A satellite-style topographic map of a mountainous region, likely the Alps. The terrain is shown in shades of brown and tan, with significant snow cover in the higher elevations and mountain passes. Several dark blue lakes are scattered across the landscape, particularly in the central and lower-right areas. The overall scene is a detailed view of a natural catchment area.

Root-zone storage and snow cover effects

A catchment study on the dynamic behaviour of the root-zone moisture capacity related to changing snow cover patterns

Christian Bouman

Delft University of Technology

Front cover: Snow cover in the Sierra Nevada, as observed by NASA's MODIS instrument on the Aqua satellite on March 27, 2010 (Allen, 2015).

Root-zone storage and snow cover effects

A catchment study on the dynamic behaviour of the
root-zone moisture capacity related to changing
snow cover patterns

by

C. (Christian) Bouman

in partial fulfillment of the requirements for the degree of

Master of Science
in Civil Engineering

at the Delft University of Technology,
to be defended publicly on ??, 2018 at ??.

Thesis committee:	Dr. M. Hrachowitz	TU Delft, supervisor
	Prof. dr. ir. S.C. Steele-Dunne	TU Delft
	Dr. ir. O. Morales Napoles	TU Delft

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Preface

Three years ago, I decided to say goodbye to the hard engineering masters and to follow the 'soft' Watermanagement master instead. I think it was an attempt to avoid difficult mathematics, but it resulted in graduating with a thesis consisting of a multivariate and a hydrological model. Mathematics came like a boomerang in my face!

However, it was not that dramatic but quite fun! Hydrological modeling is like puzzling; it's simple when you have the right piece. And I want to thank Markus Hrachowitz for tip over the box of puzzling pieces and encourage me to find the right one. Markus ideas on hydrology are those puzzling pieces and the box seems endless when I had a meeting. His enthusiasm made me even more curious and I couldn't stop searching. So I definitely owe this result to him. In addition to this, I want to thank Susan Steele-Dunne and Oswaldo Morales Napoles to limit the search area a little bit. Your inputs were vital for completing this thesis. I like the way this committee complemented each other and how I could learn from the different fields of interest, which are all included in this thesis. The combination of remote sensing, statistical analysis and hydrological modeling made it a complete story.

Furthermore I like to thank Stefan Fugger, for helping me out with the RSLE analysis. It was fun working together and to push each other to a higher level of programming and analysis. An off course all the other hard working people of room 4.93. You made the graduation room a nice combination of a good work ethic, jokes and coffee. Lots of coffee.

At last, I want to thank my wife Esther. She supported me all year long, even when we had to prepare a wedding instead of graduating. I talked so many times about my thesis to her that I am not afraid to let her present on October 18th! Combining a wedding and a thesis made this period more dynamic and a time never to forget.

*Christian Bouman
Delft, October 2018*

Contents

1	Introduction	3
1.1	Research aim and contribution	4
1.2	Report structure	5
2	Methodology	7
2.1	General approach	7
2.2	Study area	7
2.3	Data sources	9
2.3.1	Digital elevation map	9
2.3.2	NOAA COOP station data	10
2.3.3	MODIS snow cover	10
2.3.4	USGS discharge station data	14
2.4	Trend analysis.	14
2.4.1	Mann-Kendall test	14
2.4.2	Slope and intercept	15
2.4.3	P-value.	16
2.4.4	Anomalies	16
2.5	Multivariate regression	16
2.6	Scenarios	17
2.6.1	Scenario selection	17
2.6.2	Variables	18
2.7	Hydrological modeling	20
2.7.1	Modeling scheme	20
2.7.2	Snow storage.	20
2.7.3	Unsaturated zone	21
2.7.4	Calibration	22
2.7.5	Analysis	22
2.8	Conclusion	22
3	Results	23
3.1	Trends	23
3.1.1	Winter precipitation and temperature	23
3.1.2	Snow-free days per elevation zone.	23
3.1.3	Weighted snow-free days.	27
3.2	Relation between snow-free days, precipitation and temperature.	29
3.2.1	Precipitation and temperature dependence	29
3.2.2	Scenarios.	30
3.3	Hydrological processes	34
3.3.1	Soil moisture deficit	34
3.3.2	Snow storage differences.	36
3.4	Conclusion	38
4	Discussion	39
4.1	Trends	39
4.1.1	Data	39
4.1.2	Mann-Kendall test	40
4.2	Relation between snow-free days, precipitation and temperature.	41
4.3	Hydrological processes	41
4.3.1	Simplified model structure.	41
4.3.2	Transpiration.	42

4.4 Application to the SMD method	42
5 Conclusions and further research	43
5.1 Conclusions	43
5.2 Further research.	44
Bibliography	45
List of Symbols	49
List of Figures	51
List of Tables	53
A Calibration results	55
B Long term snow-free days	57
C Sensitivity analysis for interval of rolling maximum	59

Summary

Extreme weather events seem to happen more often nowadays. One of these extreme events was the California Drought between 2012 and 2016. The socioeconomic and environmental impacts of this drought were enormous: thousands square kilometers of agricultural land fallowed, thousands of lost jobs, salinization problems and forest fires. As a result of the drought, multiple research is conducted related to the hydrological and environmental impacts. However, the role of the root-zone storage capacity during the drought period is not investigated well. This root-zone storage capacity is the water available for plants to transpire and to grow. It influences the partitioning between the transpiration and the run-off rates, which controls the fundamental processes in ecosystem functioning (such as floods, droughts and groundwater recharge). Scientific evidence is growing that this root-zone storage behaves dynamically. The underlying assumption is that plants adapt their root system to climatic and environmental changes. One of these changes related to drought is the availability of snow, which is an important source of the rivers in California. It is important to know how the root-zone is changing due to the snow cover changes during the California Drought to understand the hydrological behaviour of the Californian rivers. For that, the Merced River basin in the Sierra Nevada is analyzed on the relationship between snow cover and root-zone storage capacity.

To analyze these relationship, this research consists of two parts: trend analysis and hydrological modeling. The purpose of the first is to see whether there is a change in snow cover, temperature and precipitation and to learn about the dynamics between precipitation, temperature and snow cover. This is done by conducting a trend analysis with the Mann-Kendall test for precipitation, temperature and snow-free days. These snow-free days are calculated based on snow cover data from a MODIS satellite product (MOD10A1) and a method of regional snow-line elevation. Additional to this, a multivariate regression analysis was executed to related snow-free days, precipitation and temperature to each other. Furthermore, scenarios for snow-free days were conducted to analyze the effect of the California Drought. The purpose of the second part is to relate the change in snow cover to the root-zone storage capacity. This is done by building a hydrological model calibrated on the snow cover data. By doing this, the maximum soil moisture deficit over a 15 year interval could be calculated, which is an estimation for the root-zone storage capacity.

The trend analysis showed a clear downward trend in winter precipitation, but not a significant trend in temperature. The snow-free days are increasing at all elevations, but only significant up to 3300 m.a.s.l. The maximum trend related to the observed snow-free days was 7 days/year at 2300 m.a.s.l. and the average for the whole catchment was 4 days/year. The multivariate regression analysis showed that the temperature influence is more important than the precipitation input, however the maximum R^2 was around 0.7. At higher elevations it was not so clear whether precipitation or temperature described snow-free days well enough. The scenarios revealed the enormous effects of the California Drought on the total amount of snow-free days and confirmed the importance of temperature during this extreme event. At 2300 m.a.s.l. it differs 40 snow-free days with a situation with normal winter temperatures. The effect of temperature is also visualized in the hydrological model, where the maximum soil moisture deficit was not found in the drought period but slightly before it. This is caused by the combination of small precipitation amounts and a low temperature (no melt). In the drought period the soil moisture deficit was limited by the melt from the snow. This is also confirmed by the analysis of the snow storage in the hydrological model. However, the maximum of the soil moisture deficit is gradually increasing of the past 30 years which hints to the adaptive behaviour of the root-zone.

The trend analysis and the hydrological model showed that the amount of snow in the Merced River basin is decreased due to the California Drought. At 2300 m.a.s.l., the snow-free days are changed from 170 to over 250. The maximum difference in snow storage before and after 2010 is 100 mm averaged over the whole basin. The decrease of snow during this period had a limiting effect on the soil moisture deficit. Due to the melting water, no maximum values of soil moisture deficit were found. However, slightly before the California Drought a maximum was found. This maximum is related to a cold but dry period, where not a lot of melting water was generated. Overall, the maximum of the soil moisture deficit is increasing over the past 30 years. This hints to the adaptive behaviour of plants. In summary, the California Drought caused a decrease of the amount of snow and resulted in limiting of the soil moisture deficit. Simultaneously the soil moisture deficit is increasing. The limiting effect of snow melt on the soil moisture deficit gives plants the time to cope with the changes.

However, with the found trends there will be less snow in the future to limit these effects and the soil moisture deficit will rise. It is highly uncertain whether the Californian ecosystems are able to cope with this.

1

Introduction

All around us it seems to happen more often: extreme weather events. The news shows often videos about floods, forest fires, droughts and other climate related disasters.

Californian drought One of these extreme events was the Californian Drought between 2012 and 2016. This drought was the longest drought ever recorded in the area and caused over-pumping of groundwater and led to significant land subsidence and related loss of farmland. On a single-year level, it was estimated that 17100 jobs were lost and about 170 km² of farmland were fallowed in the 2014 drought. The monetary losses to all economic sectors were estimated at about \$2.2 billion in the same year (Howitt et al., 2015).

After the drought and during the last years of the drought a lot of scientific and social interest was focused on the California drought. A lot of papers about the drought were recently published with different angles of approach: the economic effects for the agricultural sector (Howitt et al., 2015) and hydropower electricity (Glick, 2017), social effects related to crime (Goin et al., 2017) and off course hydrological impacts. He et al. (2017) showed that the reservoir operation enlarged the drought related problems in the San Joaquin Valley in 2014. Bales et al. (2018) carried out a water balance study for the upper Kings River and showed how evaporation, transpiration and runoff interacts with each other during a long drought period. Luo et al. (2017) investigated the temperature and precipitation anomalies and how these are related to drought on basin scale. He states that the warmer temperatures over the high-elevation areas during the wet season have contributed equally or more than the precipitation deficits to the reduction of snow. The latter is interesting since a change in snow accumulation in the mountainous areas of California is already predicted (Lundquist and Cayan, 2007). Due to climate change it is likely that drought periods can enlarge this effect. These developments have large potential consequences for the streamflow in the rivers, since most of the Californian rivers are dependent on snow.

Root-zone storage capacity Unfortunately, it is not clear how snow cover changes are related to the streamflow. Due to that, it is necessary to investigate the hydrological functioning of snow dependent catchments. An important aspect of the hydrological functioning of a catchment is related to vegetation. The presence of vegetation increases the infiltration capacity in water-limited ecosystems (Scheffer et al., 2005; Thompson et al., 2010) and the roots of plants affect the water retention (Ankenbauer and Loheide, 2017). These roots interact with the hydrological system to retain and extract the water for transpiration within a dynamic equilibrium with the available soil moisture to avoid water shortage. One can say that the roots create moisture storage volumes to extract water that is stored between field capacity and wilting point (Nijzink et al., 2016). It is clear that the root-zone storage capacity¹) highly influences the partitioning between the transpiration and the run-off rates, which controls the fundamental processes in ecosystem functioning (such as floods, droughts and groundwater recharge) (Milly, 1994). This makes the root-zone storage capacity an important aspect for modeling and increased understanding the behaviour of hydrological systems.

Although it is an important aspect of the dynamics within a catchment, it is impossible to observe the root-zone storage directly because of its natural nonlinear and inhomogeneous behaviour. This behaviour is caused by interaction with vegetation, living organisms and the surface. Because of this nonlinear behaviour, the root zone moisture storage capacity is often used as a calibration parameter in hydrological models or obtained from available soil parameters with an estimated rooting depth (Saxton and Rawls, 2006; Grayson et al., 1992).

¹This moisture storage volume is often called plant available moisture or plant available water holding capacity, but for this thesis I will use root-zone storage capacity with the symbol S_r .

In these cases, the root zone storage capacity is often assumed as a constant which excludes the dynamic patterns. However, growing scientific and experimental evidence indicates that vegetation dynamically adapts its root system to the climatic and environmental changes (Schymanski et al., 2008). To investigate this dynamic behaviour, Gao et al. (2014) introduced a method to estimate S_r based on change of soil moisture deficit within a catchment. Later on, this method is successfully applied for cases with deforestation effects (Nijzink et al., 2016) and for different climates (de Boer-Euser et al., 2016). To strengthen the theory and the method, other applications of the soil moisture deficit method are desirable. For example extreme events like droughts.

It is clear that more knowledge about the hydrological behaviour of snow dependent catchments in California is necessary to cope with future extreme droughts. Furthermore, the previous states that the root-zone moisture capacity is a vital part in obtaining this knowledge. For that, it is necessary to investigate the root-zone moisture capacity and its relation to snow cover changes in a Californian snow dependent catchment.

1.1. Research aim and contribution

This study focuses on the influence of snow cover changes on the root-zone storage capacity. The objective of this study is (1) to contribute to the existing knowledge about the S_r and the adaptive behaviour of vegetation related to snow cover dynamics, (2) to investigate the hydrological responses during the California Drought and (3) to relate these results to a future perspective of the study area. Basically, the research consists of two parts: data analysis and hydrological modeling. The main research question for the data analysis part is:

What is the effect of precipitation and temperature changes on the amount of snow-free days in a catchment in the Southern Sierra Nevada?

In this context, snow-free days are defined as the amount of days per year without snow. In this study, the snow-free days are analyzed per elevation zone of the catchment. In order to answer this question, the data analysis has been conducted with a focus on:

- Trend analysis of separate climate parameters: temperature, precipitation and snow cover.
- Comparative assessment of different multivariate models which relate snow-free days to combinations of temperature and precipitation.
- Prediction of snow-free days based on the multivariate models.

For the hydrological modeling part, the main question is:

What is the effect of a change in the number of snow-free days on the root zone moisture storage capacity in a catchment in the Southern Sierra Nevada?

This question is answered by:

- Assessing the change of soil moisture deficit in relation to S_r with the method by Gao et al. (2014).
- Obtaining modeled snow-free days with inverse modeling and compare this with the observed snow-free days.
- Comparative assessment of the hydrological storages and fluxes before and after the California Drought.

This study will provide more insight in the response of the S_r during the California Drought. With the knowledge of the S_r , the total knowledge about the hydrological storages and fluxes in the study area can be enhanced. Furthermore, most of the scientific work related to the California Drought focuses on the long term effects and whether or not the human-induced climate change is the cause (Seager et al., 2017) or how the drought relates to the CO₂ footprint (Hardin et al., 2017). These human effects are important and can't be seen separately from the hydrological and environmental effects. With this study an attempt is made to widen the view on the effects of the California Drought from a hydrological perspective.

1.2. Report structure

In the next chapter, the methodology will be described. Within this chapter the general approach, the study area and the data sources are introduced first. Next to the methodology of precipitation and temperature data handling, the calculation of the snow-free days from MODIS snow cover data is an important topic in the data sources section. The second part of Chapter 2 describes the methods for statistical analysis, respectively trend analysis and multivariate regression analysis. The combination of the two results in different scenarios, which are discussed in Section 2.6. The methodology will be concluded with the hydrological modeling part, which consists of the modeling scheme, introduction of the water balance equation and the inverse modeling part. In Chapter 3 the results are presented. First from the statistical analysis, thereafter from the hydrological modeling part. The results of the inverse modeling are obtained to link the hydrological modeling to the statistical analysis. This is introduced in the last paragraph. Chapter 4 discusses the results of the statistical analysis and the hydrological modeling part. Chapter 5, at last, presents the conclusions based on the results and discussion. The conclusions are focused on the hydrological responses and the future perspective of the study area specifically and the modeling of the S_r in general.

2

Methodology

2.1. General approach

This research consists of a data analysis part and a hydrological modeling part. The results of the data analysis are the inputs for the hydrological modeling, but the hydrological modeling is also validated with the analyzed data. Figure 2.1 visualizes the modeling structure. The data analysis part consists of a trend analysis of the temperature (T), precipitation (P) and snow cover data. The snow cover data is handled in such a way that elevation and snow cover is related and translated to snow-free days (SFD). Furthermore, the same data is considered to find a relation between SFD, P and T. This is explained in Section 2.5. The data of the data analysis, potential evaporation (E_p) and observed discharge (Q) are used as inputs for the hydrological model (Section 2.7). The most important outputs from the hydrological model are an estimation of S_r based on soil moisture deficit calculation and changes in snow storage. The snow storage is considered to calibrate the model on the observed SFD and to validate the trends found in the data analysis part. Both the data analysis and the hydrological model complement each other in this way to results in a better understanding of the processes occurring in the study area.

2.2. Study area

One of the most vulnerable areas in California, both from an environmental and an economic perspective, is the San Joaquin Valley. Due to large scale agricultural activities, sufficient fresh water of good quality is necessary. However, droughts cause not only water shortages, but also land subsidence and salinization problems (Schoups et al., 2005). The main fresh water source of the valley is the San Joaquin River, which streamflow is highly dependent on the snow accumulation in the Sierra Nevada. One of its tributaries located in the Sierra Nevada is the Merced River (Figure 2.2). This catchment has been selected based on four criteria:

- Located in the Southern Sierra Nevada
- Timespan for hydrological and climate data has to be until 2017
- Relative small catchment (<1000 km²)
- Relative large snow influence (elevation > 1000 m.a.s.l.)

Location The Merced River is located in the Southern Sierra Nevada, California, USA. According to the Köppen-Geiger climate classification (Kottek et al., 2006) this area consists mainly of a warm Mediterranean climate(Csa). However, some parts in the south are classified as semi-arid climate(BSh). In both climates, the snow cover is an important water source which acts as a buffer with high run-off rates in the spring due to snow melting. It is likely that snow cover changes will affect the stream flows and that the results of these changes have already been visible during the California Drought. The possibility to analyze the effects of this drought on the hydrological behaviour in the chosen catchment is a valuable addition to this study.

Timespan Two aspects are of importance related to the timespan of available data in the catchment. Since this study focusses on climate change, the available data has to contain observations for over 30 years. Secondly, the data has to incorporate the drought period (2012-2016). Fortunately, the snow cover data of MODIS

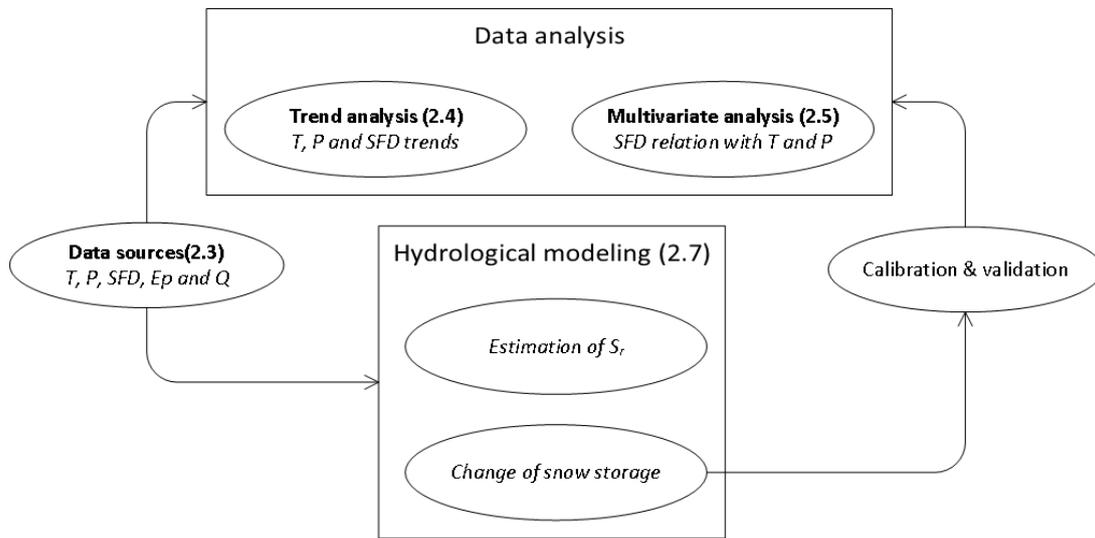


Figure 2.1: Schematic view of the general approach used in this study. The methodology is divided in two parts: data analysis and hydrological modeling. The bold text refers to sections, the related section number is shown between brackets. The italic text depict outputs of a process.

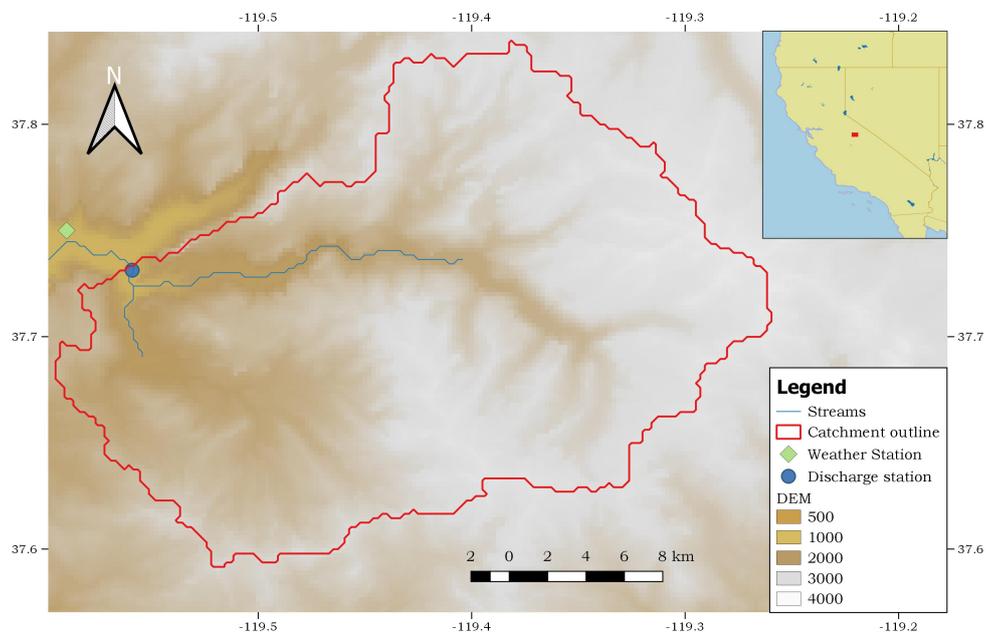


Figure 2.2: The Merced River basin in the Southern Sierra Nevada (CA) with the locations of the weather station and the discharge station.

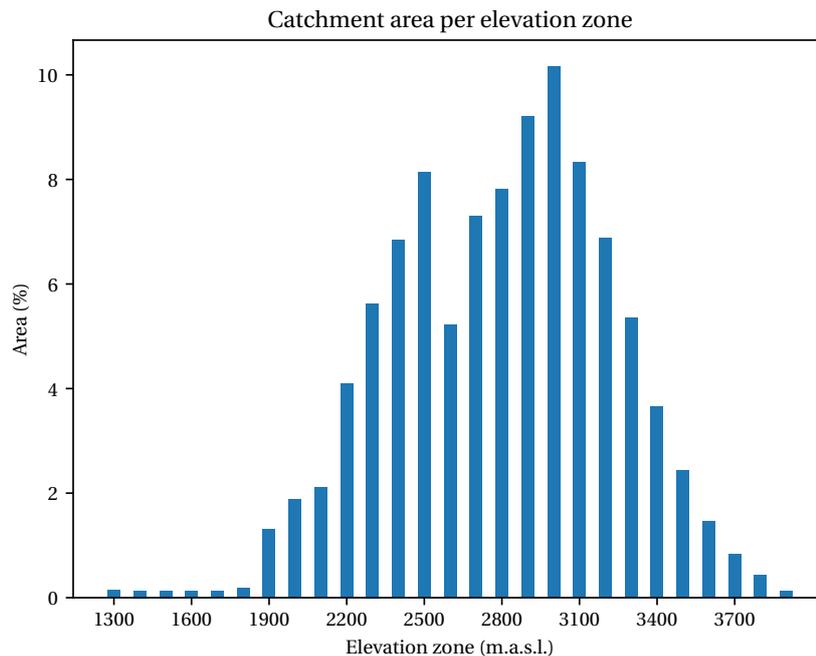


Figure 2.3: The hypsometric distribution of the Merced River basin. This distribution shows the percentage of catchment area within a elevation zone of 100 meter.

(see Section 2.3) covers the timespan 2000-2017 so the focus is on reliable discharge and climate data. The USGS discharge station located downstream of the catchment observes already for over 100 years. Unfortunately, no weather station was located within the catchment, but there is a station slightly outside of it. This weather station of the NOAA COOP network has climate data from over 50 years, however only data from 1982 until present is suitable for scientific purposes.

Size For the size of the catchment it is important that the catchment is not too large, since the focus is on quick responses. This can also be noticed in the modeling scheme of the hydrological model (Section 2.7), where no deep ground water reservoir is present. In this research a catchment size below 1000 km² is chosen. With a size of 467 km² the Merced River basin satisfies this condition.

Elevation Since the focus of the research is in snow cover changes, it is important to analyze catchments which are at least parts of the year covered with snow. To satisfy this condition, a minimum catchment altitude of 1000 m.a.s.l. is chosen. The Merced River basin satisfies this criterion, which can be seen in the hypsometric distribution in Figure 2.3.

2.3. Data sources

Several data sources are used for this research, of which especially climate data (precipitation and temperature) is of importance. The MODIS satellite product is obtained for the snow part of this research. An overview of the used data sources is given in table 2.1.

2.3.1. Digital elevation map

The digital elevation map is obtained from GMTED2010 (Danielson and Gesch, 2011). The GMTED2010 product contains several raster elevation products for 30-, 15-, and 7.5-arc-second spatial resolutions. It uses eight aggregation methods, namely: minimum elevation, maximum elevation, mean elevation, median elevation, standard deviation of elevation, systematic subsample and breakline emphasis. The global aggregated vertical accuracy of GMTED2010 can be summarized in terms of the resolution and root mean squared error (RMSE) of the products with respect to a global set of control points. For the choice in this research, 7.5 arc-seconds,

Process	Product name	Unit	Temporal range	Timespan	Spatial range
Fractional snow cover	MOD10A1	-	daily	2000-2017	500 m x 500 m
Elevation	GMTED2010	m	-	-	7.5 arc second
Mean precipitation	NOAA COOP	mm	daily	1983-2017	point
Stream flow discharge	USGS	Cf/s	daily mean	1982-2017	point
Temperature	NOAA COOP	°C	daily mean	1982-2017	point

Table 2.1: Overview of the consulted data sources during this research

the RMSE range is between 26 and 30 meters.

2.3.2. NOAA COOP station data

The station at Yosemite Park HQ (see figure 2.2) is a station of the NOAA COOP network. The station, located at the headquarters of Yosemite Park, collects temperature, precipitation and even snowfall data. However, the snowfall data is very scarce and only used for short term projects. The temperature and precipitation data in the timespan 1983-2017 is checked on the quality and suitable for academic purposes. In this research, the temperature and precipitation data is used for the winter period, which is defined from November until March. This winter period is based on recent research on snow related topics in the Sierra Nevada (Musselman et al., 2017; Jepsen et al., 2018). This choice is also made to connect this research with the broader discussion about climate change and its relation to snow in the Sierra Nevada. Furthermore, the winter period is generally defined as the period where historically 80-90% of the precipitation occurs in the mountains of semi-arid areas, most of the time in the form of snow (Pandey et al., 1999). To correspond the data in the winter period with the water year definition of the USGS (USGS, 2016), the end year of the winter period is leading. For example, the winter period of 2011 is from November 2010 until March 2011.

Winter temperature

The winter temperature (T_w) is the average temperature over the winter period in degrees Celcius. Since only one weather station is considered, it is necessary to correct for elevation. This correction based on the environmental lapse rate (Martinec and Rango, 1986):

$$T_h = T_s + \frac{(h - h_s)}{1000} \cdot -6.49 \quad (2.1)$$

With the temperature T_h at elevation h and station altitude h_s . With this corrected temperature it is possible to calculate T_w at every elevation zone.

Winter precipitation

The precipitation data is used to determine winter precipitation (P_w). This is the accumulated amount of precipitation over the winter period in millimeters.

Snowfall

An estimation of snowfall is made based on a temperature threshold. This temperature threshold is determined by combining the historical snow fall and temperature data. The median temperature is chosen as threshold, because the average temperature would be affected too much by the outliers in the data. This is shown in figure 2.4. The precipitation over a day with a daily average temperature below this threshold is defined as snow. In this research the winter snowfall (P_s , in millimeters) is defined as the accumulated snowfall over the winter period. With the environmental lapse rate (equation 2.1) it is possible to determine the amount of snowfall per elevation zone. Moreover, it is also possible to relate P_s to P_w with a ratio (P_s/P_w) and analyze whether this ratio is changing over time. This can be an useful addition to the trend analysis, since recent studies point to a shift from snowfall to precipitation (Berghuijs et al., 2014; Irannezhad et al., 2017; Easterling, 2017).

2.3.3. MODIS snow cover

As mentioned in 2.3, the analysis of the snow cover is executed with the MODIS MOD10A1 product (Hall and Riggs, 2016). This product is generated using different MODIS products related to radiance data, geolocation and cloud masking and identifies snow-covered land and even snow-covered ice on inland water (Hall et al., 2001, 2002). A 500m x 500m daily generated grid is chosen for snow cover. Nowadays the snow cover

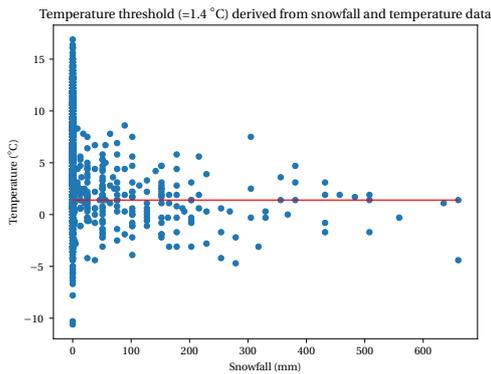


Figure 2.4: The relation between temperature and snowfall, recorded at the weather station. The red line shows the median, which is the temperature threshold used in the hydrological model.

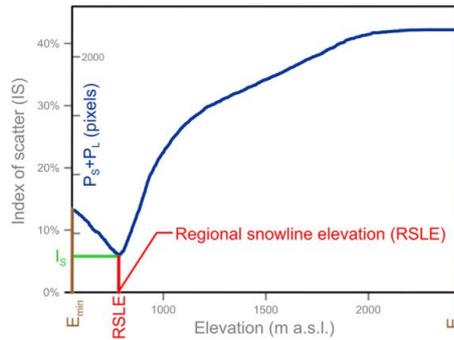


Figure 2.5: Figure from Krajci et al. (2014) where the error-value (I_s) is plotted against elevation. The RSLE is found at the point where I_s is minimum

in MOD10A1 is calculated with the normalized snow difference index (NDSI), which generates a percentage of snow cover. The NDSI is obtained by the reflection in the visible spectrum between $1.6 \mu\text{m}$ (not reflective for snow) and $0.66 \mu\text{m}$ (very reflective for snow):

$$\text{NDSI} = \frac{R_{0.66} - R_{1.6}}{R_{0.66} + R_{1.6}} \quad (2.2)$$

This method is very effective, since clouds can be clearly distinguished from snow at $1.6 \mu\text{m}$. The algorithm of MODIS is adjusted to transform the NDSI to a percentage, which is averaged over a 500×500 meter grid cell.

For linking the temperature and precipitation data to a snow related parameter, the extent of the data sources has to be similar, both in time and in space. The time scale is already similar for climate and MODIS data, which is daily. Spatially, the precipitation and temperature data is determined per catchment and the snow accumulation is determined per elevation zone. Thus, the MODIS data has to be transferred to a parameter which is equal for the whole catchment (regional) and dependent on elevation. This can be done by calculating the regional snow-line elevation (RSLE) per day and create a time series of this. This is done with the following steps:

1. Cropping and reprojection of the DEM and the MODIS files
2. Splitting of the digital elevation map (DEM) in smaller tiles
3. Derivation of the regional snow line elevation (RSLE) from the error-value of the land and snow covered pixels
4. Correction for data gaps and spikes in the timeseries
5. Translation of RSLE to snow-free days (SFD)

Cropping and reprojection The digital elevation map (DEM) from GMTED2010 (Danielson and Gesch, 2011) is used to find the elevation related to the snow cover. This DEM is cropped to the extent of the catchment. The same is done for the 6427 daily MODIS raster files.

Splitting of the digital elevation map (DEM) In the analysis by Krajci et al. (2014), the DEM is divided in smaller tiles for the sake of calculation speed. However, this is done to analyze large mountain ranges. In this research the catchment is so small that the whole basin is assumed as one region. This corresponds also with the climate data, which are also averaged over the catchment. In summary, the DEM is not splitted and this step is skipped.

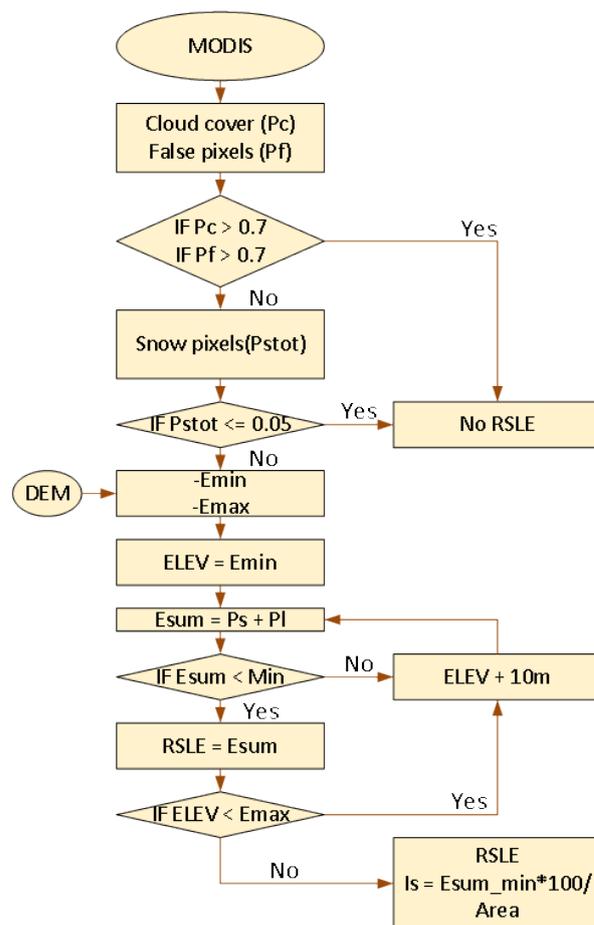


Figure 2.6: Flowchart of the RSLE methodology

Derivation of the regional snow line elevation(RSLE) The MODIS rasters and the DEM are combined in this step. The flowchart of the methodology is shown in figure 2.6. The different values in a MODIS image distinguish between snow, land, missing or cloud pixels. The missing pixels are defined as the total of no data values, missing values and values related to a saturated detector. From this, the percentages of snow (P_s), cloud (P_c), land (P_l) and missing values (P_m) can be calculated. When P_m and P_c are larger than 70 % or when the total amount of snow pixels(P_s) is smaller or equal than 5 % of the total no RSLE calculation is executed and the output will be a NaN-value. The RSLE will be calculated in all the other cases. This is done by summation of P_s and P_l between the minimum and maximum elevation of the catchment in steps of 10 meters. Mathematically, this is a variation problem, with the purpose to find the optimal elevation (RSLE) that minimizes the objective function:

$$f(\text{RSLE}) = P_s(\text{RSLE}) + P_l(\text{RSLE}) \quad (2.3)$$

Where the sum of P_s and P_l is called scattering by Krajci et al. (2014) and is formulated as:

$$I_s = P_s + P_l[\%] \quad (2.4)$$

Because the term 'scattering' is a confusing term related to remote sensing, there is chosen for the term 'error-value' for I_s . From Equation 2.4 all the I_s values per elevation can be calculated. Next, the RSLE can be found as the corresponding elevation where I_s is minimum and the successive value is larger. Visually it looks like Figure 2.5. In this way the RSLE can be found for every day in the timespan from 2000 until 2017, which is the output of this step.

Correction of data gaps and spikes The time series of the RSLE consists of data gaps (when too much clouds and missing data points were detected) and spikes (which relate to faulty zero observations). Because of this, a low-pass filter is added, which cuts off high peaks with a larger difference than 1500 meter between previous and successive days with an interval of 7 days.

From RSLE to SFD The corrected timeseries of the RSLE is now used to derive the snow-free days(SFD). This is done with several steps:

1. For the calculation it is easier to work with similar years. For that, the timeseries of leap years are corrected.
2. The time series of the RSLE is integrated for one hydrological year (from October until September).
3. The results are distributed per elevation zone of 100 meter
4. The integrated RSLE results per elevation zone are divided by the height of the elevation zone to convert from days/m² to days (or snow-free days). This can also be expressed as:

$$\text{SFD}_h = \frac{\int_{t=0}^{365} \text{RSLE} dt}{\bar{h}} \quad (2.5)$$

With the RSLE at time t (in days in the water year) and the average elevation zone height \bar{h} . This is also visualized in Figure 2.7. The figure shows an example of the corrected RSLE timeseries during the water year 2014. For this example the elevation zone of 2200-2300 m.a.s.l. is shown. The area of this elevation zone is accumulated when it is located below the RSLE, since below the snow line there is no snow present (so snow-free). This accumulated area is called 'snow-free area' and is derived by the

The methodology is also visualized in Figure 2.7. The figure shows an example of the RSLE timeseries during the water year 2014. For this example the elevation zone of 2200-2300 m.a.s.l. is shown. The area of this elevation zone is accumulated when it is located below the RSLE, since below the snow line there is no snow present (snow-free). For that, this area is called 'snow-free area', which has *daysm* as unit. To obtain snow-free days the snow-free area has to be derived by 100, the magnitude of the elevation zone. The SFD is the number of days per water year without snow in a particular elevation zone. The methodology enables to change the RSLE time series to a more robust measure of SFD for every elevation zone per water year. Basically the continuous timeseries is changed to a discrete number per year. This has some benefits for the analysis, since it is possible to compare years and elevation zones.

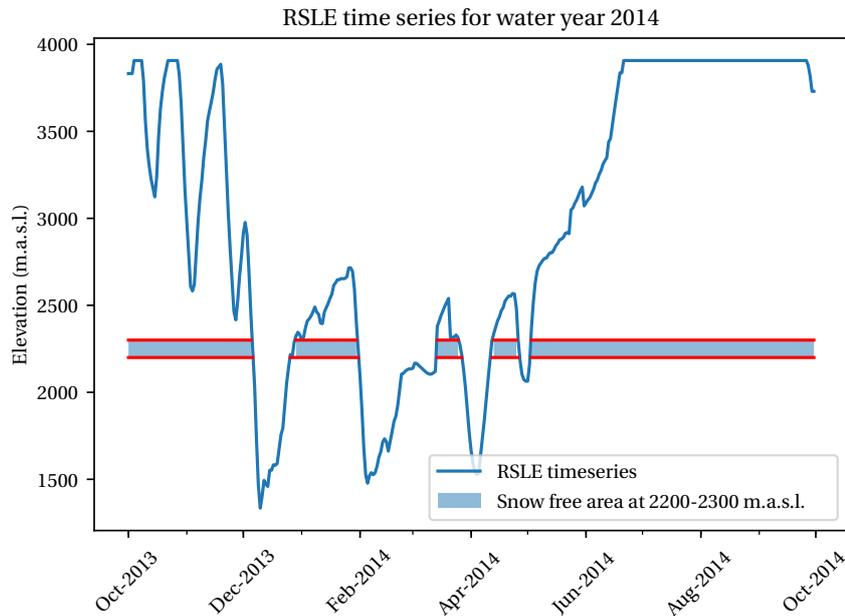


Figure 2.7: The correct regional snow line elevation (RSLE) timeseries for the water year 2014. This figure is an example for the elevation zone 2200-2300 m.a.s.l. with the blue area as the snow-free area. The derivation of this area by the magnitude of the elevation zone provides snow-free days.

2.3.4. USGS discharge station data

The discharge data is obtained from the USGS water database (USGS, 2017). The station is located at the Merced river just downstream of the conjunction with Illioutte Creek. The discharge station collects discharge data (Cf/s) from 1915 until present and measures every 15 minutes. The data is adjusted to daily averages for the period 1883-2017 and converted to m^3/s .

2.4. Trend analysis

Scatter plots are created for respectively T_w and P_w . For SFD and S_w/P_w -ratio the trends are calculated per elevation zone. Three important statistical parameters are estimated for the whole dataset, respectively: the H_0 -test, the slope and the p-value. Additionally, the data is also visualized with anomalies.

2.4.1. Mann-Kendall test

The statistical significance of the trends is tested with the Mann-Kendall test (Mann, 1945; Maurice G Kendall, 1975). This test is often used in hydrological research to discover a linear monotonic trend in a time series dataset. The test is closely related to the concept of Kendall's correlation coefficient, which is widely used and easily understood in terms of tendency of two variables (Noether, 1981). Furthermore, no rule for the data distribution is necessary (it does not require a normal distribution) and it is not affected by missing data or irregular spacing of time points. Within the test it is assumed that the datapoints are independently and identically distributed when no trend exists. Furthermore, the measurements represent the true states of the observed at the times of measurements and the instrumental measurements and data handling are unbiased. The test cannot be executed when the data consists of seasonal data. For that reason, the data is aggregated over a winter period. Another limitation is the length of the dataset. The Mann-Kendall test tends to generate more negative results for shorter time series, this is important during the analysis of the results. As common in statistical tests, a null hypothesis (H_0) is tested against an alternative hypothesis (H_a). H_0 states that no monotonic trend exists. Three options can be used for H_a :

1. An upward monotonic trend exists.
2. A downward monotonic trend exists.

3. Either an upward monotonic trend or a downward monotonic trend exists.

To test this, a test statistic is derived based on the difference between two observations. From this, it is possible to distinguish the three situations; upward, downward or either upward or downward. This is formulated as $sgn(x_i - x_j)$:

$$sgn(x_i - x_j) = \begin{cases} 1, & x_i - x_j > \epsilon \\ 0, & |x_i - x_j| \leq \epsilon \\ -1, & x_i - x_j < -\epsilon \end{cases} \quad (2.6)$$

With x_i and x_j for the observations at time i or j respectively. ϵ is the least count error, which allows small differences. Now, $sgn(x_i - x_j)$ can be substituted in the function S :

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n sgn(x_j - x_i) \quad (2.7)$$

From this function both the mean ($E[S]$) and the variance ($VAR(S)$) can be calculated with a method described by Gilbert (1987). Both $E[S]$ and $VAR(S)$ are needed to calculate the test statistic:

$$Z_{MK} = \begin{cases} \frac{E[S]-1}{\sqrt{VAR(S)}}, & E[S] > \epsilon \\ 0, & |E[S]| \leq \epsilon \\ \frac{E[S]+1}{\sqrt{VAR(S)}}, & E[S] < -\epsilon \end{cases} \quad (2.8)$$

This test equation is used to test whether a variant of H_a has to be rejected or not. Each variant is tested separately at a significance level $\alpha \leq 0.05$ (Krzywinski and Altman, 2013):

- H_a : An upward monotonic trend exists
If $Z_{MK} \geq Z_{1-\alpha}$ then H_a is not rejected.¹
- H_a : A downward monotonic trend exists
If $Z_{MK} \leq Z_{1-\alpha}$ then H_a is not rejected.
- H_a : Either an upward or a downward monotonic trend exists
If $|Z_{MK}| \geq Z_{1-\alpha/2}$ then H_a is not rejected.

2.4.2. Slope and intercept

The slope and the intercept are calculated based on an ordinary least squares approach:

$$x = \alpha + \beta \cdot t \quad (2.9)$$

With the intercept α and the slope β . Furthermore, x is the value of a certain data point (e.g. precipitation) and t is the variable time in years. The slope of a dataset of x with number of observations n is calculated with:

$$\beta = r_{x,t} \frac{\sigma_x}{\sigma_t} = \frac{n \sum tx - \sum t \sum x}{n \sum (t^2) - (\sum t)^2} \quad (2.10)$$

With $r_{x,t}$ as the Pearson's cross-correlation coefficient between x and t and standard deviations σ_x and σ_t . The intercept is calculated according to:

$$\alpha = \mu_x - \beta \mu_t = \frac{\sum x \sum (t^2) - \sum t \sum tx}{n \sum (t^2) - (\sum t)^2} \quad (2.11)$$

With μ_x and μ_t as the mean of both x and t .

¹With $Z_{1-\alpha}$ as the $100 - (1 - \alpha)$ -th percentile of the standard normal distribution

2.4.3. P-value

The estimation of the p-values is used as an extra validation of the trend in this context. The p-value describes the calculated probability of the observed and is separately estimated with the probability-density function $f(u)$ for every variation of H_a :

- H_a : An upward monotonic trend exists:

$$P_{Z_{MK}} = \begin{cases} \int_{Z_{MK}}^{\infty} f(u) du, & |E[S]| > \epsilon \\ 0.5, & |E[S]| \leq \epsilon \end{cases}$$

- H_a : A downward monotonic trend exists:

$$P_{Z_{MK}} = \begin{cases} \int_{-\infty}^{Z_{MK}} f(u) du, & |E[S]| > \epsilon \\ 0.5, & |E[S]| \leq \epsilon \end{cases}$$

- H_a : Either an upward or a downward monotonic trend exists:

$$P_{Z_{MK}} = 0.5 \begin{cases} \int_{Z_{MK}}^{\infty} f(u) du, & E[S] > \epsilon \\ 1, & |E[S]| \leq \epsilon \\ \int_{-\infty}^{Z_{MK}} f(u) du, & E[S] < -\epsilon \end{cases}$$

2.4.4. Anomalies

In addition to the linear regression, anomalies are plotted to visualize the dynamic behaviour of the trends. For this study, the anomalies are deviations from the mean values until 2010. The mean until 2010 is chosen to link this research to other climate related research, where often the mean between 1981 and 2010 is chosen to calculate anomalies. Furthermore, the World Meteorological Organization suggests to use climate normals for either 1961-1990 or 1981-2010 (WMO, 2017). Since the input data is annually, the anomalies are also annually.

2.5. Multivariate regression

To draw conclusion about future behaviour of SFD it is necessary to make predictions. For these predictions it is necessary to have long term datasets. However, this is not the case for snow-cover from MODIS. To avoid this problem, it is necessary to relate SFD to parameters with long term datasets (larger than 30 years). For choosing these parameters it is important to understand the snow cover dynamics. For this research, it is assumed that T_w and P_w are the only parameters which describe SFD. The most straightforward way to describe SFD with these parameters is with a multivariate regression function ². The multivariate function for this case can be expressed as:

$$SFD = \alpha + \beta_{T_w} \cdot T_w + \beta_{P_w} \cdot P_w \quad (2.12)$$

With SFD as the amount of snow-free days, α as the intercept, β_{T_w} as the coefficient related to the mean winter temperature (T_w) and β_{P_w} as the coefficient related to the accumulated winter precipitation (P_w). Within this multivariate regression function, several assumptions are made:

- T_w and P_w are independent at all time steps
- T_w and P_w are assumed as the only variables of influence on SFD. One can think about solar radiation as another variable with influence on SFD, but this is strongly dependent on the direction of the slope. Since a catchment is analyzed, with both North- and South-facing slopes, it will not add more information to the model.
- T_w and P_w are useful input variables at all elevations. Which means that both variables add enough information to describe SFD at all elevations. This is also tested with the calculation of the coefficient of determination (Equation 2.15).

²A similar method is successfully applied by Sospedra-Alfonso et al. (2015) to relate snow-water equivalent (SWE) to temperature and precipitation.

To see the influence of the variables (T_w and P_w) two other models are created with the variables separately:

$$\text{SFD}_{T_w} = \alpha + \beta_{T_w} \cdot T_w \quad (2.13)$$

$$\text{SFD}_{P_w} = \alpha + \beta_{P_w} \cdot P_w \quad (2.14)$$

Next, Equations 2.12, 2.13 and 2.14 are compared with the observed snow free days (SFD_{obs}). SFD_{obs} is obtained with the RSLE-method as shown in section 2.3.3. The comparison is executed by calculation of the coefficient of determination for the three models:

$$\begin{aligned} R_{\text{SFD}}^2 &= \frac{\sum_{i=1}^n (\text{SFD}_i - \overline{\text{SFD}_{\text{obs}}})^2}{\sum_{i=1}^n (\text{SFD}_{\text{obs},i} - \overline{\text{SFD}_{\text{obs}}})^2} \\ R_{\text{SFD}_T}^2 &= \frac{\sum_{i=1}^n (\text{SFD}_{T,i} - \overline{\text{SFD}_{\text{obs}}})^2}{\sum_{i=1}^n (\text{SFD}_{\text{obs},i} - \overline{\text{SFD}_{\text{obs}}})^2} \\ R_{\text{SFD}_P}^2 &= \frac{\sum_{i=1}^n (\text{SFD}_{P,i} - \overline{\text{SFD}_{\text{obs}}})^2}{\sum_{i=1}^n (\text{SFD}_{\text{obs},i} - \overline{\text{SFD}_{\text{obs}}})^2} \end{aligned} \quad (2.15)$$

With R^2 as the coefficient of determination. Equation 2.15 gives the opportunity to quantify and compare the differences in SFD, SFD_T and SFD_P . Moreover, the simple linear regression models obtained in Equations 2.12, 2.13 and 2.14 are also useful input for different scenario's.

2.6. Scenarios

In the previous, the multivariate regression formula (Equation 2.12) is introduced and used to analyze the influence of precipitation and temperature. The same formula can also be used for scenarios; e.g. the analysis of an extreme event or to predict future effects. The selection of scenarios is introduced in Section 2.6.1. Secondly, the input variables per scenario are introduced and discussed.

2.6.1. Scenario selection

Effect of the California Drought

In the previous the whole dataset from 2000 until 2017 was used to visualize the overall trends of SFD. However, the extreme drought period is also incorporated in this. To see whether the drought influences the total trend of the SFD. Furthermore, it is interesting to see which parameter (T_w , P_w or a combination of the two) is of largest influence during a long drought period. According to Flint et al. (2018), the severe environmental effects of the California Drought are difficult to recover by nature due to the length and the temperature. However, the snow cover is not incorporated in this analysis. For that, this scenario can be an addition to these conclusions. Since the multivariate formula describes SFD related to P_w and T_w this can easily be implemented by manipulating these two variables between 2012 and 2016. The methodology for this is discussed in Section 2.6.2. In summary, the purpose of performing a scenario related to the California Drought is twofold:

- Analyzing the effects of the California Drought on SFD and whether this is corresponding with the conclusions of Flint et al. (2018)
- Investigation of the separate effects of T_w and P_w on SFD during the California Drought

Extrapolation until 2025

The second scenario is related to the investigation of future effects. This is done to analyze the development of SFD in the future. For that, the purpose is not to predict a fixed number of snow-free days, but to obtain general understanding about the development and the order of magnitude per elevation zone. The trends of P_w and T_w found with the methodology of Section 2.4 can be used as inputs to calculate future SFD. Another option is to consider the trends from the observed SFD and extrapolate this. The application of both methods can give valuable information on the performance of the multivariate model and the effect of P_w and T_w on SFD per elevation. For this scenario extrapolation until 2025 is used. This choice is made because the focus is on the change of SFD for the near future, which is in this context not more than 10 years. It is important to understand that these outcomes are not one-to-one transferable for the whole Sierra Nevada or as an prediction for this catchment. For this scenario, the purpose is twofold:

- Analyzing the development of modeled SFD (with inputs T_w and P_w) and extrapolated SFD
- Quantification of SFD per elevation zone in 2025 as an indication for a future perspective on SFD for the Merced River basin.

2.6.2. Variables

Effect of the Californian Drought

In the drought scenario, equation 2.12 is applied to calculate SFD for different situations in the MODIS timespan(2001-2017). The data sets of the winter temperature and winter precipitation inputs (T_w and P_w) are divided in two periods, respectively 2001-2010 and 2010-2017. The observed T_w and P_w are used in the period 2001-2010 for every combination. For the period 2010-2017, four combinations are considered:

- SFD obtained with the observed T_w and P_w
- $SFD_{\overline{P_w}}$ obtained with the observed T_w and the mean winter precipitation until 2010($\overline{P_w}$)
- $SFD_{\overline{T_w}}$ obtained with the mean winter temperature until 2010($\overline{T_w}$) and the observed P_w
- \overline{SFD} obtained with $\overline{T_w}$ and $\overline{P_w}$

This can be mathematically expressed with four expressions of the multivariate model:

$$\begin{aligned}
 SFD &= \alpha_1 + \beta_1 \cdot T_w + \beta_2 \cdot P_w \\
 SFD_{\overline{P_w}} &= \alpha_2 + \beta_3 \cdot T_w + \beta_4 \cdot \overline{P_w} \\
 SFD_{\overline{T_w}} &= \alpha_3 + \beta_5 \cdot \overline{T_w} + \beta_6 \cdot P_w \\
 \overline{SFD} &= \alpha_4 + \beta_7 \cdot \overline{T_w} + \beta_8 \cdot \overline{P_w}
 \end{aligned} \tag{2.16}$$

The subscripts for α and β indicate that these are unique combinations. The combinations make it possible to show whether T_w or P_w is the dominant driver of the change of SFD in this particular period. The SFD is calculated for elevation zones of 100 meter (see section 2.3.3), thus these scenario's are also obtained for every elevation zone.

The above will generate four different time series of SFD. However, the trends are more useful in quantifying differences in a particular year. To obtain these trends, a linear regression is applied. This is done in a similar way as explained in section 2.4 with equation 2.10. For $t \leq 2010$, the trend line is similar for all the combinations according to:

$$y_{SFD} = \alpha + \beta \cdot t \tag{2.17}$$

With the time t in years and α and β expressed as:

$$\alpha = \frac{\sum(SFD) \sum(t^2) - \sum t \sum(t \cdot SFD)}{n \sum(t^2) - (\sum t)^2} \tag{2.18}$$

$$\beta = \frac{n \sum(t \cdot SFD) - \sum t \sum SFD}{n \sum(t^2) - (\sum t)^2} \tag{2.19}$$

By calculation of α and β the observed values of SFD in the period 2001-2010 are used. Thus, $2001 \leq t \leq 2010$ and $n = 10$. Next, the slopes for the four combinations can be calculated for $2010 \leq t \leq 2017$, according to:

$$\begin{aligned}
 \beta_{SFD} &= \frac{n \sum(t \cdot SFD) - \sum t \sum SFD}{n \sum(t^2) - (\sum t)^2} \\
 \beta_{\overline{P_w}} &= \frac{n \sum(t \cdot SFD_{\overline{P_w}}) - \sum t \sum SFD_{\overline{P_w}}}{n \sum(t^2) - (\sum t)^2} \\
 \beta_{\overline{T_w}} &= \frac{\sum(t \cdot SFD_{\overline{T_w}}) - \sum t \sum SFD_{\overline{T_w}}}{n \sum(t^2) - (\sum t)^2} \\
 \beta_{\overline{SFD}} &= \frac{\sum(t \cdot \overline{SFD}) - \sum t \sum \overline{SFD}}{n \sum(t^2) - (\sum t)^2}
 \end{aligned} \tag{2.20}$$

With $2010 \leq t \leq 2017$ and $n = 8$. Now, the trend line can be obtained for the same period. The starting point is assumed as the SFD value in 2010, which is determined with Equation 2.17, according to:

$$y_{2010} = \alpha + \beta \cdot 2010 \quad (2.21)$$

With this starting point, the four combinations for the trend line of $2010 \leq t \leq 2017$ can be expressed as:

$$\begin{aligned} y_{\text{SFD}} &= y_{2010} + \beta_{\text{SFD}} \cdot t \\ y_{\text{SFD}_{\overline{P_w}}} &= y_{2010} + \beta_{\text{SFD}_{\overline{P_w}}} \cdot t \\ y_{\text{SFD}_{\overline{T_w}}} &= y_{2010} + \beta_{\text{SFD}_{\overline{T_w}}} \cdot t \\ y_{\overline{\text{SFD}}} &= y_{2010} + \beta_{\overline{\text{SFD}}} \cdot t \end{aligned} \quad (2.22)$$

The combinations in Equation 2.22 can be used to discover the absolute difference in SFD in 2017 due to different T_w and P_w inputs.

Extrapolation until 2025

The extrapolation scenario is created by use of linear extrapolation. This is done with four combinations:

- SFD obtained with linear extrapolation of the slope of the observed SFD dataset.
- $\text{SFD}_{\overline{T_w} + \hat{P_w}}$ obtained with the mean of T_w (over the whole T_w dataset) and a predicted value of P_w based on the slope of the P_w dataset as inputs.
- $\text{SFD}_{\hat{T_w} + \overline{P_w}}$ obtained with the predicted value of T_w based on the slope of the T_w dataset and the mean value of P_w as inputs
- $\text{SFD}_{\hat{T_w} + \hat{P_w}}$ obtained with the predicted values of T_w and P_w as inputs.

Which can be mathematically written as:

$$\begin{aligned} \text{SFD} &= \alpha_1 + \beta_1 \cdot t \\ \text{SFD}_{\overline{T_w} + \hat{P_w}} &= \alpha_2 + \beta_2 \cdot \overline{T_w} + \beta_3 \cdot \hat{P_w} \\ \text{SFD}_{\hat{T_w} + \overline{P_w}} &= \alpha_3 + \beta_4 \cdot \hat{T_w} + \beta_5 \cdot \overline{P_w} \\ \text{SFD}_{\hat{T_w} + \hat{P_w}} &= \alpha_4 + \beta_6 \cdot \hat{T_w} + \beta_7 \cdot \hat{P_w} \end{aligned} \quad (2.23)$$

Again, the subscripts of α and β indicate unique combinations. Besides, these combinations are able to show whether T_w or P_w are dominant in the future, based on the data of the period 2001-2017. Note that these values of SFD are also calculated for every elevation zone of 100 meter. Next, the trend lines can be determined in the same way as the previous section. In this case, the whole dataset is observed, which is period 2001-2017. For this period, the intercept and slope are similar to Equation 2.21 for all combinations. However, y_{SFD} , α and β are now calculated with $2001 \leq t \leq 2017$ and $n = 17$. Secondly, the values for β are also found in the same way, now for $2017 \leq t \leq 2025$ and $n = 9$:

$$\begin{aligned} \beta_{\text{SFD}} &= \frac{n \sum (t \cdot \text{SFD}) - \sum t \sum \text{SFD}}{n \sum (t^2) - (\sum t)^2} \\ \beta_{\overline{T_w} + \hat{P_w}} &= \frac{n \sum (t \cdot \text{SFD}_{\overline{T_w} + \hat{P_w}}) - \sum t \sum \text{SFD}_{\overline{T_w} + \hat{P_w}}}{n \sum (t^2) - (\sum t)^2} \\ \beta_{\hat{T_w} + \overline{P_w}} &= \frac{\sum (t \cdot \text{SFD}_{\hat{T_w} + \overline{P_w}}) - \sum t \sum \text{SFD}_{\hat{T_w} + \overline{P_w}}}{n \sum (t^2) - (\sum t)^2} \\ \beta_{\hat{T_w} + \hat{P_w}} &= \frac{\sum (t \cdot \text{SFD}_{\hat{T_w} + \hat{P_w}}) - \sum t \sum \text{SFD}_{\hat{T_w} + \hat{P_w}}}{n \sum (t^2) - (\sum t)^2} \end{aligned} \quad (2.24)$$

At last, the values of β can be substituted in the trend line formula:

$$\begin{aligned} y_{\text{SFD}} &= y_{2017} + \beta_{\text{SFD}} \cdot t \\ y_{\text{SFD}_{\overline{T_w} + \hat{P_w}}} &= y_{2017} + \beta_{\overline{T_w} + \hat{P_w}} \cdot t \\ y_{\text{SFD}_{\hat{T_w} + \overline{P_w}}} &= y_{2017} + \beta_{\hat{T_w} + \overline{P_w}} \cdot t \\ y_{\text{SFD}_{\hat{T_w} + \hat{P_w}}} &= y_{2017} + \beta_{\hat{T_w} + \hat{P_w}} \cdot t \end{aligned} \quad (2.25)$$

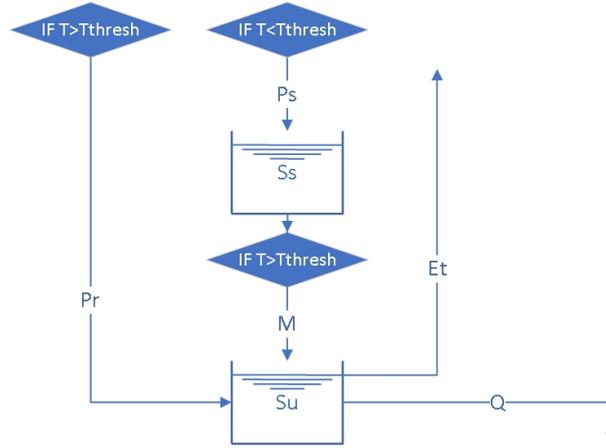


Figure 2.8: The modeling scheme with the storage reservoirs (blocks), fluxes (arrows) and conditions (diamonds).

With

$$y_{2017} = \alpha + \beta \cdot 2017 \quad (2.26)$$

The equations are for $2017 \leq t \leq 2025$ and $n = 9$. Similar as for the drought scenario, this method generates one trend line before 2017. The trend line will split in four separate trend lines after 2017, which shows an absolute difference between the four cases in 2025. In this way, it is possible to compare the temperature and precipitation influences until 2025.

2.7. Hydrological modeling

In the previous sections, the data was analyzed and related to SFD. However, the aim of this research is to discover a relation between SFD and root-zone moisture capacity. This can be done by using a conceptual hydrological model, as used by Gao et al. (2014), Nijzink et al. (2016) and de Boer-Euser et al. (2016). The modeling scheme with its reservoirs and fluxes is explained in section 2.7.1. The reservoirs will be clarified separately thereafter.

2.7.1. Modeling scheme

The hydrological model consists of two reservoirs, respectively a snow storage (S_s) and the unsaturated zone (S_u). Precipitation is the input, which is divided in rain (P_r) and snowfall (P_s). The division is based on the temperature threshold defined in section 2.3.2. The same threshold is to determine melt from the snow storage to the unsaturated zone (M), which implies that snow fall and melt can never occur at the same time. Furthermore, transpiration (E_t) from the unsaturated zone and discharge (Q) is included. No interception or deeper ground water reservoirs are included. The scheme is shown in figure 2.8.

2.7.2. Snow storage

The snow storage represents the water in the snow cover areas of the catchment according to:

$$S_s(t) = S_{s_{\text{initial}}} + P_s dt - M dt \quad (2.27)$$

Where the initial snow storage is defined as S_{s_{initial}}. The input is snow fall (P_s), which is defined as the precipitation when the temperature is lower than the threshold, as shown in section 2.3.2. The temperature is corrected with the elevation with the environmental lapse rate. This holds that the snow fall input is weighted on the areal contribution of the 100 meter elevation zone. This gives for the snow fall:

$$P_s dt = \begin{cases} 0, & T_h \geq T_{\text{thresh}} \\ \sum_{h=0}^{27} P dt \cdot WF_h, & T_h < T_{\text{thresh}} \end{cases} \quad (2.28)$$

With the 100 meter elevation zone number h , starting at 1300 meter. Furthermore, the recorded precipitation $P dt$, the areal weighing factor per elevation zone WF_h , the elevation zone related temperature T_h and the

threshold temperature T_{thresh} . Melt(M) occurs when the temperature exceeds the temperature threshold and the snow storage is larger than zero. Obviously, melt depends on the elevation because of the temperature, so this flux is also weighted on the areal contribution of the elevation zone. The rate of melting is dependent on the difference from the temperature threshold and a calibration parameter, which is called the degree-day factor (Martinec, 1975). With this information the melt flux per time step can be expressed as:

$$Mdt = \begin{cases} \min\left(\sum_{h=0}^{27} (\text{DDF} \cdot (T_h - T_{\text{thresh}}) \cdot \text{WF}_h), S_s(t)\right), & T_h \geq T_{\text{thresh}} \\ 0, & T_h < T_{\text{thresh}} \end{cases} \quad (2.29)$$

With the 100 meter elevation zone number h , starting at 1300 meter. Furthermore, the degree-day factor DDF and the snow storage at t time t defined as $S_s(t)$. The snow fall and melt cannot occur simultaneously in this model and the output is only generated by melt. This means that no sublimation effects are taken into account.

2.7.3. Unsaturated zone

In the hypothesis of the root-zone moisture capacity it is assumed that plants adjust their roots to changing climatic conditions (Collins and Bras, 2007; Milly, 1994; Schenk, 2008). This process occurs in the unsaturated zone. The changing climatic conditions can be found by calculating the soil moisture deficit, which is defined as the difference between transpiration and effective precipitation. The effective precipitation (P_e) is the input of the unsaturated reservoir, which is for this case rainfall and melt:

$$\text{SMD}(t) = E_t dt - P_e = E_t dt - P_r dt - Mdt \quad (2.30)$$

The rainfall is defined as the precipitation when the temperature is higher than the temperature threshold. This gives:

$$P_r dt = \begin{cases} \sum_{h=0}^{27} P dt \cdot \text{WF}, & T_h \geq T_{\text{thresh}} \\ 0, & T_h < T_{\text{thresh}} \end{cases} \quad (2.31)$$

To estimate the transpiration (E_t), the long-term water balance of the unsaturated reservoir is applied in the same way Nijzink et al. (2016) incorporated this. This long-term water balance is expressed as:

$$\overline{E_t} = \overline{P_r + M} - \overline{Q} \quad (2.32)$$

With the long term actual transpiration $\overline{E_t}$, long-term mean effective precipitation $\overline{P_r + M}$ and the long-term discharge \overline{Q} .

Next, the transpiration is scaled with the ratio of long-term mean daily potential evaporation E_p over the mean annual potential evaporation $\overline{E_p}$:

$$E_t(t) = \frac{\sum_{h=0}^{27} E_{p_h}(t)}{\sum_{h=0}^{27} E_{p_h}} \cdot \overline{E_t} \quad (2.33)$$

The potential evaporation $\sum_{h=0}^{27} E_{p_h}(t)$ and the long-term mean of the potential evaporation $\overline{\sum_{h=0}^{27} E_{p_h}}$ are derived with the method of Hargreaves and Samani (1985) according to:

$$E_{p_h} = c_H \cdot 0.408 R_0 \cdot (T_h + 17.8) \sqrt{T_{h_{\text{max}}} - T_{h_{\text{min}}}} \cdot \text{WF}_h \quad (2.34)$$

Where c_H is the Hargreaves constant equal to 0.0023 compensating for the errors. Furthermore R_0 is the extraterrestrial radiation, which is calculated based on the location of the weather station and the date. T , T_{max} and T_{min} are respectively the daily average, maximum and minimum temperature adjusted for the average elevation of the catchment with the environmental lapse rate according to equation 2.1. Thus this flux is also corrected with areal weighing factor WF_h . This explains also the addition of the subscripts to the potential evaporation.

Now, the soil moisture deficit can be calculated with the addition of the transpiration in equation 2.30, which will generate a time series over the period 1983-2017. From this, the maximum soil moisture deficit is calculated over a rolling window with an interval of 10 years. This return period is chosen based on the previous work on this topic, where a return period between 10-20 years is applied. (Gao et al., 2014; de Boer-Euser et al., 2016; Nijzink et al., 2016). By using a rolling maximum, it is easy to visualize the gradual changes in the root zone.

2.7.4. Calibration

Only one calibration parameter is considered for this model, DDF. This parameter is calibrated based on the RSLE. First, the snow reservoirs at a certain elevation which contain snow are separated from the empty snow reservoirs. This value is the modeled snow reservoir, $S_{sh_{mod}}$:

$$S_{sh_{mod}} = \begin{cases} 1, & S_{sh} > 0 \\ 0, & S_{sh} = 0 \end{cases} \quad (2.35)$$

S_s is defined as the snow storage and the subscript h refers to the elevation zone. So, the above distinguishes whether a modelled snow storage reservoir is empty or filled with snow. This modelled snow storage has to be compared with observed snow storage which is directly linked with the RSLE. Snow is present in an elevation zone when the RSLE is lower than the lower boundary of the elevation. This is called the observed snow reservoir, $S_{sh_{obs}}$:

$$S_{sh_{obs}} = \begin{cases} 1, & RSLE < h - 100 \\ 0, & RSLE \geq h - 100 \end{cases} \quad (2.36)$$

Now, an objective function for testing can be constructed. This is a maximization problem:

$$\max \left(1 - \frac{\sum_{t_0}^{t_{end}} \sum_{h=0}^{27} (S_{sh_{obs}} - S_{sh_{mod}})}{(t_{end} - t_0) \cdot 27} \right) \quad (2.37)$$

With the timespan 1983-2017 expressed in days with t_0 and t_{end} . The elevation zone is expressed as h . The modelled snow reservoir, $S_{sh_{mod}}$, depends on DDF. So, DDF can be estimated with a Monte Carlo simulation based on 1000 random samples. The outcomes are evaluated with the objective function (equation 2.37) and the related optimal DDF is found.

2.7.5. Analysis

The outputs of the hydrological model are analyzed with a focus on different aspects: soil moisture deficit calculation, the calculation of SFD with the hydrological modeling (inverse modeling) and the comparison of the snow storage before and after 2010. The soil moisture deficit calculation is already explained in Subsection 2.7.3 by visualizing the soil moisture deficit with a rolling maximum over an interval of 10 years. Furthermore, the hydrological modeling part is analyzed by calculating the SFD with the hydrological model and subsequently compare this with the observed SFD from the RSLE method (Section 2.3.3). This will give more insight in the performance of the hydrological model compared to the observed SFD. Lastly, a small scenario is executed related to the snow storage reservoir. The outputs of the model are separated in a part from before and after 2010. For both parts the daily means are calculated for the snow reservoir at different elevation zones and for the total amount of snow. From this figures, one can compare the starting days of the accumulation period, slopes of the accumulation period, the last days of the melting period and the slopes of the melting periods. To visualize this in one figure per elevation zone, the effect of the drought on these parameters can be discovered.

2.8. Conclusion

In this chapter the different methodologies are discussed. First, the general approach, study area and its related data sources were discussed. The available data was the input for trend analysis and a multivariate regression model. Both the trend analysis as the multivariate model were the basis for scenarios, which were introduced and discussed in Section 3.2.2. At last, the hydrological model is introduced as link between the analysis of the SFD and the root-zone storage moisture capacity. The same order will be maintained in the next chapter, where the results of this methodology will be presented.

3

Results

In this chapter, the results obtained by the statistical analysis, multivariate regression analysis and the hydrological model will be discussed. The results of the trends are the long term trends of P_w , T_w and SFD obtained with the methodology described in Section 2.4. These trends are presented first to see the general trends happening in the area of interest. Furthermore, the aim of incorporating the T_w and P_w trends is to discover their influence on SFD and whether this is elevation dependent. The relation between SFD, P_w and T_w is obtained with the multivariate regression analysis, which will be discussed in Section 3.2 based on calculation of the R^2 and two different scenarios related to the trends found in Section 2.4. The last part of this chapter contains the discussion of the hydrological modeling outputs. These outputs are threefold: calculation of the soil moisture deficit, differences in snow storage per elevation and the inverse modeling part where a long term data set for SFD is calculated.

3.1. Trends

As a start of the research, it is important to know whether the temperature, precipitation and snow-free days are changing. To discover trends for the observed data a statistical analysis is conducted. The observed data of interest is the accumulated winter precipitation (P_w) and mean winter temperature (T_w) from the weather station located at Yosemite Park HQ and the snow-free days (SFD) from the MODIS satellite product. P_w and T_w are used as input in the multivariate regression model and in the hydrological model. For that, it is interesting to analyze the trends separately beforehand. Moreover, trends of P_w and T_w are important inputs when extrapolation is used in the scenarios. The trends of the SFD are used to validate the hydrological model and to discuss the future perspective on snow cover in the study area.

3.1.1. Winter precipitation and temperature

The trends of P_w and T_w are visualized with scatter plots and anomalies (Figures 3.1 and 3.2).

Looking at the scatter plots, there is no significant upward trend for T_w observed. This is contrary to the expectation, since the increase of temperature is already observed at several other places in California (Cayan et al., 2007). Without any trend in the T_w data this seems useless, however the large upward extremes (anomalies of 2 °C in 2014 and 3 °C in 2015) are interesting to analyze. These extremes are found in the drought period, which means that the drought period is also visible as a larger temperature in these years. With these extremes it is possible to distinguish the effect of temperature in the multivariate model.

For P_w , a clear downward trend is observed with a slope of almost -11 mm/year. The anomalies show the extent of the two most recent drought periods, respectively 2007-2009 and 2012-2016. Interestingly, the first drought period has large anomalies over -400 mm. The second period is extreme related to the extent of the period of four years. However, the anomalies between -100 and -300 mm are not as spectacular as the previous. According to this two sorts of droughts are distinguished: a short very dry drought (2007-2009) and a long very warm drought (2012-2016). Flint et al. (2018) analyzed these two drought phenomena and found the long very warm drought period as the most severe for the hydrological system. This has to be also visible in the hydrological model later on.

3.1.2. Snow-free days per elevation zone

The SFD trends are based on the RSLE method from Section 2.3.3. The output of this method is similar to the climate data, with the addition of elevation zones. This gives the opportunity to show the SFD trends per

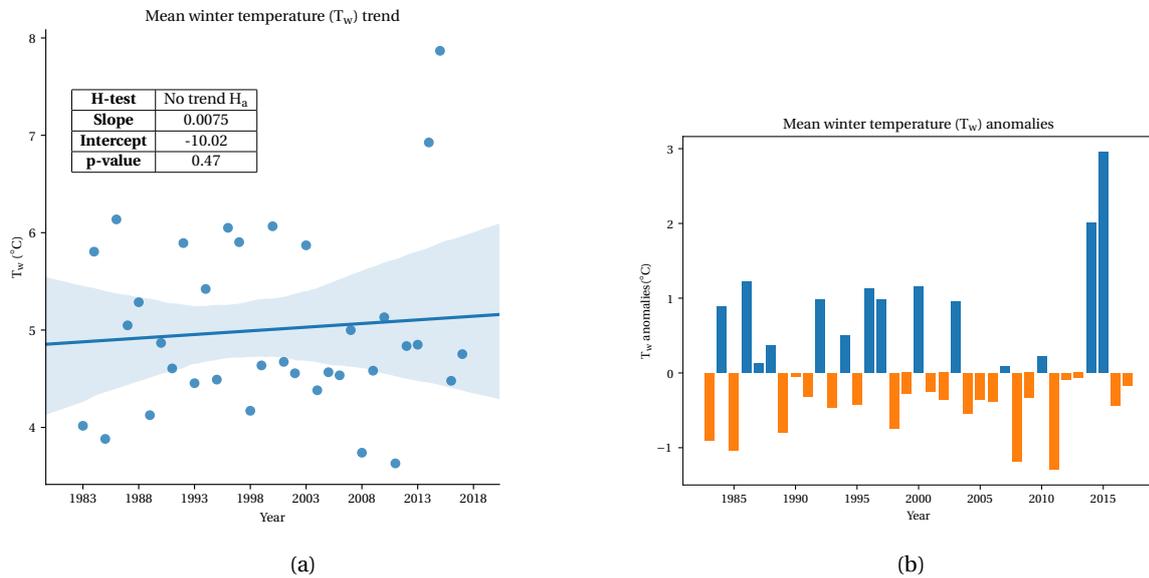


Figure 3.1: Mean winter temperature (T_w) trends at the weather station over the timespan 1983-2017. Figure (a) is a scatter plot with regression line and 95 % confidence interval as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.

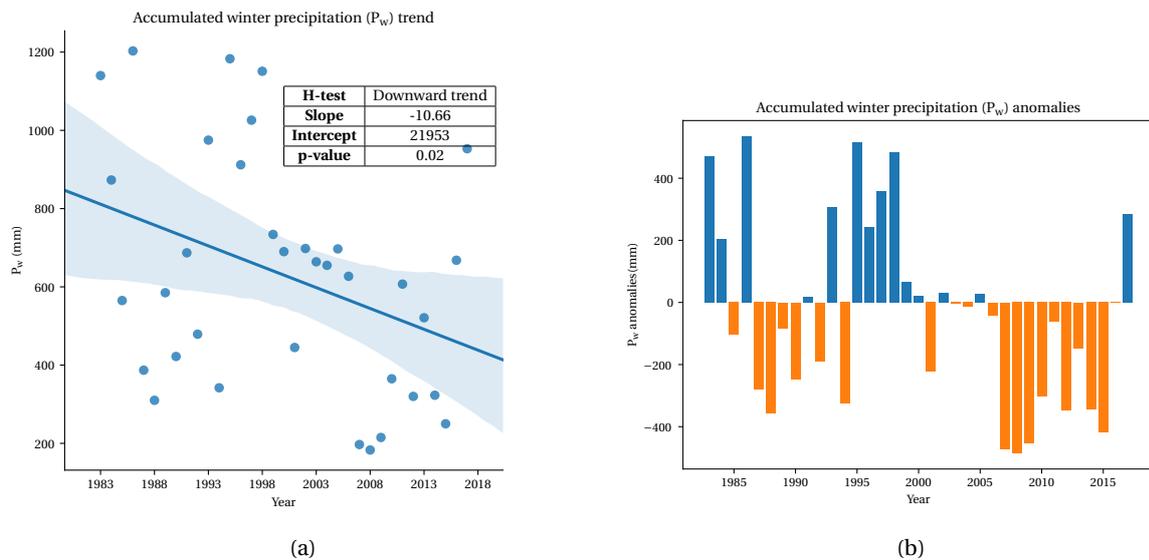


Figure 3.2: Accumulated winter precipitation (P_w) trends at the weather station over the timespan 1983-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.

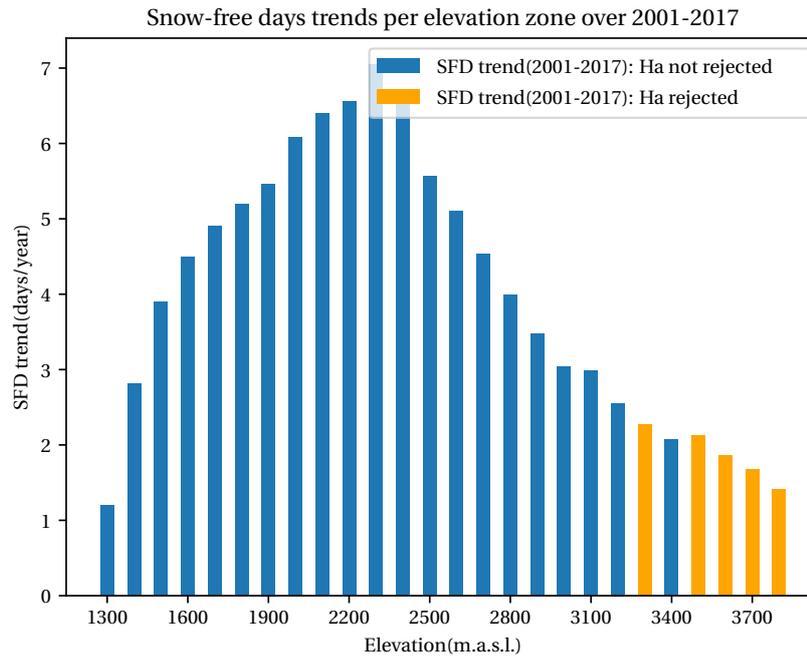


Figure 3.3: Trends of snow-free days per elevation over the MODIS timespan 2001-2017. The blue bars refer to significant trends according to the Mann-Kendall test and the orange bars show non significant trends.

elevation zone in one figure (Figure 3.3). Due to the timespan of the MODIS satellite product, trends are only calculated over 2001-2017. Figure 3.3 shows significant trends up to an elevation of 3400 m.a.s.l. There is a clear difference between elevation zones, with the maximum significant trend of 7 days per year at 2300 m.a.s.l. Furthermore, the lowest significant trend of 1.2 days per years is found at an elevation of 1300 m.a.s.l. and the lowest non-significant trend of 1.5 days per year is found at 3800 m.a.s.l. The figure gives a nice overview of the trends per elevation zone, but it is not directly clear how the SFD is developing over time. For that, the scatterplots and anomalies of 1300, 2300 and 3800 m.a.s.l. are presented in Figures 3.4, 3.5 and 3.6.

Figures 3.4 and 3.5 show significant but different trends. At 1300 m.a.s.l. the relative small upward trend can be related to the initial situation which consists of already more than 330 days. So initially, this elevation zone is not often covered with snow. However, the upward trend of 1.2 days per year shows that this elevation zone will be snow free during the whole year soon. Even more alarming are the developments in the areas which are often snow covered, like at 2300 m.a.s.l. Half of the year was snow-free at the start of the SFD time series at 2300 m.a.s.l. This is already changed to 300 snow-free days at 2300 m.a.s.l. in 2016. The anomalies (see Figure 3.5b) show the enormous impact of the drought period between 2012 and 2016. No significant trend was found at 3800 m.a.s.l. This can be explained by looking at the anomalies, Figure 3.6b. Compared to 1300 and 2300 m.a.s.l., the influence of the drought period is not so clear. The positive anomalies are slightly larger between 2012 and 2016, but in 2013 a small negative anomaly is found. Furthermore, the positive anomalies in the period 2000-2004 are relatively large (between 5 and 15 SFD) compared to the drought period (between 20 and 30). Due to this, it is not clear whether there is an upward trend in this high elevation zone. However, the maximum amount of snow-free days is found at the end of the drought period (2016) around 110 snow-free days. In summary, the significant trends at 1300 and 2300 m.a.s.l. are highly influenced by the drought period. This will eventually lead to snow-free years at an elevation of 1300 m.a.s.l. and a rapid decrease of the amount of snow at 2300 m.a.s.l., according to this analysis 7 days/year. The effect of the drought is less visible at 3800 m.a.s.l. and due to that no significant upward trend of SFD was found.

Similar as for the trends of P_w and T_w , the effect of the drought period is clearly visible in the lower elevation zones (below 3400 m.a.s.l.). This suggests that P_w and T_w can describe SFD well in the drought period below 3400 m.a.s.l. This hypothesis will be tested in Section 3.2.

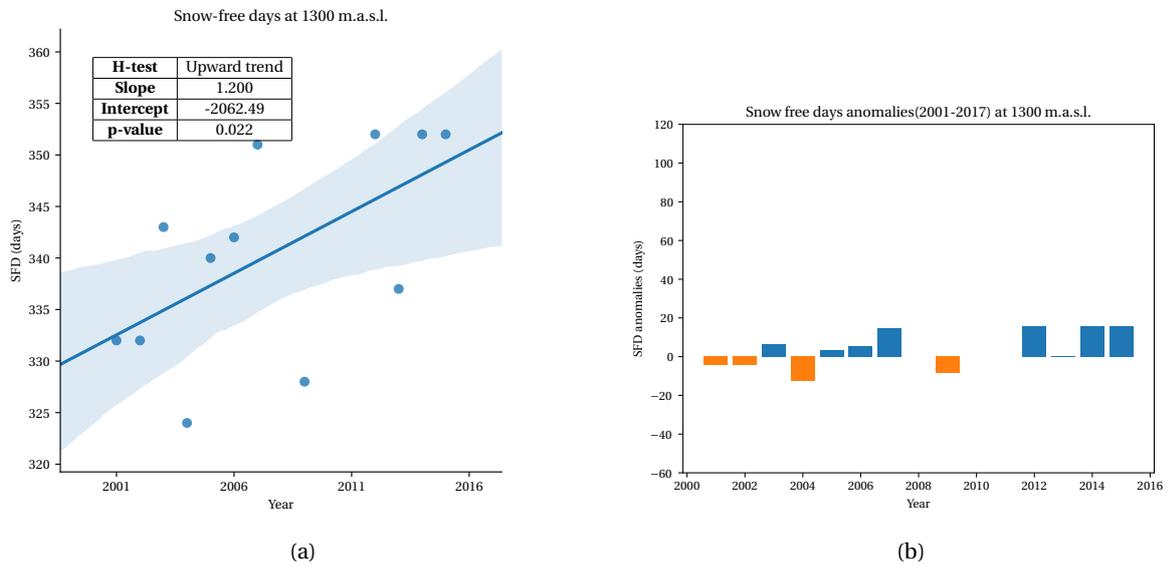


Figure 3.4: Snow-free days (SFD) trends at an elevation of 1300 m.a.s.l. over the timespan 2000-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.

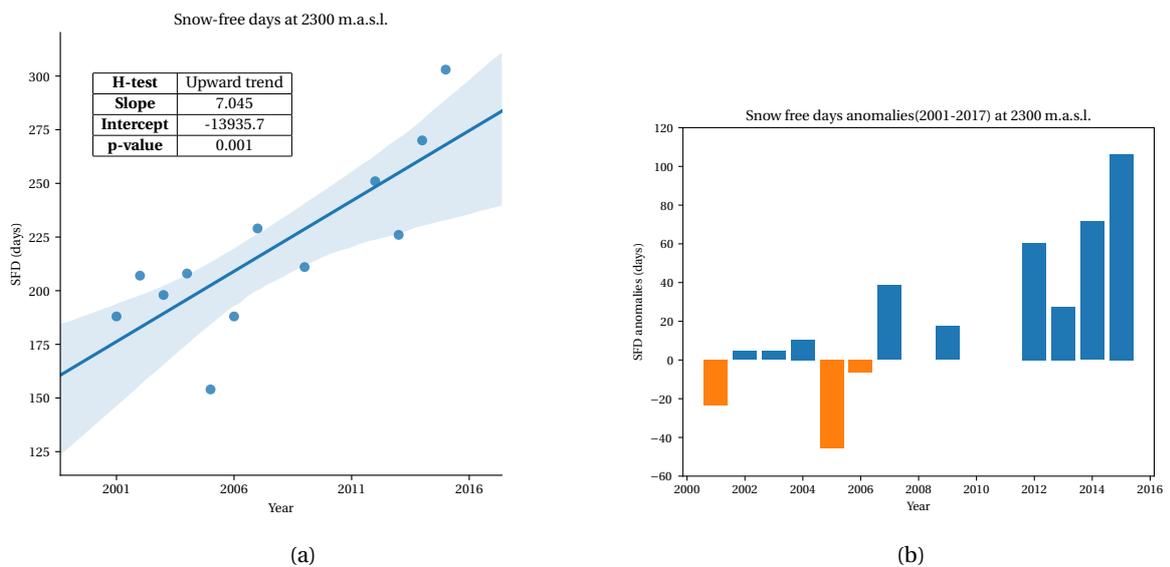


Figure 3.5: Snow-free days (SFD) trends at an elevation of 2300 m.a.s.l. over the timespan 2000-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.

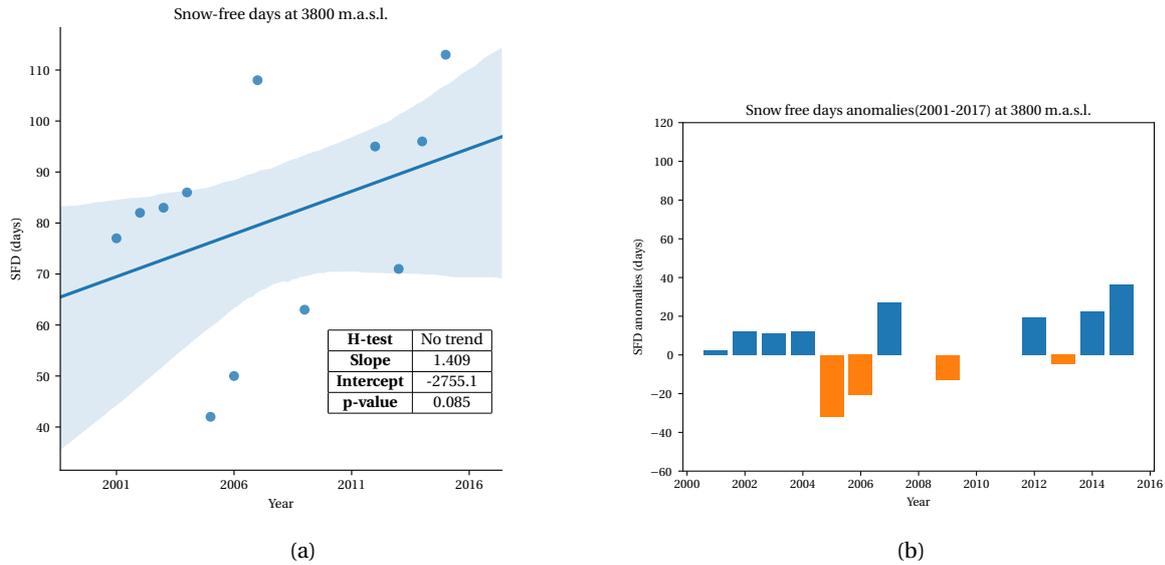


Figure 3.6: Snow-free days (SFD) trends at an elevation of 2300 m.a.s.l. over the timespan 2000-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.

3.1.3. Weighted snow-free days

As an addition, the elevation dependent results are weighted with the hypsometric curve (Figure 2.3) to show the average behaviour of SFD for the whole catchment. The results of this are shown in Figure 3.7.

According to the figure, the average trend in the catchment is positive, 4.18 days/year. This is not surprising since Figure 3.3 showed only positive trends. The other results of the Mann-Kendall test show that the trend is significant with a p-value of 0.003. This shows that the elevation zones between 2200 and 2500 m.a.s.l. have a large influence on the total SFD. These areas contribute for 25% to the catchment and have significant trends of more than 5 days/year (maximum around 7 days/year). The combination of area and a relative high trend shows that the amount of snow in this catchment is rapidly decreasing. This needs to be confirmed by the hydrological model.

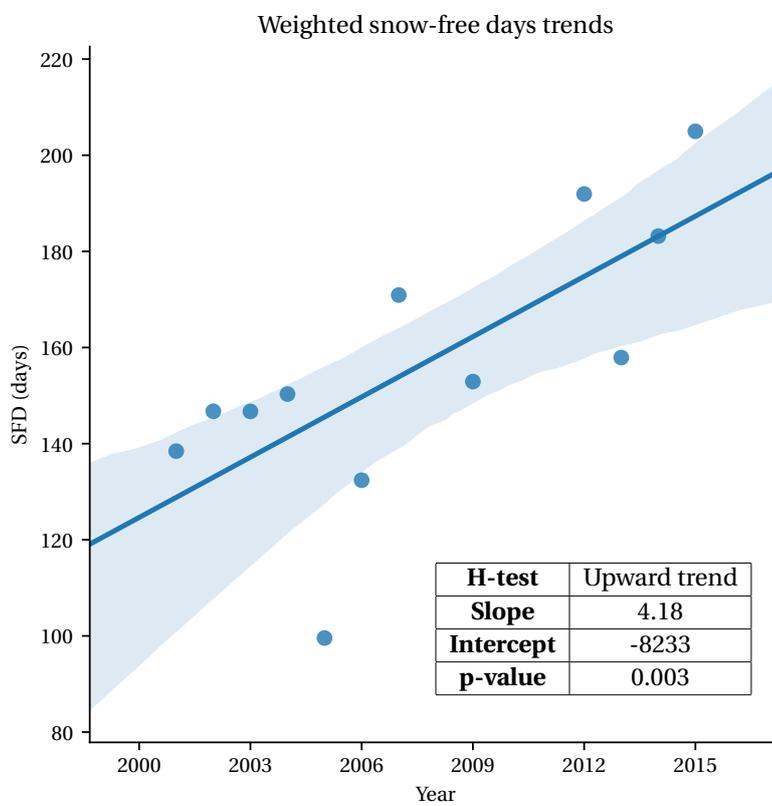


Figure 3.7: The weighted snow-free days over the Merced river basin. The SFD of all elevation zones are weighted on the hypsometric curve (Figure 2.3) with weighted SFD as result. The trend gives an indication of the average behaviour of SFD in the Merced River basin. The table shows the results of the Mann-Kendall test related to weighted SFD.

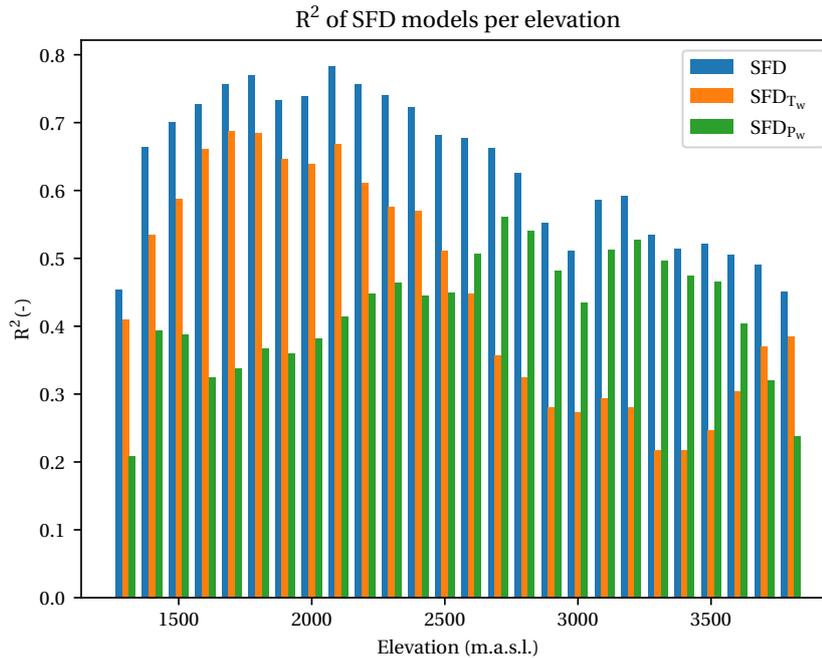


Figure 3.8: R^2 per elevation zone for different models of SFD. The blue bars show R^2 with both P_w and T_w as inputs, orange bars for SFD dependent on T_w only and green bars for SFD dependent on P_w only. These models are calculated with the method described in Section 2.5.

3.2. Relation between snow-free days, precipitation and temperature

To analyze the influence of P_w and T_w on SFD, a multivariate regression formula is created in Section 2.5. From this relation the R^2 can be calculated at different elevations and for different combinations of SFD, respectively:

- SFD calculated with P_w and T_w as inputs
- SFD_{P_w} calculated with only P_w as input
- SFD_{T_w} calculated with only T_w as input

This will enable to test the influence of P_w and T_w . To analyze the influence further, scenarios are created based on distinction of the drought period and extrapolation (Section 2.6).

3.2.1. Precipitation and temperature dependence

Since the temporal range of the SFD data is limited and in some periods highly inaccurate (due to clouds or failing detectors of the satellite) it is necessary to find a more robust way of calculating the SFD. This can be done by finding a relation between SFD, P_w and T_w . Figure 3.8 enables to test which parameter (P_w or T_w) is performing better at a certain elevation zone. The SFD with both parameters as input shows values of R^2 higher than 0.5 at elevation zones lower than 3300 m.a.s.l. (with exception of 1300 and 2900 m.a.s.l.). This is in line with the significant trends found in Section 3.1.2. The lower areas (between 1300 and 2500 m.a.s.l.) are well described by T_w inputs (SFD_{T_w}) with an maximum R^2 of 0.7 at 1700 m.a.s.l. For SFD_{P_w} it is not so clear, however above 2500 m.a.s.l. SFD_{P_w} performs better than SFD_{T_w} . In this region the maximum R^2 of SFD_{P_w} is equal 0.52 at 3200 m.a.s.l. The maximum difference between SFD_{T_w} and SFD_{P_w} above 2500 m.a.s.l. is found at 3300 m.a.s.l. where SFD_{T_w} performs poorly with a R^2 of 0.23. In the same region, SFD_{P_w} shows values of R^2 higher than 0.5 only at 2700, 2800, 3100 and 3200 m.a.s.l. From this, it is better to conclude that SFD_{T_w} performs worse than SFD_{P_w} above 2500 m.a.s.l. In summary, the calculation of SFD can be divided in a temperature influenced (between 1300 and 2500 m.a.s.l.) and non-temperature influenced (between 2600 and 3600 m.a.s.l.) part. To test whether this is correct, similar behaviour related to precipitation and temperature has to be observed in the scenarios.

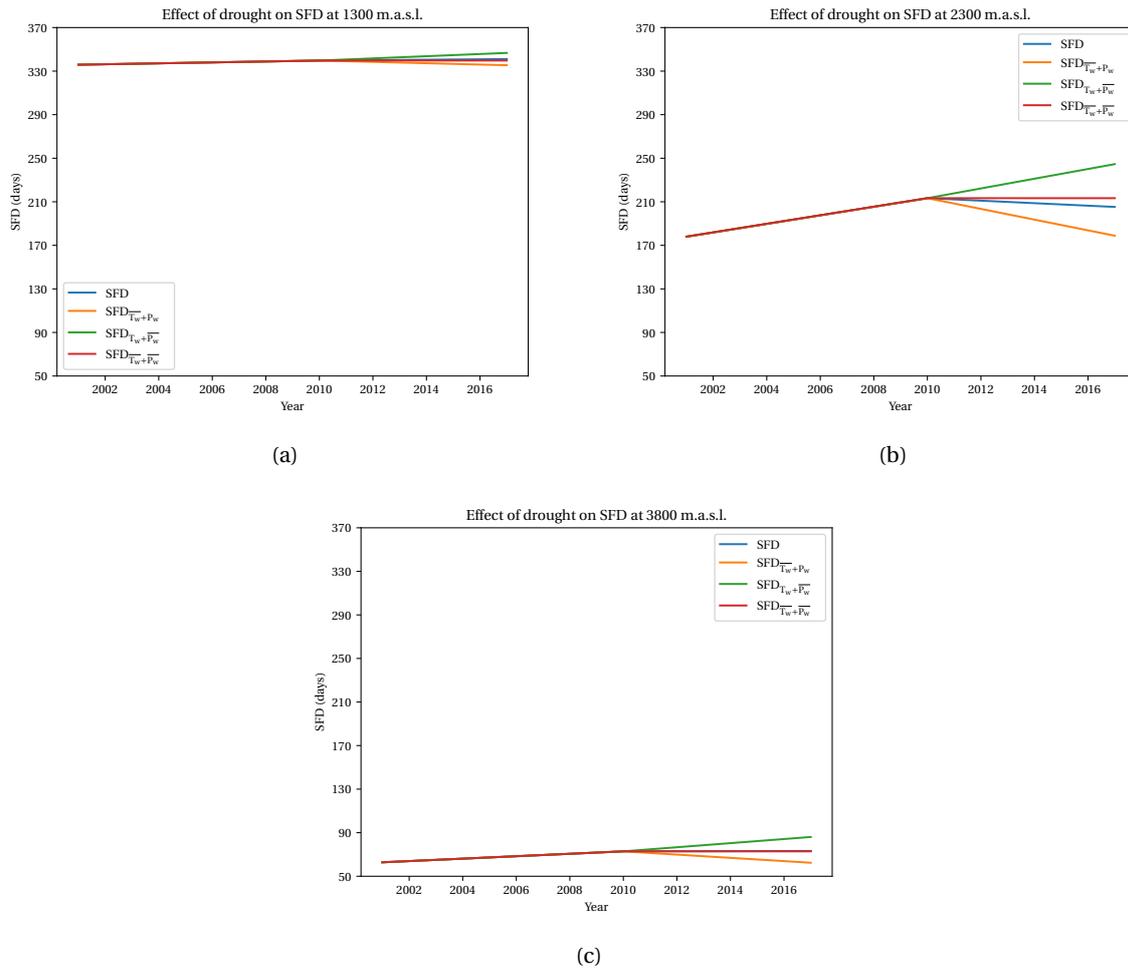


Figure 3.9: The effect of the drought period on snow-free days at elevations of 1300 (a), 2300 (b) and 3800 (c) m.a.s.l. The blue line shows the trend line of SFD during the drought with both P_w and T_w incorporated, orange shows the effect of P_w only, green show the effect of T_w only and red shows the trend line if there was no drought at all. The methodology behind this is described in Section 2.6.

3.2.2. Scenarios

Two scenarios are analyzed: the drought effect and extrapolation until 2025. The way of modelling is explained in Section 2.6.

California drought

Before the presentation of the results, the goal of this scenario is repeated as a reminder. The purpose of the implementation of this scenario is:

- Analyzing the effects of the California Drought on SFD and whether this is corresponding with the conclusions of Flint et al. (2018) (He states that the longest drought period has the most severe environment effects, however snow was not incorporated)
- Investigation of the separate effects of T_w and P_w on SFD during the California Drought

The scenario is implemented for every elevation, however according to Figure 3.3 the most interesting elevation zones are at 1300, 2300, 3800 m.a.s.l. These elevation zones are also considered in the results of the drought scenario (Figure 3.9). From the figures it can be seen that the maximum SFD is found at the green line at all elevations. This line describes the snow-free days with the temperature of the drought period and the mean precipitation. Thus, the temperature of the drought period causes the largest increase of SFD. This is also confirmed by the orange line which shows the lowest SFD at all elevations. This line describes the SFD without the temperature of the drought period.

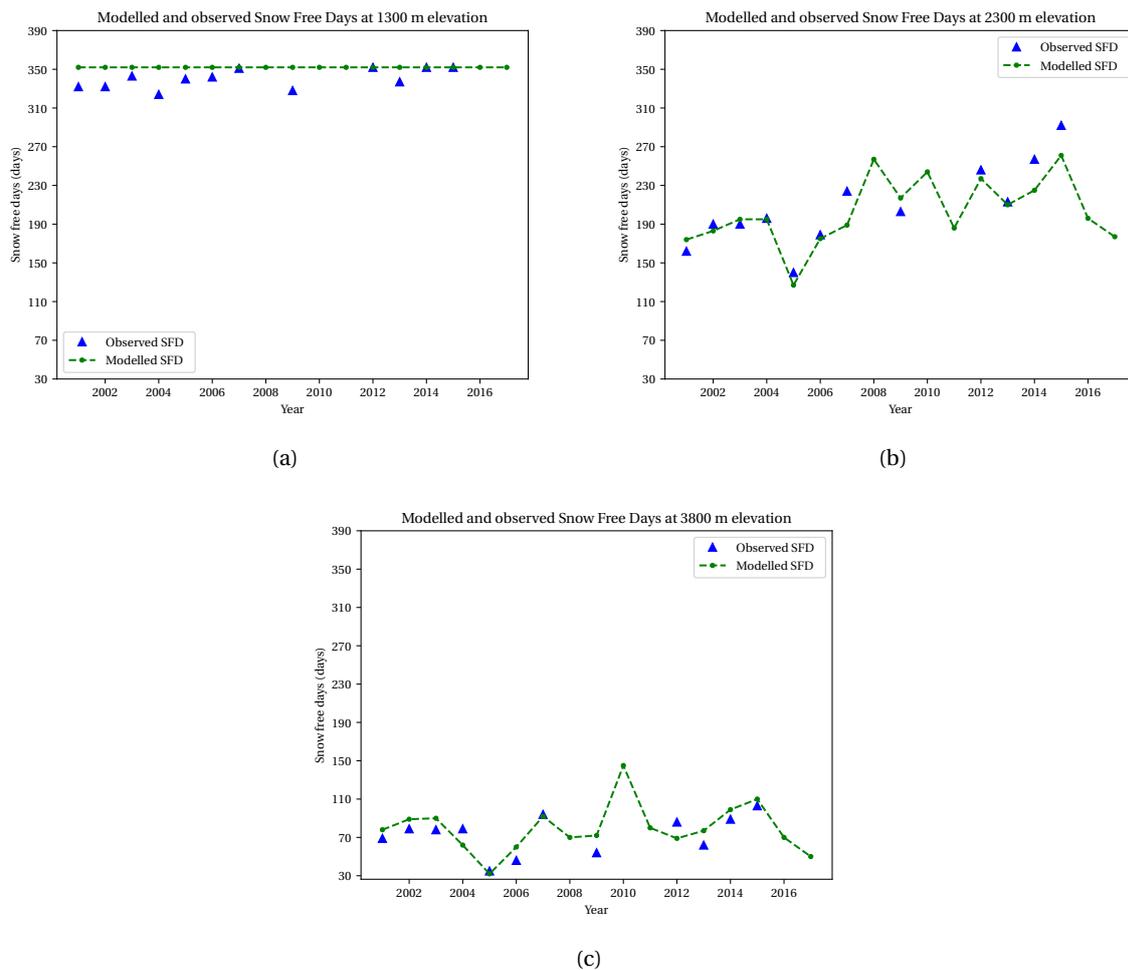


Figure 3.10: Comparison of the observed and modelled snow-free days at elevations of 1300 (a), 2300 (b) and 3800 (c) m.a.s.l. There are more modelled points than observed. This is because of the removal of outliers and SFD determined with more than 50 % interpolated values.

Furthermore it is interesting to look at the absolute values in SFD per elevation in 2017. The SFD is respectively 340 at 1300 m.a.s.l., 250 at 2300 m.a.s.l. and 90 at 3800 m.a.s.l. This is in almost all cases an underestimation of the trend lines from Section 3.1.2. To understand this difference between the trends of the observed SFD and the trends of the drought scenario from Figure 3.9 it is important to know how the Ordinary Least Squares model is built. First, the slope is determined of the model until 2010. This is in all the cases the same slope. Afterwards, four slopes are determined based on the data from 2010 until 2017. These slopes are used to create the lines as shown in Figure 3.9. Only the original model (blue line) is calibrated on the observed SFD, but that does not mean that the slopes are the same since there is a difference between modeled and observed. This difference is described with the R^2 in Figure 3.8. From that, it was clear that in the best performing elevation zones 70% of the observed SFD was described by the model. From Figure 3.10 it is clear that the model tends to underestimate the last years. Due to that, the slope of the last years is also underestimated. For that, it is not possible to compare these results in absolute sense with the trends found in Section 3.1.2.

However, the difference between models at the three chosen elevation zones is still valuable for this research and can provide more information on the temperature and precipitation effects. The maximum difference between the model with the drought temperature and without the drought temperature is 10 SFD at 1300 m.a.s.l., 70 SFD at 2300 m.a.s.l. and 25 at 3800 m.a.s.l. Again, this is corresponding with the trends of Figure 3.3. The temperature effect is clearly visible at 2300 m.a.s.l. This is less visible at 1300 m.a.s.l. due to the initial situation where the amount of SFD is already very high. Moreover, according to Figure 3.10a the temperature and precipitation do not affect the SFD much in the model at this elevation, which results in low trends before and after 2010 for every drought scenario.

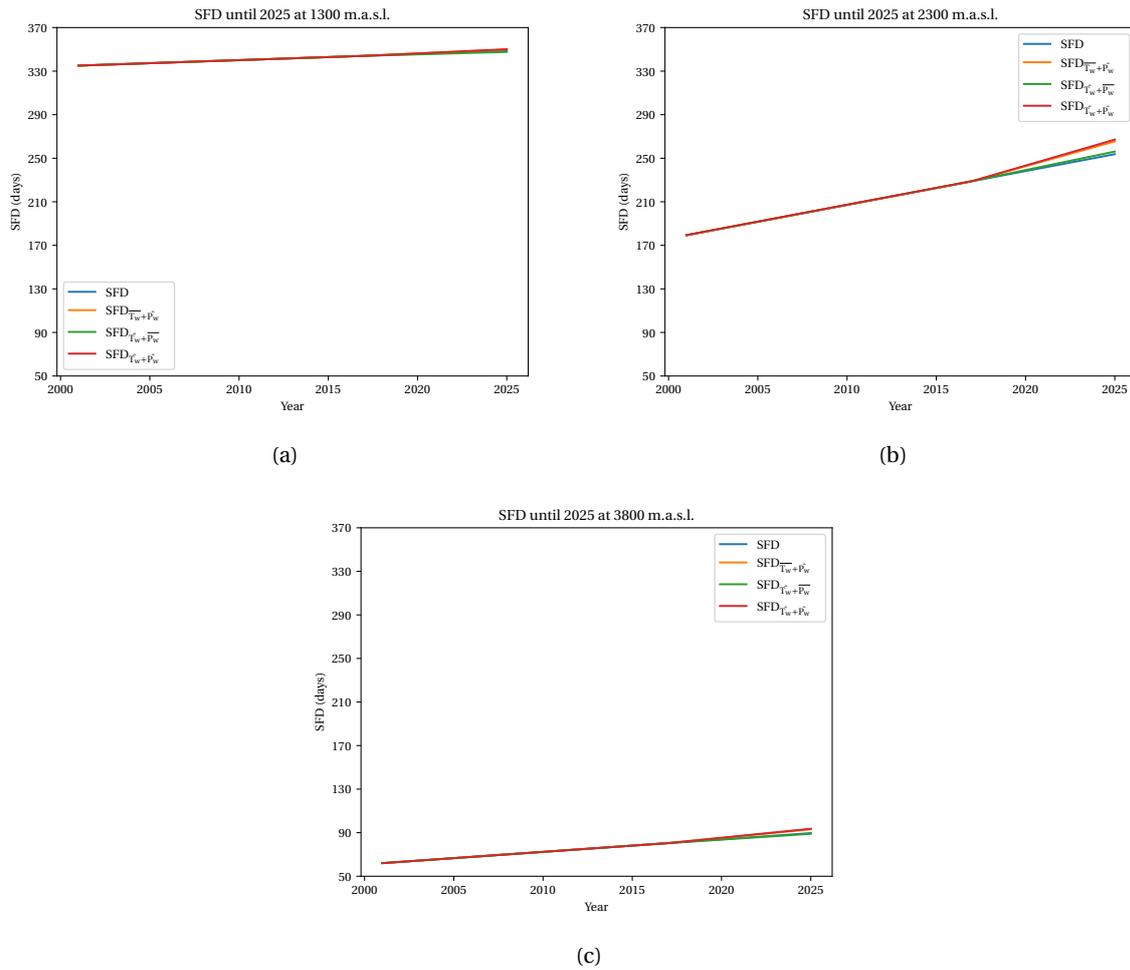


Figure 3.11: The extrapolation of snow-free days until 2025 at elevations of 1300 (a), 2300 (b) and 3800 (c) m.a.s.l. The blue line shows extrapolation of the SFD trend, orange the extrapolation with predicted values of P_w as input only, green with predicted values of T_w as input only and red with predicted values of both parameters. The methodology behind this is described in Section 2.6

So can this be related to the conclusions of Flint et al. (2018)? He distinguishes between short dry drought periods (2007-2009) and long warm drought periods (2012-2016) of which the latter is more destructive for the environment. Despite of the fact that the short dry period is incorporated in the data until 2010, it can be stated that the temperature input was more important related to SFD in the period of 2012-2016 than the precipitation input. The results show that the 2012-2016 drought was long and warm and had a very large effect on the SFD development. The combination of the two confirm the conclusions of Flint et al. (2018) related to SFD. However, the effect on snow storage has to be investigated in the hydrological model (Section 3.3).

Extrapolation until 2025

The second scenario is extrapolation until 2025. The purpose of this scenario was also twofold:

- Analyzing the development of modeled SFD (with inputs T_w and P_w) and extrapolated SFD
- Quantification of SFD per elevation zone in 2025 as an indication for a future perspective on SFD for the Merced River basin.

The same way of modelling is used as in the previous section, but now with the predicted values of temperature and precipitation (see Section 2.6.2). The results for the already mentioned elevations zones (1300, 2300 and 3800 m.a.s.l.) are shown in Figure 3.11.

Comparing Figures 3.11a, 3.11b and 3.11c shows the largest difference between models at 2300 m.a.s.l. This largest difference between the maximum (predicted temperature as input) and the minimum (extrapolation

of the SFD model) is 20 SFD. At 1300 m.a.s.l. this is not more than 2 SFD and at 3800 m.a.s.l. a difference of 5 SFD is found. It is clear that these differences are way smaller than in the drought scenario; almost two times smaller at 2300 m.a.s.l. and even 5 times smaller at 1300 and 3800 m.a.s.l. This shows that the drought period has an enormous impact on the total amount of SFD according to this multivariate model.

The extrapolation scenario shows upward trends at all elevations with the largest trend at 2300 m.a.s.l. This is conform the results in Figure 3.3. For all elevations the SFD trend shows the lowest result. This means that the incorporation of P_w or T_w always show higher results and may imply that the addition of temperature and precipitation effects may even enlarge the current trend observed in snow cover changes. The highest result is always obtained with \hat{P}_w and \hat{T}_w as inputs (the red line in Figure 3.11) which is logical since this is the worst possible combination of temperature rise and precipitation decrease. In this maximum case there will be 355 SFD at 1300 m.a.s.l., 270 SFD at 2300 m.a.s.l. and 90 SFD at 3800 m.a.s.l. This means that eventually almost snow will be found at 1300 m.a.s.l., but (related to the other results) the highest impact in terms of snow availability will be at 2300 m.a.s.l. However an indication, an increase of 100 SFD within 25 years is alarming. Especially for an elevation zone with a contribution of 7 % over the whole catchment.

Again, the results show smaller slopes than the trend analysis of Section 3.1.2 which is related to the calibration of the SFD model. The results from Figure 3.11 are generated with this model, which underestimates the observed SFD values which is similar with the results of the drought scenario.

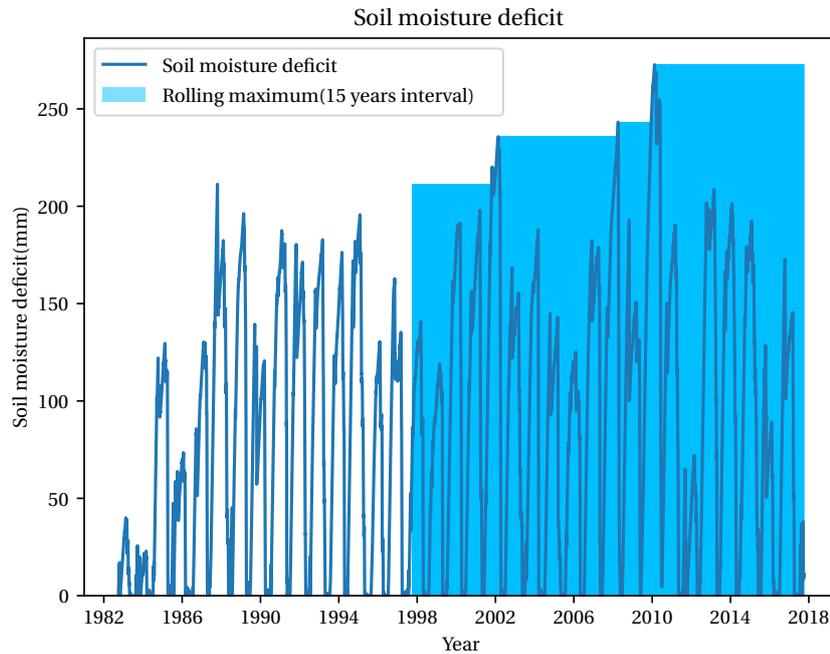


Figure 3.12: Time series of the soil moisture deficit together with a rolling maximum with an interval of 15 years. The rolling maximum is an approximation for the changing S_r .

3.3. Hydrological processes

The hydrological model is created to analyze the influence of the climate parameters and the SFD on the hydrological system. The focus is on the root-zone moisture storage capacity S_r and the snow storage. By analyzing this, it is possible to discover their interactions and the behaviour of the hydrological system. First, the results of the unsaturated zone are presented, which are calculated with the calculation of soil moisture deficit (2.7.3). Secondly, the results derived with the method from Section 2.7.2 are presented and discussed. This is done with the comparison of mean values before and after 2010. At last, it is possible to derive SFD with inverse hydrological modeling. This is useful since the relatively small dataset from MODIS(2001-2017) can be expanded to the timespan of the weather station (1983-2017) in this way, which can strengthen the conclusions related to SFD.

3.3.1. Soil moisture deficit

The soil moisture deficit is defined as the difference between effective precipitation and transpiration, shown in Section 3.3.1. The maximum over a rolling mean with an interval of 15 years can be seen as the approximation of the root-zone moisture storage capacity (S_r). The interval is chosen based on literature (Gao et al., 2014; de Boer-Euser et al., 2016) and an extreme value analysis (see Appendix C). The time series of the soil moisture deficit and the rolling maximum are presented in Figure 3.12. The soil moisture deficit is visualized in absolute values, which means that the soil moisture deficit is negatively changed when effective precipitation exceeds transpiration until a minimum value of zero. The light blue bars show the rolling maximum over the interval of 15 years, which is the approximation of the root-zone storage capacity (S_r). The increase of the maximum soil moisture deficit indicates a gradual increase of S_r , which is in line with the assumption that plants adjust their roots to extreme situations. The final maximum value for soil moisture deficit (265 mm) is a reasonable value for S_r in this area with the corresponding return period of 15 years (Wang-Erlandsson et al., 2016; Gao et al., 2014). Interestingly, this maximum is not found at the end of the latest drought period (around 2016) but at the end of the previous drought period ending in 2010. To understand this, it is important to note that the soil moisture deficit is dependent on effective precipitation, which consists of melt and rainfall. In Figure 3.13, melt and rainfall are visualized together with the effective precipitation. From this figure it is clear that most of the effective precipitation is generated from melt. Furthermore, the transpiration part is only dependent on temperature (see Equation 2.34). This means that the temperature has a large influence on both the effective

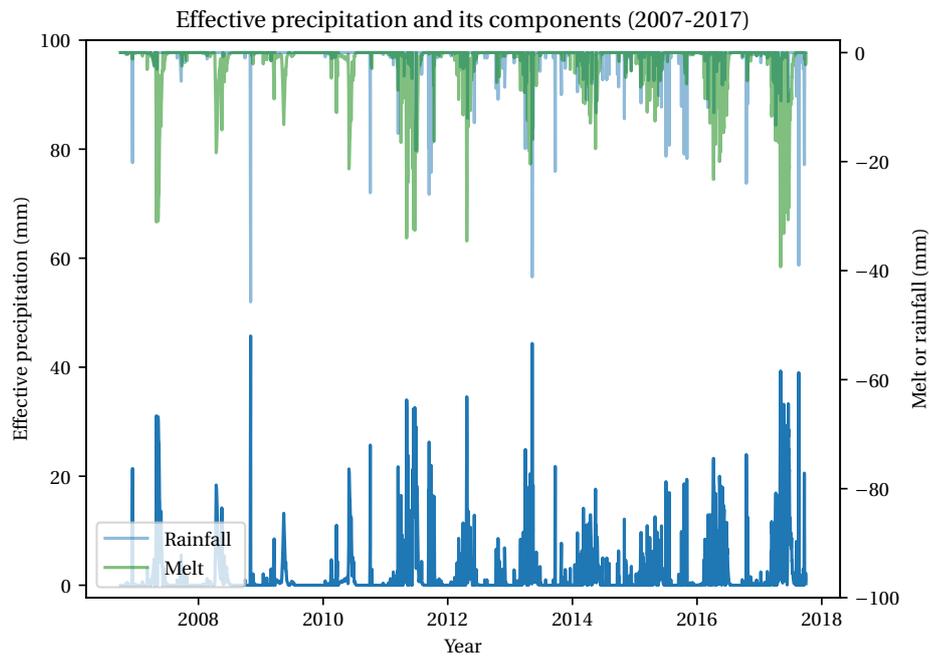


Figure 3.13: Visualization of the effective precipitation and its components. The upper bars show the rainfall and melt part, the lower bars show the total effective precipitation.

precipitation and the transpiration. This can also be related to the figures of the T_w trends (Figures 3.1a and 3.1b) which show large positive anomalies during the drought of 2012-2016, but only large negative anomalies during the drought of 2007-2009. The high temperatures of 2015 and 2016 cause a high melting rate which suppressed the development of a larger soil moisture deficit. The opposite happened in the (relative short) drought period of 2007-2009; the melting rates were not higher than normal, however the precipitation was lower than normal and caused an exceptionally high soil moisture deficit. If this is true, lower snow storage has to be found after 2010. This is analyzed in the next section.

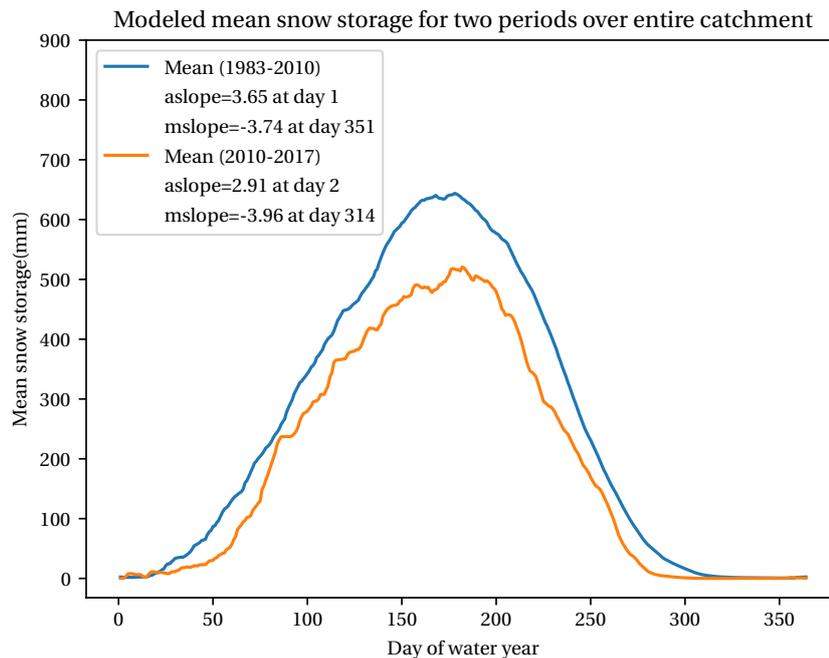


Figure 3.14: Modeled mean snow storage of the period 1983-2010 and 2010-2017. The legend shows the slope of the accumulation period, the starting day of accumulation, the slope of the melting period and the last day of the melting period. The slopes are calculated between the day of minimum (start of accumulation or end of melting period) and the day of maximum.

3.3.2. Snow storage differences

To analyze the snow storage difference, the daily means of the period 1983-2010 and 2010-2017 are calculated and compared. This is done for the whole catchment (Figure 3.14) and the same three elevation zones which are analyzed for the scenarios and SFD anomalies (Figure 3.15).

Figure 3.14 shows a clear difference between the snow storage before and after 2010. First, the maximum is approximately 100 mm lower, but still around the same day of the water year. The accumulation starts almost at the same moment, however the slope of the accumulation after 2010 is much lower than before 2010, which is logical because of the lower maximum. However, the slope of the melting period is larger after 2010. This is caused by the higher temperatures in this period (Figure 3.1b) which causes a higher melting rate. In conclusion, the mean snow storage after 2010 is 37 days earlier zero than before 2010.

To relate this to elevation, Figure 3.15 is created. The largest difference in maximum is found at an elevation of 2300 m.a.s.l. This mainly caused by the hypsometry of the catchment (Figure 2.3), where the elevation zone of 2300 m.a.s.l. covers the largest area of the chosen elevation zones. Another behaviour is discovered at 3800 m.a.s.l. In this case (Figure 3.15c) the accumulation slope is equal, but the maximum is reached earlier after 2010 due to less precipitation. Due to the combination with a larger slope during the melting period, the difference between snow free days before and after 2010 are quite large; 43 days. At last it has to be noticed that the contribution at 1300 m.a.s.l. is so low, that it is difficult to say something about these results.

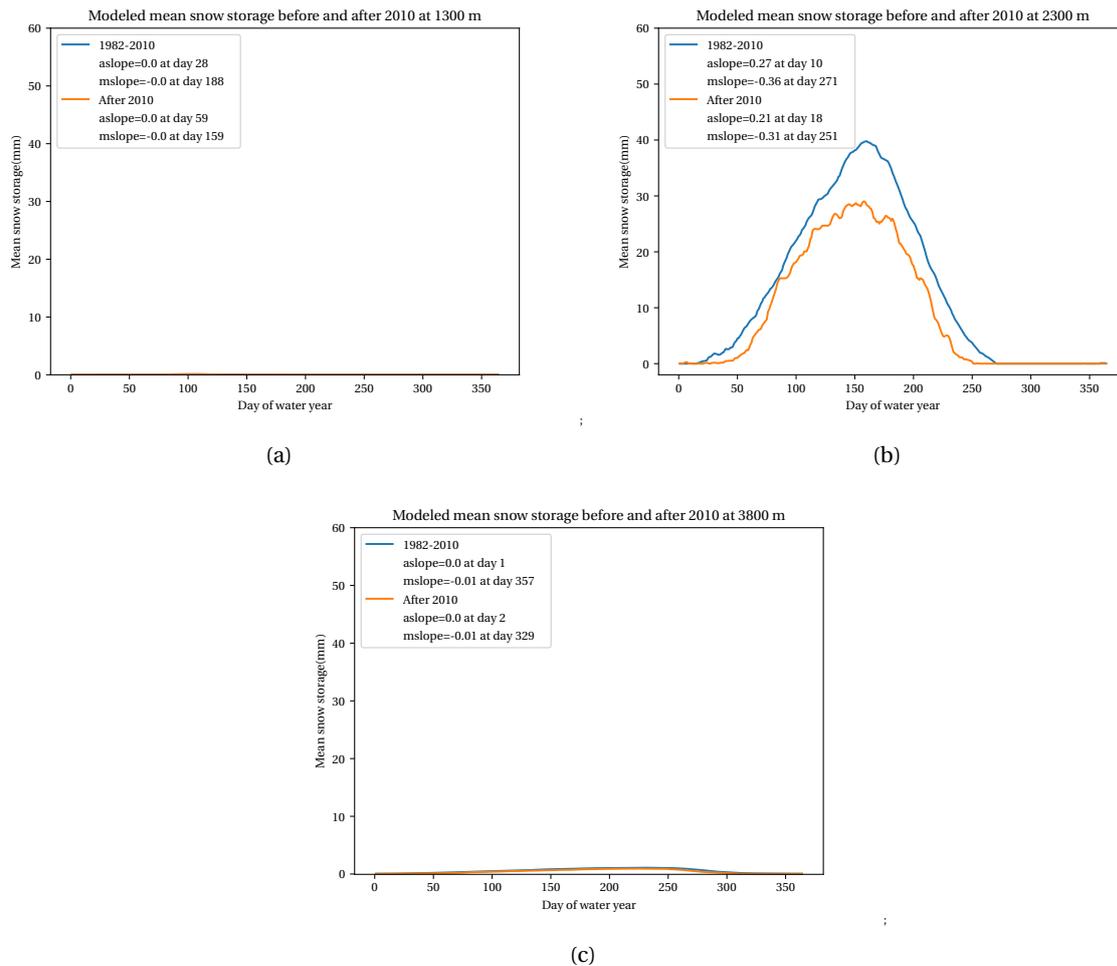


Figure 3.15: The difference between mean snow storage in periods 1982-2010 and 2010-2017 at elevations of 1300 (a), 2300 (b) and 3800 (c) m.a.s.l. The legend shows the accumulation slope (aslope), starting day of the accumulation period, melting period slope (mslope) and the last day of the melting period. The slopes are calculated between the minimum (accumulation starting day or last day of melting period) and the maximum.

3.4. Conclusion

The results for the trend analysis, the multivariate regression and the hydrological model were presented in the previous sections. The order of these sections was used to explain the large influence of trends of P_w and T_w on the calculation of SFD and the outcomes of the scenarios and the hydrological model. The conclusion on these results and the link between the results of SFD and root-zone storage moisture capacity are presented in Chapter 5. But before this is done, the methodology and the results has to be discussed. This will be done in the next chapter.

4

Discussion

This chapter presents a reflection on the results and evaluates the methodology, respectively the statistical analysis and the hydrological model. Additionally, the applicability of this study related to the soil moisture deficit method is discussed in the last section.

4.1. Trends

The statistical analysis consists mainly out of a trend analysis with linear regression. The input data and the test for the trend analysis are discussed in the first two sections. After these, the multivariate analysis including the scenarios are evaluated.

4.1.1. Data

Weather station data The most important input data is obtained via the weather station at Yosemite Park HQ and from MODIS. The weather station is located just outside the catchment. Especially in mountainous areas this can be a problem, since the terrain is inhomogeneous and the weather station does not represent the patterns in the catchment. Despite these arguments, there is chosen for this station since other weather stations located inside the catchment were only measuring for short term projects, making them less suitable for a long term trend analysis. Furthermore, no quality control on the data obtained via these short-term stations was applied. The station at Yosemite Park HQ was quality controlled and suitable for scientific purposes from 1982 onwards. However, the snowfall data consisted of many data gaps and outliers. To compensate for this, the snowfall was obtained via the environmental lapse rate (Section 2.3.2). This method is often used for temperature, but to relate it to precipitation and snowfall is a little bit blunt, especially in mountainous areas. However, when no other data source is available it is a way to relate the precipitation, temperature and snowfall to the elevation. Another point of discussion is related to the spatial distribution of the data. The input data is generated from one weather station and not spatially distributed. This will have an effect on the accuracy of the calculation since the topography influences climate data in mountainous areas (Buytaert et al., 2006). Another option is to use satellite data instead of a weather station, however this data is scarce for large daily time series and the quality is questionable. TRMM was considered because of its small temporal range (3 hourly) and a long temporal coverage (1997-2015) related to other satellite products. However, TRMM has a very low spatial resolution (4 by 4 kilometres) which does not incorporate the topographic effects either (Greenbelt, 2011). By taking this in consideration, the weather station is still the best option to use, since the data quality is guaranteed and real data is used.

MODIS and RSLE method The SFD are calculated with the MODIS satellite data which consists of snow cover data with a spatial resolution of 500 by 500 meter. This is quite coarse, but the daily temporal range is very useful for this research. The version 6 of MOD10A is the first MODIS product with the NDSI for snow cover, which gives better results in terms of cloud filtering. The combination with MODIS and the RSLE method (which incorporates a filter for images with 70% cloud cover) increases the accuracy in terms of errors between cloud and snow pixels. The accuracy of the MOD10A1 L3 version 6 is between 88% and 93% (Hall and Riggs, 2016). However, this strict procedure also causes a lot of data gaps and a necessity for interpolation. In the end, the procedure caused outliers and unusable data for some years, which had to be removed from the results. The gaps can also be seen in the Figures 3.3 and 3.10. Despite of these drawbacks, the MODIS data is very useful in

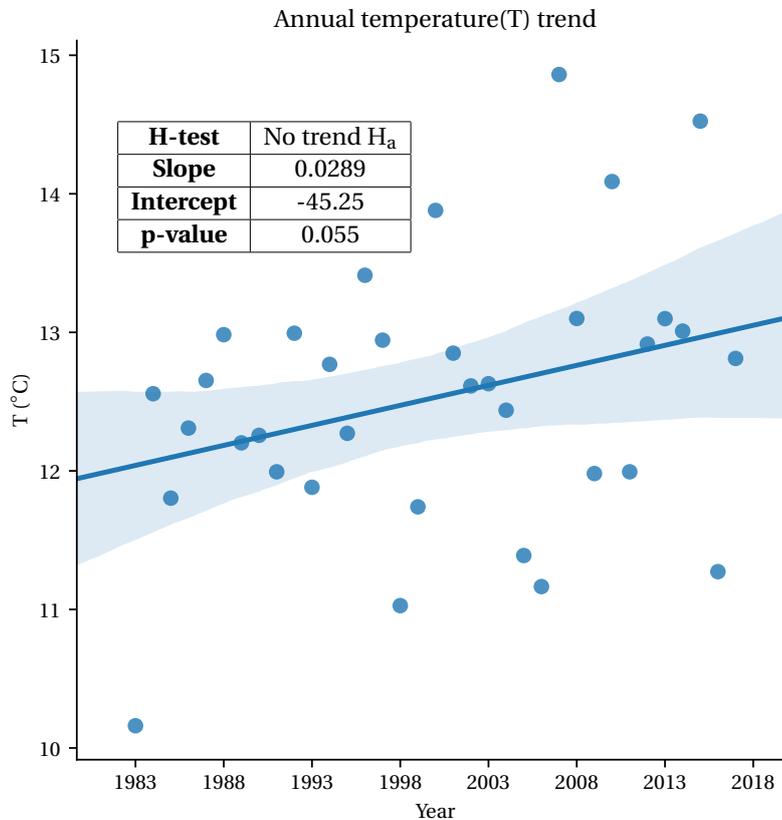


Figure 4.1: The temperature trend of mean temperature over a water year. The table shows the results of the Mann-Kendall test: test result, slope, intercept and p-value.

terms of further research concerning the RSLE and its trends. The MODIS data is available free of charge for all parts of the world which gives an opportunity to repeat and improve the methodology used in this study (Hall et al., 2001).

4.1.2. Mann-Kendall test

During the trend analysis, the datasets were tested with the Mann-Kendall test. This test is often used as test for trends in climate related research. For these kinds of datasets a timespan of at least 30 years is necessary. This is also the nature of the Mann-Kendall test, which tends to reject the alternative hypothesis more often for shorter data sets. For this study, this problems did not occur because significant trends were also found for SFD (Figure 3.3). Another aspect of importance within the Mann-Kendall test is seasonality. Since the input of the Mann-Kendall test is not a normally distributed dataset, the method is sensitive for seasonality. With the definition of a winter period this sensitivity is by-passed. The determination of the winter period is based on literature of (hydrological) research which investigated snow fall or snow water equivalent. However, many other researches use a different winter period which makes the definition a little bit arbitrary. For this study the winter period is defined between November and March to incorporate accumulation and melt, but avoiding the snow-free periods. This is favourable to detect precipitation trends (which are clear according to the results), but temperature differences are quite high in this period which caused the non-significant trends for T_w . When the temperature over the whole water year is taken, then the trends are more distinct and almost significant ($p=0.055$, see Figure 4.1).

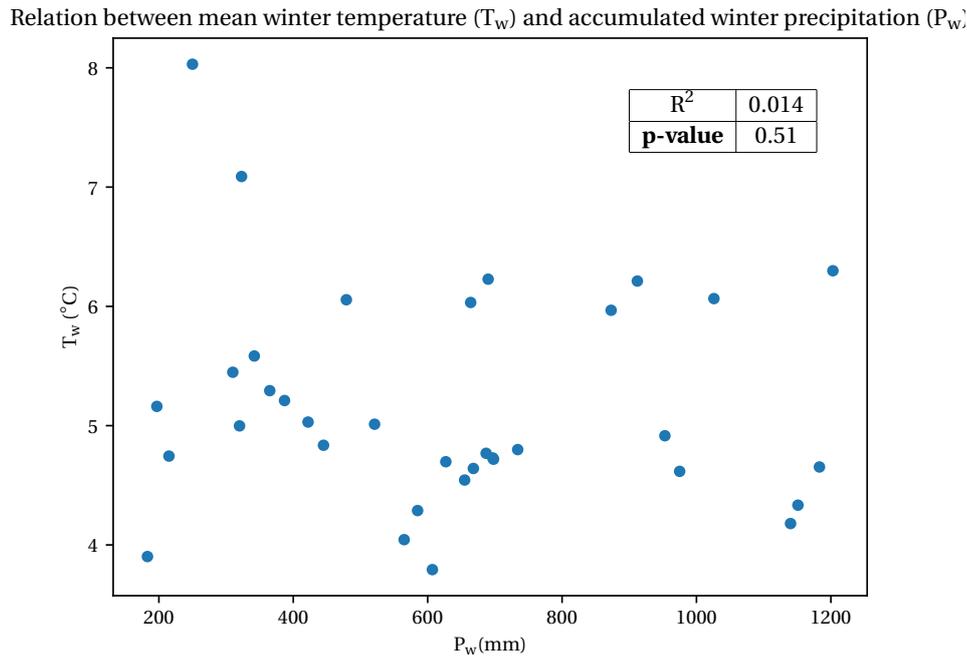


Figure 4.2: The relation between mean winter temperature (T_w) and accumulated winter precipitation (P_w). The table shows their inter-correlation (R^2) and the p-value.

4.2. Relation between snow-free days, precipitation and temperature

The multivariate regression model is based on linear regression (Section 2.5). One of the requirements for linear regression is that the two datasets have to be linear independent, for this case P_w and T_w . To check this, P_w and T_w are plotted in Figure 4.2. The R^2 shows a very low value, which means that there is no inter-correlation present and T_w and P_w are properly used in the multivariate regression analysis. Another discussion point with this analysis is already mentioned in Section 2.5 and is related to the amount of variables included in the model. It is assumed that P_w and T_w describe SFD sufficient enough, however the results show very low values of R^2 (smaller than 0.5) for all combinations of the multivariate model at high elevations. This holds that SFD is not calculated sufficiently in less than half of the cases by the multivariate model. This can be caused due to the absence of a variable in the multivariate model or by errors related to observed data and the corresponding elevation zone. The second is more likely since the higher elevations do represent small areas (see Figure 2.3), which increases the relative error. The non-significant trends at higher elevations (Figure 3.3) also hint in this direction.

4.3. Hydrological processes

4.3.1. Simplified model structure

The hydrological model consists of only two reservoirs: snow and unsaturated zone. The interception and deep ground water is not included, to reduce the amount of calibration parameters and avoid equifinality (Beven and Freer, 2001). The only calibrated parameter is the day-duration factor (DDF) which is directly related to the observed RSLE. This gives a unique combination for DDF over 2000-2017. The hydrological model can be perceived as way too simple, since some processes are not included. Looking at the interception, the contribution is assumed to be low compared to the unsaturated zone and snow reservoir in these high elevated zones. Furthermore, the interception will produce a high uncertainty since forest fires change the interception capacity of the catchment. For that, the addition of the interception reservoir would cause more noise in the model. Another process which is left out of the hydrological model is the deep ground water reservoir. This is done because relative slow responses of the deep ground water are more relevant in larger and less steep catchments. The underlying assumption of the exclusion of the deep ground water reservoir is that the catchment responds rapidly. Another important factor for the modeling is related to the purpose of the model. It is tried to model

the soil moisture deficit related to snow cover. To focus on snow cover and the unsaturated zone it is logical to investigate their direct relation. When other processes are added, it is less clear which process is influencing the outcomes. By keeping the model compact in this way, better understanding and conclusions about snow and soil moisture deficit can be obtained.

4.3.2. Transpiration

The transpiration in the model is calculated with an empirical model based on temperature and solar radiation (which is location dependent). This model of Hargreaves and Samani (1985) is the best performing one according to literature (Tukimat et al., 2012), but not presenting the physical behaviour of the transpiration in the catchment. It is important to be aware of this because the transpiration is an important process within the calculation of soil moisture deficit. For instance, the ability of vegetation to turn dormant during drought periods and transpire less (Gao et al., 2014). These kind of effects are not included in the transpiration for this study. Due to this, it is likely that the soil moisture deficit is overestimated. However, the Hargreaves method is usable for this catchment since no other reliable transpiration data is found.

4.4. Application to the SMD method

The addition of MODIS data to the SMD method to investigate snow storage effects could be promising, but has to be further evaluated. Such an evaluation could consist of the application of the model in several other mountain catchments. In one way, the method calculates soil moisture deficit but in the other way the snow storage in the catchment is also calculated. This gives an alternative to the Snow Water Equivalent (SWE), which is widely used in climate research. The SWE is calculated as the product of snow depth and snow density and represents the amount of water column that will be generated during melting (Takala et al., 2011). The SWE is based on the interpolation of ground station data combined with satellite data and for that calculated based on observations. The calculation of snow storage in the hydrological modeling is based on a physical process and validated with satellite observations. For that, this hydrological model can be a valuable addition to the existing research related to SWE because hydrological models can reveal the physical processes behind these trends. This does not mean that hydrological modelling performs better, because the performance of the model is dependent on the data on which it is calibrated. MODIS is widely accessible and applicable for a lot of areas, but has also flaws which are highlighted in Section 4.1.1. Furthermore, the RSLE is dependent on the surface area of the region. For that, this is only applicable for small catchments. To see the most suitable surface area size it is also important to test this method for more catchments of different sizes.

5

Conclusions and further research

5.1. Conclusions

Two research questions provided the basis for this study. They represent a statistical analysis part and a hydrological modeling part:

What is the effect of precipitation and temperature changes on the amount of snow-free days in a catchment in the Southern Sierra Nevada?

What is the effect of a change in the number of snow-free days on the root zone moisture storage capacity in a catchment in the Southern Sierra Nevada?

The first research question is answered by creating a multivariate linear regression model with temperature and precipitation as inputs. From the results, it can be concluded that temperature is a dominant factor until 2900 m.a.s.l. and less dominant at higher elevation zones. However, significant trends are mostly found in the lower elevation zones and the non-temperature influenced zones showed relative low values for R^2 , lower than 0.6. This means that, according to this model, the temperature is a better predictor for the SFD than precipitation. This behaviour corresponds with research related to snow water equivalent behaviour in mountainous areas at the West Coast (Sospedra-Alfonso et al., 2015) and with temperature and precipitation anomalies related to the California Drought (Luo et al., 2017). To relate this to a change of SFD, the scenarios show all the similar direction: due to temperature rise and precipitation decrease the SFD will increase and less snow will be available in the future. The impact of the last years is clearly visible in the scenarios and the hydrological model, which all show an increase of SFD. The largest increase is found at 2300 m.a.s.l. of 7 SFD/year based on the observations of SFD. For this elevation zone, a total of 270 SFD is predicted in 2025 in the maximum situation as an indication. More evident is the difference of 100 SFD between 2000 and 2025, which is an average increase of 4 SFD/year. This is also found as average when the SFD trends are weighted on the hypsometry of the catchment. In summary:

- SFD increases in elevation zones up to 3200 m.a.s.l. with significant trends
- SFD are better described by temperature at elevations lower than 2500 m.a.s.l. At higher elevations, precipitation describes SFD better although not significantly.
- The R^2 of the multivariate model is lower than 0.5 above 3500 m.a.s.l., which means that neither temperature nor precipitation described SFD well in these elevation zones. However, between 1500 and 2500 m.a.s.l., R^2 around 0.7 is found which can be qualified as a good approximation.
- An upward trend of the SFD is observed in all cases; precipitation or temperature influenced, at higher or low elevation or for the short or long term. All results show an increase of SFD with a maximum trend of 7 SFD/year at 2300 m.a.s.l. and an average of 4 SFD/year.

The second question is answered by the hydrological model (see Section 2.7, especially Figure 3.12). It can be concluded from the rolling maximum that the soil moisture deficit is gradually increasing. The assumption behind the soil moisture deficit is that the deficit occurs due to an increase of the root-zone storage capacity, thus the water is stored due to the adaptive behaviour of the vegetation. The gradual increase of the soil

moisture deficit suggests that this pattern is happening in the Merced River catchment. The results show that the maximum soil moisture deficit of 273 mm is caused by a deficit of melting water (Figure 3.13). The low amounts of snow melt are also the reason why the maximum soil moisture deficit is found before the Californian Drought (2012-2016) and not during this period. During the Californian Drought, the melting rates were high which caused not only a positive feedback against maximum soil moisture deficit on one side, but also an increase of SFD on the other side. This can also be seen as a difference of maximum snow storage of 100 mm in Figure 3.14 and the differences in the ending of melting periods at different elevation zones (Figure 3.15). For example at 2300 m.a.s.l., from 2010 the melting stops 25 days earlier than previously. When the pattern of increasing SFD will continue, as shown in the scenarios (See Section 2.6), eventually less snow is available for melting and the future soil moisture deficit will increase too.

The summarized hydrological modeling part conclusions are:

- The calculation of the soil moisture deficit shows a gradual increase of the rolling maximum up to 265 mm. This suggests adaptive behaviour of the vegetation to increase their root-zone storage capacity in the Merced river basin.
- Water originated from the snow melt prevents high soil moisture deficits, but increases the SFD.
- Eventually lower amounts of snow will be available and soil moisture deficits will increase. This is already visible, the maximum snow storage after 2010 is already 100 mm less than before 2010 and the melting period ends 37 days earlier. These drought effects corresponds with recent research about the environmental effects of drought. (Flint et al., 2018)

5.2. Further research

At first sight, the results are promising for the methods used, though the expected catchment behaviour is alarming. However, as has been discussed in the previous chapter, the data sources used contain some drawbacks, which make it hard to draw definitive conclusions.

For example, the data for the multivariate model and the input for the hydrological model are both obtained from one weather station which was located slightly outside the catchment. Due to spatial weather patterns, a distance between the weather station and the catchment could make the data less reliable. Therefore it is advised to use this methodology with data of a network of weather stations as input. Especially precipitation data is important, since the spatial distribution of precipitation is highly influenced by the topography of the mountains (Buytaert et al., 2006).

Another important data-related recommendation is focused on the transpiration calculation. The method of Hargreaves is empirical and does not incorporate physics, which can cause problems since long term weather patterns are changing. When enough data is available, it is recommended to use physical methods which can incorporate the dormancy of vegetation during dry periods.

Furthermore, the methodology has to be tested for more catchments to validate the results. For that, it is necessary to find catchments at the same scale in similar climates to see whether the results match. As an additional step to this, it can also be tested whether there are similarities and differences between catchments located in other climatic regions to see the influence of snow cover on root zone moisture capacities globally.

The recommendations as discussed above can be summarized as follows:

- It is important to use a network of weather stations, especially for precipitation data. This is mainly due to the influence of topography on precipitation in mountainous catchments.
- When enough data is available, prefer physically based transpiration calculation methods above empirical.
- To validate the results of this study, it is important to apply this methodology at other catchments with a similar climate.
- When validated, this method can be applied in other climates to investigate how snow is affecting soil moisture capacity in these areas.

Bibliography

- Allen, J. (2015). Diminished Snow Pack in the Sierra Nevada.
- Ankenbauer, K. J. and Loheide, S. P. (2017). The effects of soil organic matter on soil water retention and plant water use in a meadow of the Sierra Nevada, CA. *Hydrological Processes*, 31(4):891–901.
- Bales, R. C., Goulden, M. L., Hunsaker, C. T., Conklin, M. H., Hartsough, P. C., O’Geen, A. T., Hopmans, J. W., and Safeeq, M. (2018). Mechanisms controlling the impact of multi-year drought on mountain hydrology. *Scientific Reports*, 8(1):690.
- Berghuijs, W. R., Woods, R. A., and Hrachowitz, M. (2014). A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature Climate Change*, 4(7):583–586.
- Beven, K. and Freer, J. (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249(1-4):11–29.
- Buytaert, W., Celleri, R., Willems, P., Bièvre, B. D., and Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*, 329(3-4):413–421.
- Cayan, D. R., Maurer, E. P., Dettinger, M. D., Tyree, M., and Hayhoe, K. (2007). Climate change scenarios for the California region. *Climatic Change*, 87(1 SUPPL).
- Collins, D. B. G. and Bras, R. L. (2007). Plant rooting strategies in water-limited ecosystems. *Water Resources Research*, 43(6).
- Danielson, J. and Gesch, D. (2011). Global Multi-resolution Terrain Elevation Data 2010(GMTED2010). *U.S. Geological Survey Open-File Report 2011–1073*, 2010:26.
- de Boer-Euser, T., McMillan, H., Hrachowitz, M., Winsemius, H., and Savenije, H. (2016). Influence of soil and climate on root zone storage capacity. *Water Resources Research*, pages 1–16.
- Easterling, D. (2017). Precipitation Changes in the United States. Climate Science Special Report: Fourth National Climate Assessment, Volume I. *U.S. Global Change Research Program*, 1:207–230.
- Flint, L. E., Flint, A. L., Mendoza, J., Kalansky, J., and Ralph, F. M. (2018). Characterizing drought in California: new drought indices and scenario-testing in support of resource management. *Ecological Processes*, 7(1):1–13.
- Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, E., Sriwongsitanon, N., and Savenije, H. H. G. (2014). Climate controls how ecosystems size the root zone storage capacity at catchment scale. *Geophysical Research Letters*, 41(22):7916–7923.
- Gilbert, R. O. (1987). *Statistical Methods for Environmental Pollution Monitoring*, volume 30. John Wiley & Sons.
- Gleick, P. (2017). *Impacts of California’s Five-Year (2012-2016) Drought on Hydroelectricity Generation*. Pacific Institute.
- Goin, D. E., Rudolph, K. E., and Ahern, J. (2017). Impact of drought on crime in California: A synthetic control approach. *PLOS ONE*, 12(10).
- Grayson, R. B., Moore, I. D., and McMahon, T. A. (1992). Physically based hydrologic modeling: 1. A terrain-based model for investigative purposes. *Water Resources Research*, 28(10):2639–2658.

- Greenbelt, M. (2011). Tropical Rainfall Measuring Mission (TRMM).
- Hall, D. K. and Riggs, G. A. (2016). MODIS/Terra Snow Cover Daily L3 Global 500m Grid, Version 6, NDSI_Snow_Cover. *NASA National Snow and Ice Data Center Distributed Active Archive Center*.
- Hall, D. K., Riggs, G. A., and Barton, J. S. (2001). MODIS Snow and Sea Ice-Mapping Algorithms. *Modis Atbd*, the MODIS(September):45.
- Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGirolamo, N. E., and Bayr, K. J. (2002). MODIS snow-cover products. *Remote Sensing of Environment*, 83(1-2):181–194.
- Hardin, E., AghaKouchak, A., Qomi, M. J., Madani, K., Tarroja, B., Zhou, Y., Yang, T., and Samuelsen, S. (2017). California drought increases CO₂ footprint of energy. *Sustainable Cities and Society*, 28:450–452.
- Hargreaves, G. H. and Samani, Z. A. (1985). Reference Crop Evapotranspiration from Temperature. *Applied Engineering in Agriculture*, 1(2):96–99.
- He, X., Wada, Y., Wanders, N., and Sheffield, J. (2017). Intensification of hydrological drought in California by human water management. *Geophysical Research Letters*.
- Howitt, R., Macewan, D., Medellín-Azuara, J., Lund, J., and Sumner, D. (2015). Economic Analysis of the 2015 Drought For California Agriculture. Technical report, UC Davis Center for Watershed Sciences.
- Irannezhad, M., Ronkanen, A.-K., Kiani, S., Chen, D., and Kløve, B. (2017). Long-term variability and trends in annual snowfall/total precipitation ratio in Finland and the role of atmospheric circulation patterns. *Cold Regions Science and Technology*, 143:23–31.
- Jepsen, S. M., Harmon, T. C., Ficklin, D. L., Molotch, N. P., and Guan, B. (2018). Evapotranspiration sensitivity to air temperature across a snow-influenced watershed: Space-for-time substitution versus integrated watershed modeling. *Journal of Hydrology*, 556:645–659.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3):259–263.
- Krajci, P., Holko, L., Perdige, R. A. P., and Parajka, J. (2014). Estimation of regional snowline elevation (RSLE) from MODIS images for seasonally snow covered mountain basins. *Journal of Hydrology*, 519(PB):1769–1778.
- Krzywinski, M. and Altman, N. (2013). Points of significance: Significance, P values and t-tests.
- Lundquist, J. D. and Cayan, D. R. (2007). Surface temperature patterns in complex terrain: Daily variations and long-term change in the central Sierra Nevada, California. *Journal of Geophysical Research*, 112(D11):D11124.
- Luo, L., Apps, D., Arcand, S., Xu, H., Pan, M., and Hoerling, M. (2017). Contribution of temperature and precipitation anomalies to the California drought during 2012-2015. *Geophysical Research Letters*, 44(7):3184–3192.
- Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, 13(3):245.
- Martinec, J. (1975). Snowmelt - runoff model for stream flow forecasts. *Nordic Hydrology*, 6(3):145–154.
- Martinec, J. and Rango, A. (1986). Parameter values for snowmelt runoff modelling. *Journal of Hydrology*, 84(3-4):197–219.
- Maurice G Kendall (1975). *Rank Correlation Methods*. Hodder Arnold, 4 edition.
- Milly, P. C. D. (1994). Climate, soil water storage, and the average annual water balance. *Water Resources Research*, 30(7):2143–2156.
- Musselman, K. N., Molotch, N. P., and Margulis, S. A. (2017). Snow melt response to simulated warming across a large elevation gradient, southern Sierra Nevada, California. *The Cryosphere Discussions*, 11(December):1–47.

- Nijzink, R., Hutton, C., Pechlivanidis, I., Capell, R., Arheimer, B., Freer, J., Han, D., Wagener, T., McGuire, K., Savenije, H., and 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.
- Noether, G. E. (1981). Why Kendall Tau? *Teaching Statistics*, 3(2):41–43.
- Pandey, R., Cayan, R., and Georgakakos, P. (1999). Precipitation structure in the Sierra Nevada of California during winter. *Journal of Geophysical Research: Earth Surface*, 104(D10):12019–12030.
- Saxton, K. E. and Rawls, W. J. (2006). Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions. *Soil Science Society of America Journal*, 70(5):1569.
- Scheffer, M., Holmgren, M., Brovkin, V., and Claussen, M. (2005). Synergy between small- And large-scale feedbacks of vegetation on the water cycle.
- Schenk, H. J. (2008). The Shallowest Possible Water Extraction Profile: A Null Model for Global Root Distributions. *Vadose Zone Journal*, 7(3):1119.
- Schoups, G., Hopmans, J. W., Young, C. a., Vrugt, J. a., Wallender, W. W., Tanji, K. K., and Panday, S. (2005). Sustainability of irrigated agriculture in the San Joaquin Valley, California. *Proceedings of the National Academy of Sciences of the United States of America*, 102(43):15352–6.
- Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., Hutley, L. B., Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., and An, L. B. H. (2008). An optimality-based model of the coupled soil moisture and root dynamics To cite this version : An optimality-based model of the coupled soil moisture and root dynamics. *Hydrology and Earth System Sciences*, 12(3):913–932.
- Seager, R., Henderson, N., Cane, M. A., Liu, H., and Nakamura, J. (2017). Is there a role for human-induced climate change in the precipitation decline that drove the California drought? *Journal of Climate*, 30(24):10237–10258.
- Sospedra-Alfonso, R., Melton, J. R., and Merryfield, W. J. (2015). Effects of temperature and precipitation on snowpack variability in the Central Rocky Mountains as a function of elevation. *Geophysical Research Letters*, 42(11):4429–4438.
- Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J. P., Koskinen, J., and Bojkov, B. (2011). Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements. *Remote Sensing of Environment*, 115(12):3517–3529.
- Thompson, S. E., Harman, C. J., Heine, P., and Katul, G. G. (2010). Vegetation-infiltration relationships across climatic and soil type gradients. *Journal of Geophysical Research: Biogeosciences*, 115(G2).
- Tukimat, N. N. A., Harun, S., and Shahid, S. (2012). Comparison of different methods in estimating potential evapotranspiration at Muda Irrigation Scheme of Malaysia. *Journal of Agriculture and Rural Development in the Tropics and Subtropics (JARTS)*, 113(1):77–85.
- USGS (2016). USGS–Water Resources of the United States.
- USGS (2017). USGS Water Data for the Nation, accessed December 15, 2017.
- Wang-Erlandsson, L., Bastiaanssen, W. G., Gao, H., Jägermeyr, J., Senay, G. B., Van Dijk, A. I., Guerschman, J. P., Keys, P. W., Gordon, L. J., and Savenije, H. H. (2016). Global root zone storage capacity from satellite-based evaporation. *Hydrology and Earth System Sciences*, 20(4):1459–1481.
- WMO (2017). WMO Guidelines on the Calculation of Climate Normals. *World Meteorological Organization*, 1203.

List of Symbols

Symbol	Units	Description
α	-	Intercept in a linear regression analysis
β	-	Slope in a linear regression analysis
H_0	-	Null hypothesis for statistical testing
H_a	-	Alternative hypothesis for statistical testing
h	m.a.s.l.	Elevation
h_s	m.a.s.l.	Elevation of the weather station
I_s	%	Scattering of land and snow pixels in MODIS imagery
μ	-	Mean of a data set
n	-	Number of observation in a data set
p	-	Calculated probability
P_c	%	Cloud cover pixel percentage in MODIS imagery
P_l	%	Land cover pixel percentage in MODIS imagery
P_m	%	Missing pixel percentage in MODIS imagery
P_s	%	Snow cover pixel percentage in MODIS imagery
P_w	mm	Accumulated precipitation in a winter period
r^2	-	Coefficient of determination for statistical analysis
σ	-	Standard deviation of a data set
S_w	mm	Calculated accumulated snowfall in a winter period
T_h	°C	Temperature at a certain elevation h
T_s	°C	Temperature at the weather station
T_w	°C	Mean temperature in a winter period
Z	-	Test statistic

List of Figures

2.1	Schematic view of the general approach used in this study. The methodology is divided in two parts: data analysis and hydrological modeling. The bold text refers to sections, the related section number is shown between brackets. The italic text depict outputs of a process.	8
2.2	The Merced River basin in the Southern Sierra Nevada(CA) with the locations of the weather station and the discharge station.	8
2.3	The hypsometric distribution of the Merced River basin. This distribution shows the percentage of catchment area within a elevation zone of 100 meter.	9
2.4	The relation between temperature and snowfall, recorded at the weather station. The red line shows the median, which is the temperature threshold used in the hydrological model.	11
2.5	Figure from Krajci et al. (2014) where the error-value (I_s) is plotted against elevation. The RSLE is found at the point where I_s is minimum	11
2.6	Flowchart of the RSLE methodology	12
2.7	The correct regional snow line elevation (RSLE) timeseries for the water year 2014. This figure is an example for the elevation zone 2200-2300 m.a.s.l. with the blue area as the snow-free area. The derivation of this area by the magnitude of the elevation zone provides snow-free days.	14
2.8	The modeling scheme with the storage reservoirs (blocks), fluxes(arrows) and conditions (diamonds).	20
3.1	Mean winter temperature (T_w) trends at the weather station over the timespan 1983-2017. Figure (a) is a scatter plot with regression line and 95 % confidence interval as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.	24
3.2	Accumulated winter precipitation (P_w) trends at the weather station over the timespan 1983-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.	24
3.3	Trends of snow-free days per elevation over the MODIS timespan 2001-2017. The blue bars refer to significant trends according to the Mann-Kendall test and the orange bars show non significant trends.	25
3.4	Snow-free days (SFD) trends at an elevation of 1300 m.a.s.l. over the timespan 2000-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.	26
3.5	Snow-free days (SFD) trends at an elevation of 2300 m.a.s.l. over the timespan 2000-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.	26
3.6	Snow-free days (SFD) trends at an elevation of 2300 m.a.s.l. over the timespan 2000-2017. Figure (a) is the scatterplot with regression line and 95 percent confidence interval drawn as an area around the line. The table shows the results of the Mann-Kendall test. Figure (b) shows the anomalies calculated as explained in Section 2.4.	27
3.7	The weighted snow-free days over the Merced river basin. The SFD of all elevation zones are weighted on the hypsometric curve (Figure 2.3) with weighted SFD as result. The trend gives an indication of the average behaviour of SFD in the Merced River basin. The table shows the results of the Mann-Kendall test related to weighted SFD.	28
3.8	R^2 per elevation zone for different models of SFD. The blue bars show R^2 with both P_w and T_w as inputs, orange bars for SFD dependent on T_w only and green bars for SFD dependent on P_w only. These models are calculated with the method described in Section 2.5.	29

3.9	The effect of the drought period on snow-free days at elevations of 1300 (a), 2300 (b) and 3800 (c) m.a.s.l. The blue line shows the trend line of SFD during the drought with both P_w and T_w incorporated, orange shows the effect of P_w only, green show the effect of T_w only and red shows the trend line if there was no drought at all. The methodology behind this is described in Section 2.6.	30
3.10	short	31
3.11	short	32
3.12	Time series of the soil moisture deficit together with a rolling maximum with an interval of 15 years. The rolling maximum is an approximation for the changing S_r	34
3.13	Visualization of the effective precipitation and its components. The upper bars show the rainfall and melt part, the lower bars show the total effective precipitation.	35
3.14	Modeled mean snow storage of the period 1983-2010 and 2010-2017. The legend shows the slope of the accumulation period, the starting day of accumulation, the slope of the melting period and the last day of the melting period. The slopes are calculated between the day of minimum (start of accumulation or end of melting period) and the day of maximum.	36
3.15	short	37
4.1	The temperature trend of mean temperature over a water year. The table shows the results of the Mann-Kendall test: test result, slope, intercept and p-value.	40
4.2	The relation between mean winter temperature (T_w) and accumulated winter precipitation(P_w). The table shows their inter correlation (R^2) and the p-value.	41
A.1	Calibration results for the day-duration factor (DDF). The optimal value is found at DDF = 5.40 with an performance of 89.7 %.	56
B.1	The modelled snow-free days per elevation over the period 1983-2017. The results are tested with the Mann-Kendall test (2.4), which is highlighted with the blue (H_a not rejected) and the orange (H_a rejected).	58
C.1	Boxplots of soil moisture deficit numbers per interval of 20, 15, 10 or 5 years. The green line shows the median, the boxes show the 25th/75th percentile and the whiskers show the 5th/95th percentiles.	60

List of Tables

2.1	Overview of the consulted data sources during this research	10
C.1	The 4 return periods and their corresponding extreme values (Maximum SMD) according to the Gumbel distribution. The calculation is done with the formula $x = \bar{x} + KS$. With \bar{x} as the mean SMD, S as standard deviation and the Gumbel frequency factor K.	59

A

Calibration results

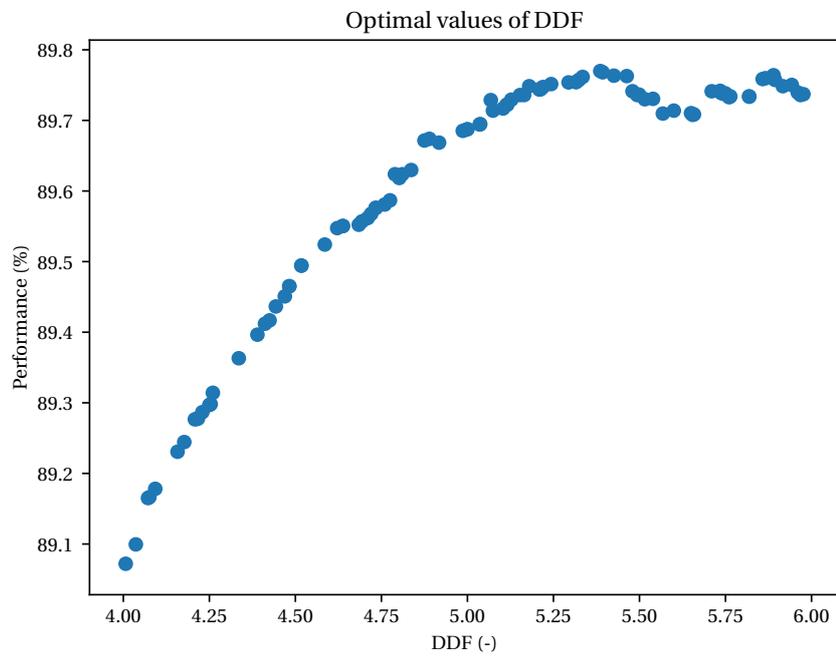


Figure A.1: Calibration results for the day-duration factor (DDF). The optimal value is found at DDF = 5.40 with an performance of 89.7 %.

B

Long term snow-free days

The hydrological model is calibrated on the observed RSLE according to the method described in 2.7.4. With this calibrated model it is possible to calculate SFD from 1983 until 2017. Before doing this, it is important to know how the model performs in the calibration period, 2001-2017. This is already presented in the results (Section 2.6, Figure 3.10).

From this figure it is clear that the model performs quite well, however it overestimates the SFD at 1300 m.a.s.l., especially before the drough period. Also the SFD in 2009 is over estimated at almost all elevation zones. Furthermore, there are some data gaps in the observed SFD. This is due to outliers and a large percentage of interpolated value (> 50%). According to the calibration values (Figure A.1) the current choice of DDF in the model shows a performance of 89.7 %.

The performance is high enough to use the model for extension of the SFD dataset and execute the same trend analysis as explained in 2.4. The results of the trend analysis with the Mann-Kendall test for every elevation is shown in Figure B.1. Comparing this figure with Figure 3.3 shows differences. First, the absolute values of the trends are much smaller by using a larger dataset. This is expected, since the most extreme values are present in the last 17 years. For that, the long term dataset will give a more realistic view on the development of the SFD trends. Another difference is the non-significant results between 2400 and 3300 m.a.s.l., however at the higher elevations significant and larger trends were found. According to this, the higher elevation zones are affected more than 3.3 suggests.

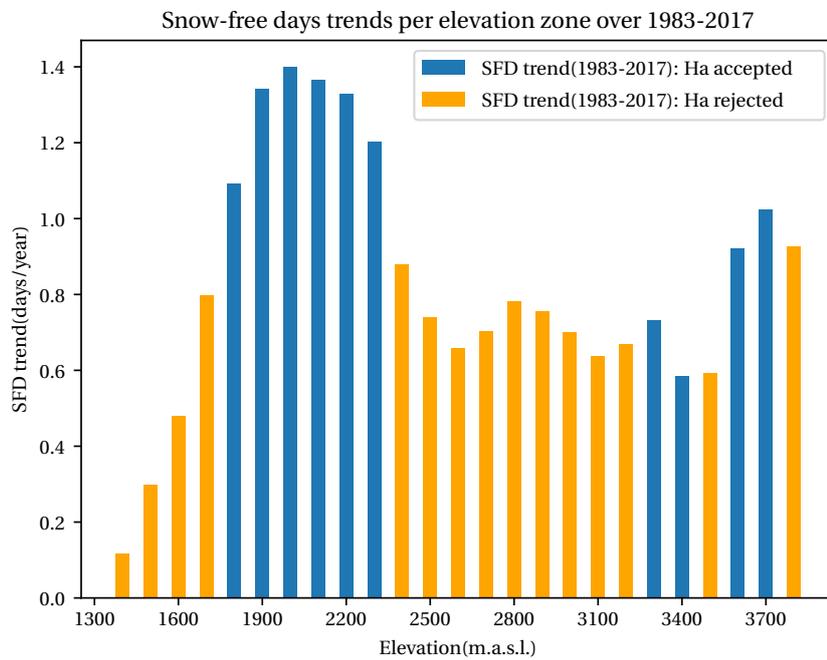


Figure B.1: The modelled snow-free days per elevation over the period 1983-2017. The results are tested with the Mann-Kendall test (2.4), which is highlighted with the blue (H_a not rejected) and the orange (H_a rejected).



Sensitivity analysis for interval of rolling maximum

To draw conclusions about the development of the root-zone the soil moisture deficit is calculated with a hydrological model. This method (described in Section 2.7) generates a time series of the soil moisture deficit. Within this model it is assumed that plants are able to survive these periods of soil moisture deficit because they adjust their roots to extreme situations. With increasing extreme drought events it means that the root-zone storage capacity has to increase too and for that the California Drought is a good event to test this hypothesis. However, plants do not change their root zone capacities instantly. Gao et al. (2014) showed that this happens with return periods between 10 and 20 years for the Seasonal Western Region where the $S_{r_{max}}$ is larger than 200 mm, exactly the case of the time series presented in Figure 3.12. This is also in line with Wang-Erlandsson et al. (2016). However, in mountainous areas there is a lot of short vegetation with not large root systems thus it might be that the vegetation of this catchment can adapt their roots faster. For that, a sensitivity analysis is conducted with intervals of 20, 15, 10 and 5 years. This is done by calculating the rolling maxima and presenting their statistics in box plots.

The boxplots show the statistics of rolling maxima with different intervals. As stated before, according to literature the root-zone moisture capacity has to be above 200 mm with a return period between 10 and 20 years. According to the boxplots, the maxima above 200 mm are only obtained with return periods of 15 and 20 years. The 5 years interval shows too much spreading (5th and 95th percentile differs with 100 mm and the 25th percentile is lower than 200 mm) which is not valuable to show a long-term process.

From the boxplots, it is clear that a 5 years interval is not suitable for this research. A better approach is to use a Gumbel approach to determine the most suitable interval. The maximum in the soil moisture deficit timeseries is found at 265 mm. The data can be fitted to the Gumbel distribution and from that the extreme value can be calculated related to the return period. For the return periods 5, 10, 15 and 20 years is used. This is shown in Table C.1.

Return period (years)	Mean SMD (mm)	S (mm)	K (-)	KS	Maximum SMD (mm)
5	176	50.1	0.845	42.320	218
10	176	50.1	1.506	75.424	251
15	176	50.1	1.879	94.105	270
20	176	50.1	2.139	107.127	283

Table C.1: The 4 return periods and their corresponding extreme values (Maximum SMD) according to the Gumbel distribution. The calculation is done with the formula $x = \bar{x} + KS$. With \bar{x} as the mean SMD, S as standard deviation and the Gumbel frequency factor K.

The most extreme situation regarding soil moisture in this dataset was 265 mm which corresponds with a return period of 15 years. Under the estimation that plants adapt their roots to these extreme situations the interval of 15 years is chosen to describe the development of the root zone.

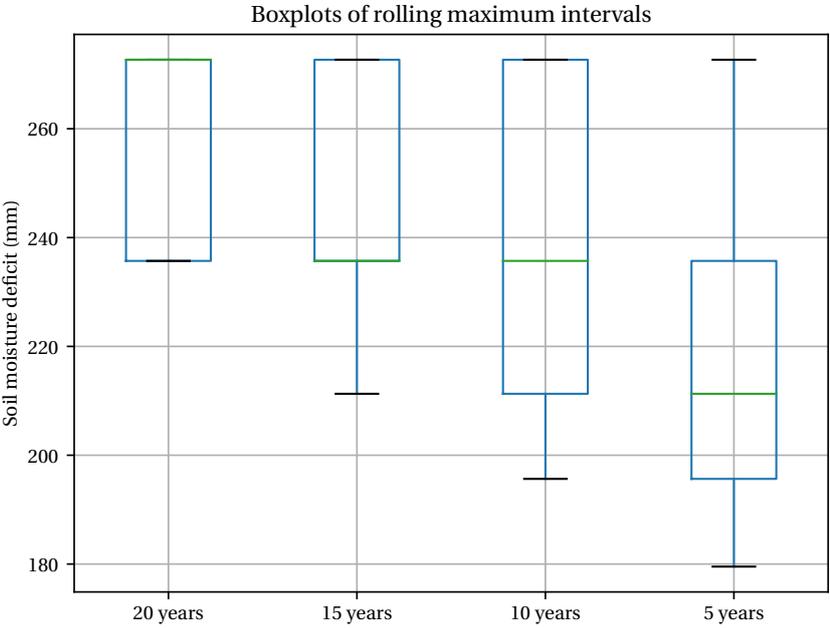


Figure C.1: Boxplots of soil moisture deficit numbers per interval of 20, 15, 10 or 5 years. The green line shows the median, the boxes show the 25th/75th percentile and the whiskers show the 5th/95th percentiles.