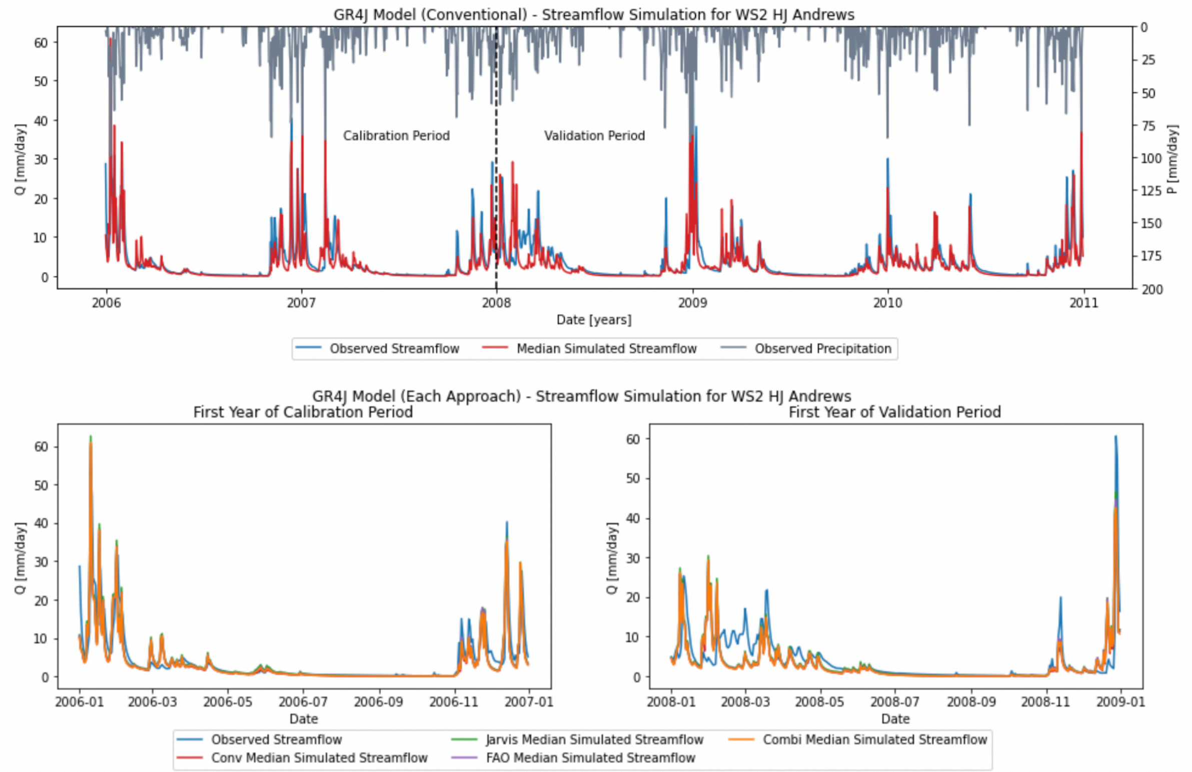


Figure 4.27: GR4J Model - Streamflow Simulations for WS2 in HJ Andrews. The upper graph consists of the daily observations of streamflow and precipitation together with the median of the daily simulations of streamflow by the conventional GR4J model. Lower graphs consist of the daily observations of streamflow together with the medians of the daily simulations of streamflow by each of the four structures of the GR4J model. Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The daily streamflow simulations of the four model structures do not deviate much from each other and almost no difference can be perceived in the graphs. However, a difference can be noticed in the performances of the streamflow simulations quantified by the objective functions (figure 4.28). For this, the streamflow simulations during the calibration periods perform better than during the validation periods. Additionally, the performances of the median streamflow simulations of the Jarvis-based and FAO-based model structures perform better than the other two model structures for the log Nash-Sutcliffe efficiency.

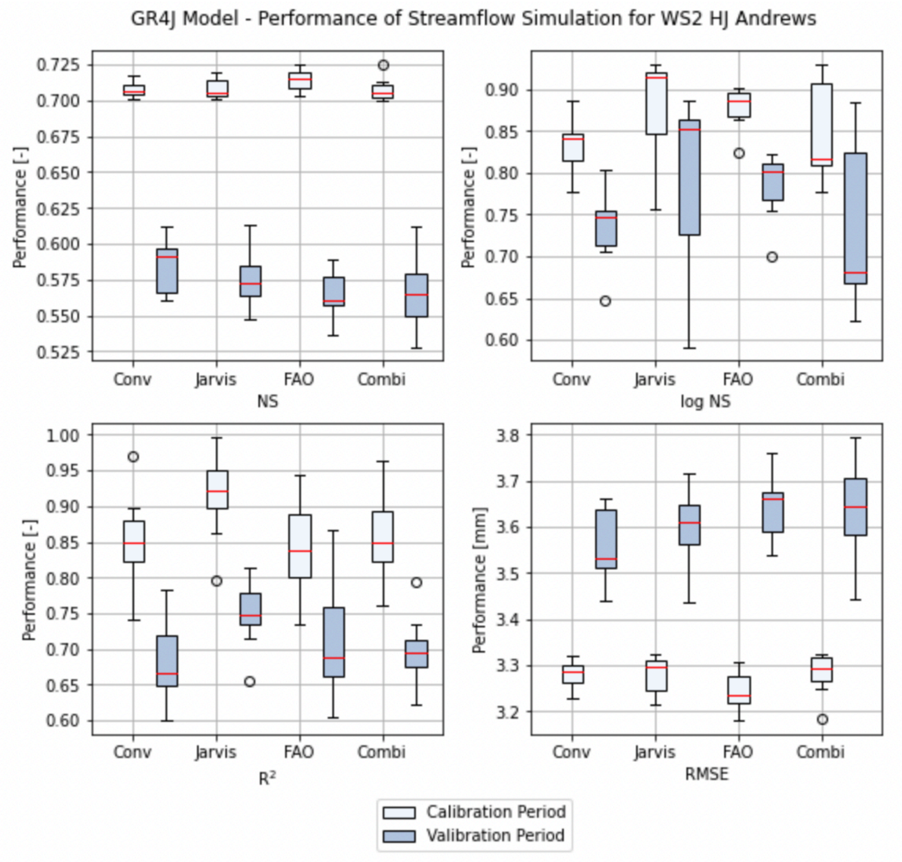


Figure 4.28: GR4J Model - Performance of Streamflow Simulations for WS2 in HJ Andrews. Performances of all behavioral streamflow simulations by each of the structures of the GR4J model. The performances are quantified by the objective functions of the Nash-Sutcliffe efficiency (NS), the log Nash-Sutcliffe efficiency (logNS), the coefficientof determination (R2), and the root mean square error (RMSE). Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

There is not much difference between the mean seasonal sums of streamflow as in figure 4.29 for the seasons mam, spring, and djf, winter. For the seasons jja, summer, and son, fall, the mean seasonal sums of the streamflow simulations are significantly lower than of the streamflow observations. The GR4J model thus underestimates the sum of streamflow during the summer and fall, which is not the case for the FLEX model.

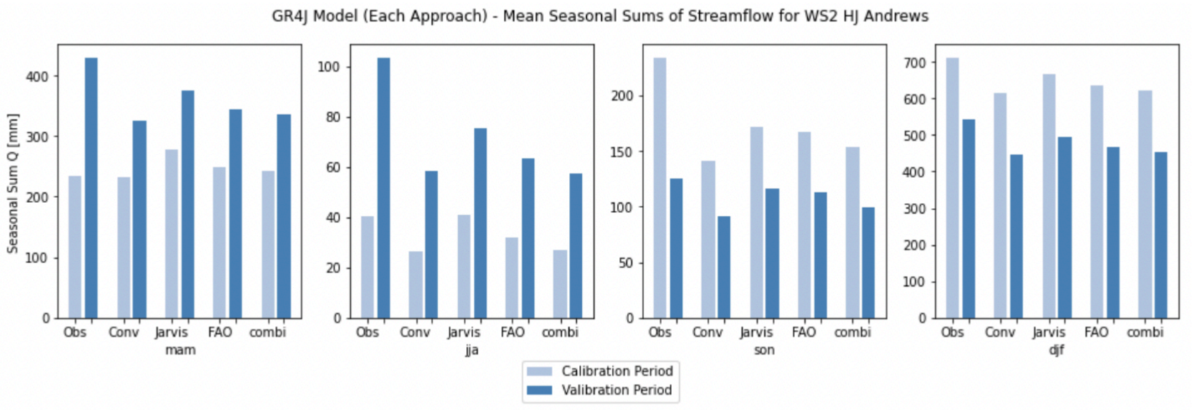


Figure 4.29: GR4J Model - Mean Seasonal Sums of Streamflow for WS2 in HJ Andrews. Mean sums of the streamflow observations and simulations by each of the four structures of the GR4J model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December,January, and February. Obs = observations, Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

According to the flow duration curves in figure 4.30 the simulations are unable to reach or exceed streamflow values around 10 mm/day as often as the observations. However the simulations are now able to reach similar streamflow peaks as the observations for approximately the same frequency.

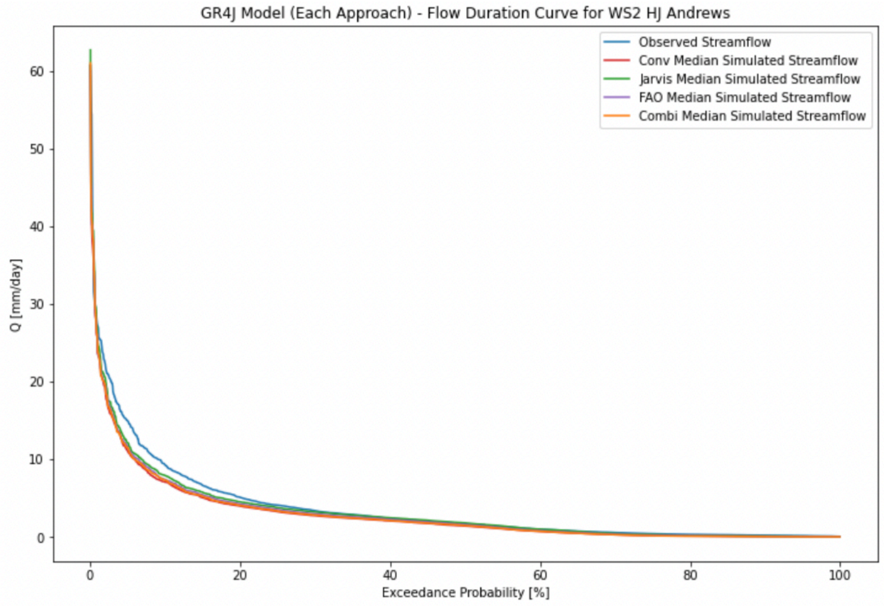


Figure 4.30: GR4J Model - Flow Duration Curve for WS2 in HJ Andrews. Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

### Evaporation

Figure 4.31 contains the graphs of the medians of the daily transpiration simulations, averaged over 8 days, by each of the four GR4J model structures together with the estimations of GLDAS-20, GLDAS-21, and PML-V3 for the catchment in HJ Andrews. The simulations are again very similar to the simulations of the FLEX models and come very close to the estimations for both the calibration and the validation periods.

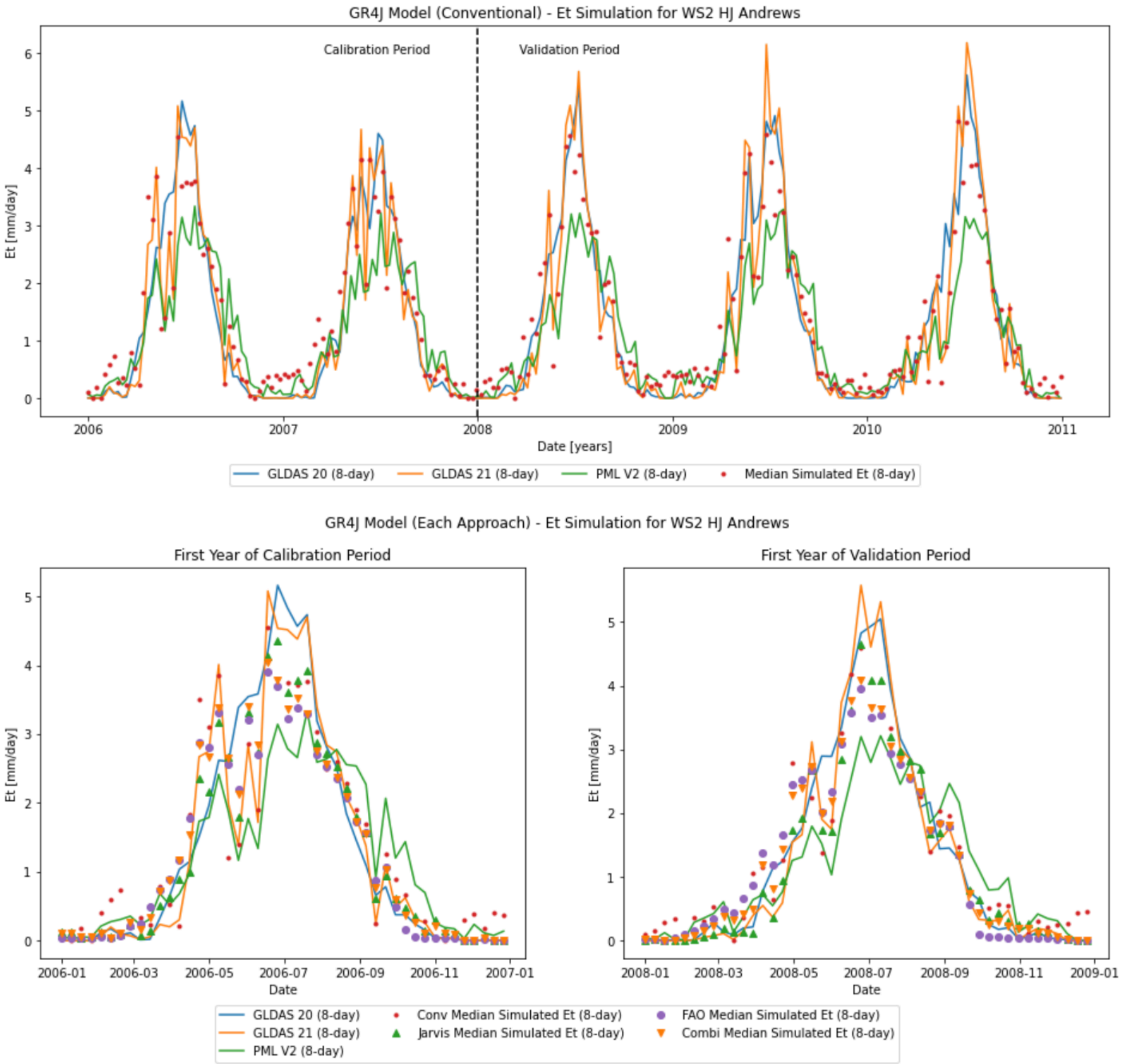


Figure 4.31: GR4J Model - Transpiration Simulations for WS2 in HJ Andrews. The upper graph consists of the estimations of transpiration together with the median of the simulations of transpiration by the conventional GR4J model. Lower graphs consist of the estimations of transpiration together with the medians of the simulations of transpiration by each of the four structures of the GR4J model. Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The transpiration simulations of each of the four model structures are similar to one another and the estimations. Overall, according to the lower graphs in figure 4.31, the transpiration simulations of the conventional model structure are higher than the other model structures during the winters and the simulations of the Jarvis-based model structure are higher than the other model structures during the summers. A much larger difference between the transpiration simulations of each model structure can be perceived by the performances according to the transpiration estimations, quantified by the root mean square error in figure 4.32. For this, the performances of the transpiration simulations of the three modified model structures perform better than those of the conventional model structure according to the transpiration estimations of GLDAS-20 and PML-V2.

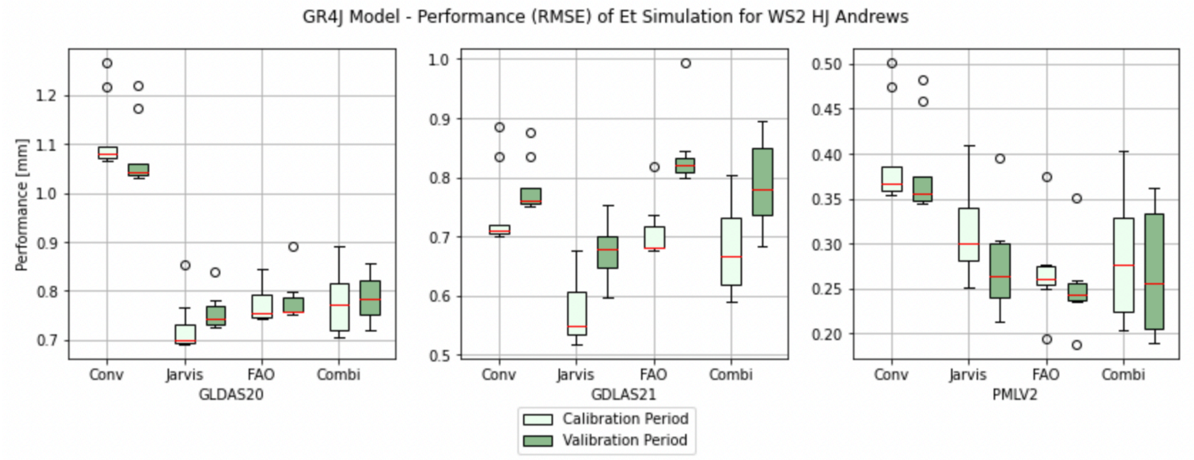


Figure 4.32: GR4J Model - Performance of Transpiration Simulations for WS2 in HJ Andrews. Performances of all behavioral transpiration simulations by each of the structures of the GR4J model. The performances are quantified by the objective function of the root mean square error (RMSE). Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

The mean seasonal sums of the transpiration and interception evaporation simulations and estimations can be seen in figure 4.33. The most noticeable difference for the seasonal sums of transpiration is that during the winters (djf) this is relatively higher for the simulations of the conventional model structure compared to those of the other model structures and the transpiration estimations. For the seasonal sums of interception evaporation, it can be perceived that the interception evaporation simulations have lower values compared to the estimations for each season and that those of the conventional model structure come closest to the estimations compared to the other model structures.

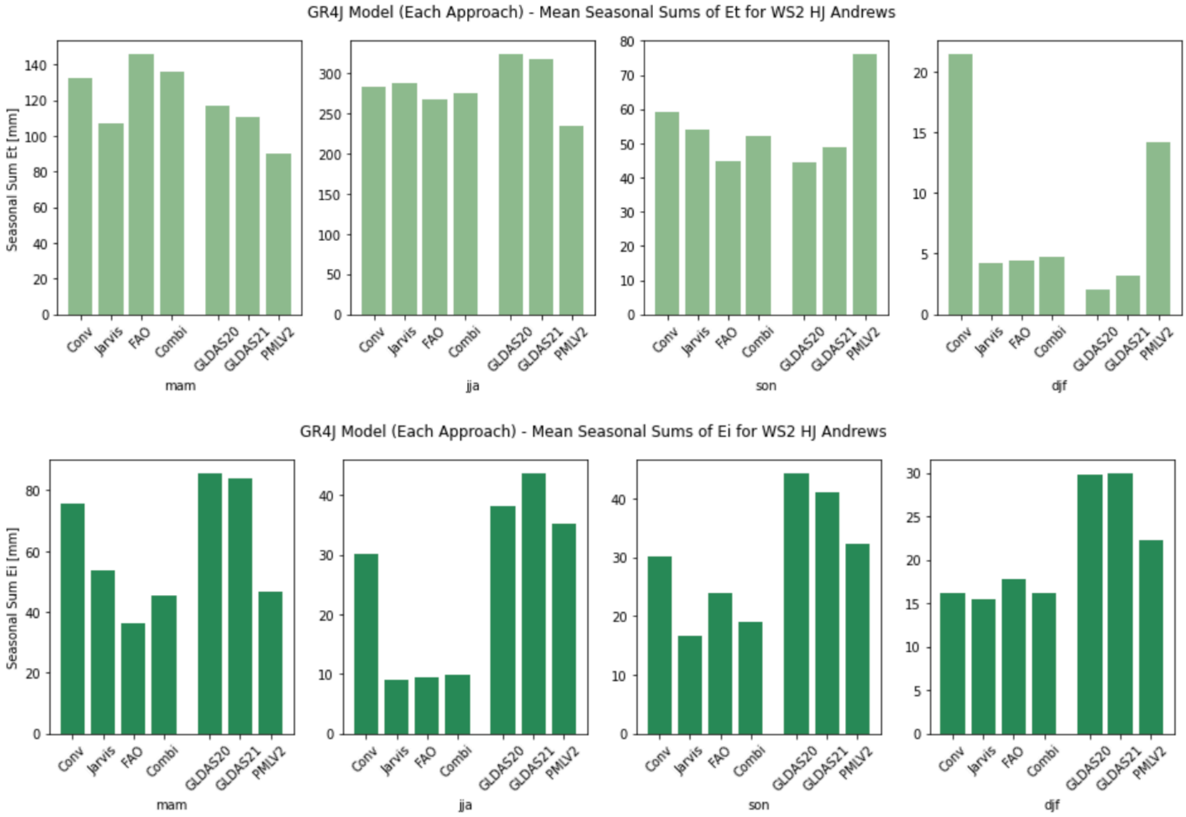


Figure 4.33: GR4J Model - Mean Seasonal Sums of Evaporation for WS2 in HJ Andrews. Mean sums of the transpiration and interception evaporation estimation and simulations by each of the four structures of the GR4J model for the seasons mam = March, April, and May, jja = June, July, and August, son = September, October, and November, djf = December, January, and February. Obs = observations, Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi = combined GR4J model.

### Root Zone Storage Capacity

None of the four GR4J model structures have median calibrated values of the maximum storage capacity of the production reservoir, which is the parameter $x_1$ in the GR4J model, that fit within the upper and lower boundaries of the Gumbel estimated root zone storage capacity at the return period of 20 years as visible in figure 4.34.

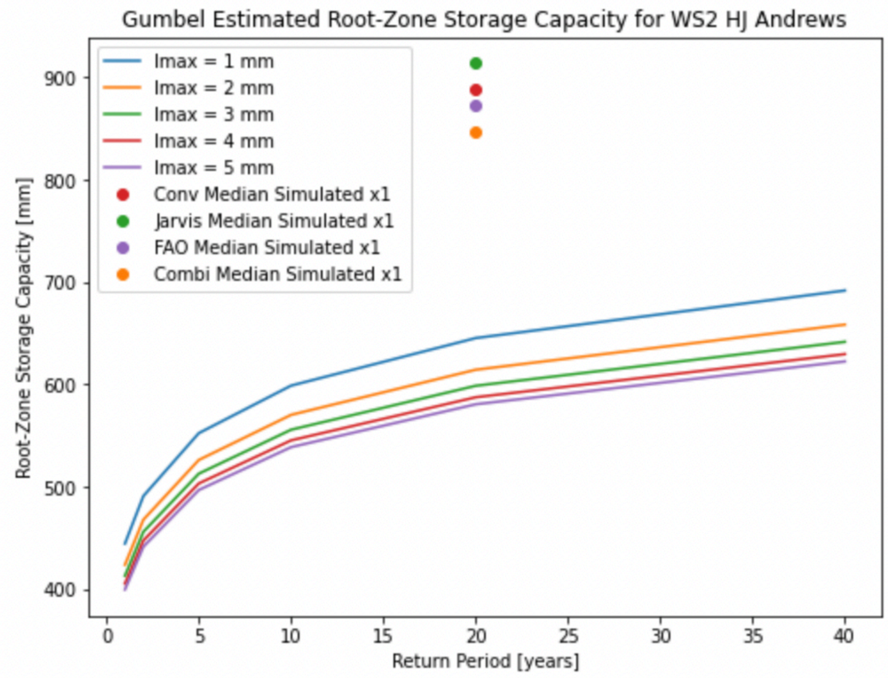


Figure 4.34: GR4J Model - Root Zone Storage Capacity Simulations for WS2 in HJ Andrews. Gumbel estimated root zone storage capacities calculated for multiple values of $I_{max}$ and for multiple return periods. The colored dots are the median simulated values of the parameter $x_1$ of each of the four GR4J model structures. Conv = conventional GR4J model, Jarvis = Jarvis-based GR4J model, FAO = FAO-based GR4J model and Combi =combined GR4J model.

# 5

# Discussion

The conventional FLEX and GR4J models only partition evaporation into two separate processes. For the FLEX model, this is in evaporation from the interception reservoir (interception evaporation) and evaporation from the unsaturated soil reservoir (transpiration). For the GR4J model, this is in evaporation from the interception reservoir (interception evaporation) and evaporation from the production reservoir (transpiration). Both models take daily precipitation and potential evaporation as system inputs. Both models use potential evaporation first for the quantification of evaporation from the interception reservoir and afterward, what remains, for evaporation from the unsaturated soil/production reservoirs. This particular order appears to be consistent with the terrestrial residence timescales of both evaporation processes. For example, interception evaporation usually occurs immediately after a precipitation event, whereas most transpiration occurs a few hours to days after the precipitation event (Wang-Erlandsson, van der Ent, Gordon, & Savenije, 2014). Interception evaporation is thus very momentary and limited by the supply from precipitation, whereas transpiration is longer lasting and can sustain dry seasons through the ability of vegetation to obtain retained soil water.

Both models, however, neglect the importance of vegetation dynamics on the partitioning of evaporation. For example, during the dormant seasons, winters, deciduous trees are leafless, and therefore evaporation from the forest floor and soil would dominate and transpiration would approximate to zero. Whereas during the peak seasons, summers, deciduous trees have fully developed leaves and therefore transpiration would dominate. This means that at the seasonal scale, vegetation is an important factor that influences the ratio of transpiration over interception and soil evaporation. This is in contrast with the conventional FLEX and GR4J models. Both models continuously prioritize interception evaporation and do not include information from the dynamic behavior of vegetation.

## 5.1. Hubbard Brook

### 5.1.1. Streamflow

The daily streamflow simulations and observations for Hubbard Brooks' watershed 3 (WS3) are very similar to each other except for the peak flows. This might be due to reasons such as the heterogeneity in a catchment, the uncertainty in the observations, or the structure of the model. Heterogeneity can be noticed anywhere within a catchment and at all scales. For example, within the structure of the soil, the distribution of the vegetation, and the elevation of the land might be completely different in one location in a catchment compared to another random location in the same catchment. Heterogeneity not only affects the physical properties of a catchment, but also the processes that generate the streamflow within a catchment. One such crucial process is the precipitation, which is highly subject to temporal and spatial heterogeneity and might be the reason for the inaccurate simulations of the peak flows in WS3. Part of the precipitation that generated these peak flows might not have been observed by the meteorological station in WS3 and therefore might have resulted in less accurate peaks in the streamflow simulations.

The daily streamflow simulations by each of the four FLEX model structures do not show many differences to each other, despite that each model structure used a different approach to partition the potential evaporation and that evaporation is highly important for the generation of streamflow. The reason for such minimal differences might be explained by the climate of WS3 and the determination of the runoff coefficient ($C_r$) in the models. Figure 5.1 contains three graphs showing the simulations of daily potential evaporation ($E_p$), daily transpiration ($E_t$) by the conventional and the FAO-based FLEX models, and the simulation of daily $C_r$ by the same models.

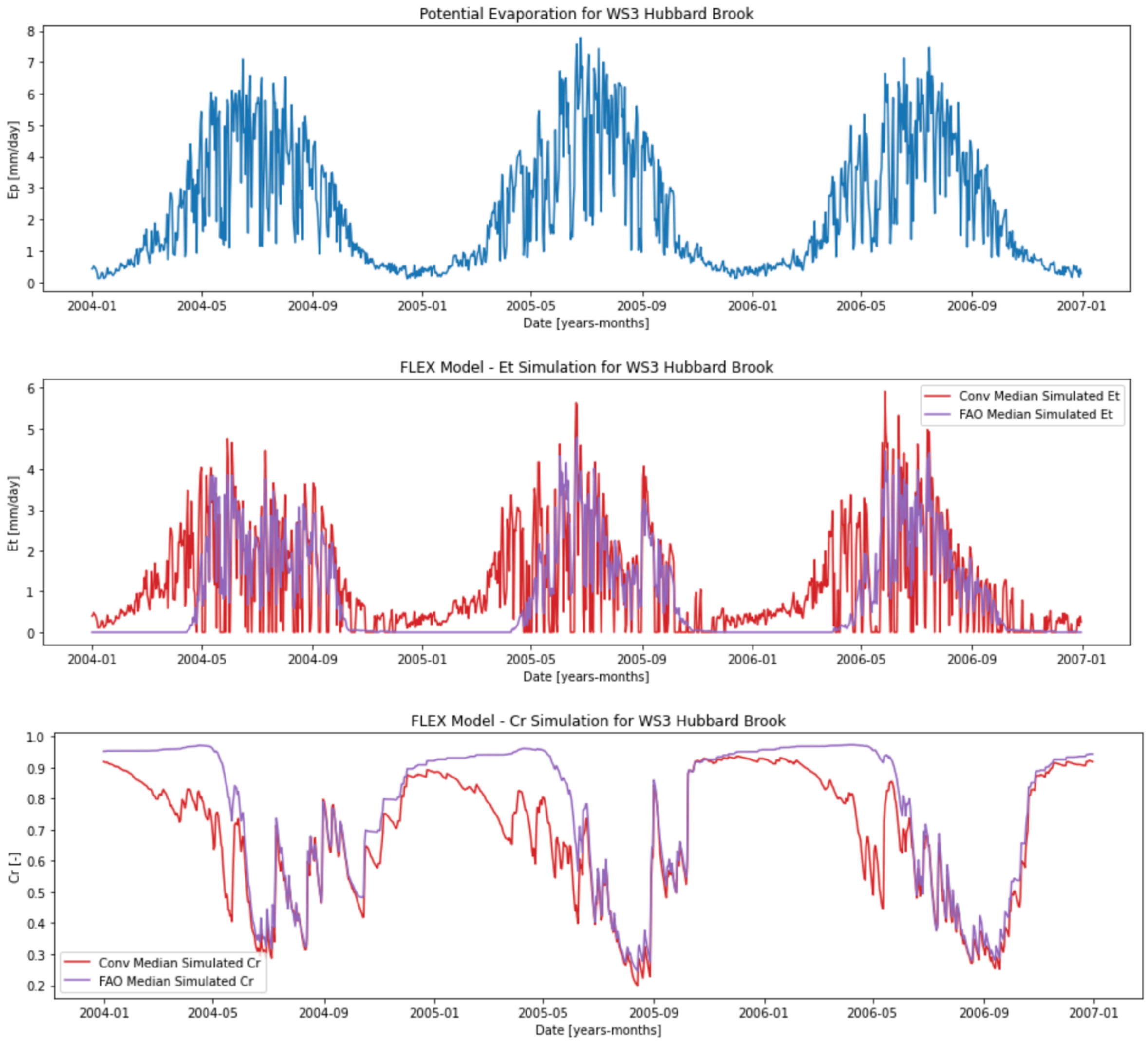


Figure 5.1: FLEX Model - Potential Evaporation ($E_p$), Transpiration ($E_t$), and Runoff Coefficient ($C_r$) Simulations for WS3 in Hubbard Brook. Conv = conventional FLEX model and FAO = FAO-based FLEX model.

From figure 5.1 can be perceived that during the winter and the start of spring, the $E_t$ simulations by the FAO-based model is zero or close to zero. Whereas the $E_t$ simulations by the conventional model have relatively higher values during these months. This difference is also visible in the $C_r$ simulations by both structures of the model. The FLEX model uses equation 3.4 to estimate $C_r$, which is a S-shaped function of the amount of water stored in the unsaturated soil reservoir ($S_u$) and the parameters of the maximum storage capacity of the unsaturated soil reservoir ($S_{u,max}$) and $\beta$. An example of this S-shaped function is visible in figure 3.2. This S-shaped function shows that relatively higher values of $S_u$ result into higher values of $C_r$ and therefore also higher values for the simulation of surface runoff in the model. Relatively lower values of $S_u$ result into lower values of $C_r$ and therefore lower values for the simulation of surface runoff. Hence during the winter and the start of spring in WS3, the higher values of the $E_t$ simulations by the conventional model, and thus less $S_u$, resulted into lower values of $C_r$ compared to the simulations by the FAO-based model. The FAO-based model included information from plant phenology through the use of LAI, which is minimal during the winter and the start of spring for WS3, and hence simulated low values or no $E_t$ at all for these months. The expectation for lower values of $E_t$ and higher values of $C_r$ by the FAO-based model simulations would be higher values of the streamflow simulations compared to the conventional model simulations. This is slightly visible in the mean seasonal sums of the streamflow simulations in figure 4.4 for the spring months, March, April, and May. During these months, the mean seasonal sum of the streamflow simulations by the FAO-based model is slightly higher than the conventional model and closer to the observed values. However, for the winter months, December, January, and February, in which the LAI of WS3 is also minimal, no significant difference can be noticed in the simulations of $E_t$ and $C_r$ by both the conventional, which does not use LAI, and FAO-based models. Hence, also no significant difference is present between the mean seasonal sums of the streamflow simulations by both structures of the model for the winter months. This is due to the very low values of the potential evaporation for WS3 during these months, which also forces the $E_t$ simulations by the conventional model to be minimal.

Considering the GR4J model for WS3 in Hubbard Brook, similar trends can be perceived for the simulations between the conventional model and the three modified models. The conventional GR4J model simulated higher $E_t$ values during the winter and the start of spring, and thus lower values of the amount of water stored in the production reservoir ($S_s$), compared to $E_t$ simulations by the modified models. However, unlike the FLEX model simulations for WS3, not many differences can be perceived in mean seasonal sums of the streamflow simulations between each of the four GR4J models (fig. 4.12). This is mostly due to the calibration of the parameter of $x_2$, which determines the addition or removal of inter-catchment groundwater flow to the total simulated streamflow. The median value of $x_2$ is 1.41 for the conventional GR4J model, indicating the addition of inter-catchment groundwater flow, and is -0.78 for the FAO-based GR4J model, which indicates the removal of inter-catchment groundwater flow. Hence the impact of the slightly higher simulations of $E_t$, and thus lower $S_s$, by the conventional model on the simulation of streamflow is compensated by the inter-catchment groundwater flow and explains the minimal differences in the mean seasonal sums of the streamflow simulations.

### 5.1.2. Evaporation

The transpiration simulations of the three modified structures by both the FLEX model and the GR4J model for WS3 in Hubbard Brook are highly similar to the $E_t$ estimations by the land surface models, integrated with remote sensing data, of GLDAS-20, GLDAS-21, and PML-V2. The transpiration simulations of the conventional structure by both models are less similar to the $E_t$ estimations. The transpiration simulations of the FAO-based structure by both models perform best according to the $E_t$ estimations and are then followed by the simulations of the combined structure and the Jarvis-based structure. The transpiration simulations of the conventional structure by both models perform worst according to the estimations. This is due to that the conventional structure does not take into account any information from plant phenology to represent the dynamics of vegetation in the catchment. Whereas GLDAS-20, GLDAS-21 and PML-V2 each do include the spatial and temporal dynamics of vegetation through the use of for example LAI. The conventional structure of both models thus considers the same amount of vegetation with leaves throughout the entire simulation periods. Hence, the conventional structure considers the presence of leaves during the winter and the possibility of $E_t$. However, the dominant land cover in Hubbard Brook's WS3 is a deciduous broadleaf forest, which has a strong seasonal pattern in its development of leaves as visible in the LAI in figure 4.1. The modified structures do take into account the dynamics of vegetation and are therefore able to simulate $E_t$ that is more consistent with the LAI of WS3.

The transpiration simulations of the three modified structures by both the FLEX and GR4J models do not deviate much from each other. However, overall it can be perceived that out of the three modified structures the $E_t$ simulations of the Jarvis-based structure have the highest values and those of the FAO-based structure have the lowest and most similar values to the $E_t$ estimations of the remote sensing integrated land surface models. The reason that the Jarvis-based structure simulates the higher $E_t$ values is due to the method it uses to partition the potential evaporation. The Jarvis-based structure partitions the potential evaporation according to a value between 0 and 1 based only on the stomatal response to air temperature. For optimal air temperatures, the Jarvis coefficient approaches 1, and therefore most of the potential evaporation is then used for $E_t$, which might have resulted in too high $E_t$ values, especially during rainy days. The FAO-based structure rather uses separate values to partition the potential evaporation according to the development stage or season of the vegetation in the catchment using information from the LAI of the catchment. The remote sensing integrated land surface models of GLDAS-20, GLDAS-21, and PML-V2 also use LAI to estimate $E_t$ and might explain why the $E_t$ simulations by the FAO-based structure are most similar to the $E_t$ estimations.

The interception evaporation ($E_i$) simulations of the three modified structures of both the FLEX and GR4J models do not deviate much from each other, whereas the $E_i$ simulations of the conventional structure generally consist of higher values. The reason for the higher $E_i$ values of the conventional structure is that apart from the potential evaporation and the supply of water there is no vegetation-related constraint on the quantification of this process, whereas the other three structures do have such a constraint. Furthermore, none of the $E_i$ simulations of the four structures by both models are similar to the $E_i$ estimations by the remote sensing integrated land surface models. This is due to the different methods used between the four structures by both models and the remote sensing integrated land surface models to determine $E_i$.

## 5.2. HJ Andrews

### 5.2.1. Streamflow

The daily streamflow simulations and observations for HJ Andrews' watershed 2 (WS2) are very similar to each other except for the period around February and March in the year 2008. The reason for this might be the same as described for the inaccuracy of the peak flow simulations for Hubbard Brook's watershed 3 (WS3). The heterogeneity in the catchment and in the precipitation observations might have caused the inaccuracy of the streamflow simulations during the period around February and March in 2008. Additionally, uncertainty in the streamflow observations might also have caused the dissimilarity between the simulations and observations.

There are almost no differences in the performances of the daily streamflow simulations of each of the four structures by both the FLEX model and the GR4J model according to the daily streamflow observations for WS2. Similarly, there are not many differences in the mean seasonal sums of streamflow simulation between each of the four structures by both the models. This is due to the minimal differences in the transpiration ($E_t$) simulations, and hence minimal differences in the surface runoff, between each of the four structures for WS2.

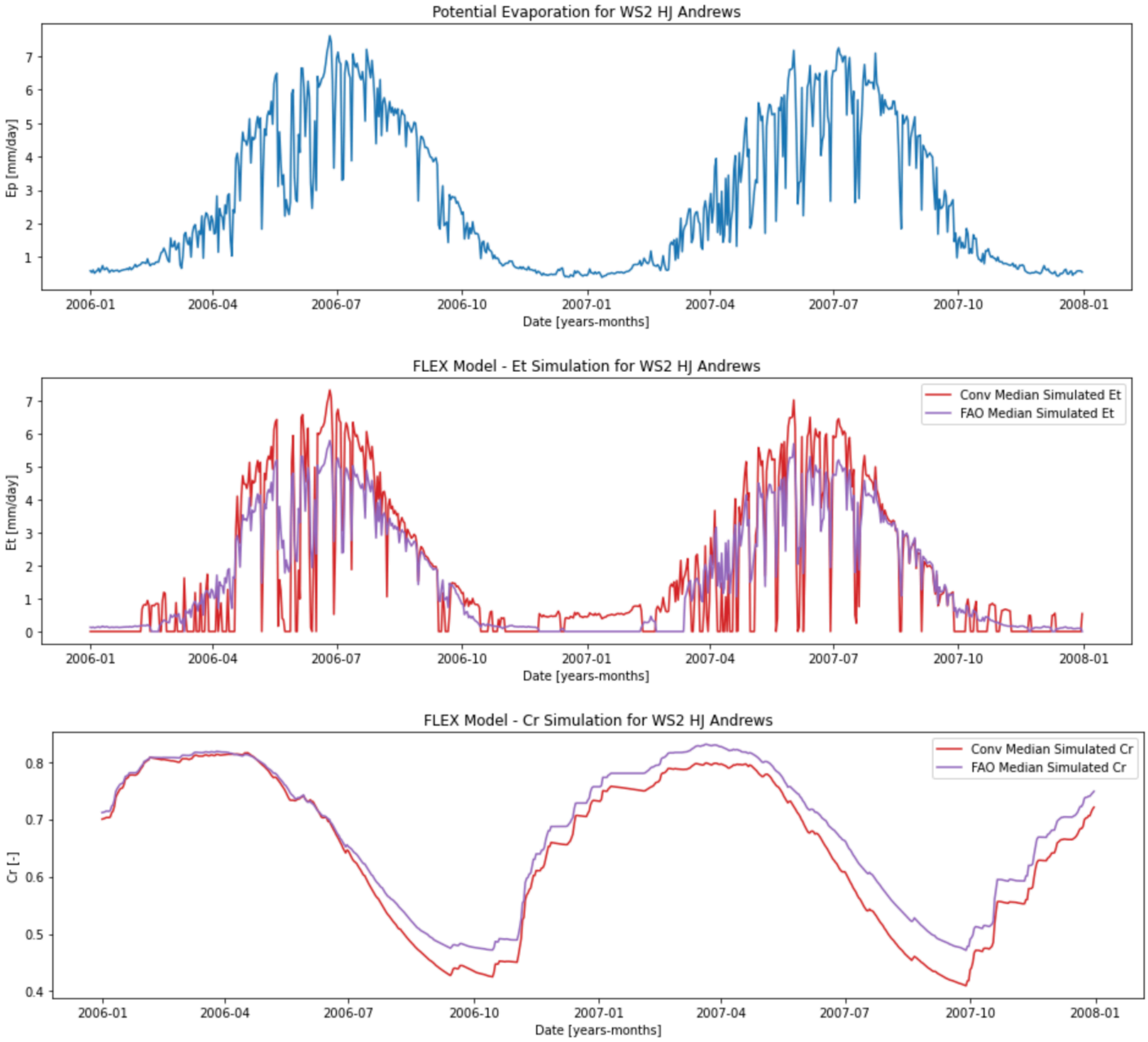


Figure 5.2: FLEX Model - Potential Evaporation ($E_p$), Transpiration ($E_t$), and Runoff Coefficient ($C_r$) Simulations for WS2 in HJ Andrews. Conv = conventional FLEX model and FAO = FAO-based FLEX model.

### 5.2.2. Evaporation

The transpiration simulations of each of the four structures by both the FLEX and GR4J models are similar to each other and to the $E_t$ estimations of the remote sensing integrated land surface models of GLDAS-20, GLDAS-21, and PML-V2. Overall, the $E_t$ simulations by the conventional structure tend to have the highest values, followed by the $E_t$ simulations of the Jarvis-based structure and the combined structure. The transpiration simulations by the FAO-based structure tend to have the lowest and the most similar to the $E_t$ estimations.

The transpiration simulations of the conventional structure for WS2 are relatively higher during the winters, just like for WS3, which is the result of the conventional models considering the same amount of vegetation with leaves throughout the entire simulation period. However, overall the differences between the $E_t$ simulations of each of the four structures by both the models and the $E_t$ estimations are less for HJ Andrews' WS2 than for Hubbard Brook's WS3. This is due to the dominant land cover in WS2, which is the evergreen needleleaf forest. This vegetation type has leaves throughout the entire year and for WS2 has less variability in its LAI, as visible in figure 4.18, compared to the LAI of WS3. Hence, transpiration in WS2 should still be possible during the winters as opposed to in WS3. The modified models used parameters to consider the dynamic behavior of the vegetation in WS2 and obtained values for these parameters through calibration. This resulted into $E_t$ simulations similar to the $E_t$ simulations of the conventional models.

The interception evaporation ($E_i$) simulations of the FAO-based and combined structures by both the FLEX model and the GR4J model do not deviate much from each other. The interception evaporation simulations of the conventional and the Jarvis-based structures by both models slightly differ from the $E_i$ simulations of the other two structures. This is due to the same reasons as described for WS3. Similar to the $E_i$ simulations for WS3, the simulations for WS2 do not show many similarities to the estimations of the remote sensing integrated land surface models. Again, this is due to the different methods used between the four structures by both models and the remote sensing integrated land surface models to determine $E_i$.

# 6

# Conclusion and Recommendations

The objective of this study was to include information from plant phenology to partition evaporation in conceptual hydrological models and this study hypothesized that such models would simulate more reliable hydrological responses compared to those that do not include plant phenology. This study included information from plant phenology in the FLEX model and the GR4J model in three different structures. The first modified structure uses a part of the Jarvis model, in which stomata respond to temperature, to partition evaporation. The second modified structure uses a method similar to the crop evaporation method of the Food and Agriculture Organization (Allen et al., 1998). Lastly, the third modified structure uses a combination of the other two structures.

Overall, the daily streamflow simulations by each of the four structures of both the FLEX model and the GR4J model perform equally well for both the selected catchments in the experimental forests of Hubbard Brook and HJ Andrews. The three modified structures of the FLEX model and GR4J model are thus able to include information from plant phenology and still provide daily streamflow simulations that perform as well as the streamflow simulations by the conventional structures of the same models. A few differences, however, can be noticed for the mean seasonal sums of the streamflow of each of the four FLEX model structures for Hubbard Brook’s watershed 3 (WS3).

The dominant land cover in Hubbard Brook’s WS3 consists of a deciduous broadleaf forest, which has a highly seasonal pattern in the development of its leaves. The vegetation in WS3 has fully developed leaves during the summers and almost no leaves during the winters as visible in the LAI in figure 4.1. As transpiration is mostly dependent on the leaves of the vegetation in a catchment, it would be expected for WS3 to have peaked transpiration values during the summers and almost no transpiration during the winters. The three modified structures of both the FLEX model and the GR4J model are able to simulate this pattern for the transpiration in WS3 and are able to simulate transpiration values that are consistent with the transpiration estimations of the land surface models of GLDAS-20, GLDAS-21, and PML-V2. The conventional structure of both the FLEX model and the GR4J model is not able to simulate this same pattern as its transpiration simulations are too high during the start of spring compared to the transpiration simulations of the three modified model structures and the transpiration estimations. During the winter, in which the LAI of WS3 is minimal, the differences between the transpiration simulations for WS3 of the conventional model structures and of the three modified model structures are not as apparent as during the start of spring. This is due to the potential evaporation of WS3, which forces the transpiration simulations of the conventional model structures also to be minimal during the winters.

The differences between the transpiration simulations for WS3 of the conventional FLEX model and of the modified FLEX models for the start of spring resulted in slight differences between the mean seasonal sums of the streamflow simulations of the conventional FLEX model and of the modified FLEX models for the spring months, March, April, and May. For these months, the mean seasonal sums of the streamflow simulations of the modified FLEX models are compared to the conventional FLEX model more similar to the mean seasonal sums of the streamflow observations. However, for the summer months, June, July, and August, the mean seasonal sums of the streamflow simulations of the conventional FLEX model are slightly more similar to the sums of the streamflow observations. The modified FLEX models are able to provide slightly more accurate streamflow simulations during the spring months for WS3 in Hubbard Brook compared to the conventional FLEX model. However, this is the opposite for the summer months, which might be the result of the uncertainties in the transpiration simulations and especially the interception evaporation simulations as these are not consistent with other interception evaporation estimations.

The differences between the transpiration simulations for WS3 of the conventional GR4J model and of the modified GR4J models resulted in almost no differences between the mean seasonal sums of the streamflow simulations of the conventional GR4J model and of the modified FLEX models for each season. This is unlike the simulations of the FLEX model for WS3 and is due to the calibration of the GR4J model parameter $x_2$, which depending on its value either adds or removes water as inter-catchment groundwater flow to the simulation of streamflow. Furthermore, there are almost no differences in the simulations between each of the four structures of both the FLEX model and GR4J models for HJ Andrews' watershed 2 (WS2) for the daily values of streamflow, transpiration, and the mean seasonal sums of streamflow. This is due to the characteristics of the dominant land cover in WS2, which is an evergreen needleleaf forest. This vegetation type has leaves throughout the year and has a less variable pattern in the development of the leaves compared to the deciduous broadleaf forest. The transpiration simulations of the modified structures for WS2 use information of the catchment's climate and its less variable pattern of the vegetation development. Hence, the transpiration simulations of the modified structures for WS2 are relatively similar to the transpiration simulations of the conventional structure, which does not include the development of vegetation at all.

This study concludes that conceptual hydrological models that include information from plant phenology as in the three modified model structures, are able to provide streamflow simulations as well as models that do not include plant phenology, and not necessarily more reliable. However, conceptual hydrological models that include information from plant phenology might still be beneficial to catchments with vegetation dynamics subject to the effects of among others climate change, land-use change, and forest fires. Before plant phenology should be used in conceptual hydrological models, such as by the modified model structures, challenges such as the uncertainty in the simulations of the evaporation components need to be addressed for future studies. Additionally, the modeler needs to consider the effects of the climate and the characteristics of the catchment before deciding to use the information from plant phenology. If for example, the combination of the precipitation and potential evaporation in a catchment would force the transpiration simulations of a model, without plant phenology, to be similar to the transpiration simulations of a model, including plant phenology, the latter model could be considered unnecessary depending on the modeler's objective. Furthermore, the modeler should consider the effects of the free parameters on the simulations of streamflow, as the calibration of such parameters could compensate for the less accurate streamflow simulations.

## 6.1. Recommendations

This study recommends future studies to investigate the performance of streamflow simulations of conceptual hydrological models that use the FAO-based model structure for catchments that are subject to significant land cover changes. Examples of such significant land cover changes are afforestation, deforestation, forest fires, and urbanization. For this, it is interesting to have a catchment with significant land cover changes during periods with high potential evaporation, such as during summers, to be able to observe significant changes in the partitioning of evaporation. The reason for the recommendation of only the FAO-based structure is that this structure includes the temporal and spatial distribution of the vegetation in a catchment, which is influenced by long-term changes in light, precipitation, and temperature. The Jarvis-based model structure, instead of the FAO-based model structure, is only interesting for significant changes in temperature during short periods.

Furthermore, this study recommends future studies to investigate the dynamic ability of vegetation to provide an additional medium from which interception evaporation can occur. This study focused on the dynamic interaction between the ability of vegetation to transpire and the partitioning of evaporation. For this reason, it is recommended to additionally focus on the dynamic interaction between the growth and abscission of leaves, the growth and decomposition of leaf litter, and the partitioning of evaporation.

# References

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). *Crop evapotranspiration: guidelines for computing crop water requirements* (Vol. 56). Food and Agriculture Organization of the United Nations.

Arora, V. (2002). Modeling vegetation as a dynamic component in soil-vegetation-atmosphere transfer schemes and hydrological models. *Rev. Geophys.*, *40*(2), 1006. doi: 10.1029/2001RG000103

Bergström, S. (1992). The hbv model – its structure and applications. *SMHI Reports Hydrology*.

Beven, K. J. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*(1), 41-51. doi: https://doi.org/10.1016/0309-1708(93)90028-E.

Beven, K. J. (2000). Uniqueness of place and process representations in hydrological modelling. *Hydrol. Earth Syst. Sci.*, *4*, 203-213. doi: https://doi.org/10.5194/hess-4-203-2000

Beven, K. J. (2001a). How far can we go in distributed hydrological modelling? *Hydrol. Earth Syst. Sci.*, *5*(1), 1-12.

Beven, K. J. (2001b). Rainfall-runoff modelling: The primer. *John Wiley and Sons, Ltd*.

Blöschl, G. (2001). Scaling in hydrology. *Hydrol. Process.*, *15*, 709-711. doi: 10.1002/hyp.432

Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the penn state–ncar mm5 modeling system. part i: Model implementation and sensitivity. *Monthly Weather Review*, *129*, 569-585. doi: 10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H., Koren, V., … Betts, A. (1996). Modeling of land-surface evaporation by four schemes and comparison with fife observations. *Journal of Geophysical Research. D. Atmospheres*, *101*, 7251-7268. doi: 10.1029/95JD02165

Coenders-Gerrits, A., van der Ent, R., Bogaard, T., & et al. (2014). Uncertainties in transpiration estimates. *Nature 506*, *E1-E2*. doi: https://doi.org/10.1038/nature12925

Deardorff, J. W. (1979). Prediction of convective mixed-layer entrainment for realistic capping inversion structure. *J. Atmos. Sci.*, *36*, 424-436.

De Bruin, H. A. R. (1989). Physical aspects of the planetary boundary layer with special reference to regional evapotranspiration. *Estimation of Areal Evapotranspiration, IAHS*, *177*.

Donohue, R. J., Roderick, M. L., & McVicar, T. R. (2007). On the importance of including vegetation dynamics in budyko's hydrological model. *Hydrol. Earth Syst. Sci.*, *11*, 983-995. doi: https://doi.org/10.5194/hess-11-983-2007

Dooge, J. C. I. (1986). Looking for hydrologic laws. *Water Resour. Res.*, *22*(9), 46S-58S.

Duursma, R. A., Barton, C. V. M., Lin, Y., Medlyn, B. E., Eamus, D., Tissue, D. T., … McMurtrie, R. E. (2014). The peaked response of transpiration rate to vapour pressure deficit in field conditions can be explained by the temperature optimum of photosynthesis. *Agricultural and Forest Meteorology*, *189-190*, 2-10. doi: https://doi.org/10.1016/j.agrformet.2013.12.007

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., … Tarpley, J. D. (2003). Implementation of noah land surface model advances in the ncep operational mesoscale eta model. *Journal of Geophysical Research*, *108*. doi: 10.1029/2002JD003296

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of National Academy of Sciences*, *114*(40), 10572-10577. doi: 10.1073/pnas.1712381114

Fenicia, F., Savenije, H. H. G., Matgen, P., & Pfister, L. (2006). Is the groundwater reservoir linear? learning from data in hydrological modelling. *Hydrol. Earth Syst. Sci.*, *10*, 139-150. doi: https://doi.org/10.5194/hess-10-139-2006

Forest, H. A. E. (2021). *Experimental watersheds and gauging stations.* https://andrewsforest.oregonstate.edu/research/infrastructure/watersheds.

Gan, R., Zhang, Y. Q., Shi, H., Yang, Y. T., Eamus, D., Cheng, L., … Yu, Q. (2018). Use of satellite leaf area index estimating evapotranspiration and gross assimilation for australian ecosystems. *Ecohydrology*. doi: https://doi.org/10.1002/eco.1974

Gao, H., Hrachowitz, M., Fenicia, F., Gharari, S., & Savenije, H. (2014). Testing the realism of a topography-driven model (flex-topo) in the nested catchments of the upper heihe, china. *Hydrology and Earth System Sciences*, *18*, 1895-1915. doi: 10.5194/hess-18-1895-2014

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., & Savenije, H. H. G. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophys. Res. Lett.*, *41*, 7916-7923. doi: 10.1002/2014GL061668

Hrachowitz, M., & Clark, M. P. (2017). Hess opinions: The complementary merits of competing modelling philosophies in hydrology. *Hydrol. Earth Syst. Sci.*, *21*, 3953–3973. doi: https://doi.org/10.5194/hess-21-3953-2017

Jarvis, P., & Mcnaughton, K. (1986). Stomatal control of transpiration: Scaling up from leaf to region. *Advances in Ecological Research*, *15*, 1-49.

Karterakis, S. M., Karney, B. W., Singh, B., & Guergachi, A. (2007). The hydrologic cycle: a complex history with continuing pedagogical implications. *Water Supply*, *7*(1), 23-31. doi: https://doi.org/10.2166/ws.2007.003

Kool, D., Agam, N., Lazarovitch, N., Heitman, J. L., Sauer, T. J., & Ben-Gal, A. (2014). Review of approaches for evapotranspiration partitioning. *Agricultural and Forest Meteorology*, *184*, 56-70. doi: https://doi.org/10.1016/j.agrformet.2013.09.003

Latham, J., Cumani, R., Rosati, I., & Bloise, M. (2014). *Global land cover share (glc-share) database beta-release version 1.0.* http://www.fao.org/uploads/media/glc-share-doc.pdf.

Mahrt, L., & Ek, M. (1984). The influence of atmospheric stability on potential evaporation. *J. Climate Appl. Meteor.*, *23*, 222-234.

Manske, L. L. (n.d.). Range plant growth and development are affected by climatic factors. *North Dakota State University, Dickinson Research Extension Center*.

McMichael, B. L., & Quisenberry, J. E. (1993). The impact of the soil environment on the growth of root systems. *Environmental and Experimental Botany*, *33*(1), 53-61. doi: https://doi.org/10.1016/0098-8472(93)90055-K

McNaughton, K. G., & Jarvis, P. G. (1991). Effects of spatial scale on stomatal control of transpiration. *Agric. For. Meteorol.*, *54*, 279-301.

Monteith, J. L. (1965). Evaporation and environment. *Symposia of the Society for Experimental Biology*.

Morton, F. (1983). Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology*, *66*(1-4), 1-74. doi: https://doi.org/10.1016/0022-1694(83)90177-4

Myneni, R., Knyazikhin, Y., & Park, T. (2015). Mcd15a3h modis/terra+aqua leaf area index/fpar 4-day l4 global 500m sin grid v006 [data set]. *NASA EOSDIS Land Processes DAAC*. doi: https://doi.org/10.5067/MODIS/MCD15A3H.006

Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., … Hrachowitz, M. (2016). The evolution of root-zone moisture capacities after deforestation: A step towards hydrological predictions under change? *Hydrology and Earth System Sciences*, *20*(12), 4775-4799. doi: https://doi.org/10.5194/hess-20-4775-2016

Nord, E. A., & Lynch, J. P. (2009). Plant phenology: a critical controller of soil resource acquisition. *Journal of experimental botany*, *60*(7), 1927-1937. doi: https://doi.org/10.1093/jxb/erp018

Organization, W. H. (2019). *Drinking-water.* https://www.who.int/news-room/fact-sheets/detail/drinking-water.

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London*, *193*, 120-145.

Penman, H. L. (1956). Evaporation an introductory survey. *Netherlands Journal of Agricultural Science*, *4*, 9-29.

Perrin, C., Michel, C., & Andreassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, *279*, 275-289. doi: 10.1016/S0022-1694(03)00225-7

Pregitzer, K. S., King, J. S., Burton, A. J., & Brown, S. E. (2000). Responses of tree fine roots to temperature. *New Phytologist*, *147*, 105-115. doi: 10.1046/j.1469-8137.2000.00689.x

Priestly, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Weather Rev.*, *100*, 81-82.

Reig, P., Maddocks, A., & Gassert, F. (2013). *World's 36 most water-stressed countries.* https://www.wri.org/blog/2013/12/world-s-36-most-water-stressed-countries.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., … Toll, D. (2004). The global land data assimilation system. *Bull. Amer. Meteor. Soc.*, *85*(3), 381-394.

Rosbjerg, D., & Rodda, J. (2019). Iahs: a brief history of hydrology. *Hist. Geo Space. Sci.*, *10*, 109-118. doi: https://doi.org/10.5194/hgss-10-109-2019

Rui, H., & H., B. (2021). Readme document for nasa gldas version 2 data products. *NASA/GSFC, Greenbelt, MD, USA, NASA Goddard Earth Sciences Data and Information Services Center (GES DISC)*.

Sitterson, J., Knightes, C., Parmar, R., Wolfe, K., Muche, M., & Avant, B. (2017). An overview of rainfall-runoff model types. *U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-17/482*.

Sivapalan, M. (2003a). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrol. Process.*, *17*, 3163-3170. doi: 10.1002/hyp.5155

Sivapalan, M. (2006). Pattern, process and function: Elements of a unified theory of hydrology at the catchment scale. *In Encyclopedia of Hydrological Sciences (eds M.G. Anderson and J.J. McDonnell)*. doi: 10.1002/0470848944.hsa012

Sivapalan, M., & Blöschl, G. (2017). The growth of hydrological understanding: Technologies, ideas, and societal needs shape the field. *Water Resources Research*, *53*, 8137-8146. doi: https://doi.org/10.1002/2017WR021396

Station, P. N. R. (2003). *H.j. andrews experimental forest.* https://andrewsforest.oregonstate.edu/sites/default/files/lter/pubs/pdf/pub3654.pdf.

Stoy, P. C., El-Madany, T. S., Fisher, J. B., Gentine, P., Gerken, T., Good, S. P., … Wolf, S. (2019). Reviews and syntheses: Turning the challenges of partitioning ecosystem evaporation and transpiration into opportunities. *Biogeosciences*, *16*, 3747-3775. doi: https://doi.org/10.5194/bg-16-3747-2019

Study, H. B. E. (2021). *Online book: A synthesis of scientific research at hubbard brook.* https://hubbardbrook.org/online-book.

Thomas, K. A., Denny, E. G., Miller-Rushing, A. J., Crimmins, T. M., & F., W. J. (2010). The national phenology monitoring system v0.1. *USA-NPN Technical Series 2010-001*.

Towers, P. C., Strever, A., & Poblete-Echeverria, C. (2019). Comparison of vegetation indices for leaf area index estimation in vertical shoot positioned vine canopies with and without grenbiule hail-protection netting. *Remote Sensing*, *11*(9), 1073. doi: 10.3390/rs11091073

van der Sleen, P. (2021). *http://atdn.myspecies.info/file/109.* http://atdn.myspecies.info/file/109.

van Dijk, A., & Bruijnzeel, L. A. (2001). Modelling rainfall interception by vegetation of variable density using an adapted analytical model. part 1. model description. *J. Hydrol.*, *247*, 230-238. doi: 10.1016/s0022-1694(01)00392-4

Walter, H., & Breckle, S. W. (2002). Walter's vegetation of the earth: The ecological systems of the geo-biosphere. *Berlin: Springer*.

Wang-Erlandsson, L., van der Ent, R. J., Gordon, L. J., & Savenije, H. H. G. (2014). Contrasting roles of interception and transpiration in the hydrological cycle – part 1: Temporal characteristics over land. *Earth Syst. Dynam.*, *5*, 441-469. doi: https://doi.org/10.5194/esd-5-441-2014

Wyngaard, J. (1985). Structure of the planetary boundary layer and implications for its modeling. *Journal of Climate and Applied Meteorology*, *24*(11), 1131-1142.

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*. doi: https://doi.org/10.1155/2017/1353691

Zhang, L., Dawes, W. R., & Walker, G. R. (2001). Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water Resour. Res.*, *37*(3), 701-708. doi: 10.1029/2000WR900325

Zhang, Y., Kong, D., Gan, R., Chiew, F., McVicar, T. R., Zhang, Q., & Yang, Y. (2019). Coupled estimation of 500m and 8-day resolution global evapotranspiration and gross primary production in 2002-2017. *Remote Sens. Environ*, *222*, 165-182. doi: https://doi.org/10.1016/j.rse.2018.12.031

Zhang, Y., Peña-Arancibia, J. L., McVicar, T. R., Chiew, F. H. S., Vaze, J., Liu, C., … Pan, M. (2016). Multi-decadal trends in global terrestrial evapotranspiration and its components. *Sci. Rep.*, *6*. doi: https://doi.org/10.1038/srep19124

Zhao, L., Xia, J., Xu, C., & et al. (2013). Evapotranspiration estimation methods in hydrological models. *J. Geogr. Sci.*, *23*, 359-369. doi: https://doi.org/10.1007/s11442-013-1015-9

# A

# Transpiration Calibrated Models for WS3 in Hubbard Brook

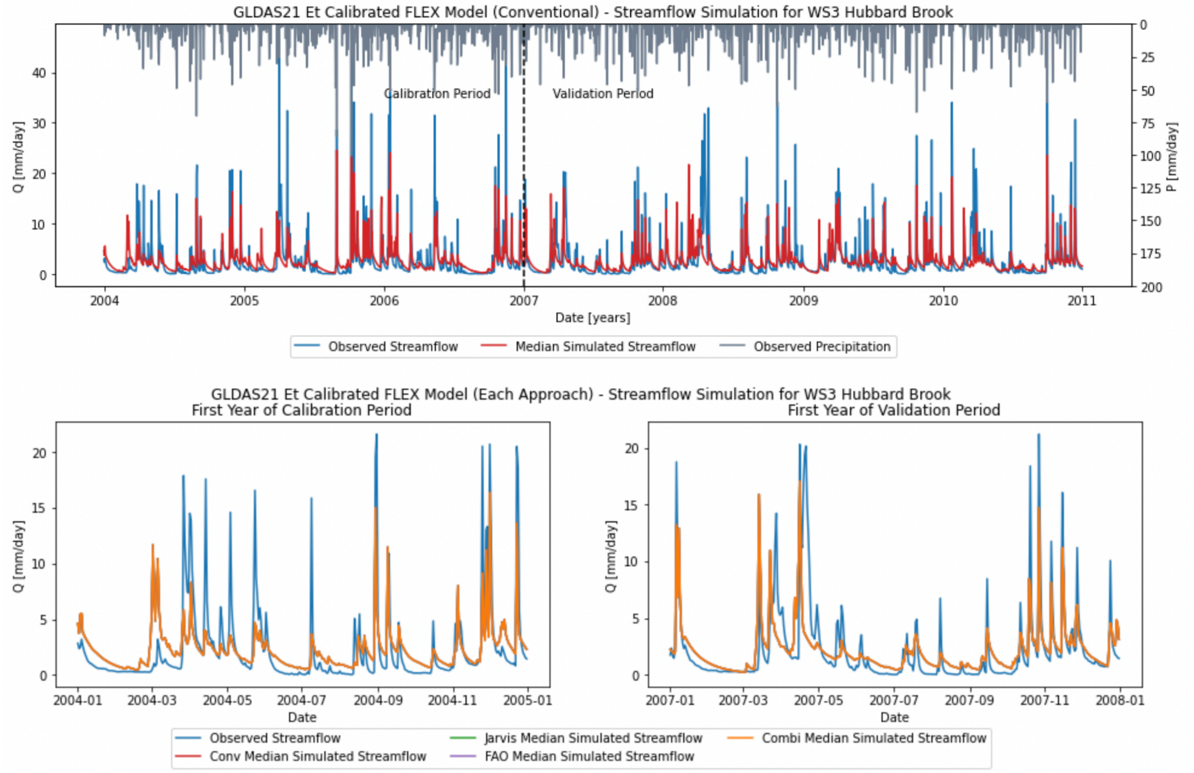


Figure A.1: FLEX Model - Streamflow Simulations for WS3 in Hubbard Brook.

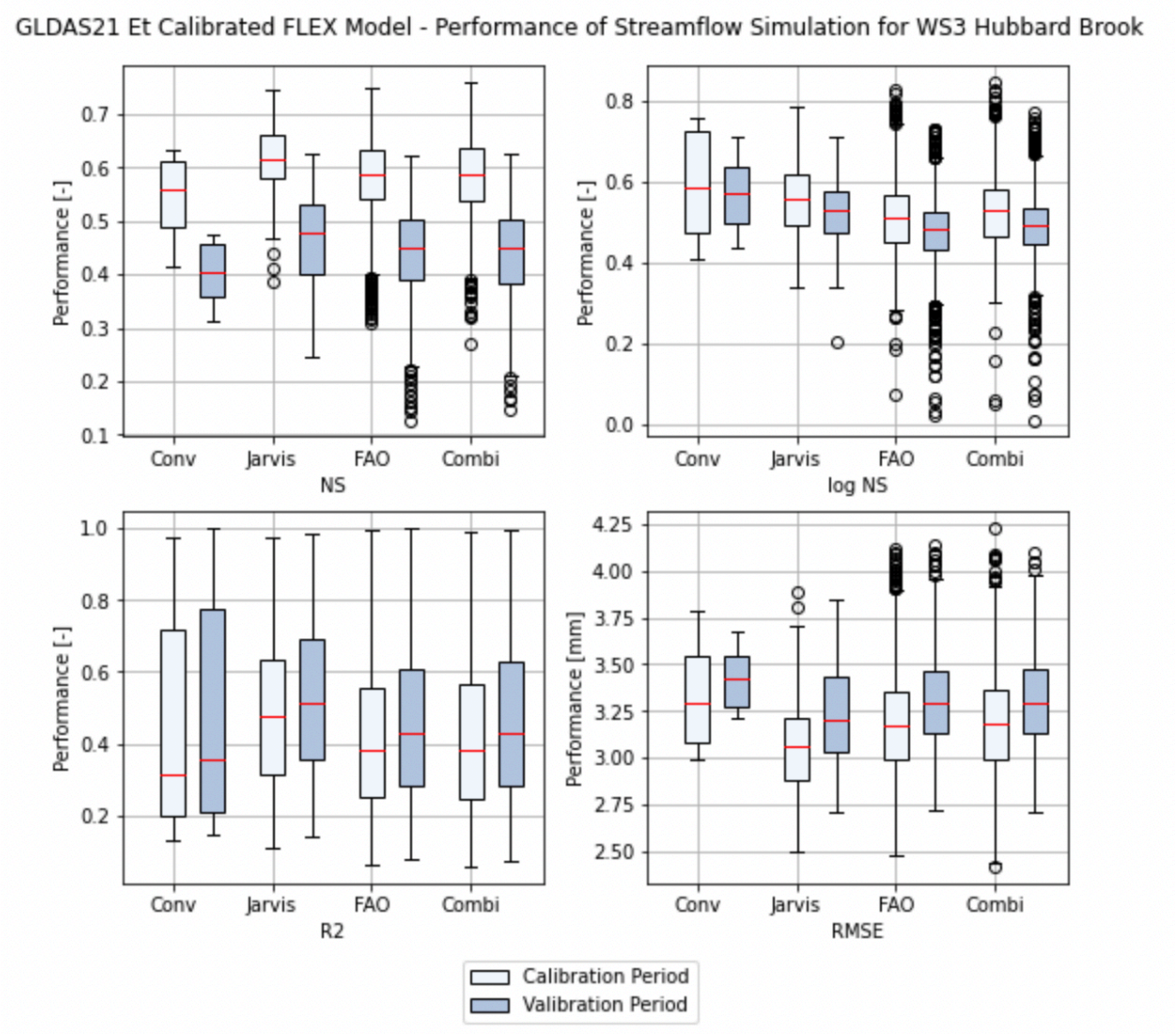


Figure A.2: FLEX Model - Performance of Streamflow Simulations for WS3 in Hubbard Brook.

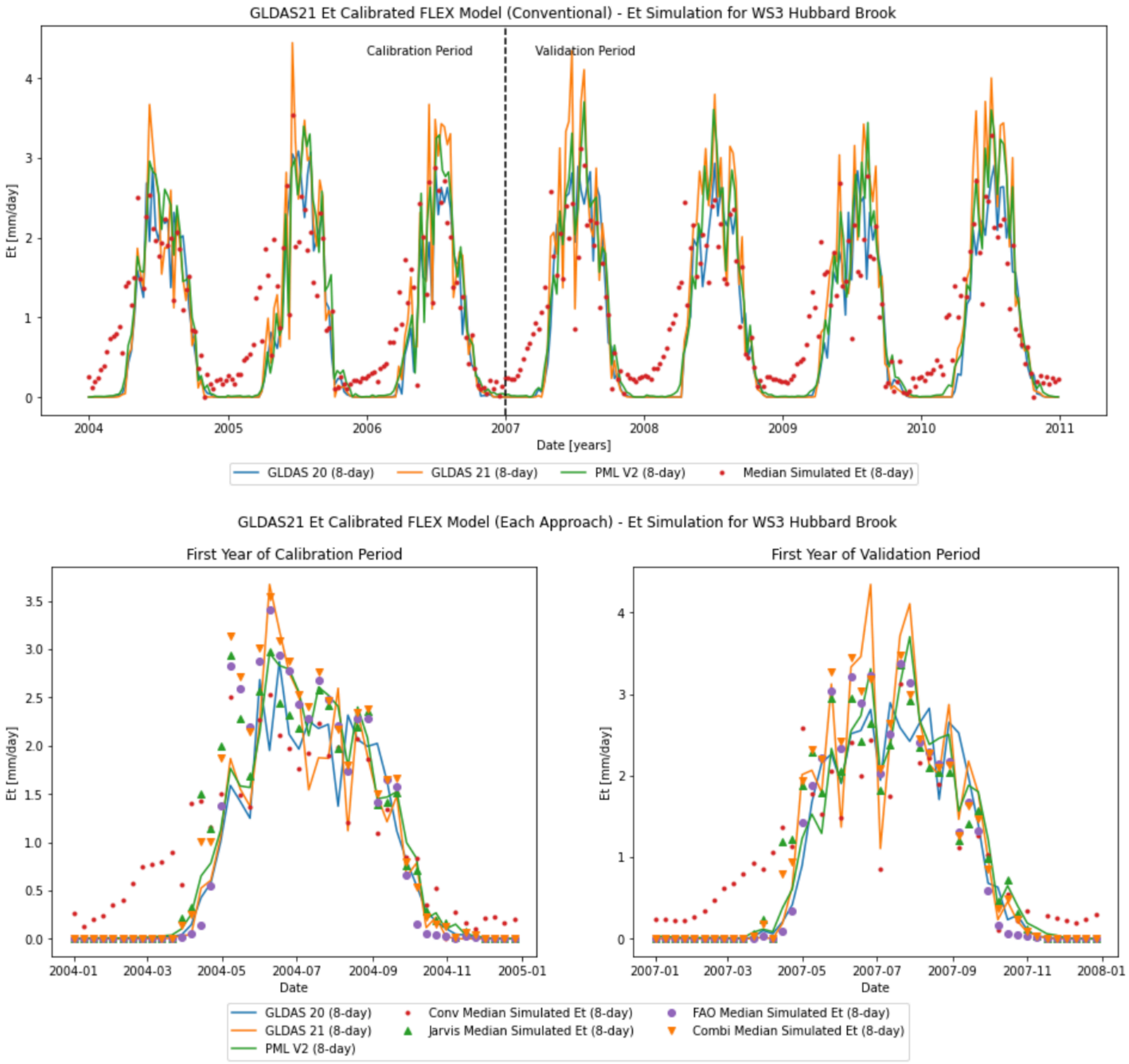


Figure A.3: FLEX Model - Transpiration Simulations for WS3 in Hubbard Brook.

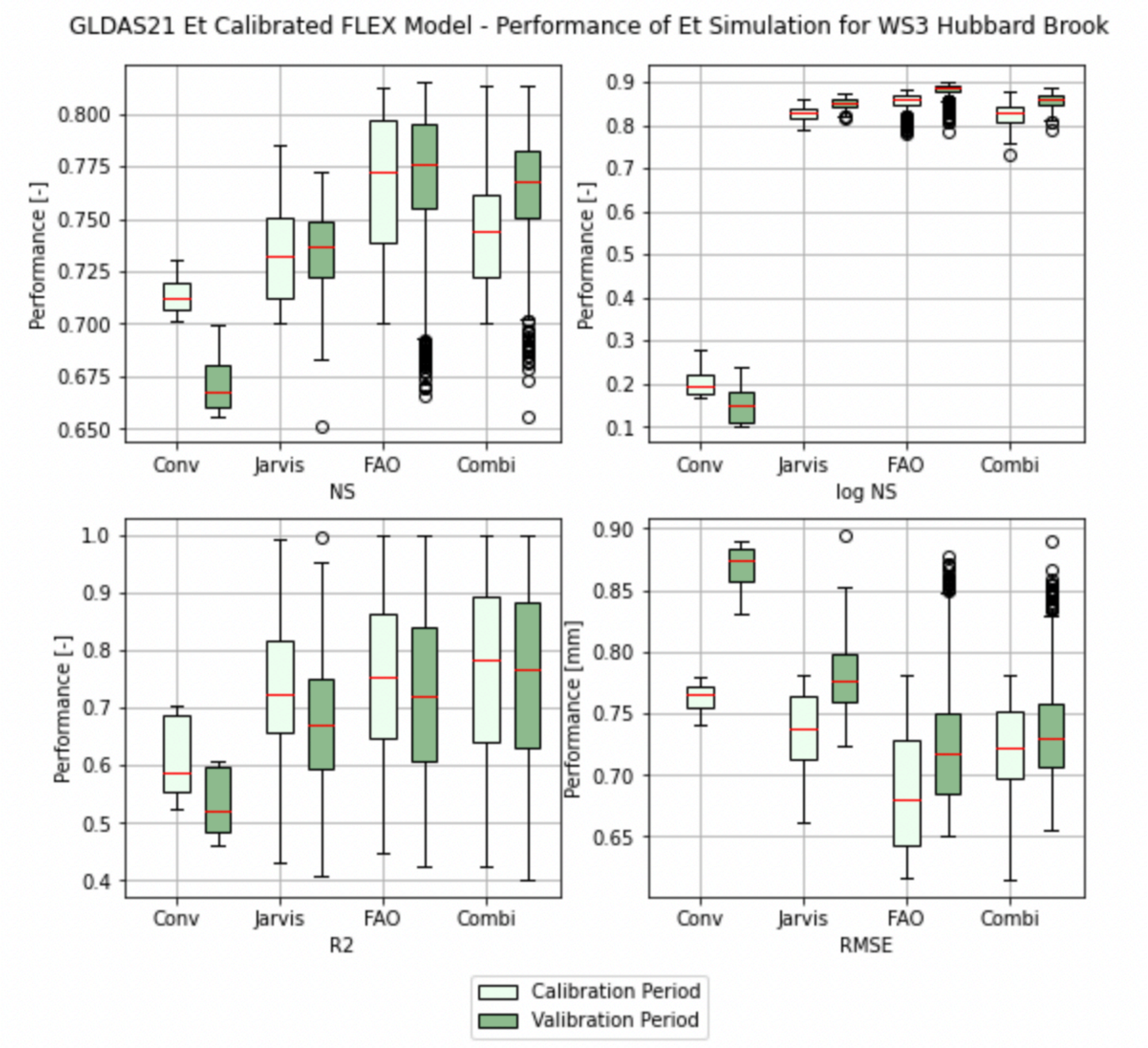


Figure A.4: FLEX Model - Performance of Transpiration Simulations for WS3 in Hubbard Brook.

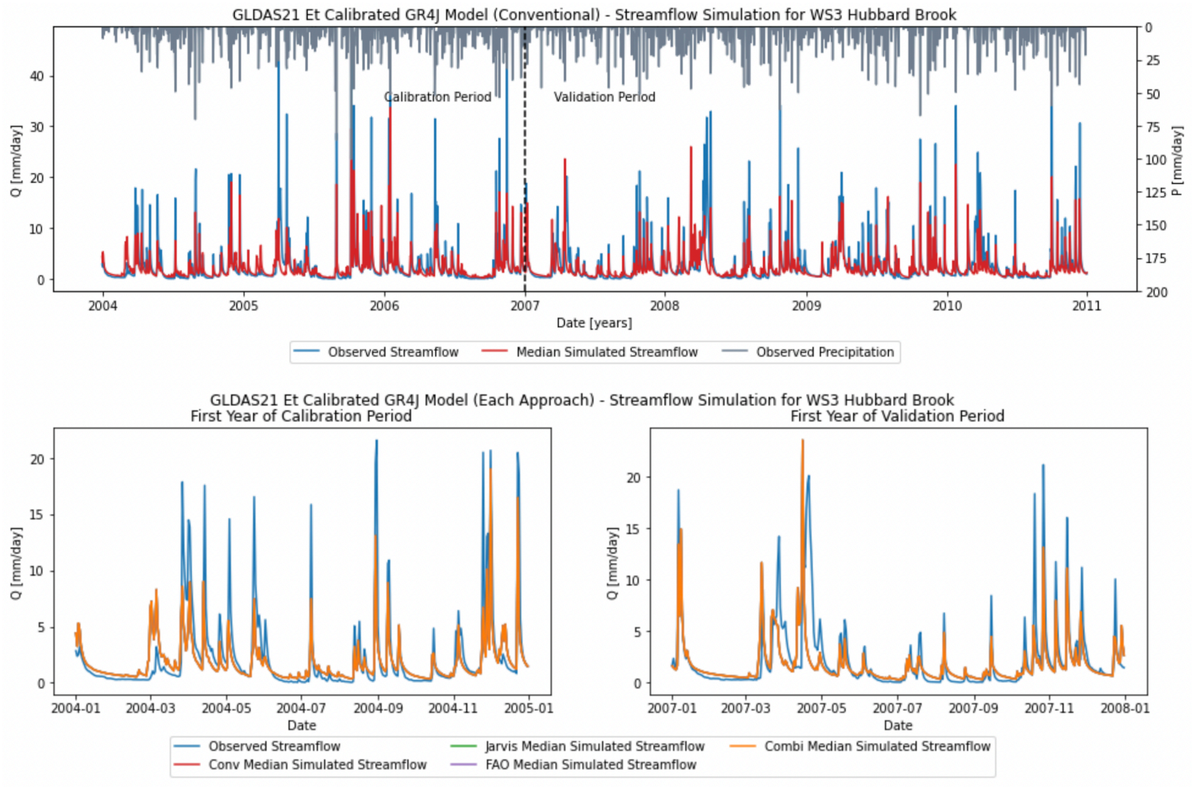


Figure A.5: GR4J Model - Streamflow Simulations for WS3 in Hubbard Brook.

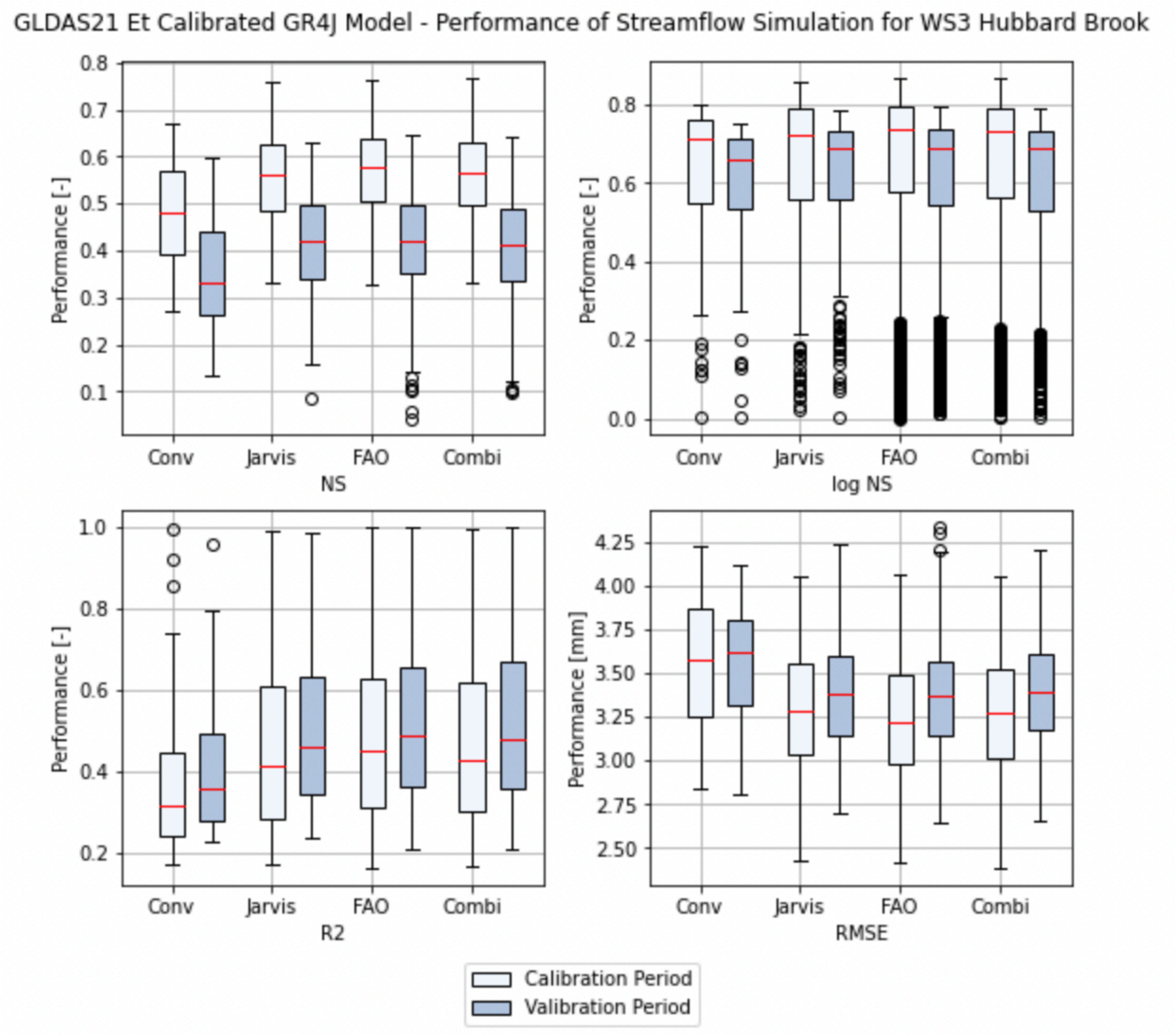


Figure A.6: GR4J Model - Performance of Streamflow Simulations for WS3 in Hubbard Brook.

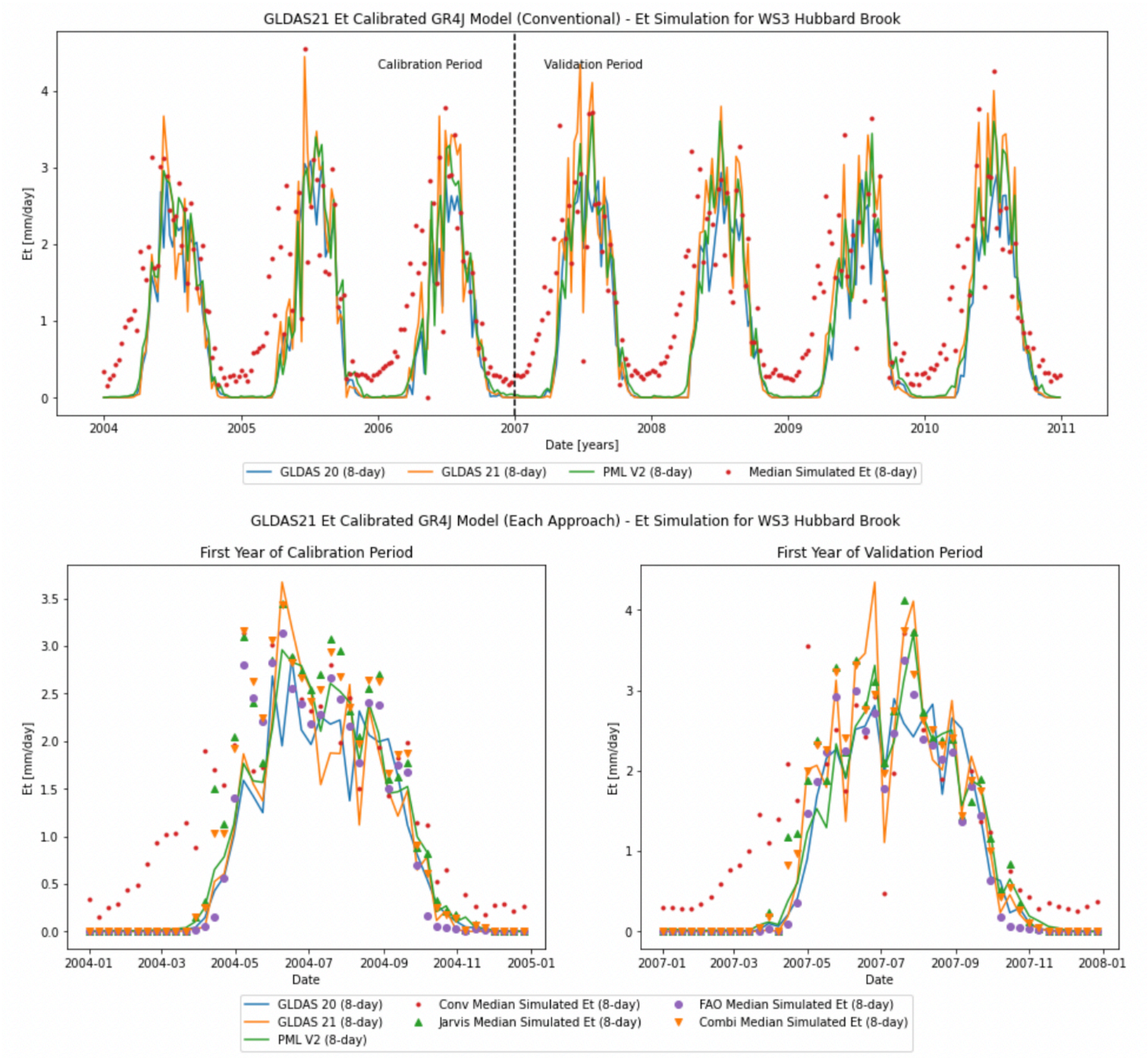


Figure A.7: GR4J Model - Transpiration Simulations for WS3 in Hubbard Brook.

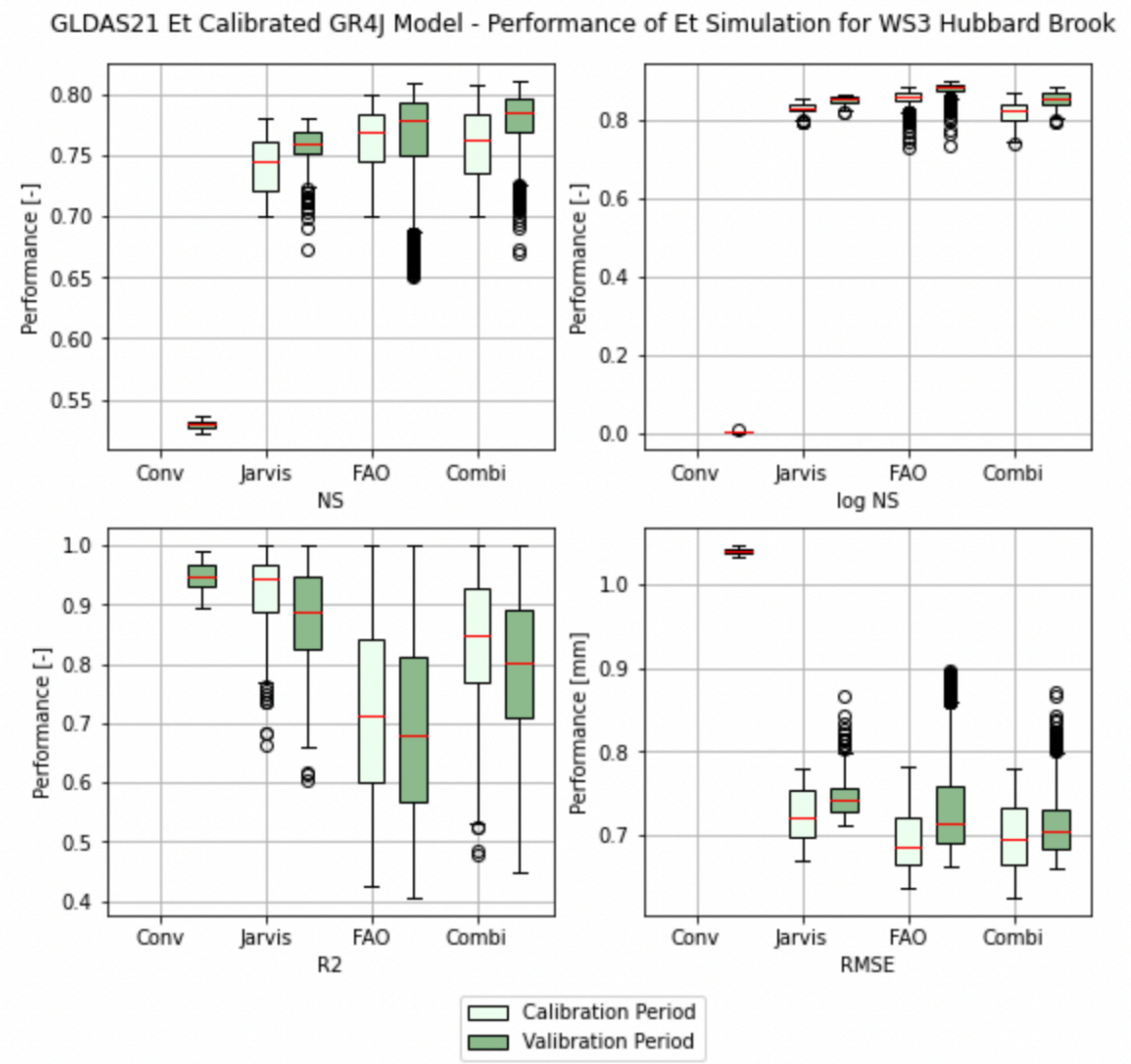


Figure A.8: GR4J Model - Performance of Transpiration Simulations for WS3 in Hubbard Brook.

# B

# Transpiration Calibrated Models for WS2 in HJ Andrews

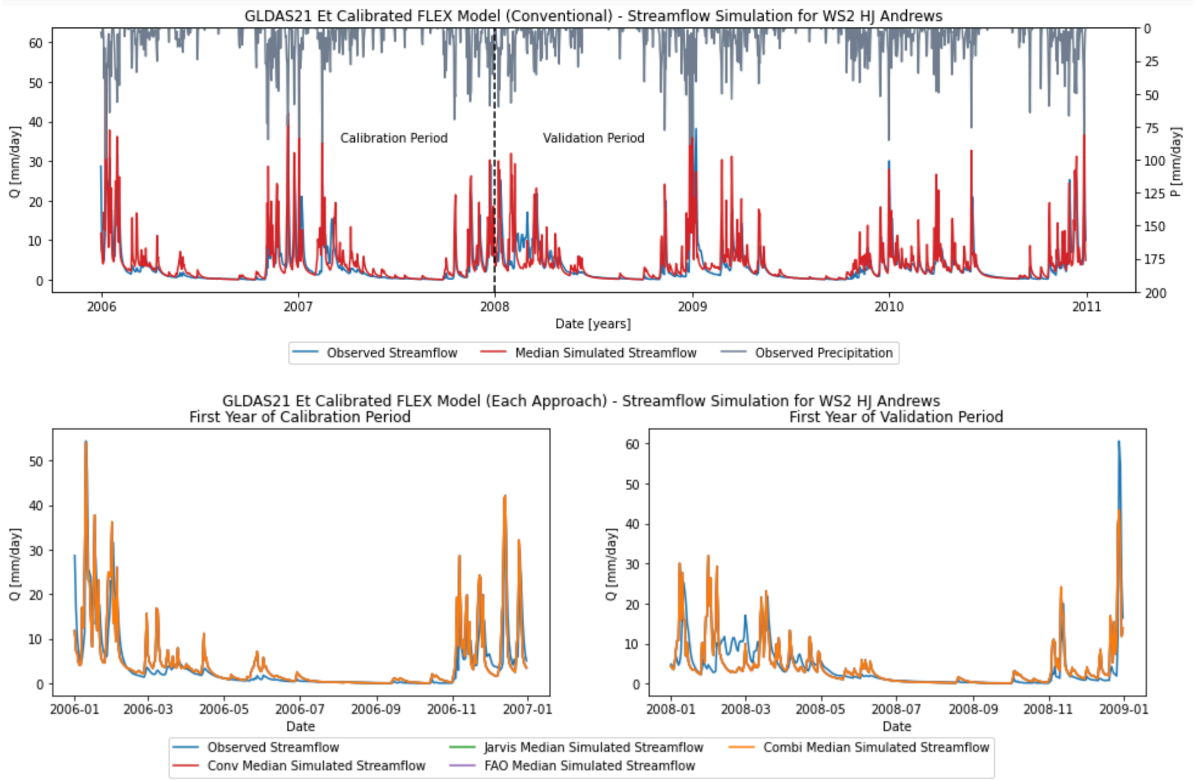


Figure B.1: FLEX Model - Streamflow Simulations for WS2 in HJ Andrews.

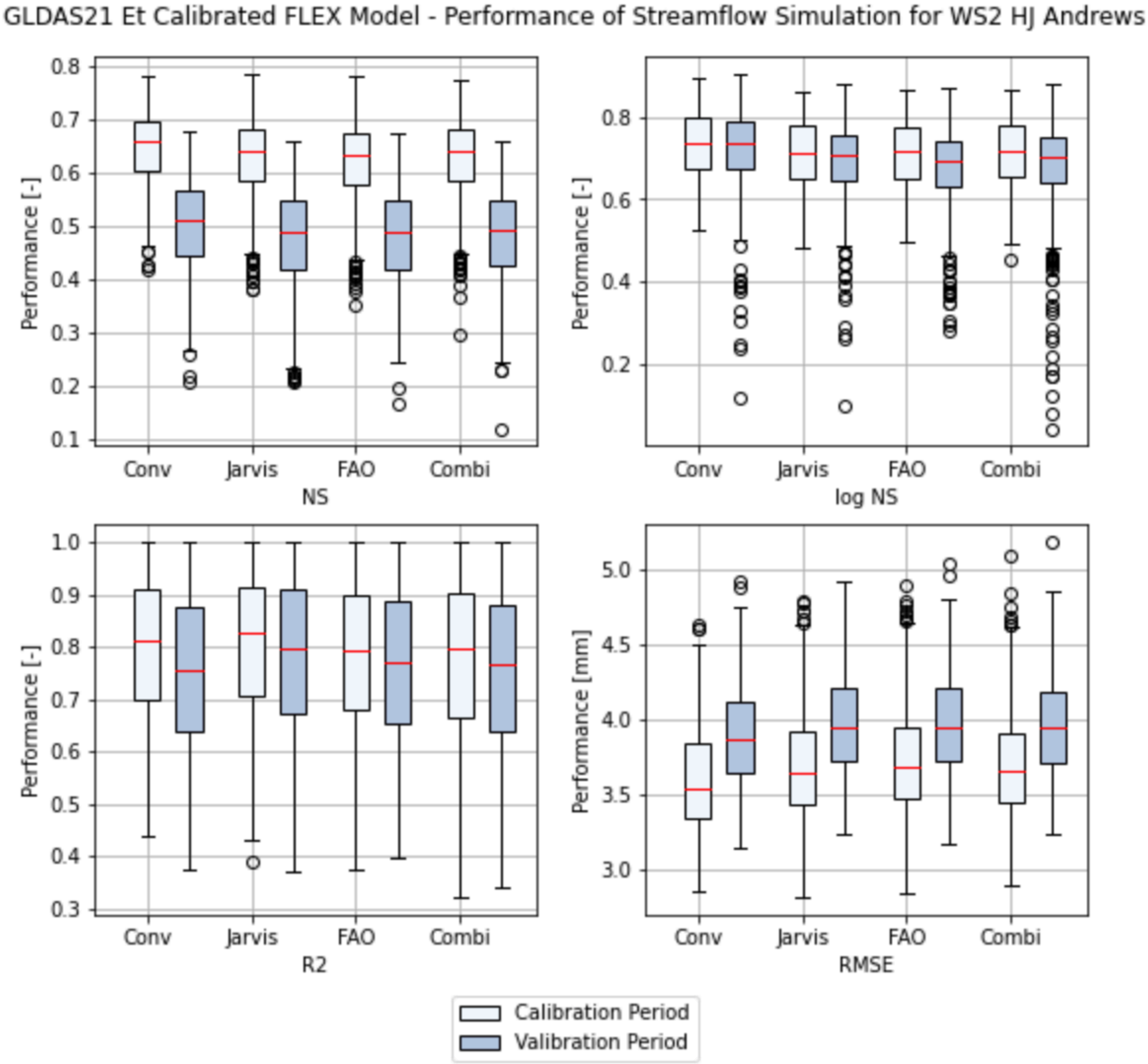


Figure B.2: FLEX Model - Performance of Streamflow Simulations for WS2 in HJ Andrews.

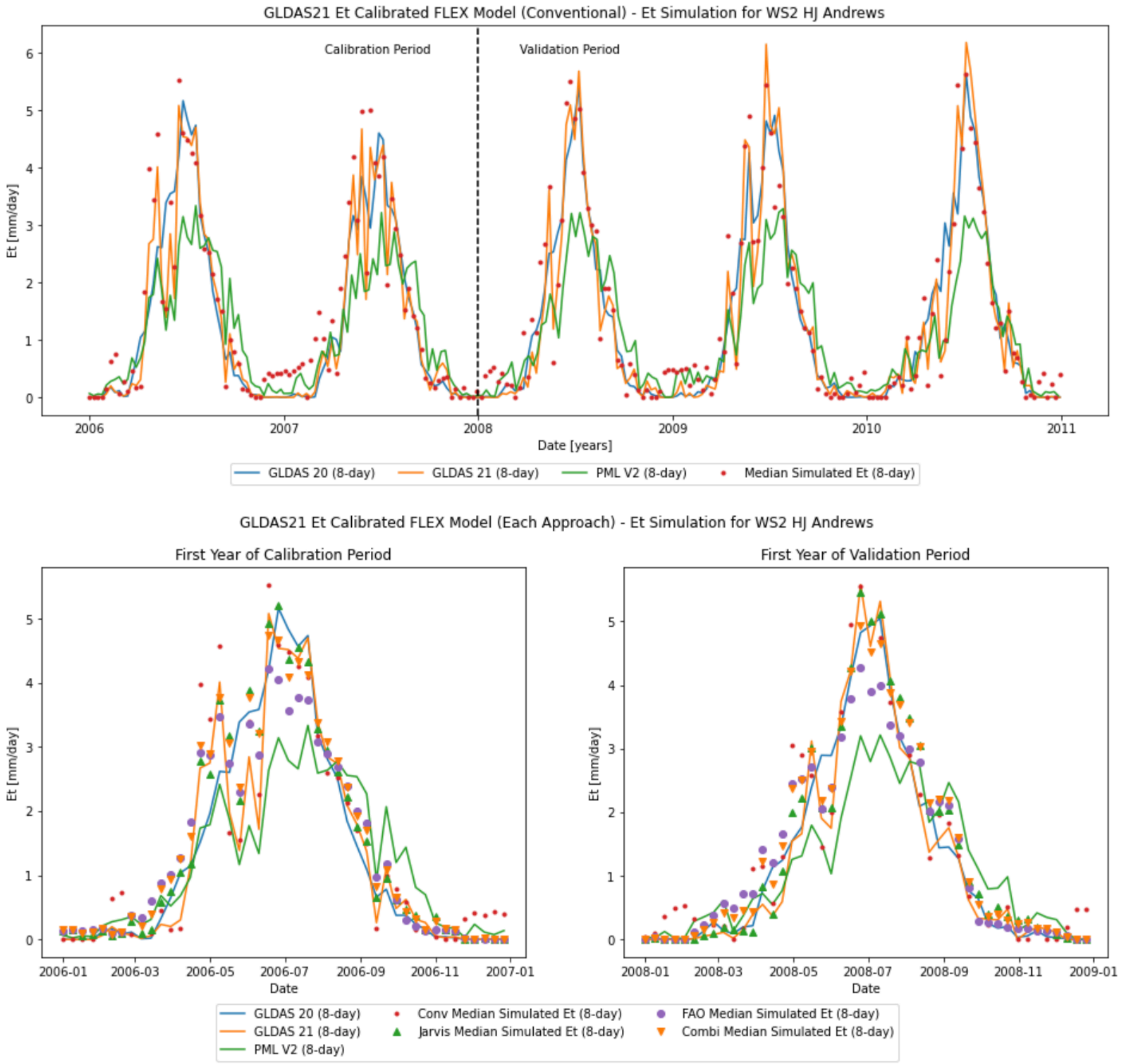


Figure B.3: FLEX Model - Transpiration Simulations for WS2 in HJ Andrews.

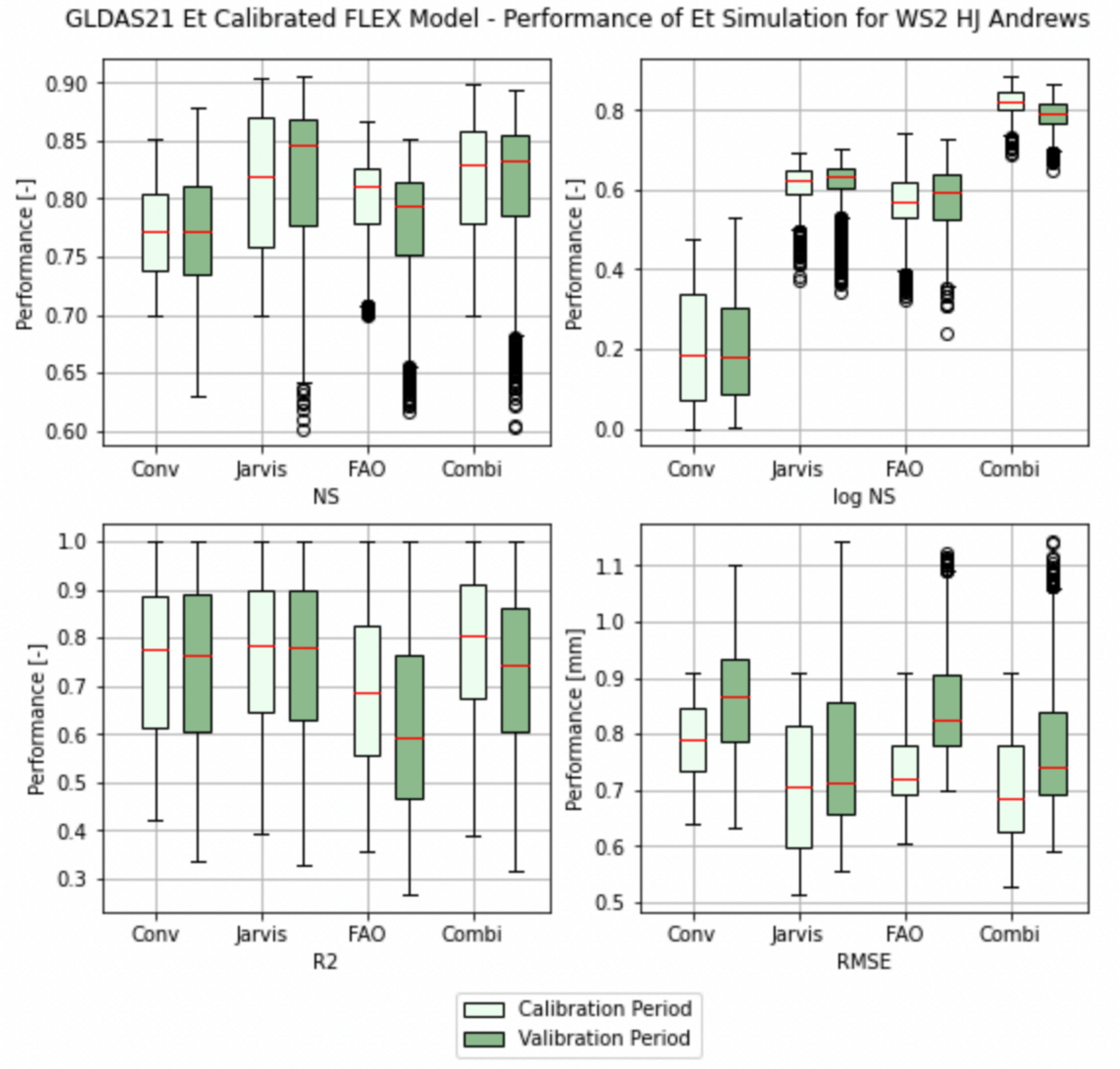


Figure B.4: FLEX Model - Performance of Transpiration Simulations for WS2 in HJ Andrews.

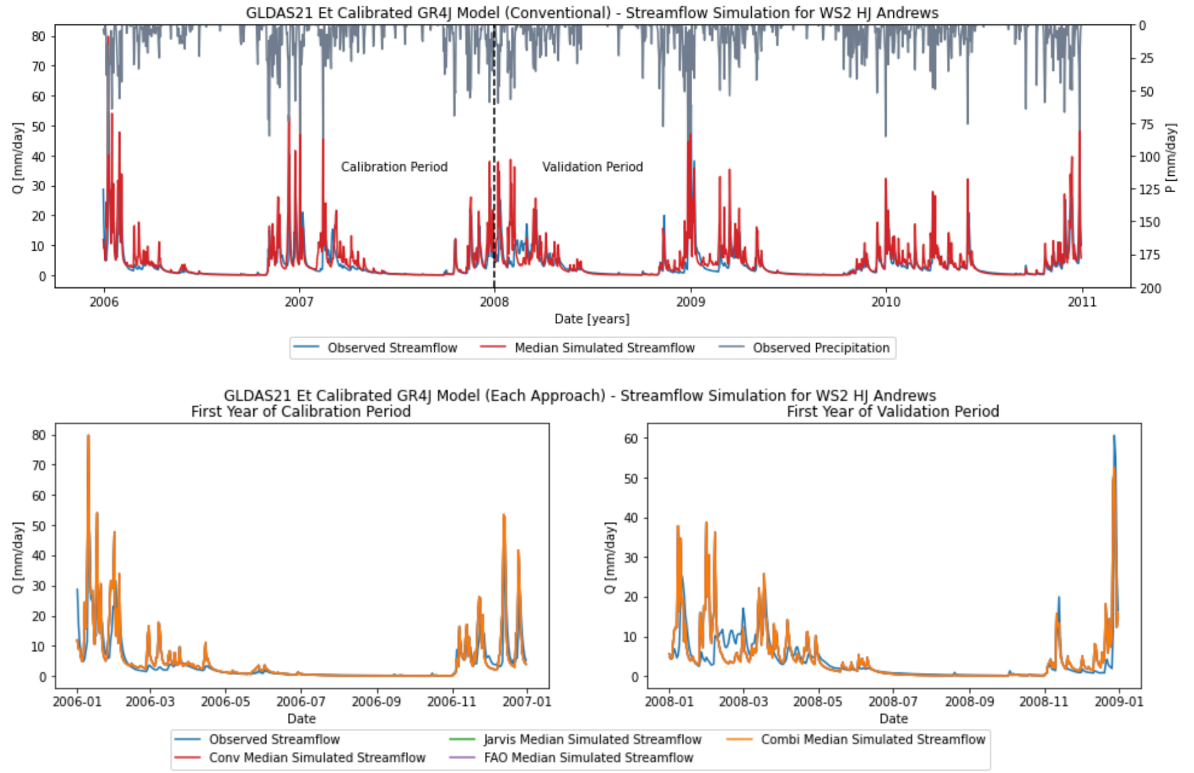


Figure B.5: GR4J Model - Streamflow Simulations for WS2 in HJ Andrews.

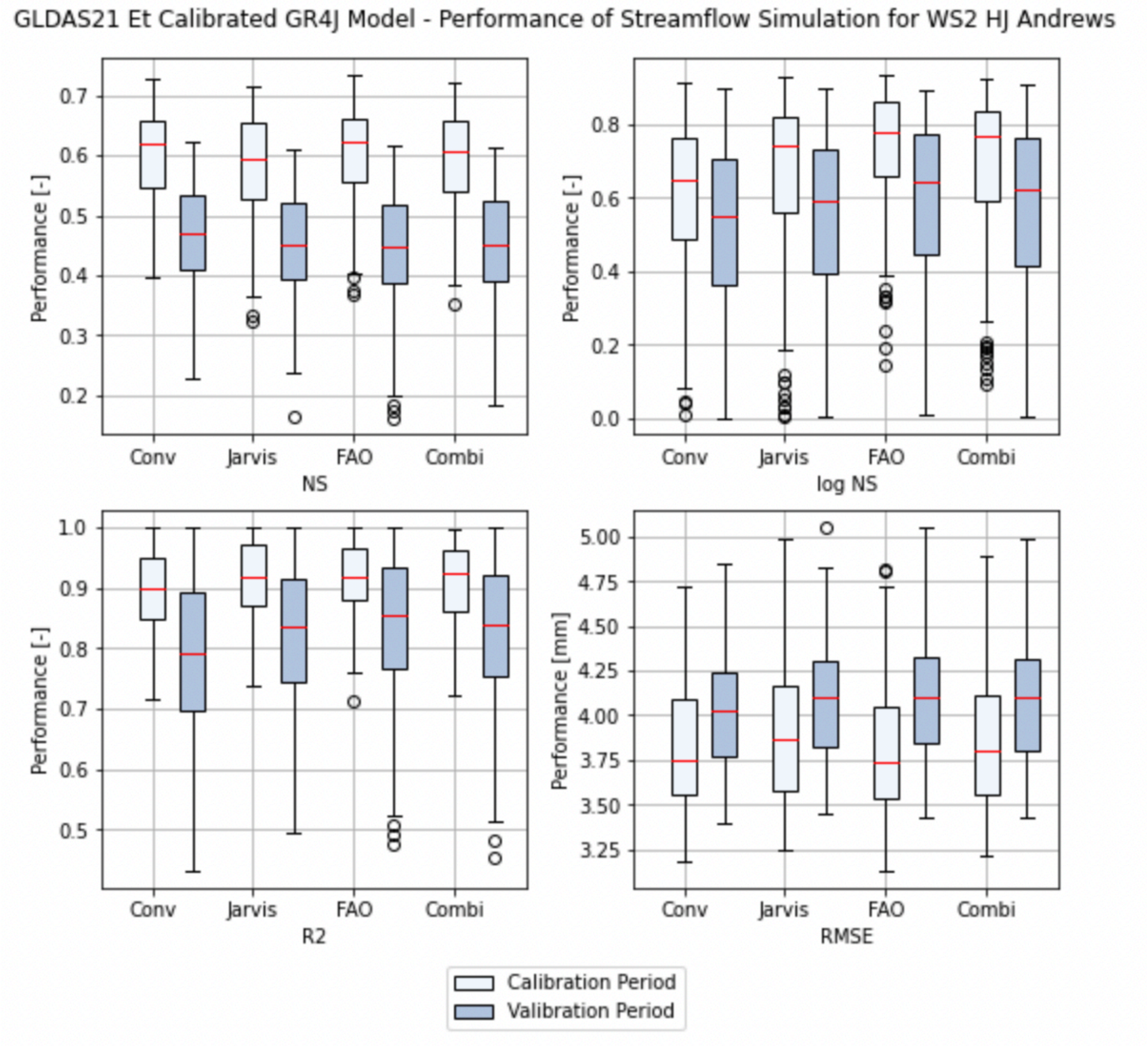


Figure B.6: GR4J Model - Performance of Streamflow Simulations for WS2 in HJ Andrews.

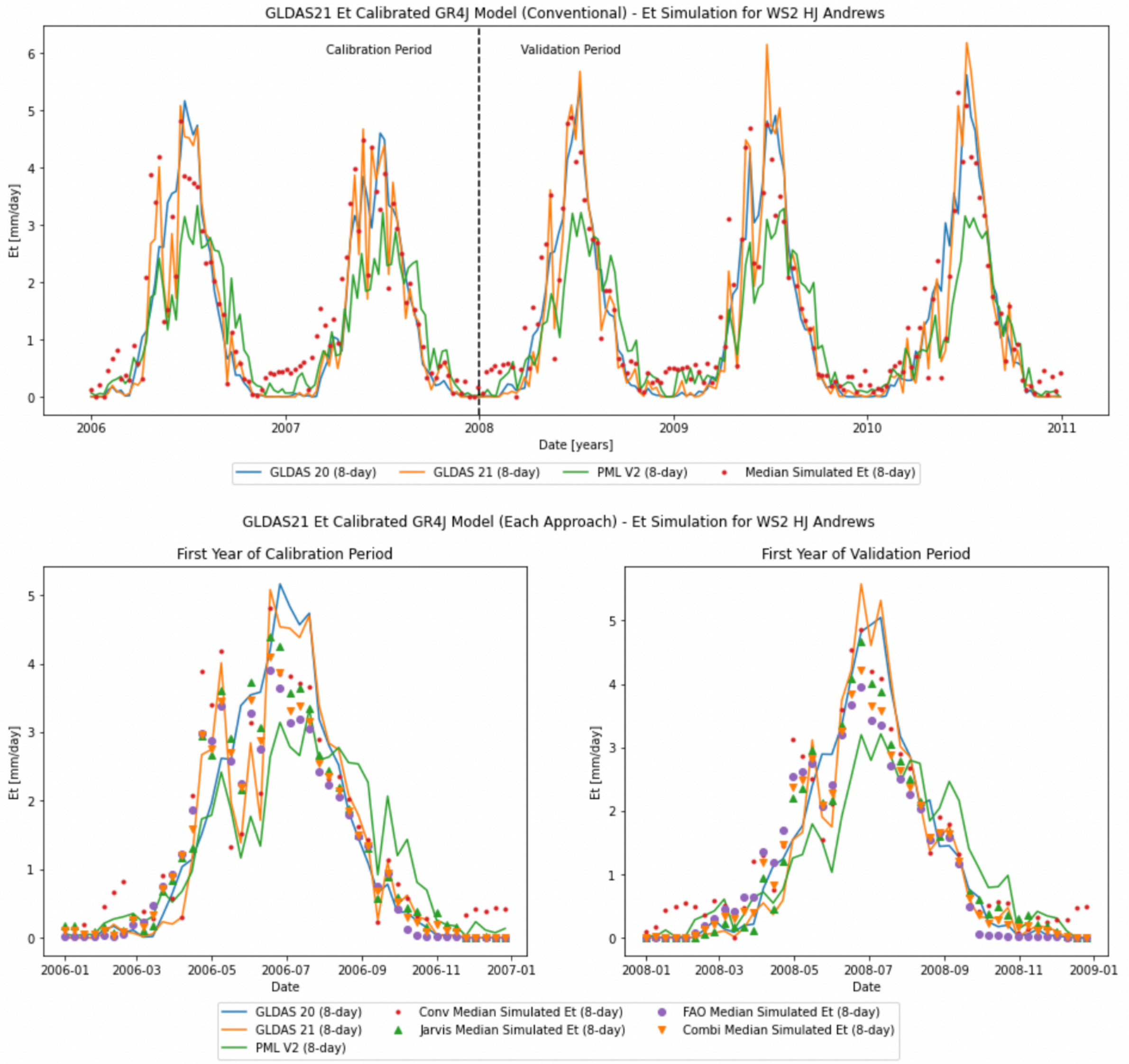


Figure B.7: GR4J Model - Transpiration Simulations for WS2 in HJ Andrews.

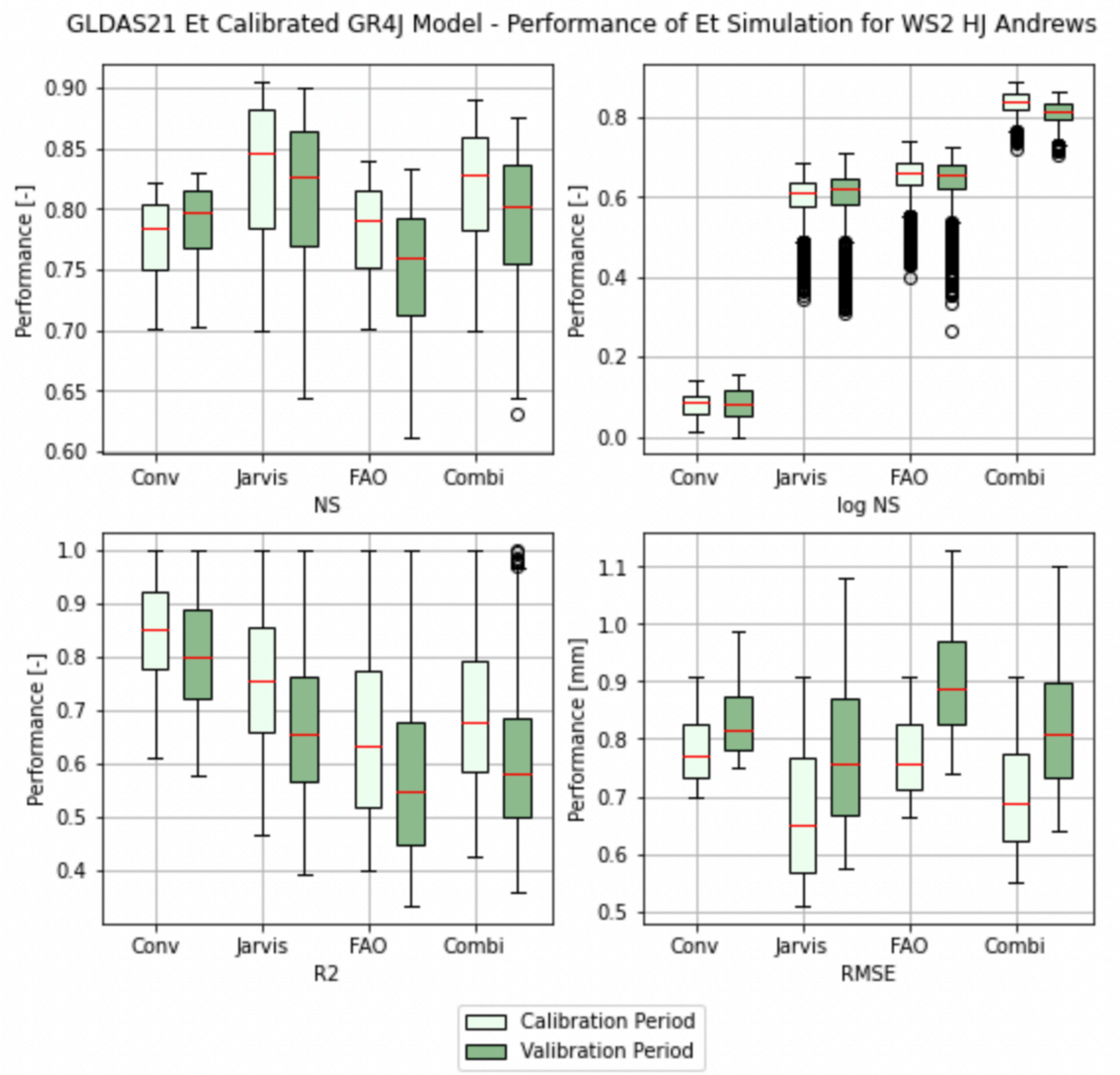


Figure B.8: GR4J Model - Performance of Transpiration Simulations for WS2 in HJ Andrews.