Effect of land-use change and climate variability on hydrology during the Holocene in Europe

Master Thesis

Utrecht University
Sustainable Development – Environmental Change and Ecosystems

Student
name: Marieke van Liere
e-mail: m.m.vanliere@students.uu.nl
number: 6211496

Supervisor
name: dr. Kees Klein Goldewijk
e-mail: C.G.M.KleinGoldewijk@uu.nl

Abstract

This research compared the land use change (LUC) of six different time periods with the Holocene, being 10.000 BCE, 1.000 BCE, 100 CE, 1300 CE, 1850 CE and 1950 CE, of the HYDE 3.2 data base to the hydrological responses in variables discharge, evaporation, evapotranspiration, groundwater recharge and runoff modelled with PCR GLOBWB 2.0. The research, using a time slice experiment method, run the hydrology model for 25 climatic year, of which observations were statistically tested against the changing land use. Overall LUC over the six time slices did impact the evaporation, groundwater recharge and evaporation, but differed greatly within regions based on cultivation type, topography and precipitation. The research focused on a spatial analysis, as well as a river flux analysis. The spatial analysis showed a strong positive correlation in runoff with the transformation to rainfed crop, and a simultaneous strong negative correlation in evaporation. Irrigated crops had the opposite effect, and had a strong negative correlation with groundwater recharge. Analyzing the rivers, their different LUC pattern greatly affected the significant response found in the hydrology variables. The river Dniepr endured not much LUC up to 1950, together with the construction of reservoirs, lead to a spiked increase in evaporation, and decrease in runoff. For the Danube and Rhine a more gradual LUC pattern occurred, and these rivers both show a gradual increase in runoff and decrease in evaporation over the time slices. For the Rhine only 1950 was an exception, as non-natural vegetation decreased under population growth, and a decrease in runoff and increase in evaporation was seen. The river Po has known maximum agricultural fractions since 100 CE, during the Roman Empire, and endured decreased amounts of evaporation and increased runoff before other rivers. The introduction of irrigated crops however changed this significantly, and evaporation increased and groundwater recharge decreased as a result. These results indicate that the extend of consequences LUC has on hydrology, and with that the climate system, is severely enough to be considered for water management and future climate modelling.

Table of Contents

1. Introduction

Large human interference of ecosystems by land use change (LUC) has occurred since early agriculture settlement, leading to threatening global effects in recent decades in terms of greenhouse gas emissions and its effects on global temperatures (Ellis et al., 2013; Pielke, 2005; Ruddiman, 2013; Houghton, 1994). Arneth et al. (2019) issued in an IPCC report that in 2015 the used land by humans is 71% of total global ice-free land surface, of which 12% is cropland, 37% grazing land and 22% forestry (Arneth et al., 2019). This is a great increase in human land use over the course of the Holocene, when comparing it to the estimated presence of 1% of cropland and pastures around 1,000 BCE (Klein Goldewijk, Beusen, van Drecht & de Vos, 2011). Land, water and atmosphere interactions are fundamental for understanding the functioning of ecosystems, and as the biogeochemical processes of these ecosystems are influenced by its vegetation, LUC makes an important factor in studying these processes (Kling et al., 2014; Pielke, 2005; Betts, 2006; Wood, Hannah & Sadler (eds.), 2008). LUC impacts terrestrial carbon, increases the use of inorganic nitrogen fertilizer and irrigation water, and decreases global biodiversity (Houghton & Doodale, 2004; Shukla, 2019). With an ever increasing population and its rising demand for fresh water and food, the conversion of natural vegetation to intensive agriculture might jeopardize water quantity and quality as we know it (Arneth eatl., 2019; Ruddiman, 2013). Looking at hydrological aspects that are influenced by drastic LUC, the water quantity and quality, erosion, siltation, rainfall and more are undesired effects of human activity (Dickinson, 1991; Calder, 1992).

Although studies have assessed the impact of LUC on hydrological properties, research on large-scale LUC in relation to hydrology during the Holocene is still not elaborately studied. Two of the main difficulties in identifying and quantifying the effects of LUC on hydrology are the only relatively recent hydrological records and the relatively high natural variability of most hydrological systems (DeFries & Eshleman, 2004). Important consequences of urbanization and LUC, such as runoff and groundwater recharge, jeopardize ecosystem health and fresh water supply (Harbor, 1994). The effect on groundwater recharge and evaporation in relation to LUC is highly dependent on the type of land cover. Different vegetation types have different leaf areas, soil exposure and root systems, which are important determinants of groundwater recharge and evapotranspiration (Zhang, Dawes & Walker, 2001; Calder, 1998). Evaporation and transpiration are relatively high responsive components in the hydrological cycle (Huntington, 2006), and thus make good indicators for a possible distortion in the water balance as a results of LUC. However, field measures of evapotranspiration (ET) are not easy. Evaporation and transpiration occur simultaneously, and when measured cannot be distinguished easily. Soil evaporation is mostly present when the soil is exposed or the crop is still small, while transpiration occurs dominantly when the crop has grown, and provides the soil with more shade (Allen et al., 1998). In tropical regions annual average evapotranspiration is 39% lower in agricultural ecosystems (Dias et al., 2015). A study on European land cover shows that the transpiration of natural vegetation can be bigger or smaller compared to agricultural fields, depending on the age and type of the forest and crop (United Nation University, sd). Generally a decrease in evapotranspiration is seen when natural vegetation is transformed to agricultural vegetation (Bonan et al., 2012). Most of the literature on the topic of LUC and hydrology, focuses on runoff and river discharge. For example, the local study of the Belgian Dijle catchment that measured an average river discharge increase by ±7% and increased discharge peaks due to LUC (Notebaert, Verstraeten, Ward, Renssen & Rompaey, 2011). Focusing on increased floods, Benito et al., (2010) presents a study in Spain that suggests an increase in frequency and magnitude of floods between 1830 and 1900 as a result of climatic variability, intensive deforestation and land use practices on the Guandalentín river. A global study by Bosmans et al., (2017) on land cover change and hydrology between 1850 and 2000 shows an overall decrease in tall and short vegetation and an increase in cropland and pasture. Simultaneously the global annual mean river discharge rate increased with 1,9%, showing clear

differences between the hemispheres, as the northern hemisphere endured only little LUC in this period (Bosmans et al., 2017). Thus, for the effects of LUC on hydrology a spatially large scale of Europe, an earlier time frame must be applied. Lastly, Bosmans et al. (2017) stresses the importance of including LUC when assessing the anthropogenic impact on the global hydrology cycle. The studied effects of LUC on the considered hydrological variables indicate that humans may have changed the hydrology cycle as soon as they started transforming natural vegetation for cultivation purposes. Looking at how the hydrological cycle has been affected by humans in Europe, an early timeframe must be applied, as deforestation started at the early beginning of the first agricultural revolution or Neolithic revolution, about 10.000 BCE (Weisdorf, 2005). Human induced effects on this planet are much more prevalent prior to the industrial era than generally assumed, partly due to the higher per-capita land-use in preindustrial era's (Ruddiman, 2013).

To understand past responses and possible future hydrologic responses to land use change on a larger scale, a reconstruction of past LUC on the hydrology cycle can indicate the extent of their interactions. In this research the hydrology properties river discharge, groundwater recharge, evaporation and evapotranspiration are considered, as they represent the basic important terrestrial hydrology components. With these components the local and regional water balance can be assessed on changes in land cover. The aim of this research is to determine to what extend land use change, the loss of non-natural vegetation, during the Holocene in Europe had an effect on the regional hydrology. By hydrologically modelling the different periods in the Holocene with their corresponding land use, the impact of land use change can be studied, leading to the following research question:

"To what extend did land use change affect hydrology in Europe during the Holocene?"

By using a time-slice experiment method, different time periods within the Holocene can be assessed on land use change and consequently hydrology. The following sub-questions have been determined to support the main research question:

A. What is the local effect of LUC on the river water balance and discharge during the Holocene?
B. What is the regional effect of LUC on the runoff, groundwater recharge, evaporation and evapotranspiration values during the Holocene?
C. What are the differences among the uncertainty estimation scenarios on LUC fractions for the above named hydrological variables?
D. Is there any difference between the non-natural vegetation types and the extent of their effects on hydrological variables?

Seen the increasing demand for water and food as a result of population growth and economic development, the pressure on land availability and water resources augments globally. This emphasizes the importance of understanding land use change effects on hydrology (Compagnucci et al., 2001; Smith et al., 2014; Shukla et al., 2019). Both LUC and hydrology are important factors in the climate systems (Pielke, 2005; Betts, 2006). Understanding past impacts creates perspective to global changes and its extent and magnitudes for future LUC and management (Ellis et al, 2013). With the expansion of information on historical land use and land cover, the understanding of long-term relationships between land use and climate is necessary for future modelling projections (Klein Goldewijk & Verburg, 2013). Additionally, when considering ecological restoration as a climate mitigation strategy, historical information is needed on past hydrological responses to land use change and could make a relevant contribution to the discussion (Harris, Hobbs, Higgs & Aronson, 2006).

To visually support the outline of this research the conceptual frame, figure 1, shows the outline of the research and relations between key concepts and variables central in this study. The key concepts and a

substantiation for the chosen time slices is further explained in chapter 2. In chapter 3 the research methods are discussed. Starting with an overview of the vegetation maps from HYDE, the results on local and regional analysis on river discharge, groundwater recharge, runoff, evaporation and evapotranspiration are presented in chapter 4. In chapter 5 and 6 the research is concluded and discussed, followed by the reference list and the appendix.

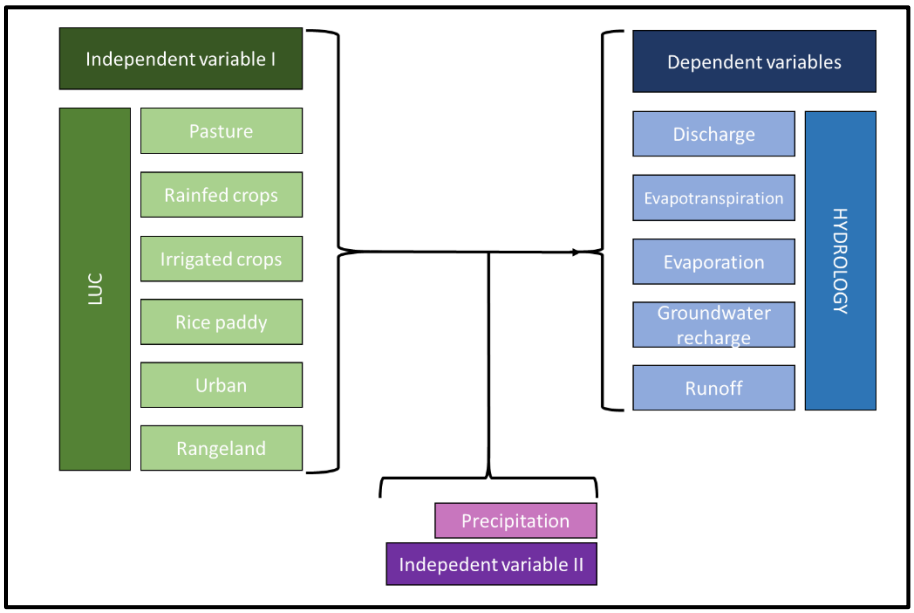

Figure 1, conceptual framework

2. Theory

First an historical introduction on the development of agriculture during the Holocene is discussed, to which six historic time-slices are highlighted that form the foundation of this research. The key concepts of the research are elaborated on, being land-use change, climate variability and hydrology, leading to the hypotheses for this research.

2.1 Background literature

The first agricultural revolution, also called the Neolithic revolution, can be traced back around 10,000 BCE, when humans first shifted from foraging towards agricultural settlement (Wirtz & Lemmen, 2003; Childe, 1935). But even before, though exact dates are hard to recover, humans altered the landscape with slash-and-burn and planting techniques to promote the growth of favorable edible species (McConnell, 1992; Gammage & Gammage, 2011; Doughty, Wolf & Malhi, 2013; Roosevelt, 2013; Levis et al., 2017). In the Early Neolithic agriculture was restricted to warm, dry climates with very fertile soils (Rösch, 2013). The fractions of these agricultural practices are small, and the land cover thus is almost all natural vegetation and rangeland (Klein Goldewijk et al., 2011).

In the late Neolithic, around 2000 BCE, the cropland and pasture is assumed to be mostly modest and extremely uncertain with 0,1 to 5 km^2 cropland and pasture per grid cell in parts of Europe (Klein Goldewijk et al., 2011). The exact extend of farming and farming methods are unknown, but deteriorated soil fertility and decreased yields were likely to occur after centuries of agriculture (Rösch, 2013). The beginning of the Bronze Age (3000 BCE to 1200 BCE) brought many mechanic tools, and this period is known for the expansion and intensification of agriculture (Rösch, 2013; Bartosiewicz, 2013). The use of fertilized cultivation methods such as slash-and-burn and husbandry systems increased yields, but also shows first strong evidence of permanent deforestation, suggesting the presence of fields or forest pasture with alternately fallow to pastoral use (Rösch, 2013). Both population and LUC were slightly increasing during these millennia (Klein Goldewijk et al., 2011; Klein Goldewijk et al., 2010), enhancing human impact on the environment. The human impact on the environment through LUC became stronger in the Iron Age. The Iron age (1200 BCE to 550 BCE) is said to see an increase from the 3% of hectares of agriculture use in the Late Neolithic to 50% while the suitable 39,000 hectares of agriculture had remained the same in Central Europe (Rösch, 2013).

During the rise of the Roman Empire, from 700 BCE to 200 CE, agriculture intensified due to climatic stability of favorable conditions (McCormick et al., 2012). The Roman economy was mainly agrarian based, meaning that agriculture and trade dominated the Roman economic fortunes (UNRV, sd-a). Study shows that the Roman government divided the colonies in an ideal way to optimize land use and they implemented irrigation infrastructures (Willi & Kolb, 2014; UNRV, sd-b). The Roman agricultural regime was intensive and diversified, and high farming practices were applied on a large scale. With the use of convertible husbandry, crop rotations, heavy manuring, drainage and irrigation, pruning and grafting, and improved fodder crops the produced yields and livestock numbers were high (Kron, 2013). The type of produce in Roman cultivation entailed a wide range of cereals, viticulture, horticulture and livestock, and the cultivation of olives, legumes and nuts on a more moderate scale (Geraghty, 2007). Between 200 and 400 CE the instable climate, with cooler and drier periods (McCormick et al., 2012), and its effect on water supply, agricultural productivity, human health and civil conflict is believed to be a big influence on collapse of the Roman Empire (Büntgen et al., 2011).

With the collapse of the Roman Empire, and the start of the Middle Ages (476 CE to 1453 CE), human impact generally increases in terms of LUC, but mid-15th century endures a slight decrease in cultivated area (Rösch, 2013). Agriculture that originated from the Roman era proceeded, with orchards, horticulture and the cultivation of various cereals, fibers, vegetables and spices found to be cultivated all over Europe (Rösch, 2008; Pearson, 1997; Crabtree, 2010). Medieval agriculture is mostly associated with a growing population, with deforestation as a result, and associated with increased yields under overall temperature rises around 1,000 CE, the so-called Medieval warm period (Astill & Langdon, 1997). In the 10th and 11th century new crops, three-field planting schemes, harness facilitating horses as draft animals, new power sources such as wind and watermills resulted in food surplus. Decreasing temperatures lead to the ending of the Medieval warm period and the start of 'The Little Ice Age', and contributed to fall of Europe into an cycle of poverty around 1250 CE (Gottfried, 2010). Around 1300 CE a highly lethal epidemic, the Black Death, reduced European population by 60% in the 14th century alone. This resulted in the destruction or depopulation of many villages, towns and cities, consequently leading to the abandonment of agricultural grounds that got reclaimed by nature (Mahoney & Nardo, 2016; Bork and Lang, 2003).

The second Agricultural Revolution, starting around 1815, consists of many changes that transitioned previous medieval European open-field farming to twentieth-century factory farming, as described by Thompson (1968). With the use of oilseed cakes in the 18th century as livestock feed or as fertilizer, and with the later economic and technological shifts the second Agricultural revolution began. The technical change consists of crop rotations and livestock improvement in relation to the demand of the market, and the use of machinery such as mowers, reapers, binders, tedders, threshers and horse labor. Additionally the range of crop variety, but also seed improvement by seed-raisers, led to increased intensity of cultivation (Thompson, 1968).

The more recent agricultural revolution is the so-called third or "green" revolution, that emerged after the second world war. The growing demand for food led to the increased use of artificial fertilizer and pesticides (Pimentel, 1996). This period is characterized by the technological revolution, regarding fertilizer and pesticide use, seed improvements, genetic modification and technological mechanization (Martín-Retortillo & Pinilla, 2015).

From this summary of agricultural periods in Europe, six historic time-slices are distinguished that historically significantly differ in theory, for which a visible change in LUC is assumed. The following six time slices represent the age of that time period, for which the vegetation fractions from the HYDE database 3.2 are retrieved:

- Time slice 1 (10,000 BCE): *Agricultural Revolution/Baseline* - The start of Neolithic revolution when foragers first transitioned to agrarian settlements.
- Time slice 2 (1,000 BCE): *The Iron age* - Most cropland and pasture area has been established between the foraging and first major civilization in Europe.
- Time slice 3 (100 CE): *Roman Empire* - During the most productive years of the Roman Empire when agriculture had strongly intensified.
- Time slice 4 (1300 CE): *Middle Ages* - When the Black Death took many lives, agricultural land had been reclaimed by nature, changing the land use significantly.
- Time slice 5 (1850): *Second Agricultural Revolution* - During the industrial revolution the mechanization of agriculture led to much change in agricultural practices.
- Time slice 6 (1950 CE): *Third Agricultural/Green Revolution* - When agriculture as the result of industrialization, rise of artificial fertilizer and pesticides, and genetic modification intensified practices and enabled agriculture on previously unsuitable land areas.

2.2 Key concepts

2.2.1 Land-use change

Changes in land use by humans are linked with population growth, economic and technological development, and environmental change (Houghton, 1994). LUC is a process induced by human activities that transform the natural landscape (Paul & Rashid, 2017), and occurs in in the form of land area and intensity of use (Houghton, 1994). It is rather uncertain how much greenhouse gas emissions LUC is responsible for, but an IPCC report estimated about a quarter of total human greenhouse gas emissions is caused by deforestation, livestock and fertilizer (Arneth et al., 2019). LUC affects soil carbon stocks, with 10% carbon loss when a pasture is changed to a plantation, 13% loss from native forest conversion to plantation, 42% loss from native forest conversion to crop, and the highest soil carbon loss is with 59% when pasture is converted to crop. Carbon stocks increase with 8% when native forests are changed to pastures, 19% increase for the conversion from crop to pasture, 18% increase for crop to plantation conversion and a LUC from crop to secondary forest increases soil carbon stock with 53% (Guo & Gifford, 2002).

Forests can impact water and climate at local scale, through the seven principles of change in water cycles: soil infiltration and groundwater recharge, recycling of precipitation, create precipitation triggers, transport of atmospheric moisture land inwards, soil moisture as solar radiation deflection and the overall dispersion of water (Ellison et al., 2017). Forest soils show faster water infiltration rates and percolation than agricultural soils due to 1) the surface organic layers, 2) large numbers of root channels extending into subsoil, and 3) a high content of large soil pores (Carmean, 1957). Twine et al., (2004) identified the consequences of converting forest and grassland to cropland. When forests are transformed to cropland, surface albedo increases, capony height decreases, Leaf Area index decreases and the rooting system becomes more shallow. For grassland to crops, the surface albedo and Leaf Area index vary greatly between seasons (Twine, Kucharik & Foley, 2004). The average ET in agroclimatic regions for a temperate region, with a cool (10°C) temperature is 1-3 mm/day, moderate temperature (20°C) is 2-7 mm/day and warm temperature (30°C) has about 4-9 mm/day, with different ranges between humid and arid regions with the temperate climate (Allen, Pereira, Raes & Smith, 1998). A study by Govindasamy et al. (2001) suggest that the long-term cooling between 1,000 and 1,900 in the northern hemisphere can partly be explained by the extensive LUC. Additionally, in contrast to studies of tropical forests, the hydrological cycle on global scale is likely not greatly affected by deforestation in Europe, as in the mid-latitudes the precipitation dynamics is less affected by local convection and evapotranspiration recycling (Govindasamy, Duffy & Caldeira, 2001). When modelling LUC Veldkamp & Lambin (2001) stress the importance of multi-level interactions and feedbacks, as only unidirectional impacts are now incorporated and assessed, preventing self-organizing behaviour within the model. However, the inclusion of such dynamic model features requires high computing power (Veldkamp & Lambin, 2001).

2.2.2 Climate variability

Although climate variability is not included in the research as a varying independent variable, the climate does play a major role in both hydrology and natural and human land-use change. Climate variability entails the variation of temperature, precipitation and the interaction between those two over time (Ray, Gerver, MacDonald & West, 2015). Average temperatures from the Holocene are well-known, figure 2, and have been fluctuating around the annual average of 15˚C compared to the current 19˚C (Brauch, Spring, Bennett & Oswald (Eds.), 2016). As studied by McCornick et al., the period of 100 BCE to 200 CE is characterized by climatic stability of favorable conditions. The relative warm temperatures compared to its preceding years were the results of two potential climate-forcing factors that have remained unusually stable during this period: solar activity and volcanic activity. Towards 200 CE there is a cooling trend, as volcanic activity increases around 150 CE. The overall wetter climate has also believed to become to an end around 200 CE. Between 200 and 400 CE the climate gets more instable with cooler and drier periods, and the division between Eastern and Western climate within Europe becomes more evident (McCormick et al., 2012). After the last Neoglacial cooling event, the Medieval warm period thrives for about a millennia, before the Little Ice Age dominates between 1300 and 1870 CE (Brauch, Spring, Bennett & Oswald (Eds.), 2016). More recent climate reconstructions, via tree ring methods, shows detailed year to year climate variations regionally across European (Briffa, 2000). Ever since the 1950's there is an increasing global warming trend (IPCC, 2013; Brauch, Spring, Bennett & Oswald (Eds.), 2016)

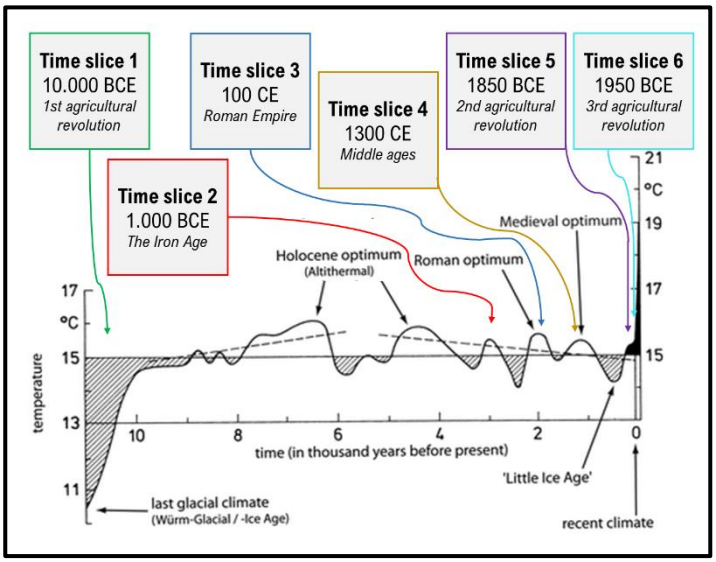

Figure 2, Reconstruction of the Holocene climatic fluctuations with time slices. Original source: Brauch, Spring, Bennett & Oswald (Eds.), 2016.

Climate variability highly interacts with hydrology and agriculture. A study by Ray et al. shows that in areas of global breadbaskets more than 60% of yield variability is explained by climate variability. Spatial patterns also show a response of yield variation explained by climate variability ranging from 0 to 75% regionally within Europe for maize and wheat (Ray, Gerber, MacDonald & West, 2015). The hydrology cycle and climate are two very entangled components, that greatly influence each other. The basic hydrology concept of water balance, is that from all precipitation fallen on land surface, the part that does not evaporate or percolates, is considered runoff. Research assessing water balance on global scale found that changes in runoff and evapotranspiration were controlled by precipitation variations (Liu et al., 2018). The groundwater levels also has corresponding changes with climate variability, but varies much more with local geology, land use and land cover (Russo & Lall, 2017). At last, principal weather parameters such as radiation, air temperature, humidity and wind speed affect evapotranspiration (Allen et al., 1998). Overall it can be concluded that climate and hydrology are highly interactive.

2.2.3 Hydrology

The hydrological cycle is connected to all biogeochemical processes in the biosphere, atmosphere, and cryosphere (Vörösmarty et al., 2001). Terrestrial hydrology is the state and dynamics concerning water at the land surface, and when testing general climate models the interaction and feedbacks with the atmosphere and vegetation is of high importance (Dickinson, 1991). Simply put, the hydrological cycle includes the evaporation from (ocean)water and evapotranspiration from land, that are mainly driven by solar radiation. Water condensates and the precipitation on land surface is either temporarily stored in as snow or soil moisture, or is runoff that form rivers and streams. The water, either by land surface in the rivers or through deep water layers in the soil by percolation, end up in the ocean, completing the hydrological cycle (Trenberth et al., 2007). DeFries and Eshleman (2004) stress the importance of hydrology as a focal point, especially now under global warming trends. A study by Huntington (2006) reviewed the intensification of the global hydrological cycle under increased global temperatures. This intensification may lead to changes in water availability, magnitude and frequency of storms, floods, droughts and shifts in seasons (Huntington, 2006). Land use change and climate change are expected to strongly impact the global water budgets and river catchments (DeFries & Eshleman, 2004). Major threats of global anthropogenic water use on the hydrologic cycle are through aquifer mining, desertification by overgrazing, wetland drainage, deforestation and dam building. These human activities lead to changes in water storage and sea level rise, distortion of continental runoff and retention of continental runoff (Vörösmarty & Sahagian, 2000). In that matter, large-scale land use change is thought to affect rainfall through evapotranspiration, runoff, sedimentation and siltation, erosion and water resource depletion (Dickinson, 1991; Calder, 1992).

When looking at the terrestrial hydrology, the basic concept of a water balance can be applied to review the effects of land use change to the hydrological cycle. A study by Li et al., (2007) simulated the total deforestation of two West African basins, which found an increased runoff and streamflow, despite the fact that only <5% of the basin area was forest. A hydrology basin simulation for the Pennar river basin, India, found increases and decreases in the basins, and recommended looking at a finer spatial scale, but overall found the hydrological components such as runoff, baseflow and evapotranspiration responding to land use changes (Garg et al., 2017). As for ET, the types of vegetation highly influence the amount of evaporation and transpiration. A study on Flemish forest and cropland showed 0.315 m/year transpiration in forests opposed to 0.261 m/year on agricultural soils. Soil evaporation was substantially larger with 0.047 m/year in forests and 0.131 m/year in croplands (Verstraeten et al., 2005). A study comparing natural and agricultural vegetation in Spain, found that a green forest had a 796 mm per hour infiltration rate in comparison to crop land with 67 mm per hour (Neris, Jiménez, Fuentes, Morillas, & Tejedor, 2012). The water infiltration rates and capacity for a Nigerian regrowth forest decreased when it had been turned into cropland. Along with the crusting that occurred, the most extreme decrease in infiltration rate was from 25,4 cm/hour to 9,1 cm/hour after three years (Wilkinson & Aina, 1976). Furthermore, a study by Thompson, Harman, Heine and Katul (2010) discovered about ~20% increase of mean infiltration capacity in hardwood forest compared to grass fields. Thus, forestation helps reduce runoff and peak flows, as vegetation increases infiltration and groundwater recharge (Zuazo & Pleguezuelo, 2009).

In this thesis the focus on the hydrological properties river discharge, groundwater recharge, runoff, evaporation and evapotranspiration were chosen as indicators. These indicators make up sufficient components of the hydrology cycle to draw conclusions on the possible of LUC on hydrology, and the literature research shows that a response to LUC is to be expected.

2.3 Hypothesis

From the conducted literature research on interaction between vegetation, especially land use change, and hydrology, the main hypothesis is that an influence is expected. Each vegetation cover has different characteristics that interact differently with hydrology, thus a change in natural vegetation to cultivated vegetation or a shift in type of cultivation, is expected to be seen. As in this research the LUC fraction for a specific time slice is used, the hydrology responses are simulated for that time slice, and the difference is hydrological response between the time slices indicate the effect.

"Over the Holocene time period, the change of a dominantly natural vegetated land to a more agricultural productive land use has an effect on the local hydrology."

This overarching hypothesis is very broad, and can be narrowed down to the specific hydrology properties in relation to the set sub questions. As found in the literature research the discharge data of a river is heavily influenced by precipitation, however, increased runoff was seen as a response to natural vegetation clearance.

"With the increase of land use change from natural vegetation to agricultural vegetation an increase in runoff is expected between the six time slices"

As river discharge is considered runoff, but with different measurement units, a similar hypothesis is set for the discharge of rivers. Additionally, the literature found increased discharge peaks in rivers. This leads to the following hypothesis.

"With the increase of land use change from natural vegetation to agricultural vegetation an increase in river discharge and discharge peaks is expected, over the six time slices"

The climate variables, such as precipitation and temperature, are held constant for each time slice, thus the change in hydrological properties is the sole consequence of the land cover difference between the time slices. From the literature evaporation and evapotranspiration decreases as a result from the loss of natural vegetation and the exposure of bare soil.

"With the increase of land use change from natural vegetation to agricultural vegetation over the time slices, a decrease in evaporation and evapotranspiration is expected."

The research also evaluates the effect of the different uncertainty scenario's of land use during the Holocene. These three different scenario's may indicate how strong the effect of LUC is, and with a lower, best and upper estimate, it is expected to see differences between these scenario's in the response values.

"Between the estimation scenarios of land use change, a difference is expected between the lower, best and upper scenario in regards to response of the hydrological variables."

Lastly, the LUC can be subdivided into the different land use types, as they all have different characteristics that influence the hydrology. As pasture most closely resembles natural vegetation, this land type is thought to have the least impact on hydrology, opposed to the other vegetation types.

"The relation between land use change and the response in hydrology is expected to be less strong for pasture, compared to the other land use types."

3. Methods

In this chapter the research design is explained. The chapter is divided into the strategy, data collection, analysis and research quality.

3.1 Research Strategy

The research design is a time-slice experiment to match the HYDE 3.2 land use change database to the PCR GLOBWB 2.0 hydrology model and retrieve possible response patterns and correlations between LUC and hydrology during the Holocene in Europe. The LUC for each time slice, in km^2 per grid cell, according to HYDE 3.2, is used in three variations: higher and lower uncertainty boundary and the best estimate. These are referred to as the estimation scenarios: best, upper and lower. The time slices represent the vegetation fraction for that year, run for a 100 year time period. With that each time slice represents its corresponding vegetation cover over the course of the same 100 climatic years. This climatic input is representative of the current climate, and the response in the hydrology model solely reflects LUC. Statistical analysis was used to indicate a possible significant difference between the hydrology variables due to LUC. The model output considered for the analysis include the variables total groundwater storage, total runoff, total evaporation, annual and monthly average river discharge and actual evapotranspiration. The analysis of data knows two strategies: the spatial analysis and the data point analysis. The spatial analysis reviews Europe in its entirety on the responses to LUC, and gives an overview on how LUC impacts hydrology. The data point analysis uses a very local approach to identify the water balance per major European river on 5 coordinates each. The research strategy, supported by the framework from figure 3, is elaborated on in the following paragraphs.

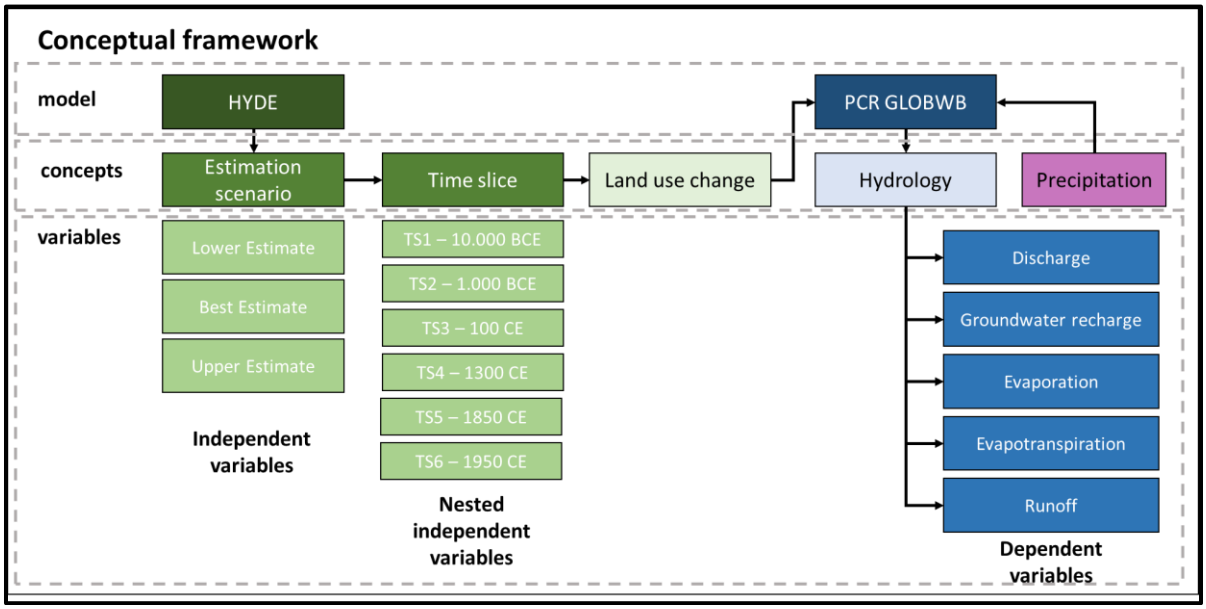

Figure 3, conceptual research framework

3.2 Data collection

The data needed for this research can be divided in different types of input data. Firstly, the input data consists of the HYDE 3.2 data for the time slices with different vegetation types, being the independent variable. The hydrological output variables from the PCR GLOBWB model are the dependent variables.

3.2.1 Data collection

The HYDE 3.2 database provides data as ascii files, of which the following were downloaded and used:

- Pasture
- Conventional rangeland
- Rangeland
- Irrigation paddy rice
- Irrigation paddy non-rice
- Total rainfed
- Urban area

The ascii files were converted to raster maps, and were divided to obtain fractional values instead of km^2/grid cell. Then the maps were rearranged into vegetation types that fit the vegetation input needed in the PCR GLOBWB model. A raster map of natural vegetation was created from the remaining fraction per grid cell resulting from compiling all of the above named raster maps in a non-natural vegetation fraction map. The natural vegetation was divided in 50% natural short and 50% natural tall for each grid cell, creating two new maps. The conventional rangeland and rangeland maps were then added to the natural short vegetation map. This resulted in the following vegetation fraction raster maps that are used as PCR GLOBWB vegetation input:

- Pasture
- Irrigated rice paddy
- Irrigated crops no-rice
- Total rainfed crops
- Natural short vegetation
- Natural tall vegetation
- Urban area

3.2.2 Model preparation

The PCR GLOBWB 2.0 model was prepared in Linux™ with Mobaxterm™. The overarching script with input files was adjusted to incorporate the vegetation maps. Each vegetation type was assigned with corresponding representative files containing values for that specific vegetation type, on variables such as soil depth, root fraction and crop coefficient. The soil, root and crop value maps of grassland were used for the pasture landcover fraction, as no specific file for pasture exists, and pasture and grassland are deemed very similar. For rainfed crop, files with values of rainfed cereal crops were used, except for the minimal and maximum soil depth fraction that was absent for rainfed. Instead the grassland value maps were used. At last, the urban area did not have files for snow and storage, thus files of pasture were used.

Aside from changing the vegetation input, the inclusion of water use for irrigation, domestic, livestock and industrial demand was disabled to solely retrieve the response of land use change between the different files. At last, only the natural waterbodies option was enabled for the first five time slices, to not take the reservoirs into account before 1900. The reservoirs were coded for the period after 1900 and thus for the last time slice the reservoirs were included. All variables from the standard PCR GLOBWB model file that were adjusted are presented in appendix section 1 where an overview of the input components is given. In total 18 different models were prepared, with three scenarios (best, lower and upper estimate scenario), that each had six time

slices (10.000 BCE, 1.000 BCE, 100 CE, 1300 CE, 1850 CE and 1950 CE). The models were run on a 100 year time period, with equal climate input data, hereafter referred to as the climatic years.

3.2.3 Output variables

When the runs were completed, the various variables produced in the form of NetCDF files were collected. From all output variables, only the following were extracted and used in this research:

- Monthly average discharge [m3/s]
- Annual average discharge [m3/s]
- Annual actual evaporation [m/year]
- Annual total evapotranspiration [m/year]
- Annual total groundwater recharge [m/year]
- Annual precipitation [m/year]

From these output files only 25 climatic years were analyzed, opposed to the initially 100 year run. This was due to a mistake in the input model, that lead to only 25 years of valuable data for analysis. This error is more elaborately discussed in the discussion. For all variables the mean, seasonal mean, standard deviation and variance were calculated, and saved as NetCDF files. Additional conversion and adjustment was needed to obtain variables that could be used to analyze the results. For example, for the analysis on the vegetation maps, the raster maps were masked for the same land area in Europe as the PCR GLOBWB model. This excludes the land area that is also not included in the variables and makes the extent of longitudes and latitudes equal.

For the river flux analysis a cell-based analysis is performed where specific grid cells are selected to be analyzed on the above named variables. Selected rivers were Danube, Dniepr, Po and Rhine. The coordinates of these grid cells were chosen based on the presence of LUC and were spread out as evenly over the length of the river as possible. Each river has five data points that together are called the series. These series were retrieved from the output variables for each time slice and each estimation scenario. The coordinates of these points in the river can be found in appendix section 3.

3.3 Analysis

PCR GLOBWB gives in a computational grid the hydrological output data in response to the input data. The goal is to explore the response of hydrological properties (dependent variables) on the changing land cover (independent variable) over six time slices for three different estimation scenarios. The time slices are nested within the scenarios and the time slices per scenario therefore must be split when analyzed. At first, the vegetation fraction as provided by HYDE 3.2 is visualized in maps, to indicate the expected LUC over the time slices and between the estimation scenarios. Two types of data are analyzed, with firstly the river flux analysis that works with continuous data over a climatic time period for one specific coordinate point. Secondly, there is a spatial analysis, that encompasses a large area and visualizes the hydrological variables in response to LUC. These (statistical) analysis are performed with *Rstudio*™. The visualization of the data collection to the analysis is shown in research framework figure 4.

3.3.1 River flux analysis

The research strategy of the river flux is based on a simple version of the water balance, in which precipitation, actual evaporation, total runoff and groundwater recharge are the main components. Additionally discharge is considered, but its different unit of measure [m^3/s] must be taken into account. The discharge of the river indicates the type of river, and might show how the discharge response to LUC over the time slices. Of all variables, the LUC is the independent variable to which the model responses. However, precipitation within the PCR GLOBWB model is also an independent variable. The precipitation is constant over the time slices

and scenarios, and is not included in statistical analysis, but is checked on multicollinearity with the dependent variables.

To analyze the output of the river data, a descriptive statistical analysis is performed for a general indication of the results. To compare the river flux of multiple time slices amongst each other, the ANOVA test is used to determine if groups statistically differ from each other. This test is only credible when the three assumptions are met, being 1) the groups are independent, 2) the values are normally distributed, and 3) that the variance within a group is significantly different from the variance of other groups. The first assumption is already violated, as the hydrological responses are confounding, therefor an ANCOVA test must be used, to control for covariance. For the second assumption an Shapiro-Wilk normality test is performed. When the test shows a p-value smaller than α=0,05, a normal distribution can be assumed. For the equal variance assumption Levene's test of homogeneity is run, that if significant with a p-value smaller than α=0,05, the assumption of equal variances is violated. If the assumption is not violated, the ANOVA can be performed, if the assumption is violated, a Welch ANOVA must be performed. The one-way (Welch) ANOVA test indicates whether the groups significantly differ from each other, but does not account for the specific time slices that differ within the group. Therefor the post hoc Games-Howell is performed to determine which time slices shows a significant difference in river flux in response to LUC. A post hoc Tukey HSD is performed when the homogeneity of variances is not violated.

3.3.2 Spatial analysis

The spatial analysis starts with the presentation of each variable (precipitation, evaporation, groundwater recharge, runoff, evapotranspiration and LUC) in a figure. When all time slices are described, a further analysis on individuals variables are conducted, being evaporation, groundwater recharge, runoff and evapotranspiration. With maps that differentiate between two consecutive time slices, show the increase or decreases of values for groundwater recharge, runoff, evaporation and evapotranspiration. Supporting the maps with descriptive statistics gives a first indication of the response of LUC. Hereafter spatial correlation analysis is performed between the vegetation maps and the dependent variables. Therefor at each time slice the vegetation map and the map of the dependent variable are compared, from which a correlation analysis and regression analysis can be retrieved.

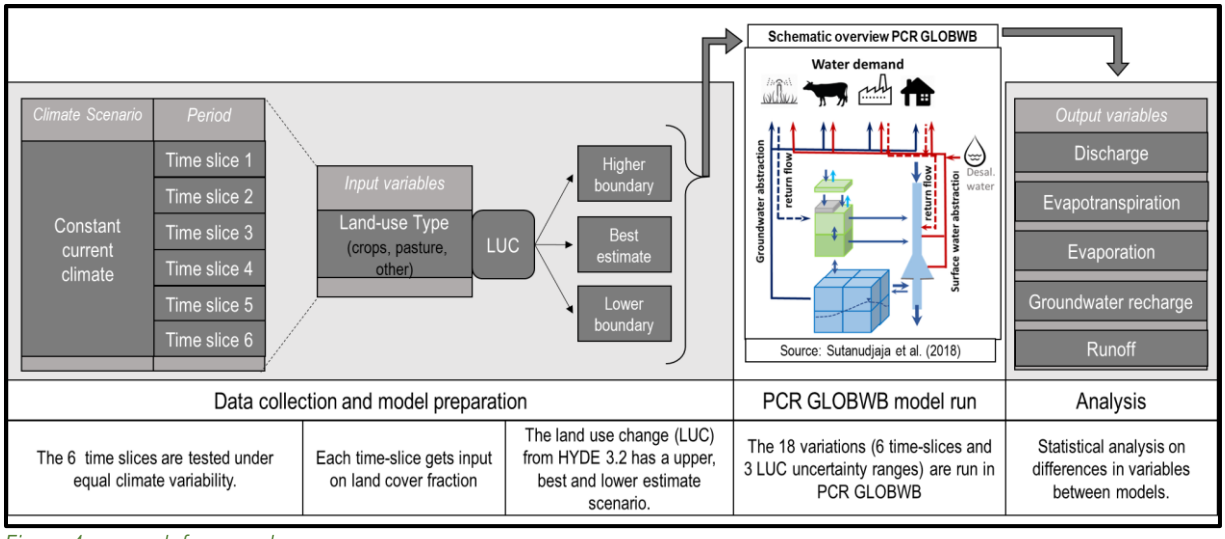

Figure 4, research framework

3.4 Research quality

To maximize the quality of this research optimization of the validity and reliability is strived for. Reliability refers to the consistency of results which the research will produce when repeated, and validity refers to the accuracy or correctness of the findings. Due to the high uncertainty in the data that will be used and the results from modelling, transparency on assumptions is key to ensure reliable and valid result interpretation. The three uncertainty scenarios are be compared to determine how sensitive the model is to LUC, and contribute to the reliability of the results. Also, findings are critically evaluated against similar studies. However, both the input data HYDE 3.2 and the PCR GLOBWB 2.0 model acknowledges uncertainties (Klein Goldewijk & Verburg, 2012; Sutanudjaja et al., 2018). In this research a simplified version of the PCR GLOBWB 2.0 is used, where many concepts and options were disabled or not used, that may fit the scope of the purpose of this research, but does not achieve the hydrological accuracy that the model can give. The focus of this research therefor restricts to the effect of LUC on hydrology, not the exact outcomes of values for the hydrology variables.

4. Results

The result chapter first reviews the land use fractions as provided by the HYDE database, to understand the land use change over the time slices and identify the magnitude of natural vegetation loss. Hereafter four rivers are further analyzed for 5 grid cells on their discharge and water balance. At last the spatial analysis is performed for the groundwater recharge, total runoff, evaporation and evapotranspiration in relation to the land use change.

4.1 Fraction non-natural vegetation

The HYDE 3.2 data provides the land use fractions per grid cell. In figure 5 the vegetation fraction of non-natural vegetation fraction is presented for each time slice for the best estimated scenario. The non-natural vegetation land use is barely present in the first time slice (deep green) and highly present in the last time slice (deep brown). Overall there is an increase in non-natural vegetation fraction in Europe over the six time slices. Standing out is the decrease in fraction of non-natural vegetation in the sixth time slice for the western European region, but the increase of non-natural vegetation in the eastern European region. The eastern European region greatly exists of natural vegetation and rangeland fractions, up until 1850.

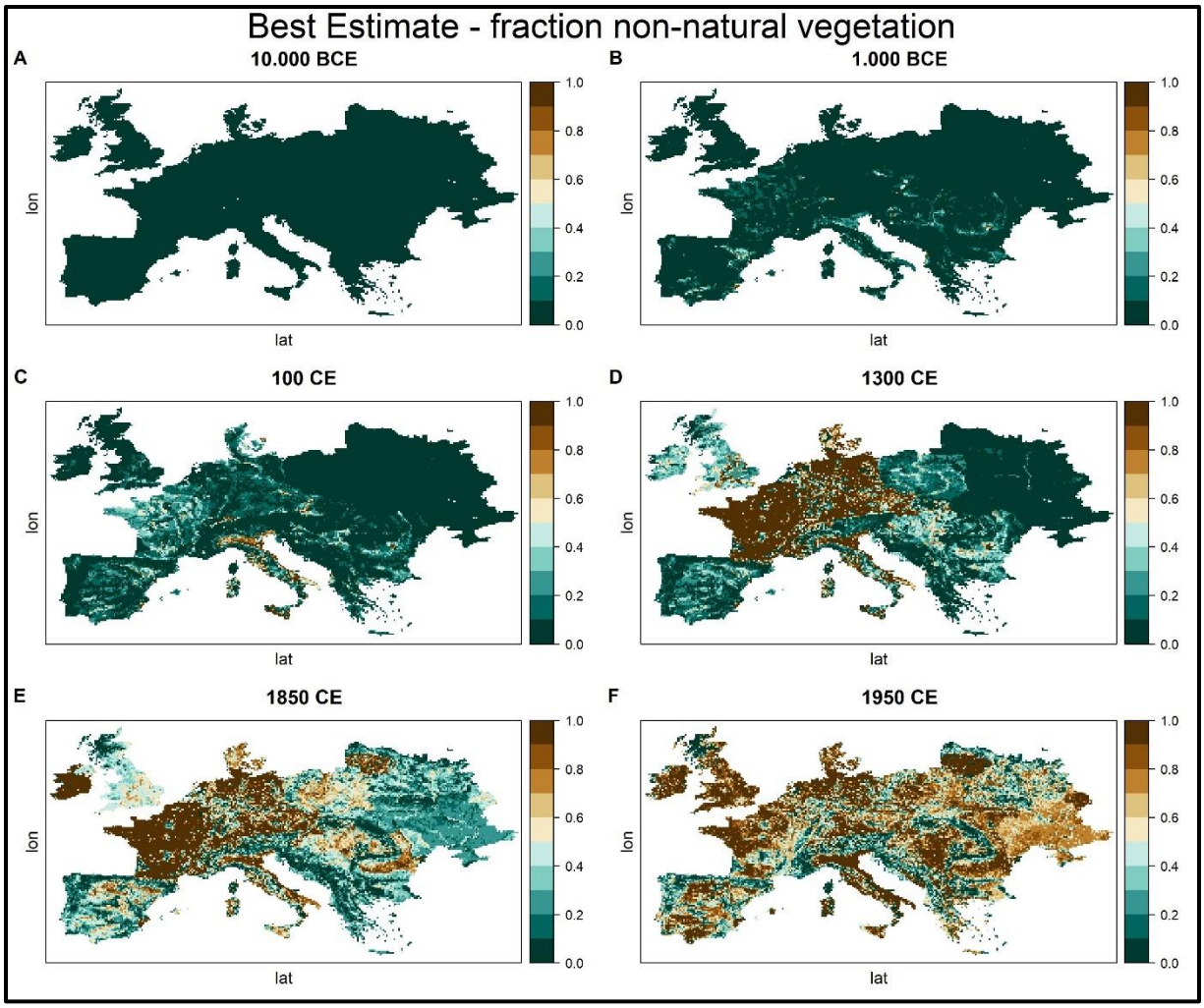

Figure 5, the vegetation maps, in fractions of non-natural vegetation, of the best estimate scenario over the six time slices.

A figure similar to figure 5 can also be found for the lower and upper scenario in appendix section 2. In the upper estimate scenario most notably is the increased non-natural vegetation fraction in Italy and eastern France in time slice 3 (100 CE), and the almost completely non-naturally dominated vegetation fraction in eastern Europe in 1300 and 1850. But overall the maps for the upper and lower estimation scenarios seemed similar to the best estimation scenario. In table 1 the mean values of non-natural vegetation for the time slice maps are presented from the descriptive statistical analysis that was performed. The non-natural vegetation fractions for the first time slice were very low (for best estimate 6.471141e-09), and the mean value increases for all the estimation scenarios steadily over the other time slices. The values align between the scenarios, with the lowest mean values for the lower estimate scenario, and the highest mean values for the upper scenario. The values show a wider uncertainty range for earlier time slices, and are more equal in the last time slice. To locate the difference between the upper and lower estimation scenarios in regards to the best estimate scenario, the different scenarios were subtracted from one another. The maps show the difference between the non-natural vegetation fractions between the Best and Lower, Upper and Best and Upper and Lower estimation scenario for the sixth time slice, figure 6. Here 0.0, the lightest brown, indicates no difference in land use change. The green color represents a gain in natural vegetation, and the brown color a gain in non-natural vegetation, compared to the other scenario. From figure 6 it becomes clear that the uncertainty range between the estimation scenario is spatially non-linear, and rather has specific regions on the map that may deviate from the best estimate scenario. The highest fraction of change between the scenarios is in central Europe, where both loss and gain of non-natural vegetation is present.

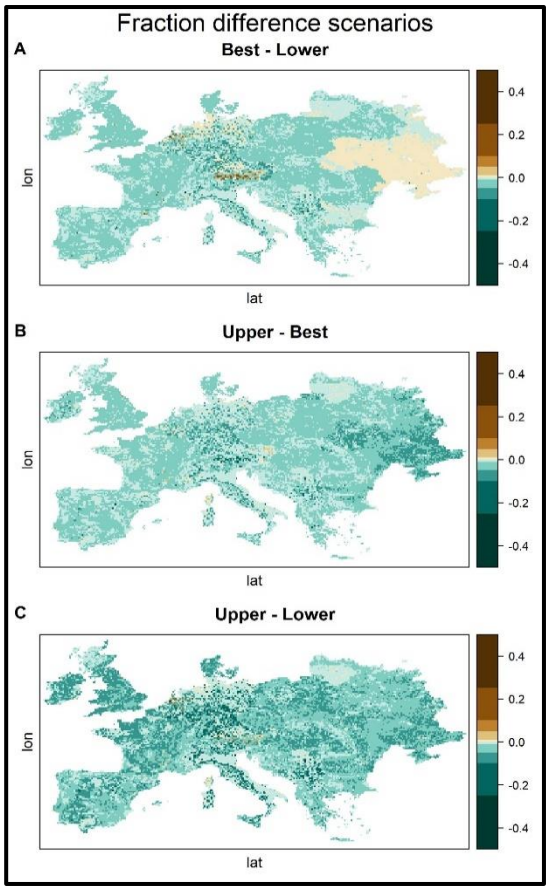

Figure 6, difference in vegetation fraction between the estimation scenarios

Table 1, mean values of non-natural vegetation area

Descriptive statistics non-natural vegetation maps						
	Time slice 1	Time slice 2	Time slice 3	Time slice 4	Time slice 5	Time slice 6
Lower	0.000	0.007	0.022	0.207	0.428	0.603
Best	0.000	0.033	0.098	0.334	0.473	0.618
Upper	0.000	0.065	0.180	0.427	0.518	0.640

The vegetation types mostly present in Europe are pasture, rainfed crops and irrigated crops (non-rice). The fraction of irrigated rice paddies and urban area were very small and even though they are included in the model, their results will not specifically be evaluated. The fraction of pasture is especially large in the British isles. The fraction of rainfed crops are very high in Iberia, northwestern Europe, Italy and large parts of eastern Europe. The irrigated crops are not present very much at all, up until 1950, and then only for northern Italy and parts in Iberia that have large irrigated crop fractions. The vegetation fractions, how it changes between the time slices, and the effect of different LUC types are further discussed in relation to the hydrologic variables in the sections below.

4.2 River flux water balance

For the four chosen rivers in Europe to be reviewed each five data points on which discharge data was gathered over the length of the river are analyzed. The coordinates of each point in the river are stated in appendix 3. The Danube and Dniepr are Europe's largest rivers, whereas the Po may be a very short river, but has known LUC in very early days of the Roman Empire. The Rhine is the largest river for western Europe, that completes the scope of the research with a total package of four important European rivers.

Figure 7 shows the average monthly discharge of one year, to indicate the discharge pattern of each river. It should be noted that the data for figure 7 was extracted from the climatic year 2014, which was a dry year with low precipitation rates. The monthly discharge over one year is generally larger, and is mentioned in following sections.

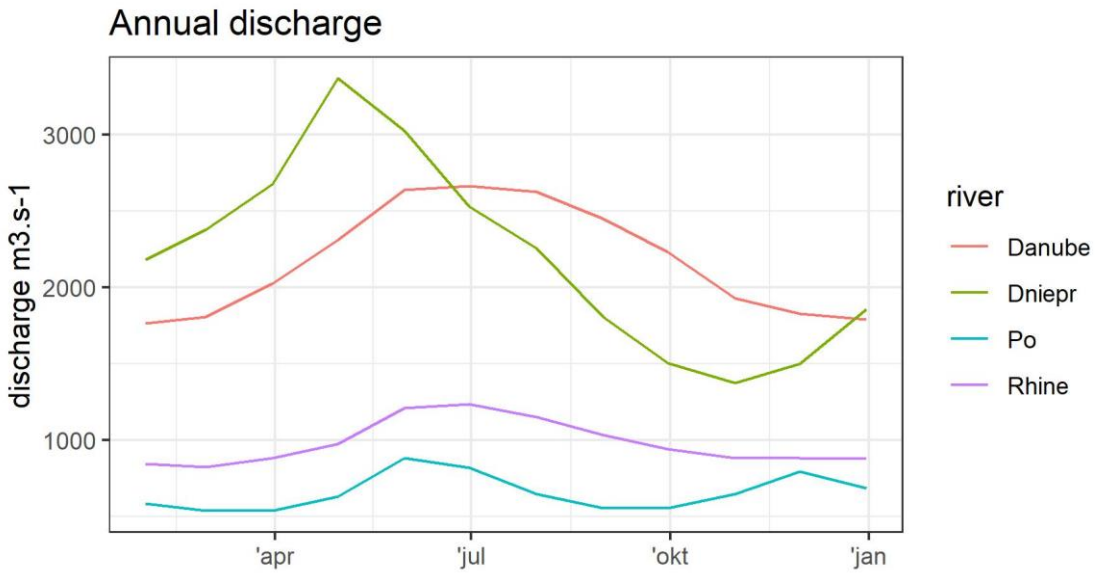

Figure 7, annual monthly mean of all rivers for climatic year 2014

4.2.1 Dniepr

The river Dniepr is a long river that starts in Smolensk Russia. The most inwards data points obtained for this research is in Ukraine, as the river discharge rates would be too dispersed if chosen on the entire length of the Dniepr. The average discharge of the Dniepr river at Smolensk was 1.500 m^3/s, opposed to the 2.000 m^3/s to 5.000 m^3/s discharge rate the current data points shows. Aside from the discharge rates, the Dniepr was under great human development in early 1900. Dam and reservoirs were constructed for drinking water and hydropower. These non-natural waterbodies are included for time slice 6 (1950 CE) in the model, and with the expectation that this would influence the data point data only data points were taken that are present in this highly constructed area. In figure 8F indeed shows a very different distribution from other time slices, most strikingly is the rise of Dniepr 4 series, that now has the highest discharge rate, as well as the higher peak discharges and overall higher discharge rates. From figure 8 the drier year compared to other years can clearly be seen.

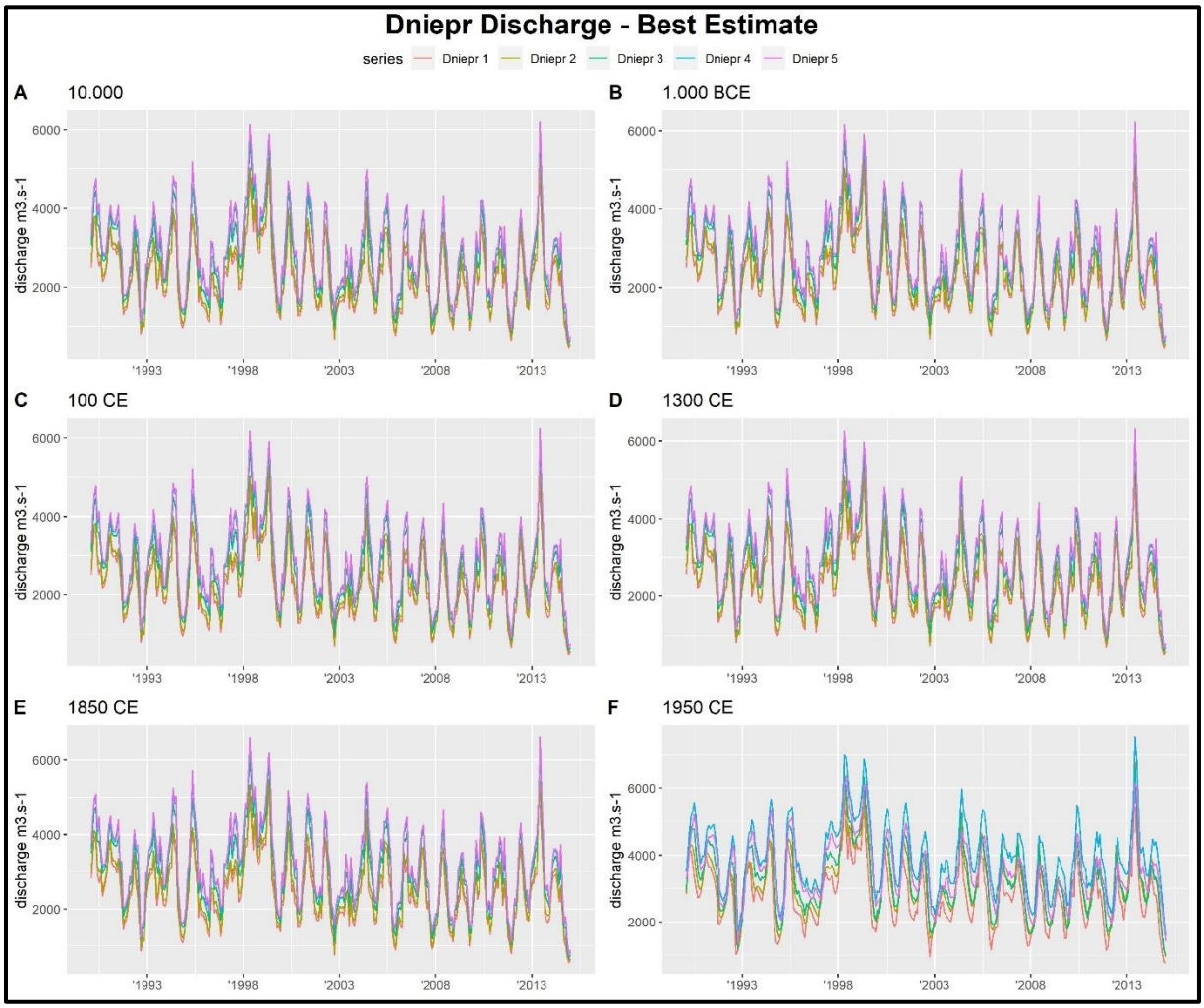

Figure 8, monthly discharge of Dniepr for the best estimate scenario

In the water balance analysis the annual river discharge is used, which shows different plots. The figures of these discharge plots can be found in appendix section 3.1.

4.2.1.1 Dniepr water balance descriptive statistics

The descriptive statistical analysis on the hydrological variables, table 2, show that the maximum land cover fraction for the non-natural vegetation is not reached, and natural vegetation is present. From the results it can be seen that evaporation is higher than the precipitation value for each, except in the 1st quartile, this does not mean that this is representative for a single observation. When in the water balance the evaporation is higher than the precipitation, the value of run off or groundwater recharge becomes negative. In the results a large range of negative values, especially for runoff can be seen.

Table 2, descriptive statistics for the Dniepr river

Descriptive statistics for river Dniepr on the variables						
variable	min	1st quartile	median	mean	3rd quartile	max
Total dataset						
Total LUC	0.000000	0.000000	0.003544	0.135256	0.199059	0.812806
Evaporation	0.3531	0.4552	0.5440	0.5861	0.6434	1.3204
Precipitation	0.3514	0.4822	0.5518	0.5577	0.6360	0.8260
Groundwater recharge	-0.036536	0.005392	0.044544	0.044253	0.073086	0.172802
Discharge	1639	2322	2679	2785	3151	5701
Runoff	-0.877850	-0.086632	0.043802	-0.005147	0.111293	0.337538
Lower estimate scenario						
Total LUC	0.000000	0.000000	0.001365	0.128191	0.184345	0.770551
Evaporation	0.3550	0.4552	0.5451	0.5864	0.6434	1.3204
Precipitation	0.3514	0.4822	0.5518	0.5577	0.6360	0.8260
Groundwater recharge	-0.036536	0.005417	0.043921	0.044025	0.072933	0.170305
Discharge	1639	2312	2661	2771	3137	5649
Runoff	-0.877850	-0.086649	0.042115	-0.005462	0.111293	0.337538
Best estimate scenario						
Total LUC	0.0000	0.0000	0.005699	0.131963	0.199059	0.768776
Evaporation	0.3545	0.4552	0.5441	0.5862	0.6434	1.3204
Precipitation	0.3514	0.4822	0.5518	0.5577	0.6360	0.8260
Groundwater recharge	-0.036536	0.005405	0.044293	0.044213	0.072969	0.170060
Discharge	1639	2321	2677	2783	3149	5650
Runoff	-0.877850	-0.086605	0.043955	-0.005185	0.110009	0.337538
Upper estimate scenario						
Total LUC	0.000000	0.000000	0.009167	0.145614	0.228531	0.812806
Evaporation	0.3531	0.4553	0.5437	0.5858	0.6434	1.3204
Precipitation	0.3514	0.4822	0.5518	0.5577	0.6360	0.8260
Groundwater recharge	0.036536	0.005406	0.045126	0.044522	0.073226	0.172802
Discharge	1639	2335	2697	2802	3165	5701
Runoff	-0.877850	-0.086560	0.044848	-0.004794	0.111293	0.337538

4.2.1.2 Dniepr water balance assumptions

To further analyze the data retrieved from the Dniepr data points, preliminary tests on the data and assumptions for a MANOVA are performed. This analysis starts with the identification of outliers, for which the raw data can be found in appendix section 4.1. Looking at the outliers per variable, the discharge data

finds outliers 53, almost all observations for climate year 1998, a year that was exceptionally wet. For the evaporation variable, a total of 225 outliers were identified, of which the extreme outliers were mostly from time slice 6. The runoff also shows 306 outliers, with some extreme outliers as well. At last, the groundwater recharge also has 54 outliers, which cannot be derived from one certain predictor group. Altogether, an analysis Mahalanobis distance outlier analysis considered 35 observations to be outliers. The observations marked as outliers had very high evaporation rates (above >1 m/year) and very low runoff rates (<0.5 m/year). However, due to the dispersed data and with only 25 years of climatic data, the removal of outliers may harm the data set in its credibility, as the outliers are no defaults of the model. The outliers will be considered with the interpretation of the results. Now that the outliers are identified, the univariate normality assumption is tested for the variables. Using the Shapiro-Wilk normality test, all variables were found to be significant, hence the assumption of normality is violated. The normality test results, along with the visualized normality distribution can be found in appendix 4.1. From the visualized normality distribution, the variable evaporation on time slice six shows that the observations are very different from earlier time slices, and shows very skewed data. This also explains the 225 outliers for evaporation. As all variables has large amount of observations, the skewedness of distribution is accepted but considered. For the assumption of truly independent variables, a Pearson correlation test is performed to check multicollinearity. From this test, see appendix 4.1, a strong negative linear correlation between evaporation and runoff was found of -0.84. As the multicollinearity was expected, the ANCOVA considers the correlation of dependent variables, and precipitation can be tested as a confounding variable. Next the linearity of the variables are evaluated, as presented in figure 9. The linearity of the lower and upper estimate can be found in the appendix section 4.1. The linearity analysis shows how the evaporation is differently distributed for time slice 6.

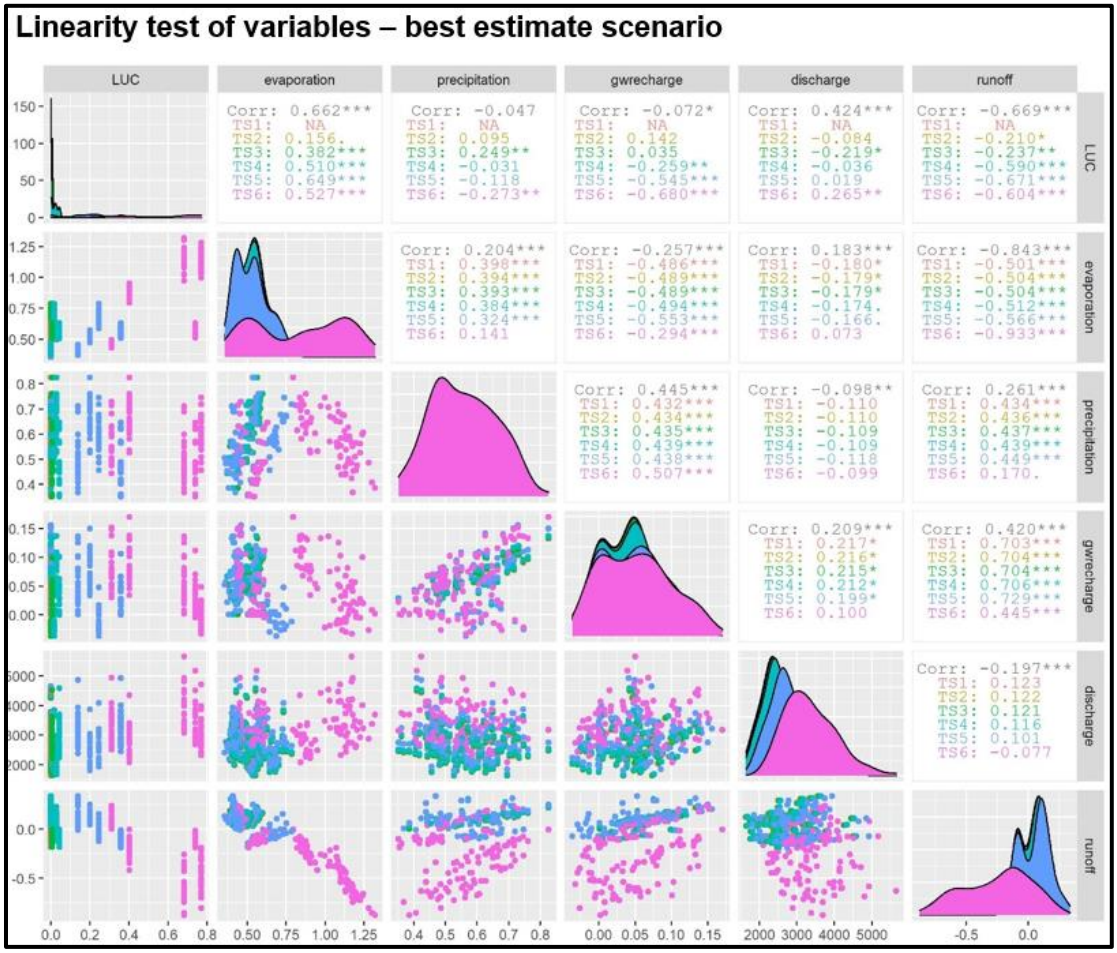

Figure 9, linearity between the variables total LUC, evaporation, precipitation, groundwater recharge and runoff for the best estimate scenario. The colors differentiate between the time slices.

The linearity figure also shows in the scatterplots that most observations are very similar, except for time slice 6 (pink), that is very different from the other time slices in both the scatterplots, histograms and correlation coefficients. Before the differences between the groups can be tested, the variance within the groups must be tested, to see if the assumption of equal variance within groups is violated. The Levene's homogeneity test, appendix section 4.1, shows very significant values for variables except discharge, with the lowest p-value in the best estimate for evaporation with F(5, 744)=1.65, p=4.53e-87 (<α=0,01), and the highest for discharge F(5, 744)=110.17 p=0.145 (<α=0,01). This means that the assumption of equal variance is violated for most variables, which was to be expected with the large amount of outliers found earlier. This means a Welch-ANOVA needs to be executed, with which the unequal variances are assumed.

4.2.1.3 Dniepr water balance statistical analysis

When there is a distortion of equal variances, a Pillai's trace is executed. In this case, the Pillai trace value is 0.557 which is between 0=no statistical difference, and 1=a statistical difference. The MANOVA Pillai's trace analysis for the best estimate scenario of Dniepr stated F(5, 744)=18.66, p= <0,000 (<α=0,01), meaning that the trace is found to be significant. The MANCOVA for Pillai's trace shows F(20, 2972)=24.05, with Pillai=0.557 for pr(>F)=<2.2e-16 for the timeslices, and shows for precipitation F(4, 740)=599.21, Pillai=0.764 for pr(>F)=<2.2e-16. This shows that precipitation is a better predictor for the variables, as expected, but still shows also a large Pillai value with significant difference for the time slices against the variables. Looking more closely into the different variables separately explain the significant difference for the Pillai's trace value of 0.557. The individual responses of the variables in the time slices were only found significantly different for evaporation F(5,744)=83.526, p=<0.000 (<α=0,01), discharge F(5, 744)=31.973, p=<0.000 (<α=0,01), and runoff F(5,744)=80.221, p=<0.000 (<α=0,01). The results can be seen in table 3.

Table 3, summary of Welch ANOVA statistics for all variables

ANOVA – Welch test for Dniepr						
		N	Statistic	DFn	DFd	P value
lower	Discharge	750	26.09	5	347	2.28e-22
	Evaporation	750	25.87	5	345	3.54e-22
	Groundwater recharge	750	1.370	5	347	2.34e-01
	Runoff	750	28.88	5	345	1.80e-24
		N	Statistic	DFn	DFd	P value
best	Discharge	750	26.5	5	347	1.14e-22
	Evaporation	750	25.8	5	345	3.77e-22
	Groundwater recharge	750	1.42	5	347	2.17e- 1
	Runoff	750	28.8	5	345	1.92e-24
		N	Statistic	DFn	DFd	P value
upper	Discharge	750	27.6	5	347	1.46e-23
	Evaporation	750	25.8	5	345	4.01e-22
	Groundwater recharge	750	1.44	5	347	2.08e- 1
	Runoff	750	28.8	5	345	2.23e-24

The results for the ANOVA without Welch assumption can be found in appendix section 3.1. Between the estimation scenarios there is no difference in significant findings or the strength of the significant difference.

With a closer inspection with the Games-Howell post hoc test, the differences between the time slices become visible. For discharge the time slices 1 to 5 were significantly different from time slice 6, and time slices 1 to 4 were significantly different from time slice 5. For evaporation and runoff the time slices 1 to 5 were significantly different from time slice 6. Between the upper and the lower estimate is no difference in results, as can be seen in figure 10.

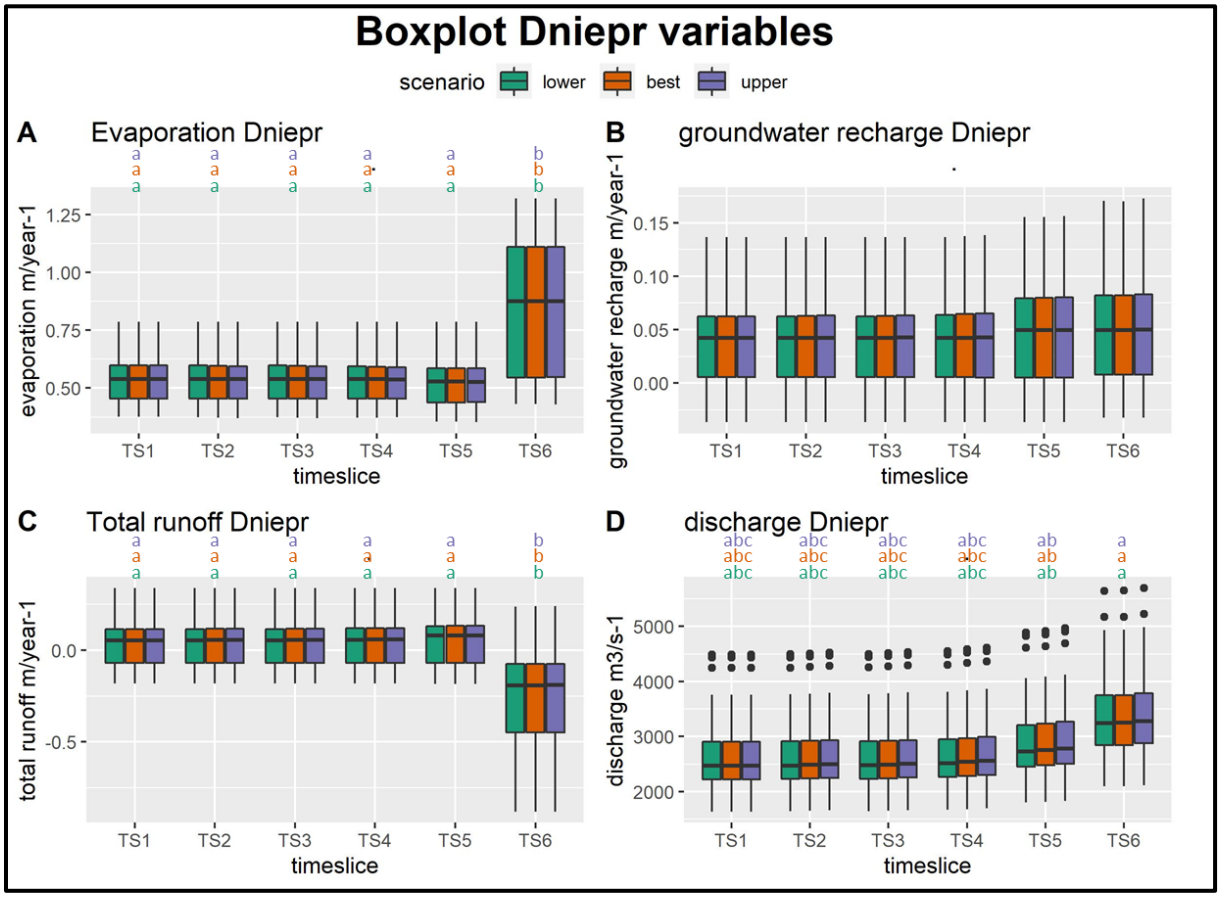

Figure 10, boxplot of hydrological variables

When we compare the found results between the time slices with the land use change that has taken place over the six time slices, figure 11, we clearly see that only in 1850 and 1950 the total non-natural vegetation fraction increases at the chosen data points. Most of the non-natural vegetation consists of rainfed crops. And between 1850 and 1950 the land area of pasture increases. The total non-natural vegetation fraction does not reach values above 0.82, meaning there still is a lot of natural vegetation present.

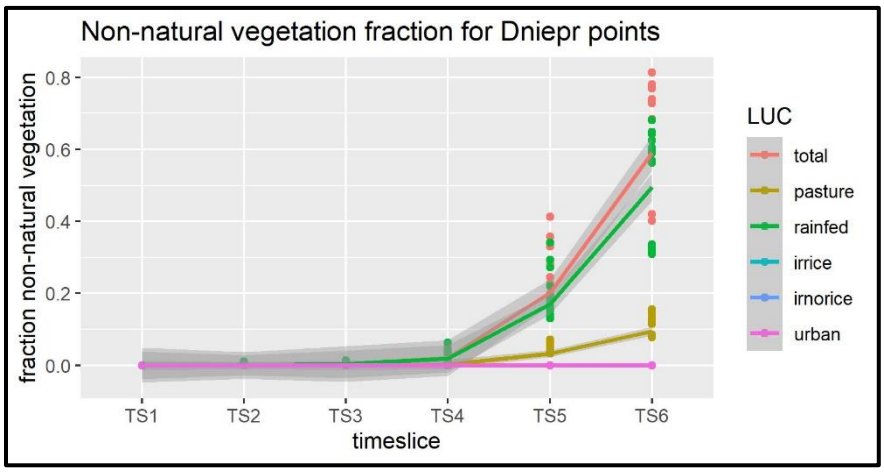

Figure 11, non-natural vegetation fraction for the Dniepr data points, for the total non-natural fraction, pasture, rainfed crops, irrigated rice paddies, irrigated non-rice paddies and urban area. The grey area represents the standard error of each vegetation fraction.

The pearson's correlation, table 4, between the variables and LUC indicate only a moderate correlation of 0.660 with evaporation (positive) and with runoff of -0.670 (negative), both significant. Also discharge was found to be significantly correlation with a moderate strength of 0.420.

Table 4, Pearson's correlation test of LUC and the variables

Pearson's correlation between LUC and the hydrology variables.						
		cor	statistic	P value	Lower conf	Higher conf
LUC	Evaporation	0.660	24.167931	8.07e-96	0.6199760	0.7005545421
LUC	Precipitation	-0.047	-1.282076	2.00e-01	-0.1180189	0.0248460438
LUC	GW recharge	-0.072	-1.971883	4.90e-02	-0.1427662	-0.0003255246
LUC	Discharge	0.420	12.809160	4.19e-34	0.3635874	0.4811166669
LUC	runoff	-0.670	-24.603735	2.14e-98	-0.7065606	-0.6272439811

4.2.2 Danube

The river Danube is large and has very far spread datapoints, which is visible from figure 12. The average annual discharge [m^3/s] pattern over a 25 year time period, as seen in figure 12 for the time slices (figure 12A to F) the patterns remain similar over the equal time period. The data point series Danube 1 is the closest to the source, and does not exceed the 5000 m^3/year values, whereas all other data points are around or above the 5000 m^3/year discharge rate. The Danube discharge rate is generally between 4000 and 10.000 m^3/s. For climatic years 2010 and 2013 there are great discharge peaks, opposed to 2012 and 2014.

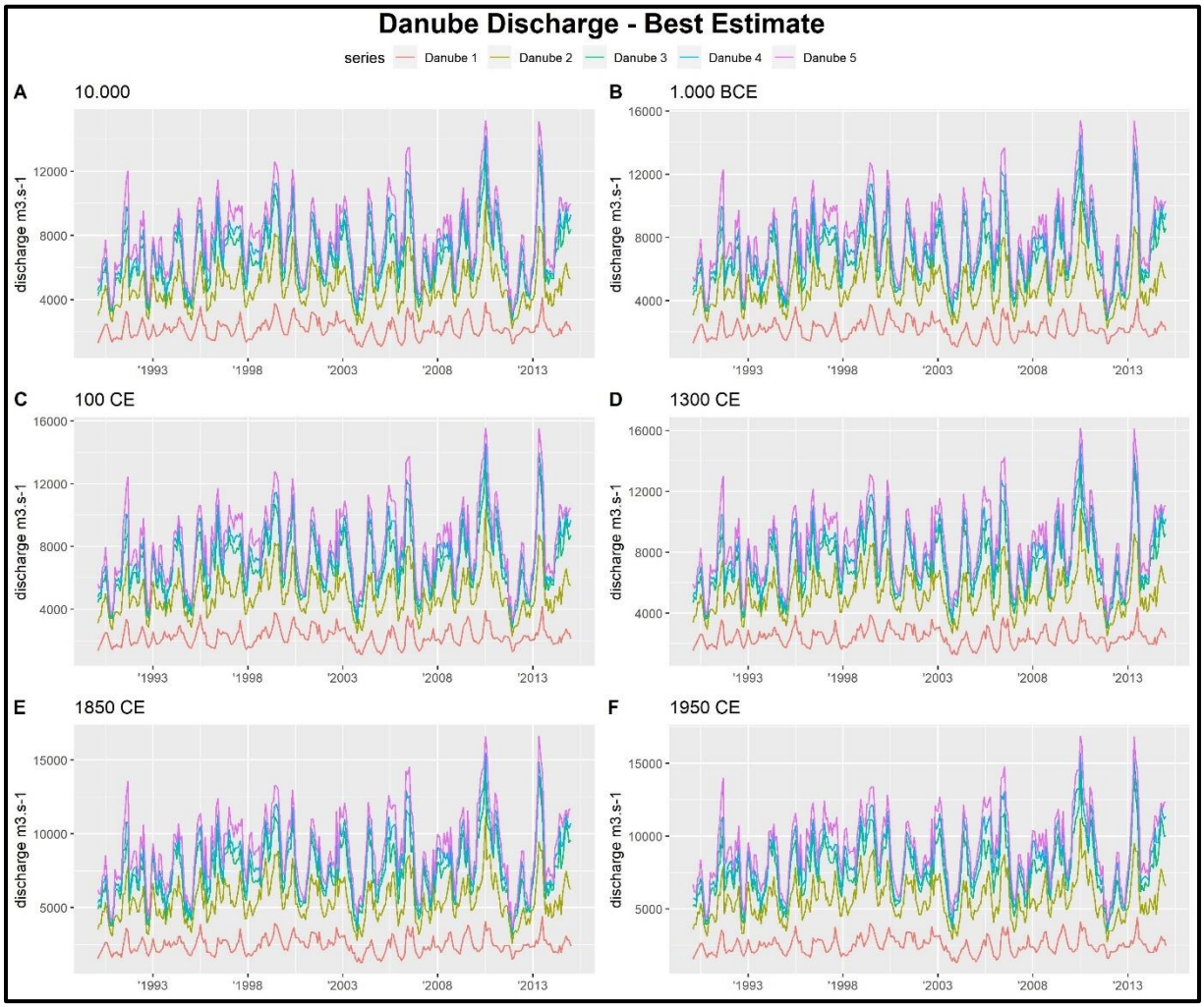

Figure 12, discharge of Danube over 25 years - Best Estimate

In the water balance analysis the annual river discharge is used, which shows different plots. The figures of these discharge plots can be found in appendix 3.2.

4.2.2.1 Danube water balance descriptive statistics

The descriptive statistical analysis on the hydrological variables, table 5, show that the maximum land cover fraction is 1, meaning at certain data points, for some time slices, there is no natural vegetation fraction present. The total LUC does not seem to be gradual, as the mean and median show a value below 0.5. Similarly to the river Dniepr, the runoff and groundwater recharge are negative for a large part. This can again be explained by the evaporation values that are mostly higher than the precipitation levels. There are only small difference between the estimation scenarios, and minimum and maximum values remain the same.

Table 5, descriptive statistics for river Danube

Descriptive statistics for river Danube on the variables						
variable	min	1st quartile	median	mean	3rd quartile	max
Total of all estimation scenarios						
Total LUC	0.00000	0.01249	0.17726	0.33504	0.65269	1.00000
Evaporation	0.4179	0.6302	0.6878	0.6627	0.7243	0.8366
Precipitation	0.2795	0.5547	0.6225	0.6326	0.7243	1.0335
Groundwater recharge	-0.01741	0.01272	0.04482	0.05760	0.09030	0.24190
Discharge	1707	4664	6441	6145	7917	13597
Runoff	-0.32271	-0.15740	-0.05601	-0.03393	0.07358	0.43224
Lower estimate scenario						
Total LUC	0.000000	0.008729	0.046396	0.273219	0.571695	1.000000
Evaporation	0.4271	0.6345	0.6917	0.6671	0.7289	0.8366
Precipitation	0.2795	0.5547	0.6225	0.6326	0.7243	1.0335
Groundwater recharge	-0.01726	0.01198	0.04344	0.05601	0.08546	0.24066
Discharge	1707	4543	6380	6069	7830	13494
Runoff	-0.32271	-0.15848	-0.06052	-0.03827	0.06435	0.41140
Best estimate scenario						
Total LUC	0.0000	0.0382	0.1544	0.3346	0.6385	1.0000
Evaporation	0.4200	0.6317	0.6884	0.6624	0.7245	0.8366
Precipitation	0.2795	0.5547	0.6225	0.6326	0.7243	1.0335
Groundwater recharge	-0.01724	0.01244	0.04490	0.05770	0.09039	0.24142
Discharge	1707	4625	6434	6145	7911	13545
Runoff	-0.32271	-0.15750	-0.05622	-0.03362	0.07191	0.42652
Upper estimate scenario						
Total LUC	0.00000	0.06921	0.29534	0.39726	0.70962	1.00000
Evaporation	0.4179	0.6275	0.6858	0.6587	0.7220	0.8366
Precipitation	0.2795	0.5547	0.6225	0.6326	0.7243	1.0335
Groundwater recharge	-0.01741	0.01285	0.04634	0.05910	0.09260	0.24190
Discharge	1707	4706	6510	6220	8000	13597
Runoff	-0.32271	-0.15662	-0.05284	-0.02992	0.07768	0.43224

4.2.2.2 Danube water balance assumptions

First looking at the outliers for the variables to examine our observations, the evaporation shows 14 outliers, all for series point 2. For the variable runoff the climatic year 2005 and 2010 are found to be outliers, a total of 12, similarly to the variable discharge, 38 outliers, that also shows climatic year 2010 to be a distortion. At last, ground water recharge has 115 outliers, not particularly assigned to one type of data. If all outliers together are evaluated, with Mahalanobis distance calculation, there are no outliers identified. For the exact results on the outliers, consult appendix section 3.2. Although no outliers were found with the Mahalabonis test, the individual test do show outliers, therefor we can expect still some violation of the assumption of normal distribution or equal variance. First, the Shapiro-Wilk normality test is performed, appendix section 3.2, that shows all variables were found to be significant. This means that all response variables can be considered skewed, which will be taken into consideration. The correlation matrix shows strong positive correlation between evaporation and discharge with a value of 0.79. Discharge also shows a moderate negative correlation with runoff of -0.59. Runoff has a moderate negative correlation with evaporation of -0.60, and a strong positive correlation with precipitation of 0.71 and groundwater recharge of 0.89. Another strong positive correlation exists between groundwater recharge and precipitation with a value of 0.73. As so many correlation have been identified, the correlation matrix is included, table 6. These results show that multicollinearity is a huge violation for the dataset, and ANCOVA is necessary for the evaluation of statistical difference between the time slices as a response to LUC, not precipitation.

Table 6, correlation matrix of total dataset. The correlation strength is classified by 1 to 0.91 = very strong, 0.90 to 0.71 = strong, 0.70 to 0.51 = moderate, 0.50 to 0.31= weak, 0.30 to 0.01 = very weak. The sign indicates a positive or negative correlation.

Multicollinearity in a correlation matrix between the variables of total dataset					
	evaporation	precipitation	Groundwater recharge	discharge	runoff
evaporation	1.000	0.016	-0.450	0.790	-0.600
Precipitation	0.016	1.000	0.730	-0.100	0.710
Groundwater recharge	-0.450	0.730	1.000	-0.400	0.890
Discharge	0.790	-0.100	-0.400	1.000	-0.590
runoff	-0.600	0.710	0.890	-0.590	1.000

There is only little differences in the values of correlation between the estimation scenarios. The correlation matrix per estimation scenario can be found in appendix section 3.2. The test of linearity for the variables, as seen in figure 13, show a similar result, with high correlation coefficients. In figure 13 more aligned patterns within the data can be seen visually, opposed to the figure of Dniepr river. This likely is due to the more gradual increase of LUC, along with the strong correlations between the variables.

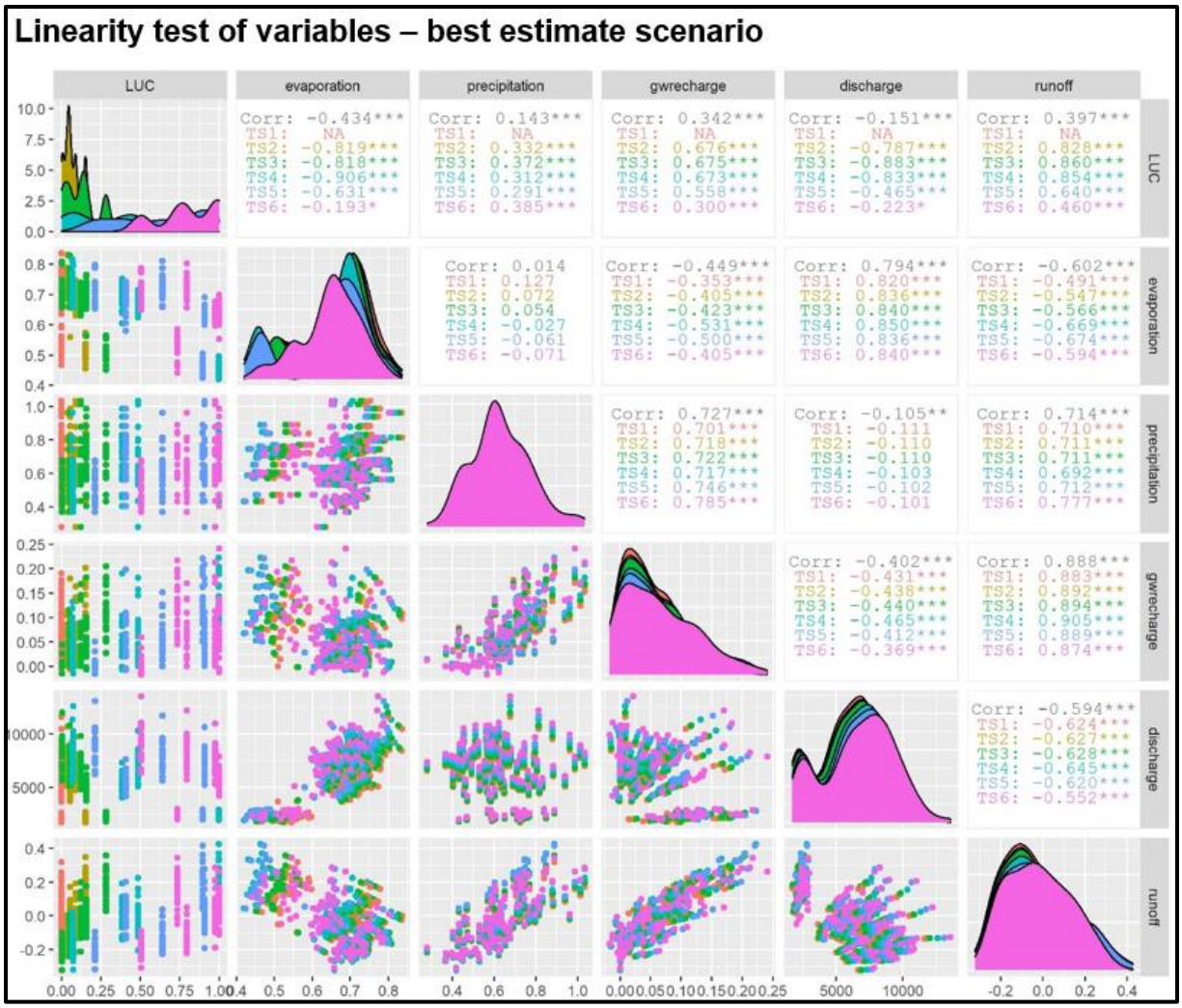

Figure 13, linearity for river Danube of the best estimate scenario, for the variables total LUC, evaporation, precipitation, groundwater recharge, discharge and runoff

In Levene's test the significant values indicate the violation of the equal variance assumption within the groups. This is true for groundwater recharge in each estimation scenario, with values F(5,774)=2.708, p=0.019<α=0.05 in lower estimate scenario, values F(5,774)=2.356, p=0.039<α=0.05 in best estimate scenario, and values F(5,774)=2.261, p=0.047<α=0.05 in upper estimate scenario. The values are close to the α=0.05, and as all other values do not violate the assumption, the ANOVA test as well as the Welch ANOVA test is run for comparison. The assumption of equal variance is violated for discharge, groundwater recharge, precipitation and runoff in the total data set with p values much smaller than α=0.05, and with this total data set a Welch ANOVA would be required. For all results of Levene's homogeneity test, see appendix section 3.2.

4.2.2.3 Danube water balance statistical analysis

The MANOVA statistical analysis shows a small Pillai trace value of 0.226, with a very small p-value (<2.2e-16), that indicates a significant small difference between the groups. The MANCOVA for Pillai's trace shows F(20, 2972)=8.26, with Pillai=0.211 for pr(>F)=<2.2e-16 for the timeslices, and shows for precipitation F(4, 740)=599.21, Pillai=0.825 for pr(>F)=<2.2e-16. This shows that precipitation is a better predictor for the variables, as expected, with a small Pillai value with significant difference for the time slices against the variables. Further analysis , ANOVA-Welch in table 7, on the specific variables shows for the best estimate scenario that discharge is statistically significant with F(5, 744)=2.92, p=0.013 (<α=0.05), and statistical

significance for evaporation with F(5, 744)=3.85, p=0.002 (<α=0.01). Groundwater recharge and runoff were not found significantly different for the time slices.

Table 7, ANOVA - Welch test results for river Danube

ANOVA – Welch test for Danube						
		N	Statistic	DFn	DFd	P value
lower	Discharge	750	2.92	5	347	0.013
	Evaporation	750	3.85	5	347	0.002
	Groundwater recharge	750	1.65	5	347	0.145
	Runoff	750	1.28	5	347	0.271
		N	Statistic	DFn	DFd	P value
best	Discharge	750	2.87	5	347	0.015
	Evaporation	750	3.52	5	347	0.004
	Groundwater recharge	750	1.55	5	347	0.174
	Runoff	750	1.22	5	347	0.300
		N	Statistic	DFn	DFd	P value
upper	Discharge	750	3.02	5	347	0.011
	Evaporation	750	3.57	5	346	0.004
	Groundwater recharge	750	1.61	5	347	0.158
	Runoff	750	1.28	5	347	0.273

Further distinction between the time slices are necessary to see on which time slices the variables vary. The Games-Howell test shows that for discharge time slice 1 significantly differs from time slice 6. For evaporation time slice 1 significantly differs from time slice 5 and 6. Figure 14 displays the variables and the groups the statistically different time slices.

Now these results for the Danube river do not show much difference between time slices for each variable, thus, a closer look into the vegetation might explain this. From figure 15 it can be concluded that the total vegetation is very dispersed. The standard error already shows a wide range (grey area) compared to the other land use lines, and other vegetation figures from other rivers. This can be explained by the length of the river of the Danube, and its position. The first data point was gathered in Germany, and the last point in Romania. The LUC between these data points, and thus also their hydrological responses, are inequal. This means that within the series there is an distortion in observations, which dominate the results for the analysis on difference between the time slice.

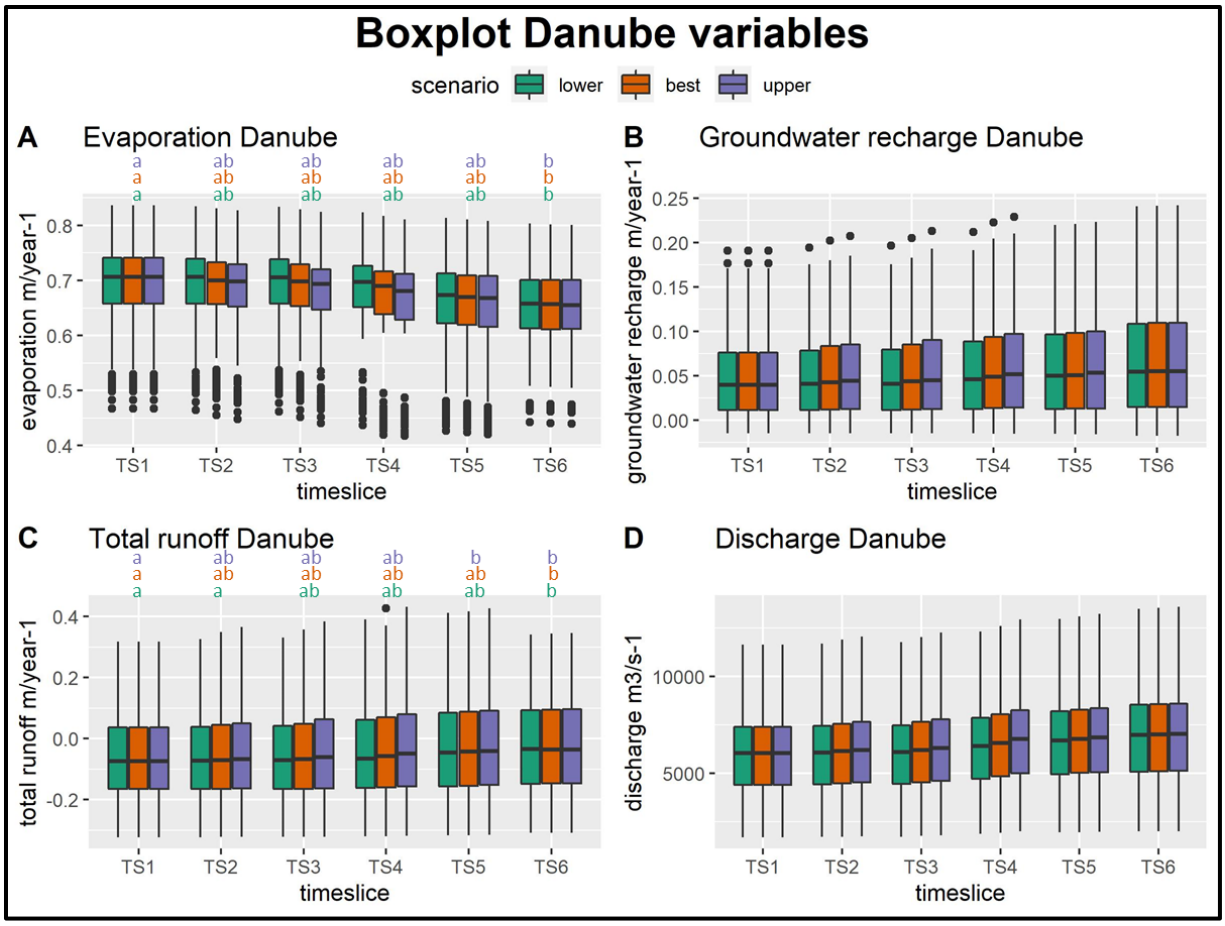

Figure 14, boxplot of hydrological variables with post-hoc Tukey test significant group indications

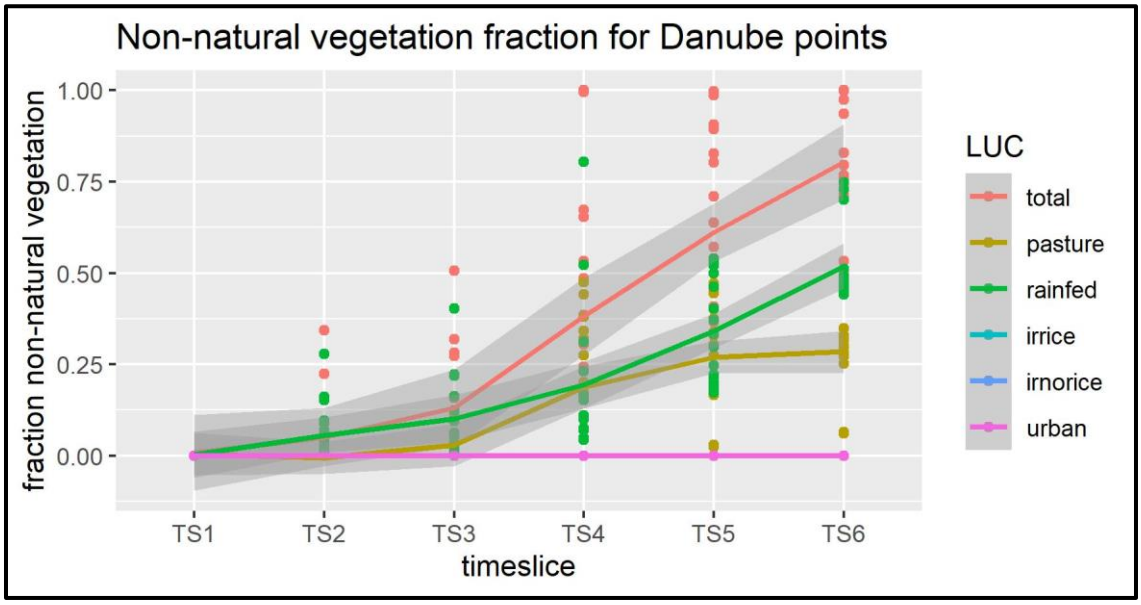

Figure 15, non-natural vegetation fraction for the Danube data points, for the total non-natural fraction, pasture, rainfed crops, irrigated rice paddies, irrigated non-rice paddies and urban area. The grey area represents the standard error of each vegetation fraction.

Lastly a correlation between land use change and the variables were performed, table 8. This correlation test, alike the Dniepr river, show significant moderate correlations between LUC and evaporation, discharge and runoff. This time all three correlations are negative.

Table 8, Pearson's correlation matrix of LUC and the hydrology variables

Pearson's correlation between LUC and the hydrology variables.						
		cor	statistic	P value	Lower conf	Higher conf
LUC	Evaporation	-0.430	-13.166824	9.26e-36	-0.49014361	-0.37379452
LUC	Precipitation	0.140	3.960032	8.21e-05	-0.07245337	0.21270537
LUC	GW recharge	0.340	9.966008	4.73e-55	0.27758611	0.40405644
LUC	Discharge	-0.150	-4.171055	3.39e-05	-0.21998001	-0.08004085
LUC	runoff	-0.400	11.847276	8.51e-30	0.33544511	0.45609853

4.2.3 Po

The river Po is shorter compared to the other rivers, and thus the five datapoint were less spread than the others. The average annual discharge [m^3/s] pattern over a 25 year time period, as seen in figure 16 for the time slices (figure 16A to F) the patterns remain similar over the equal time period. In the discharge, the peaks are less frequent, compared to the other rivers.

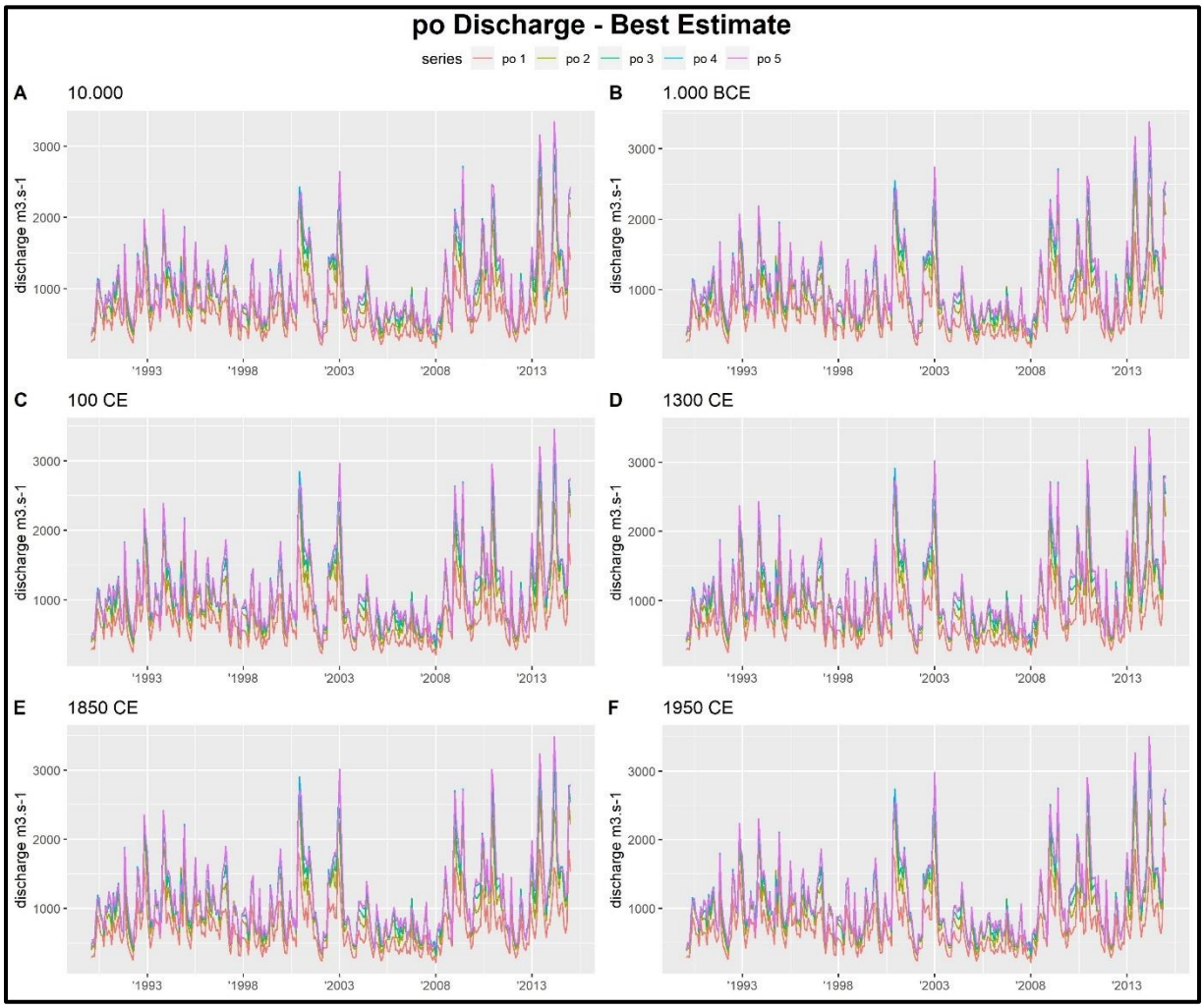

Figure 16, discharge Po of best estimate scenario for 25 years, for time slice 1 is 10.000 BCE in figure A, to time slice 6 is 1950 CE in figure F

In the water balance analysis the annual river discharge is used, which shows different plots. The figures of these discharge plots can be found in appendix 3.3.

4.2.3.1 Po water balance descriptive statistics

The descriptive analysis, table 9, shows that the LUC in the Po region reached a total fraction of non-natural vegetation of 1, making no natural vegetation present for some time slices very early on compared to the other rivers. What is remarkable, compared to the other rivers, is that evaporation rates now are lower than precipitation for the minimal value, up to the maximum value.

Table 9, descriptive statistcis for the Po river

Descriptive statistics for river Po on the variables						
variable	min	1st quartile	median	mean	3rd quartile	max
Total dataset for all estimation scenarios						
Total LUC	0.0000	0.1618	0.8593	0.6192	1.0000	1.0000
Evaporation	0.4696	0.6637	0.7156	0.7151	0.7717	0.9399
Precipitation	0.5274	0.7056	0.8436	0.8821	0.9564	1.4926
Groundwater recharge	0.001856	0.116784	0.173979	0.208507	0.249827	0.739301
Discharge	343.5	764.0	950.7	991.9	1173.4	2109.7
Runoff	-0.09534	0.05424	0.11749	0.16110	0.21795	0.70012
Lower estimate scenario						
Total LUC	0.00000	0.03993	0.40449	0.46350	0.84528	1.00000
Evaporation	0.4697	0.6791	0.7290	0.7288	0.7828	0.9399
Precipitation	0.5274	0.7056	0.8436	0.8821	0.9564	1.4926
Groundwater recharge	0.001856	0.106415	0.162715	0.194774	0.234053	0.691938
Discharge	343.5	752.0	935.9	973.9	1152.2	2043.1
Runoff	-0.09534	0.04482	0.10329	0.14684	0.19611	0.69924
Best estimate scenario						
Total LUC	0.0000	0.1940	0.9068	0.6655	1.0000	1.0000
Evaporation	0.4696	0.6578	0.7128	0.7113	0.7661	0.9399
Precipitation	0.5274	0. 7056	0.8436	0.8821	0.9564	1.4926
Groundwater recharge	0.001856	0.123109	0.178153	0.212247	0.252227	0.735146
Discharge	343.5	760.5	950.7	992.4	1173.0	2059.4
Runoff	-0.09534	0.06021	0.12220	0.16499	0.22200	0.70012
Upper estimate scenario						
Total LUC	0.0000	0.3438	1.0000	0.7286	1.0000	1.0000
Evaporation	0.4696	0.6470	0.7044	0.7051	0.7619	0.9399
Precipitation	0.5274	0.7056	0.8436	0.8821	0.9564	1.4926
Groundwater recharge	0.001856	0.127190	0.183796	0.218499	0.259510	0.739301
Discharge	343.5	774.1	968.9	1009.5	1201.0	2109.7
Runoff	-0.09534	0.06249	0.12709	0.17147	0.22931	0.70012

4.2.3.2 Po water balance assumptions

The outlier analysis on the individual variables shows that the observations know outliers. For evaporation the amount of outliers detected is 14, mostly data point series Po 1, but none of these were considered extreme outliers. In the runoff variable 218 outliers were identified, with just 1 extreme outlier, and with the observations mostly being from the climatic years 2010, 2013 and 2014. The outliers for the variable groundwater recharge amounted to a total of 263, which is a lot, considered the total observations are 2250. There were 30 extreme outliers for groundwater recharge, and all outliers had the climatic years 2010, 2013 and 2014 in common. At last, the discharge variable has 53 outliers, none extreme, and all of the outliers belong to the climatic year 2014. However, the mahalonobis distance calculated no outliers for the total data frame, and the outliers are not removed to preserve the large data set. For the tables that contain all outliers for each variable, consult the appendix section 3.3. Next the normality of the observations for the variables are calculated with the Shapiro-Wilk test to test the assumption of normal distribution. From the test it can be concluded that only evaporation is normally distributed, with very large p-values. Taking a closer look on this variable, the plot of the evaporation variable is very well distributed with a very large p-value. The other variables show very small p-values, meaning the data is not normally distributed. From the figures in appendix section 3.3 it becomes clear that runoff, discharge and groundwater recharge show a similar peak to the precipitation. As precipitation is a predictor, and not dependent on LUC, and equal on all time slices, the skewed variables are accepted, and the skewed distribution is considered when the results are interpreted. The multicollinearity analysis shows that the there is a strong positive correlation between precipitation and groundwater recharge (0.800) and precipitation and runoff (0.810). There is a very strong positive correlation between groundwater recharge and runoff (0.960). Other correlations vary from moderate to very weak, however, the table 10 with the values is presented below as a very strong correlation was found. The strong correlation between precipitation and groundwater recharge and runoff is expected, as precipitation is a known confounding variable in this analysis. But it is interesting that in the previous rivers this correlation was not as strongly found. Even more interesting is very strong linear relationship indicating the confounding variable precipitation. As both variables have strong correlations with precipitation and in theory an increase in groundwater recharge would not necessarily lead to an increase in runoff solely so strongly. A few changes within the scenarios, but the power of the correlations remain similar. For the display of all the correlation matrices, see appendix section 3.3.

Table 10, multicollinearity test in a correlation matrix plot for the total data set. The correlation strength is classified by 1 to 0.91 = very strong, 0.90 to 0.71 = strong, 0.70 to 0.51 = moderate, 0.50 to 0.31= weak, 0.30 to 0.01 = very weak. The sign indicates a positive or negative correlation.

Multicollinearity in a correlation matrix between the variables of total dataset					
	evaporation	precipitation	Groundwater recharge	discharge	runoff
evaporation	1.000	0.550	0.240	0.460	0.150
Precipitation	0.550	1.000	0.800	0.640	0.810
Groundwater recharge	0.240	0.800	1.000	0.690	0.960
Discharge	0.460	0.640	0.690	1.000	0.610
runoff	0.150	0.810	0.960	0.610	1.000

The linearity test, figure 17, for the best estimate scenario shows similar results to the correlation matrix above. The data shows clear linearity among all variables.

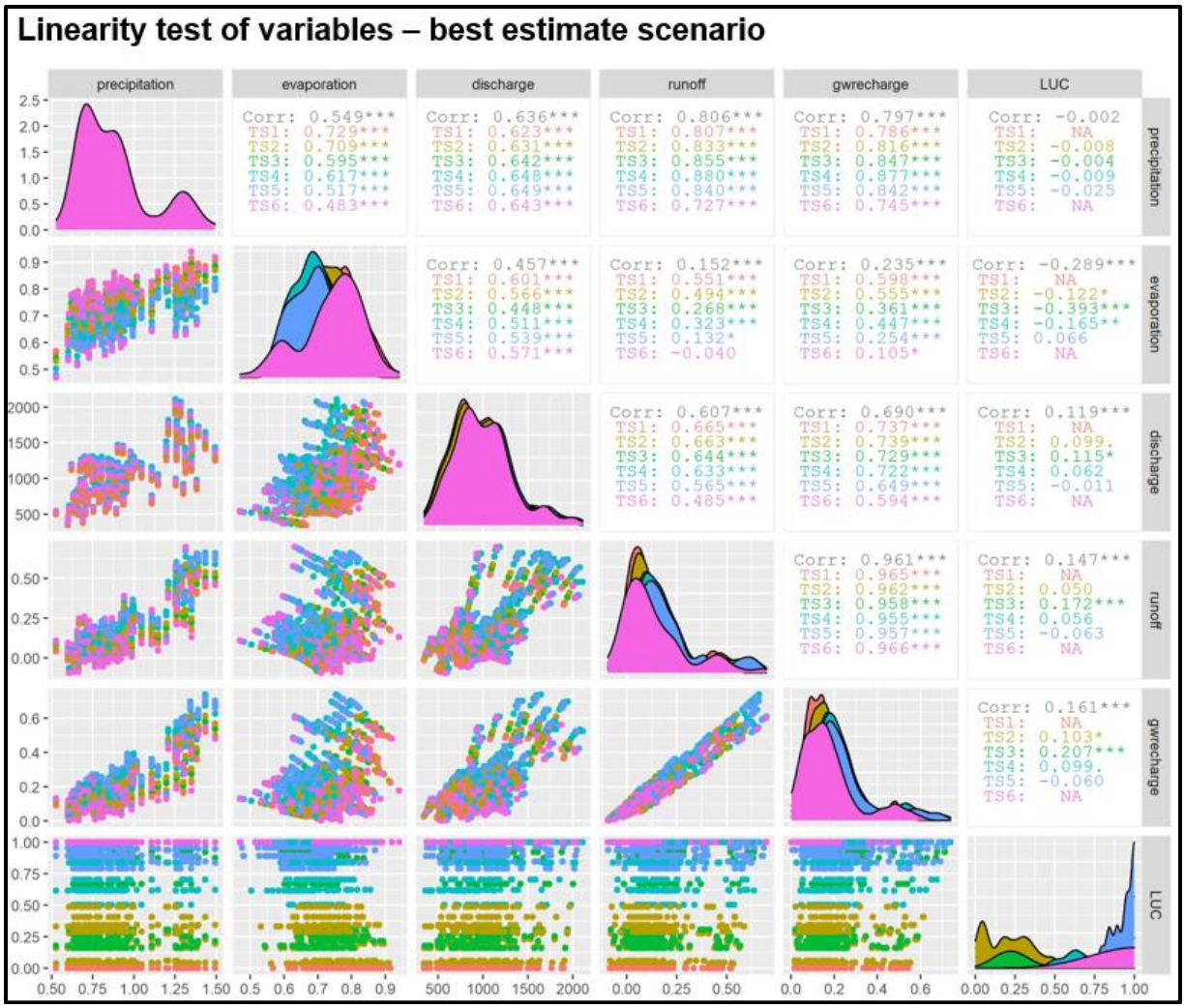

Figure 17, linearity test of the best estimate scenario for the Po river.

At last, the Levene's test of homogeneity is performed, appendix section 3.3. The test shows that there is a statistical significance found for evaporation in all estimation scenarios, meaning that for evaporation the assumption of equal population variance is violated and a Welch test must be conducted. The other variables do not violate the assumption of equal variance. For the data point series, all the variables have violated equal variance assumptions, except evaporation, meaning that the observations within the points in the river have inequal variances but not within time slices. As to these violations of equal variance for either within the data point series or time slices, the Welch ANOVA is additionally conducted for the variables.

4.2.3.3 Po water balance statistical analysis

Now the statistical analysis for the variables is conducted using a MANOVA Pillai's trace, ANOVA Welch, ANCOVA and the post hoc Games-Howell. The Pillai's trace test results shows a small value in the lower estimate of 0.26558, and a higher value in the best estimate of 0.39216 and an even higher value for the upper estimate with 0.42302, all with significant p-values. This indicates that the are differences between the time slices are stronger for the upper estimate compared to the lower estimate. The MANCOVA for Pillai's trace shows F(20, 2972)=16.15, with Pillai=0.392 for pr(>F)=<2.2e-16 for the timeslices, and shows for precipitation F(4, 740)=1027.43, Pillai=0.847 for pr(>F)=<2.2e-16. This shows that precipitation is a better predictor for the variables, as expected, with a small Pillai value with significant difference for the time slices against the variables. The ANOVA-Welch analysis, table 11, differing only a little from the ANOVA test, shows that evaporation F(5, 347)=19.4, p=<0.000 (< α=0.01), groundwater recharge F(5, 347)=5.34, p=<0.000 (< α=0.01), and runoff with F(5, 347)=4.31, p=<0.000 (< α=0.05) are statistically significant, and the reduction of variance in the residuals can be explained by the LUC over the time slices. The variable discharge only becomes significant in the upper estimate scenario with F(5, 347)=2.89, p=0.01655 (< α=0.05).

Table 11, Welch-ANOVA test of river Po for each variable in all estimation scenarios

Welch ANOVA test results						
		n	statistic	DFn	DFd	p-value
lower	Discharge	750	1.47	5	347	0.200
	Evaporation	750	19.4	5	347	4.73e-17
	Groundwater recharge	750	5.34	5	347	9.62e- 5
	runoff	750	4.31	5	347	8.10e- 4
best	Discharge	750	1.810	5	347	0.111
	Evaporation	750	36.83	5	346	3.08e-30
	Groundwater recharge	750	9.990	5	346	6.01e-09
	runoff	750	7.830	5	346	5.32e-07
upper	Discharge	750	2.89	5	347	0.0140
	Evaporation	750	40.9	5	346	5.07e-33
	Groundwater recharge	750	11.3	5	347	4.18e-10
	runoff	750	8.75	5	347	7.91e- 8

With a closer look on the post hoc results, the time slices that significantly differ from each other can be seen. Along with the visualization in boxplot figure 18, the variables show various significant differences. For discharge, the lower and best estimate scenario sees no significant difference between the time slices, but the upper estimate has shown a difference between time slice 1 and time slice 3 and 4. Evaporation has a significantly decreased in time slice 4 and 5, opposed to time slice 1, 2 and 3. Evaporation significantly increases again in time slice 6 opposed to time slice 4 and 5. In the best estimate scenario the time slices 1 and 2 significantly differ from 3, 4 and 5, and then time slice 3, 4 and 5 from time slice 6. In the upper scenario evaporation time slice 1 shows a significant difference with each scenario except time slice 6. Time slice 2 also shows a significant difference for all other time slices except time slice 6. And then time slice 3, 4 and 5 show a significant difference from time slice 6 again. For groundwater recharge values in time slice

1 and 2 were found significantly lower than for time slice 4 and 5. The groundwater recharge significantly decreases again in time slice 6 relative to time slice 4 and 5. In the best and upper scenario time slice 1 and 2 significantly differ from 3 and 4. Additionally time slice 1 differs from time slice 5, and time slice 3, 4 and 5 differ from time slice 6. The runoff in the lower estimate scenario differs in time slice 1 and 2 significantly from 4 and 5. The runoff in the best and upper scenario shows a significant difference of time slice 1 and 2 with time slice 3 and 4, and for time slice 1 also a difference with time slice 5. Additionally time slice 3 and 4 differ from time slice 6.

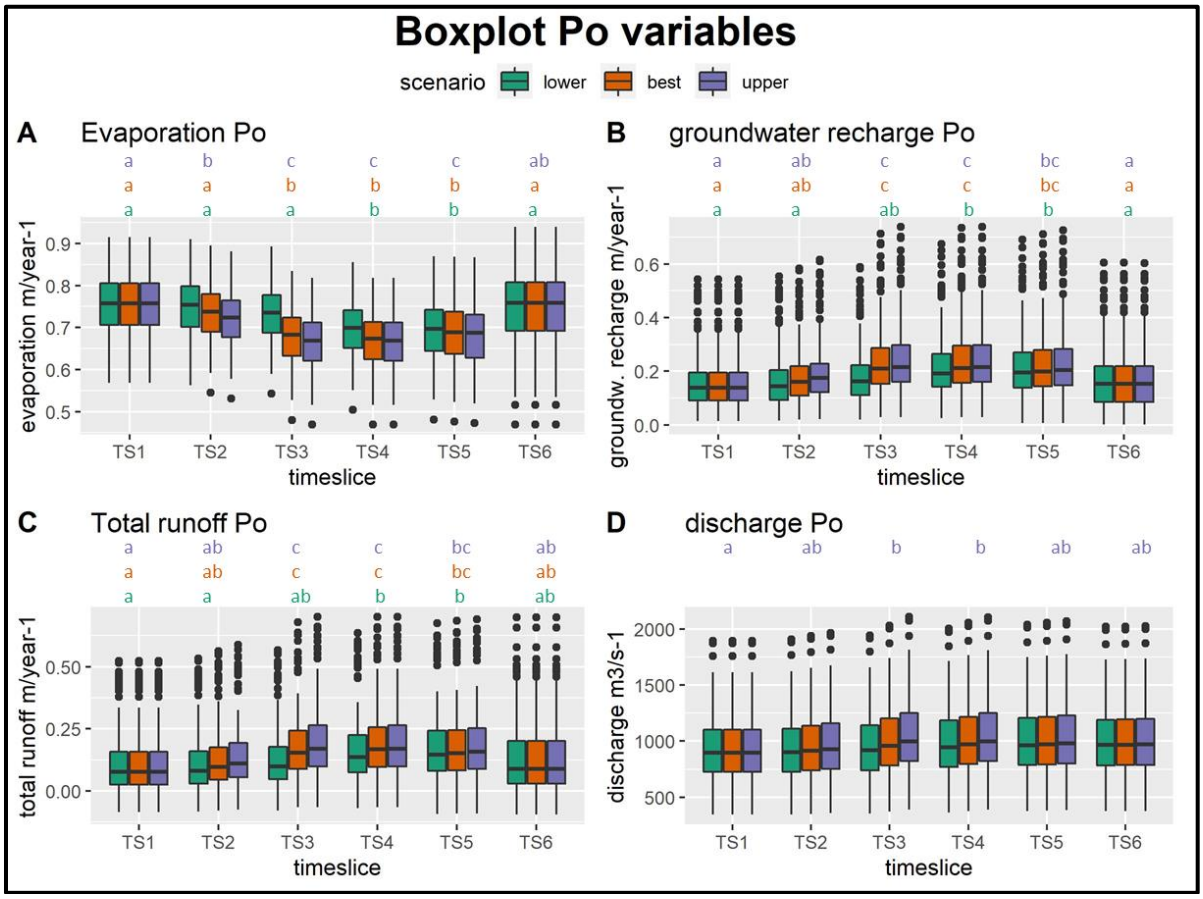

Figure 18, boxplot of hydrological variables with post hoc statistical difference.

To explain the differences found between the time slice, a closer look on the LUC for the time slices must be taken, see figure 19. Remarkably, despite the steady average increase in LUC over the time slices, the type of non-natural vegetation switches dramatically. The rainfed crops are replaced by irrigated crops (non-rice crops). Additionally barely no natural vegetation is present from a rather early time slice.

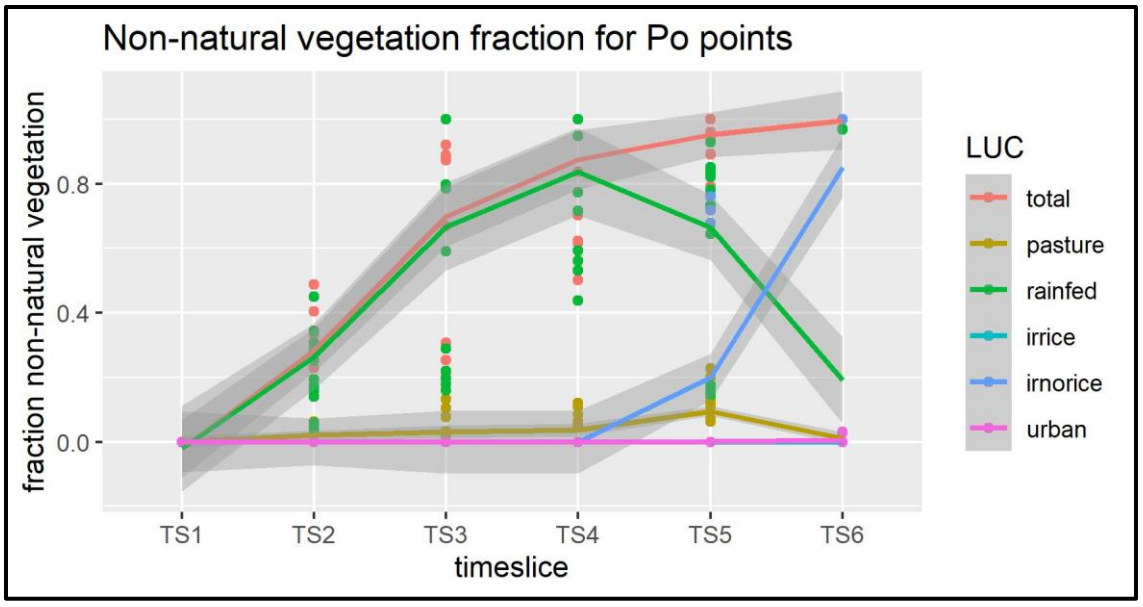

Figure 19, non-natural vegetation fraction for the Po data points, for the total non-natural fraction, pasture, rainfed crops, irrigated rice paddies, irrigated non-rice paddies and urban area. The grey area represents the standard error of each vegetation fraction.

The pearson's correlation test, table 12, between LUC and the variables is performed at last, showing only very weak correlations with the variables.

Table 12, Pearson's correlation matrix of LUC and the hydrology variables

Pearson's correlation between LUC and the hydrology variables.						
		cor	statistic	P value	Lower conf	Higher conf
LUC	Evaporation	-0.290	-8.2199984	8.96e-16	-0.35216570	-0.22079459
LUC	Precipitation	-0.002	-0.0656065	9.48e-01	-0.07397476	0.23204439
LUC	GW recharge	0.160	4.52314521	7.08e-06	0.09265990	0.23204439
LUC	Discharge	0.100	2.73900028	6.29e-03	-0.02829133	-0.17005330
LUC	runoff	0.150	4.01306385	6.60e-05	0.07436180	0.21453657

4.2.4 Rhine

The chosen points in the Rhine all show a similar average annual discharge [m^3/s] pattern over a 25 year time period, as seen in figure 20. Between the time slices (figure 20A to F) the patterns remain similar over the equal time period. The coordinate point closest to the sea (Rhine series 5) has lower values than the coordinate point closest to the starting point of the river (Rhine series 1). From the figure no obvious difference in discharge volume or peaks stand out. The discharge plots for the lower and upper discharge of the Rhine show a similar pattern, and can be consulted in Appendix section 3.4.

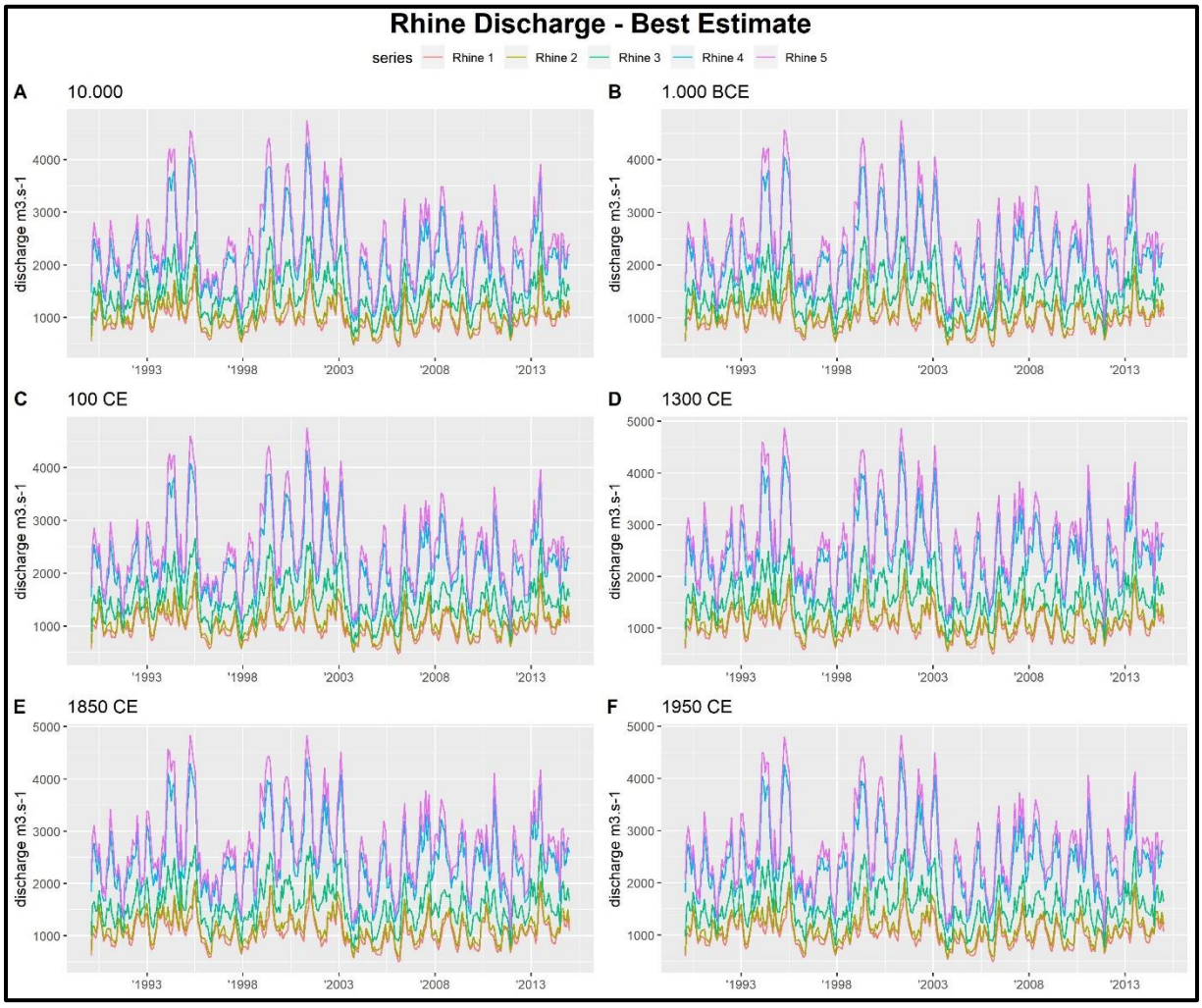

Figure 20, Rhine Discharge for 25 years of the Best Estimate Scenario, for A: 10.000 BCE, B: 1000 BCE, C: 100 CE, D: 1300 CE, E: 1850 CE, F: 1950 CE

In the water balance analysis the annual river discharge is used, which shows different plots. The figures of these discharge plots can be found in appendix 3.4.

4.2.4.1 Rhine water balance descriptive statistics

The descriptive statistics, table 13, for the Rhine river show a gradual increase of LUC. The evaporation levels are lower than the precipitation levels for the measurement. Also standing out, compared to the other rivers, that no negative numbers were found for any variable of the Rhine. The differences between the estimation scenario are very small and but in all scenario's a total fraction of 1 for non-natural vegetation is reached.

Table 13, descriptive statistic Rhine

Descriptive statistics for river Rhine on the variables						
variable	min	1st quartile	median	mean	3rd quartile	max
Total dataset for all estimation scenarios						
Total LUC	0.000000	0.009154	0.200222	0.370567	0.673603	1.000000
Evaporation	0.3918	0.4837	0.5207	0.5472	0.5972	0.7746
Precipitation	0.5736	0.7890	0.8884	0.9271	1.0505	1.4393
Groundwater recharge	0.05936	0.19990	0.26429	0.27863	0.34542	0.55731
Discharge	761.5	1122.6	1467.7	1700.1	2196.6	3481.2
Runoff	0.09344	0.25171	0.35171	0.37807	0.45978	0.95390
Lower estimate scenario						
Total LUC	0.000000	0.002087	0.097917	0.289262	0.559443	1.000000
Evaporation	0.3918	0.4955	0.5269	0.5546	0.6013	0.7746
Precipitation	0.5736	0.7890	0.8884	0.9271	1.0505	1.4393
Groundwater recharge	0.05936	0.19800	0.26017	0.27502	0.34215	0.53356
Discharge	761.5	1110.6	1449.6	1677.0	2144.5	3370.3
Runoff	0.09344	0.24638	0.34312	0.37055	0.45745	0.88815
Best estimate scenario						
Total LUC	0.00000	0.01302	0.20705	0.38538	0.71559	1.00000
Evaporation	0.3927	0.4840	0.5194	0.5464	0.5972	0.7746
Precipitation	0.5736	0.7890	0.8884	0.9271	1.0505	1.4393
Groundwater recharge	0.05936	0.20025	0.26349	0.27888	0.34625	0.54273
Discharge	761.5	1123.7	1466.9	1702.1	2193.0	3409.4
Runoff	0.09344	0.25372	0.35174	0.37889	0.46154	0.93148
Upper estimate scenario						
Total LUC	0.00000	0.02243	0.23317	0.43706	0.93550	1.00000
Evaporation	0.3926	0.4706	0.5160	0.5406	0.5941	0.7746
Precipitation	0.5736	0.7890	0.8884	0.9271	1.0505	1.4393
Groundwater recharge	0.05936	0.20219	0.26696	0.28199	0.35112	0.55731
Discharge	761.5	1136.2	1492.8	1721.2	2231.2	3481.2
Runoff	0.09344	0.25639	0.35551	0.38477	0.46483	0.95390

4.2.4.2 Rhine water balance assumptions

First the data is checked on outliers. The variable evaporation has 223 outliers, all for Rhine 2 data point series. The variables precipitation, groundwater recharge and discharge have no outliers. There were 92 outliers for runoff, that all were from the Rhine 1 data point. The Mahalanobis distance analysis that computes the distance by time slice groups shows no outliers. Following the outlier analysis, the normality of the variables is calculated using a Shapiro-Wilk test. These test and plots, appendix section 3.4, show that none of the variables are normally distributed. In the correlation matrix, the variables discharge and evaporation, precipitation and groundwater recharge, runoff and precipitation, runoff and groundwater recharge all show a strong positive linear relationship. Runoff and evaporation show a moderate negative linear relationship, just as discharge and runoff. These correlation coefficients can be found in table 14, and for the other estimation scenarios the appendix section 3.4 can be consulted.

Table 14, multicollinearity test in a correlation matrix plot for the total data set. The correlation strength is classified by 1 to 0.91 = very strong, 0.90 to 0.71 = strong, 0.70 to 0.51 = moderate, 0.50 to 0.31= weak, 0.30 to 0.01 = very weak. The sign indicates a positive or negative correlation.

Multicollinearity in a correlation matrix between the variables of total dataset					
	evaporation	precipitation	Groundwater recharge	discharge	runoff
evaporation	1.000	0.450	0.400	-0.540	-0.026
Precipitation	0.450	1.000	0.860	-0.390	0.840
Groundwater recharge	0.400	0.860	1.000	-0.130	0.770
Discharge	-0.540	-0.390	-0.130	1.000	-0.150
runoff	-0.026	0.840	0.770	-0.150	1.000

The linearity test for the best estimate scenario is displayed below, figure 21, the lower and upper scenario can be found in appendix section 3.4. The linearity test shows similar results to the correlation test, and the data is clearly correlated among the variables.

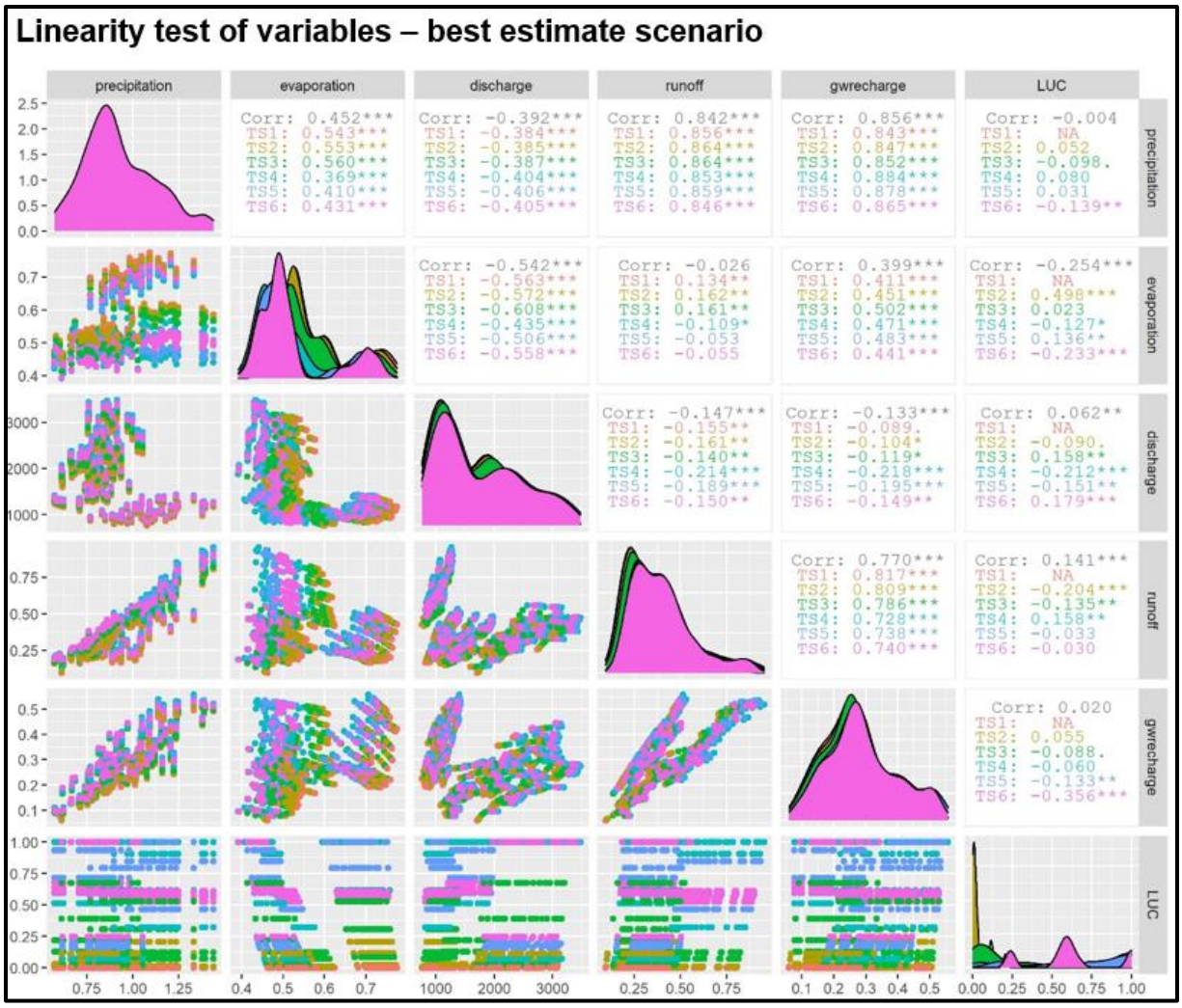

Figure 21, linearity test of the best estimate scenario for the Rhine river.

At last the Levene's homogeneity test, for the assumption on equal variance within groups. All test results can be found in the appendix section 3.4. In the Levene's test of homogeneity, first a test on the Rhine data points is executed. From this analysis it becomes evident that the series in itself carry large inequalities within the variables. The results from this analysis are all significant, which means that the variance of the series are inequal. The result from variables on the basis of the time slices are for all estimation scenarios non-significant, meaning that the assumption of variance within the variables is not violated. This means a regular ANOVA can be performed, however a Welch test is also conducted as the series do show great unequal variances. In case no large difference between the tests is found, the regular ANOVA results are used.

4.2.4.3 Rhine water balance statistical analysis

All tables with the results of the statistical analysis can be found in appendix section 3.4. The Pillai trace test has a value of 0.17862 (lower estimate), 0.22257 (best estimate), and 0.24774 (upper estimate), with all values a probability <0.000. All Pillai's trace value are low, which indicate a weak statistical difference for the groups. In the summary of the MANOVA's it becomes clear which variables were found significant. The MANCOVA for Pillai's trace shows F(20, 2971)=8.76, with Pillai=0.223 for pr(>F)=<2.2e-16 for the timeslices, and shows for precipitation F(4, 740)=2689.69, Pillai=0.936 for pr(>F)=<2.2e-16. This shows that precipitation is a better predictor for the variables, as expected, but still shows a significant difference for the time slices against the variables. The statistical analysis performed is MANOVA, table 15, that showed that in the best estimate scenario evaporation was found significant with F(5, 744)=12.505, p=1.129e-11(< α=0.05),

as well as runoff with F(5, 744)=4.9559, p=0.00018 (< α=0.05). The variables groundwater recharge and discharge were only found significant with an adjusted probability of 10%. Groundwater recharge was found significant with F(5, 744)=2.1663, p=0.056 (< α=0.10), and discharge with F(5, 744)=2.0963, p=0.064 (< α=0.10). For the lower estimation scenario the variable discharge becomes insignificant with F(5, 744)=1.7301, p=0.1253 (< α=0.10), and in the upper scenario becomes significant under α=0.05, with F(5, 744)=5.5443, p=0.0308(< α=0.05). This indicates a high influence of the LUC scenario on the data for the Rhine river.

Table 15, MANOVA test for the best scenario on each variable.

MANOVA summary on variables - Best estimate scenario							
Evaporation		df	Sum sq	Mean sq	F value	Pr(>F)	Sig
	Time slice	5	0.6203	0.124066	16.575	1.546e-15	***
	Residuals	744	5.5691	0.007485			
Groundwater recharge		df	Sum sq	Mean sq	F value	Pr(>F)	Sig
	Time slice	5	0.1290	0.025795	2.1663	0.05603	.
	Residuals	744	8.8589	0.011907			
Discharge		df	Sum sq	Mean sq	F value	Pr(>F)	Sig
	Time slice	5	5009156	1001831	2.0963	0.06395	.
	Residuals	744	355560871	477904			
Runoff		df	Sum sq	Mean sq	F value	Pr(>F)	Sig
	Time slice	5	0.6367	0.127348	4.9559	0.00018	***
	Residuals	744	19.1181	0.025696			
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1							

To further analyze which time slices statistically differ from one another a post hoc Tukey HSD is performed. Remarkably the post-hoc analysis shows that discharge and groundwater recharge is not significant between any of the time slices, for each scenario. This explains the sensitivity to the ANOVA test between the scenario's. This means that the reduction in residual variance, with other words the difference between the predicted and observed discharge and groundwater recharge, can be explained by the difference in time slices. However, looking critically at the data for groundwater recharge and discharge, the p-value were on the verge for the set alpha, and the reason for a significant difference in the ANOVA may be due to the observations itself. Despite the Levene's test showing no statistical difference, thus an equal variance within groups, one predictor group may have a mean that is significantly different from the others for the ANOVA to return a significant value. Looking at the mean values of each time slice for groundwater recharge – in sequence of time slices: 0.2622078, 0.2649721, 0.2710951, 0.293974, 0.2896006 – this could be the case. The post hoc test specifically compares two time slices per variable, and evaluates them on pooled variance estimates, rather than means and residual variance. Another possibility could be that there are pooled groups of time slices that do differ from the other (grouped) time slices that would show a significant result in the ANOVA, but not the post hoc (Huck, 2015). However, it can be concluded that for the variables ground water recharge and discharge no significant difference between the time slices were found, but that in general the time slices do significantly differ in the upper estimate scenario. For the other variables, evaporation of time slices 1, 2 and 3 were found significantly different from time slices 4, 5 and 6. For the variable runoff time slice 1 was found significantly different from time slice 4, 5 and 6, and time slice 2 was found significantly

different from time slice 4 and 5. The best and upper scenario show similar results, however the lower scenario found that the variable runoff only was significant between time slice 1 and 6. These results are visualized in boxplot figure 22.

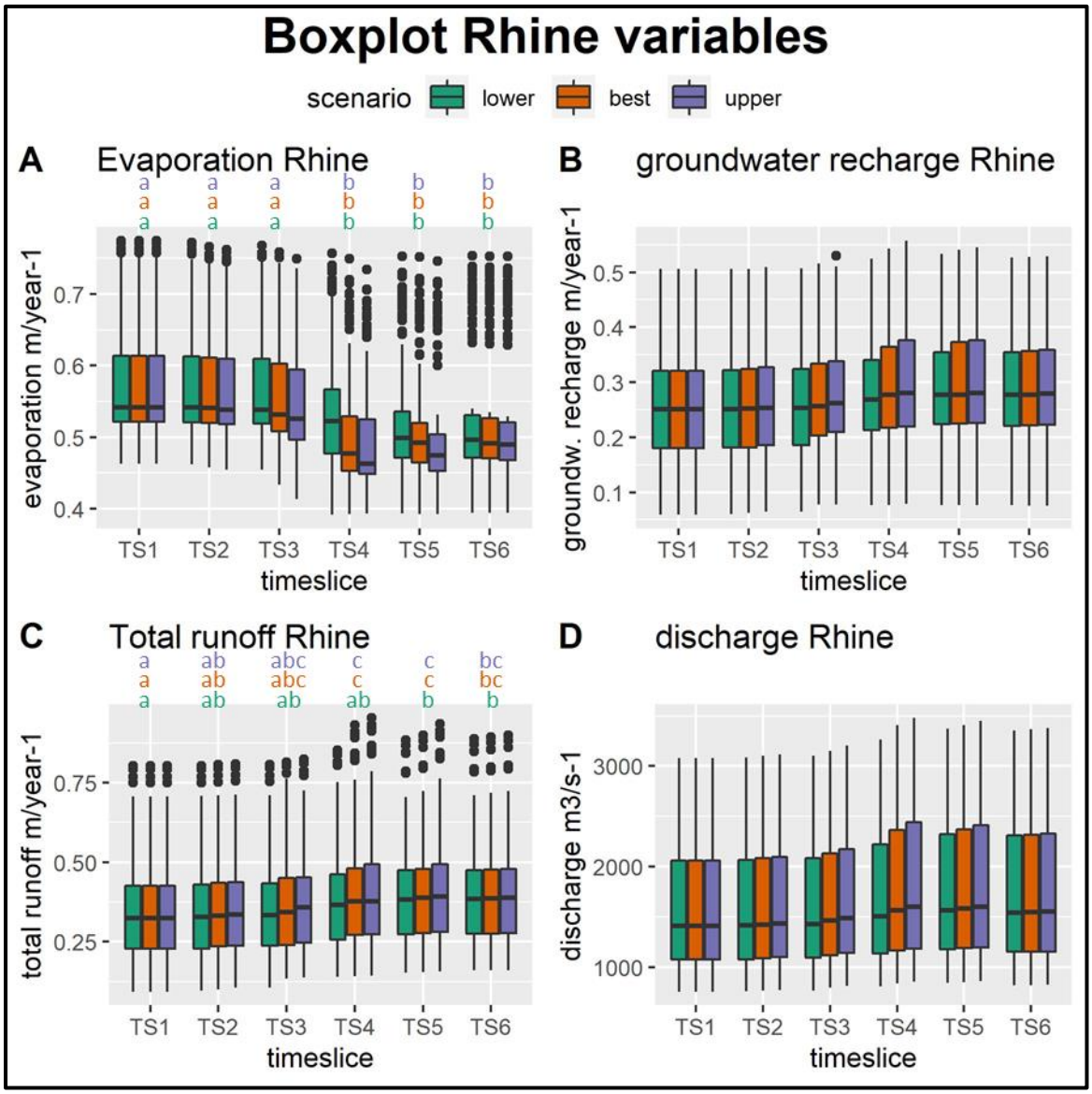

Figure 22, boxplot of hydrological variables for the Rhine with post-hoc Tukey HSD significance letter groups included

To evaluate the results, a closer look on the non-natural vegetation fractions must be done. In figure 23 it can be seen that the non-natural vegetation on average decreases in the sixth time slice. This was also visible in the maps from chapter 4.1, where in western Europe the non-natural vegetation fraction decreases. Of the total non-natural vegetation fraction, the most present cultivation type are the rainfed crops. What also stand out, is this broad standard error grey range, that indicates that the values are very dispersed for the data points.

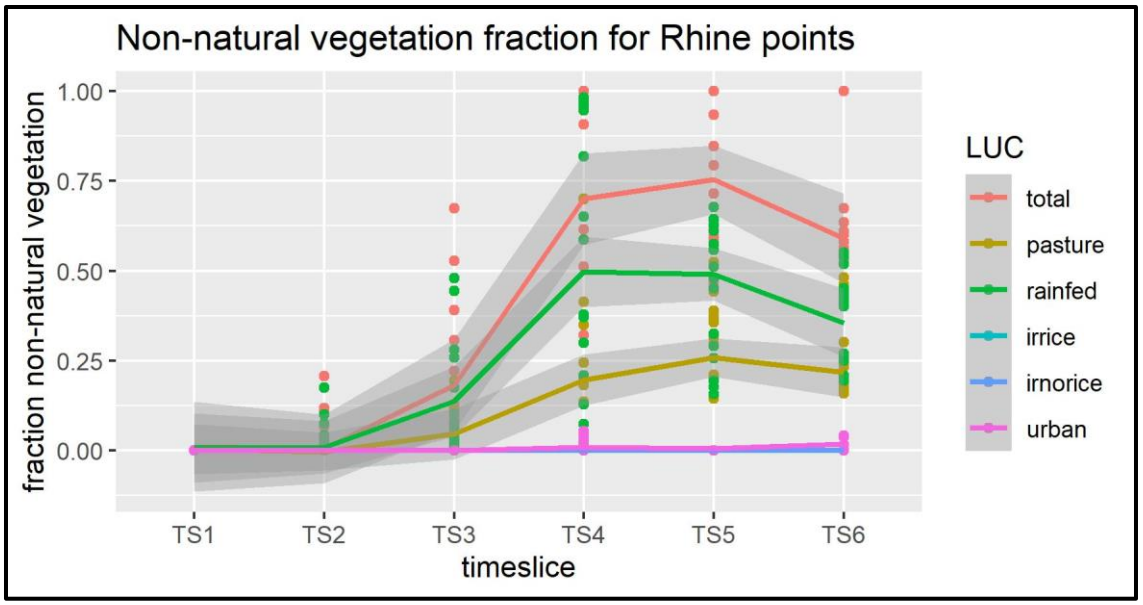

Figure 23, non-natural vegetation fraction for the Rhine data points, for the total non-natural fraction, pasture, rainfed crops, irrigated rice paddies, irrigated non-rice paddies and urban area. The grey area represents the standard error of each vegetation fraction

The correlation matrix, table 16, shows that all variables have very weak correlations with LUC, for some variables not found significant under p-value α=0.05.

Table 16, Pearson's correlation matrix of LUC and the hydrology variables

Pearson's correlation between LUC and the hydrology variables.						
		cor	statistic	P value	Lower conf	Higher conf
LUC	Evaporation	-0.230	-6.3967843	2.80e-10	-0.29452996	-0.15874272
LUC	Precipitation	0.010	0.2626631	7.93e-01	-0.06202783	0.08113635
LUC	GW recharge	0.028	0.7678562	4.43e-01	-0.04361173	0.09945339
LUC	Discharge	0.033	0.9055855	3.65e-01	-0.03858672	0.10443459
LUC	runoff	0.140	3.9093445	1.01e-04	0.07062834	0.21095326

4.3 Spatial analysis

In this spatial analysis the main components of the water balance are evaluated, being the precipitation, evaporation, runoff and groundwater recharge. Additionally the evapotranspiration is analyzed, as the hypothesis states a response to the loss of natural vegetation is expected. Firstly the components are discussed per time slice, and then in separate sections each component is analyzed in depth in relation to the land use change.

The first time slice shows in figure 24F that vegetation fraction is almost completely natural vegetation, which for each grid cell is equally divided in 50% tall vegetation and 50% short vegetation, and all fractions of rangeland are additionally added to this short vegetation fraction. Now 10.000 BCE can be seen as a baseline, to which the future LUC can be compared relatively to the hydrology variables. A high precipitation value can be seen in the northwestern tip of Iberia, around the Alpine region and the coast of southeastern Europe. In this same regions and its surrounding, positive values of total runoff and groundwater recharge are found. For all estimation scenario's a similar LUC and response in water balance variables is seen, as the fraction of non-natural vegetation is mostly zero for all estimation scenarios. Similar figures for the lower and upper estimation scenario can be found in appendix section 4.

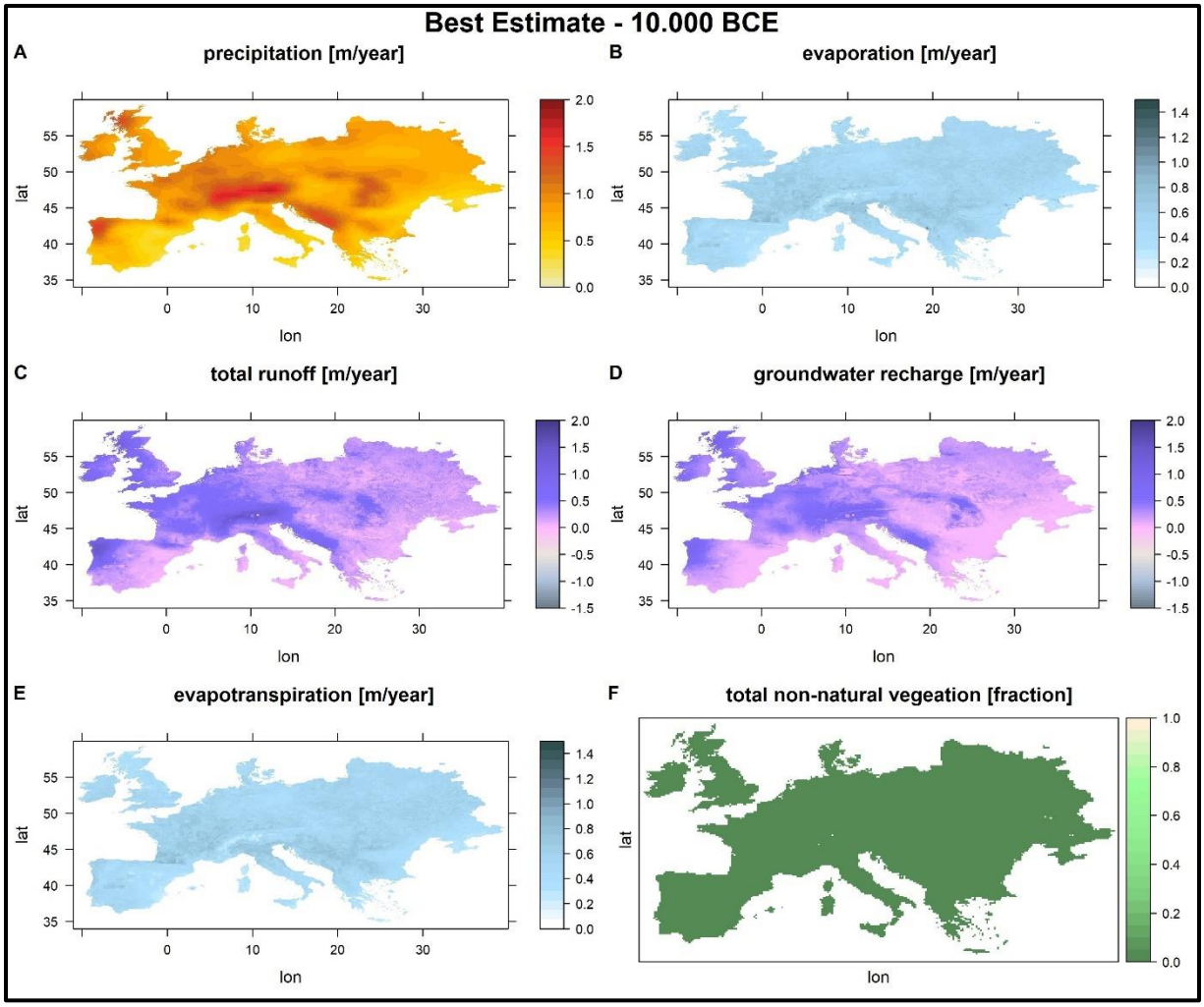

Figure 24, spatial analysis of best estimate

The second time slice, 1.000 CE in figure 25, already shows an increase of about 0.25 in Iberia, western Europe and central eastern Europe of non-natural vegetation. Visually no difference in the hydrological variables can be seen, compared to the first time slice. The northern Italian region that shows ±0.2 fraction of non-natural vegetation, also shows ±0.05 m/year decrease in evaporation. The Swiss Alpine region as well as the points in the central Europe that are light green in figure 25F, see an increase of ±0.1 m/year runoff. This indicates that a shift for natural vegetation to rainfed crops (figure 26) leads to a decrease in evaporation, an increase in runoff. The groundwater recharge is less unambiguous, showing a slight decreased (± 0.1 m/year) groundwater recharge for the LUC on the land near the black sea, the Alpine region and western France, but a slight increase (± 0.1 m/year) in parts of Switzerland, central Europe and the southeastern region. In section 4.3.1 to 4.3.3 the differences are more evidently presented and discussed.

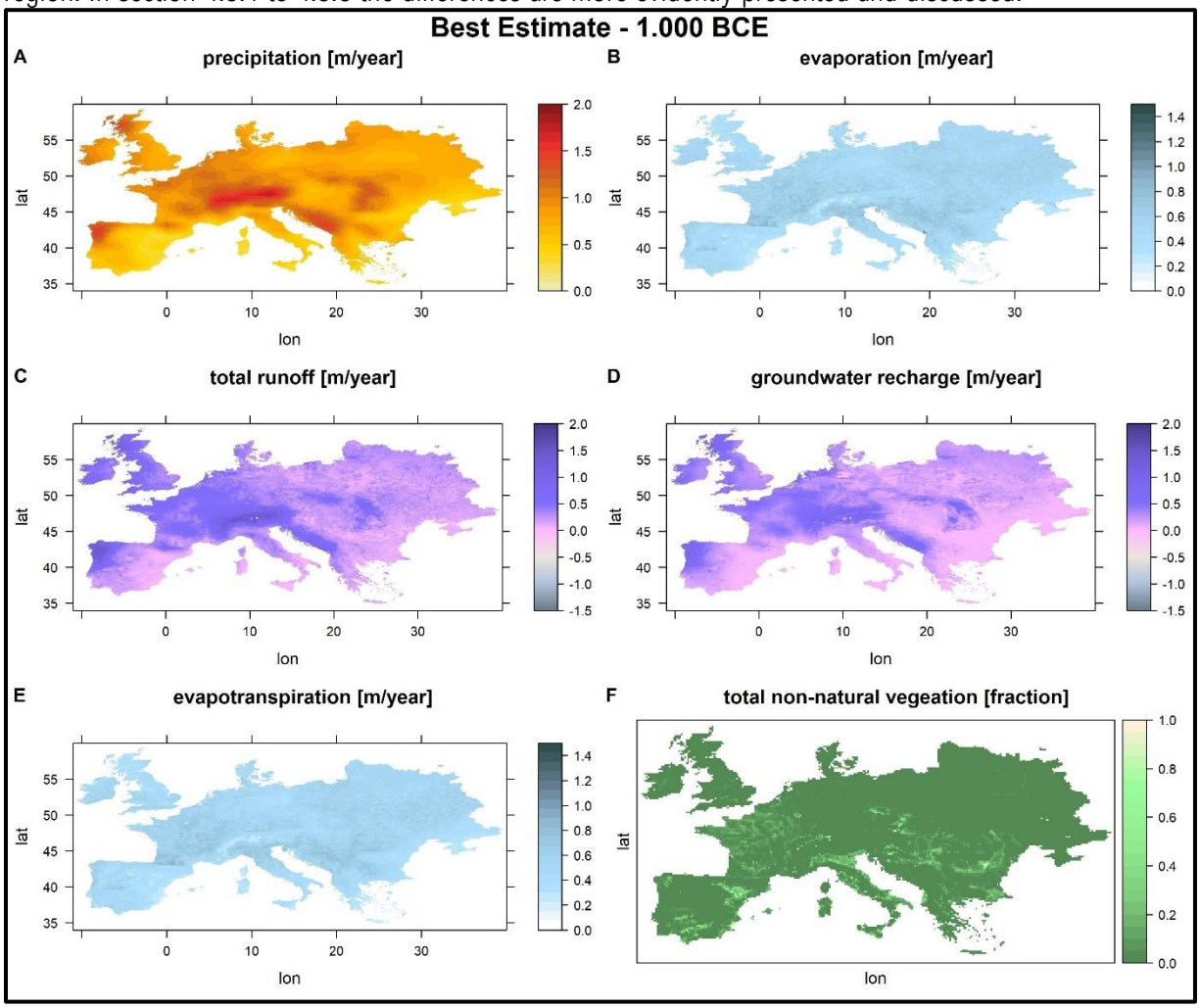

Figure 25, the hydrological variables for time slice 2, 1.000 BCE, in the best estimate scenario

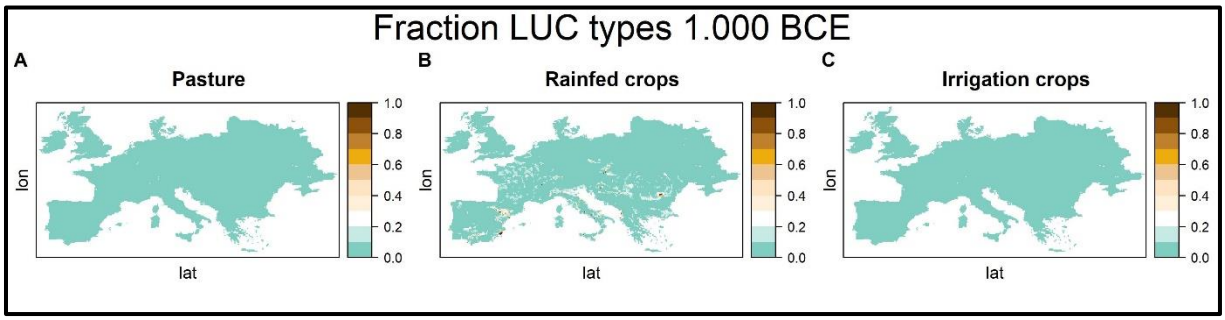

Figure 26, different land use change types for 1.000 BCE in the best estimate scenario, with A. pasture, B. rainfed crops and C. irrigated crops

The next time slice is year 100 CE, showing great changes in LUC that is visible from figure 27F. Especially in central Iberia, western France, most of Italy and East-Central Europe the amount of fraction in non-natural vegetation rises, specifically rainfed crops (figure 28). An 0.20 decrease in evaporation for the Alpine region most notably stands out, as an reaction to the even further increase of non-natural vegetation in 100 CE. The large scale of cultivation in the Roman Empire cannot just be seen in loss of vegetation, but also by the increase in runoff. Again, groundwater recharge remains indefinite and shows both gains and losses. Remarkable in this time slice is also the homogenous decrease of 0.1 m/year runoff and groundwater recharge in Hungary that does not particularly stand out in evaporation, evapotranspiration nor the LUC. This decrease in runoff for Hungary is more prominently visible in chapter 4.3.2 figure 38.

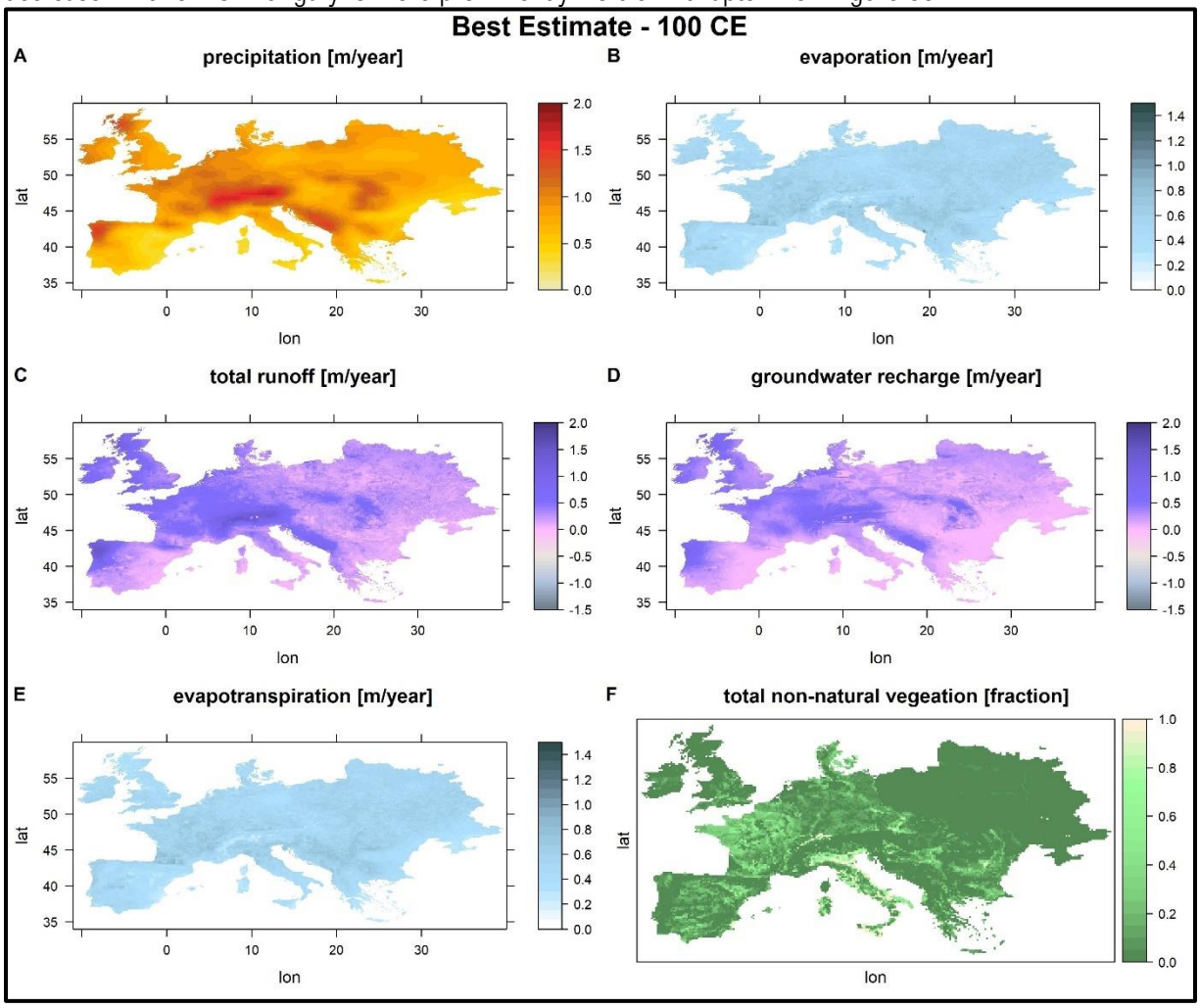

Figure 27, the hydrological variables for time slice 3, 100 CE, in the best estimate scenario

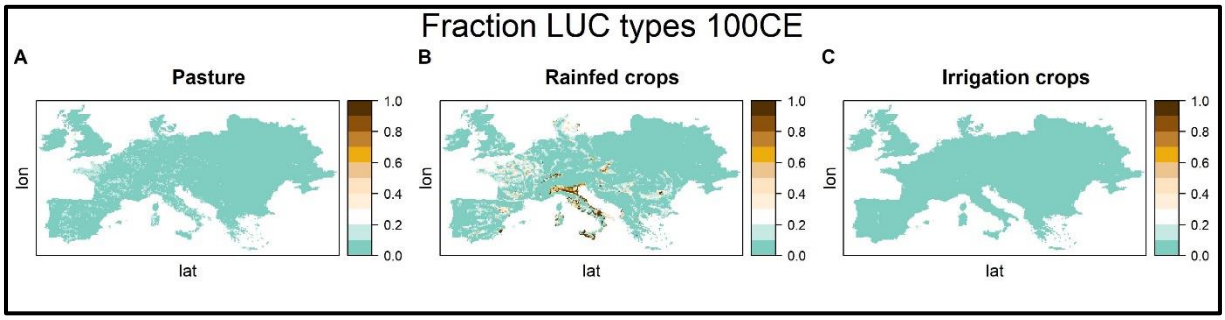

Figure 28, different land use change types for 100 CE in the best estimate scenario, with A. pasture, B. rainfed crops and C. irrigated crops

The following time slice is 1300 CE, figure 29, the year where most of western Europe was maximally occupied by non-natural vegetation, having large areas of grid cells with values of 1. The share of pasture has been explosive, compared to 100 CE from figure 28. The LUC is visible in the evaporation variable with additional decreased values 0.1 m/year in western Europe. The value of evaporation for northern Italy compared to 100 CE remains the same, as non-natural vegetation fraction was already 1. The regions that now completely exists of rainfed crops and pasture have seen a total decrease of about 0.05 – 0.10 m/year evaporation compared to 10.000 BCE, and about 0.2 m/year increase of runoff.

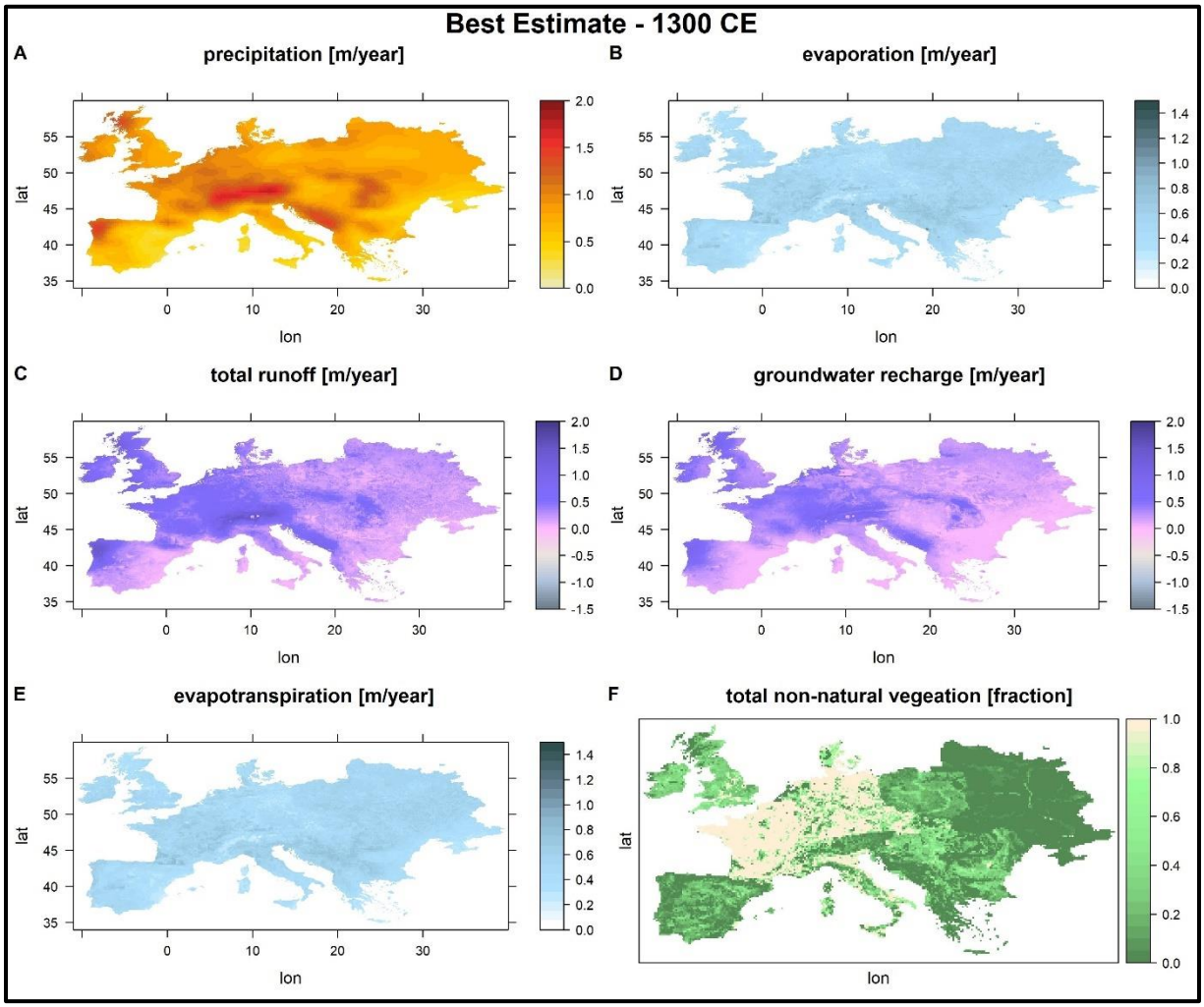

Figure 29, the hydrological variables for time slice 4, 1300 CE, in the best estimate scenario

To further analyze the changes in 1300 CE, the vegetation types should be analyzed. Showing what type of vegetation is present, as with increasing values of LUC, the type of vegetation might play a bigger role. Figure 30 already indicates the large present of rainfed crops in Italy and the sudden large introduction of pasture.

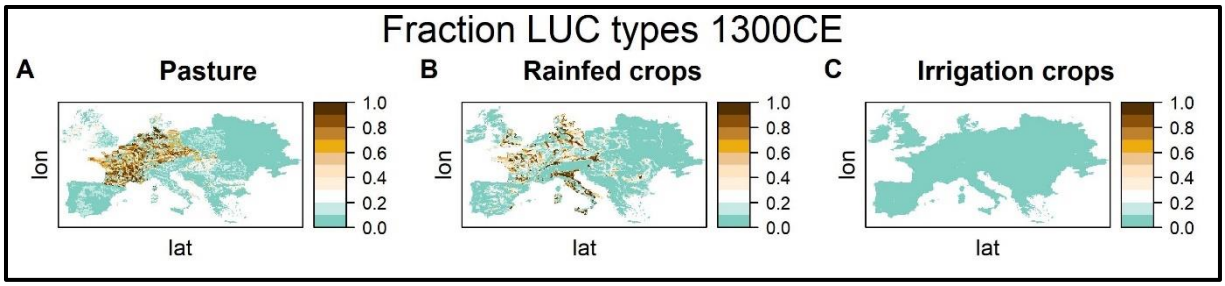

Figure 30, different LUC types for 1300 CE, with A. pasture, B. rainfed crops and C. irrigated crops.

In 1850 CE the fraction of non-natural vegetation further intensifies in regions that were first only moderately occupied by human cultivation. Figure 31 shows again in all variables, except precipitation, changes compared to earlier time slices. Looking first into the LUC, Ireland, central Europe, central Iberia and eastern Europe now mostly cross the 50% of non-natural vegetation cover per grid cell.

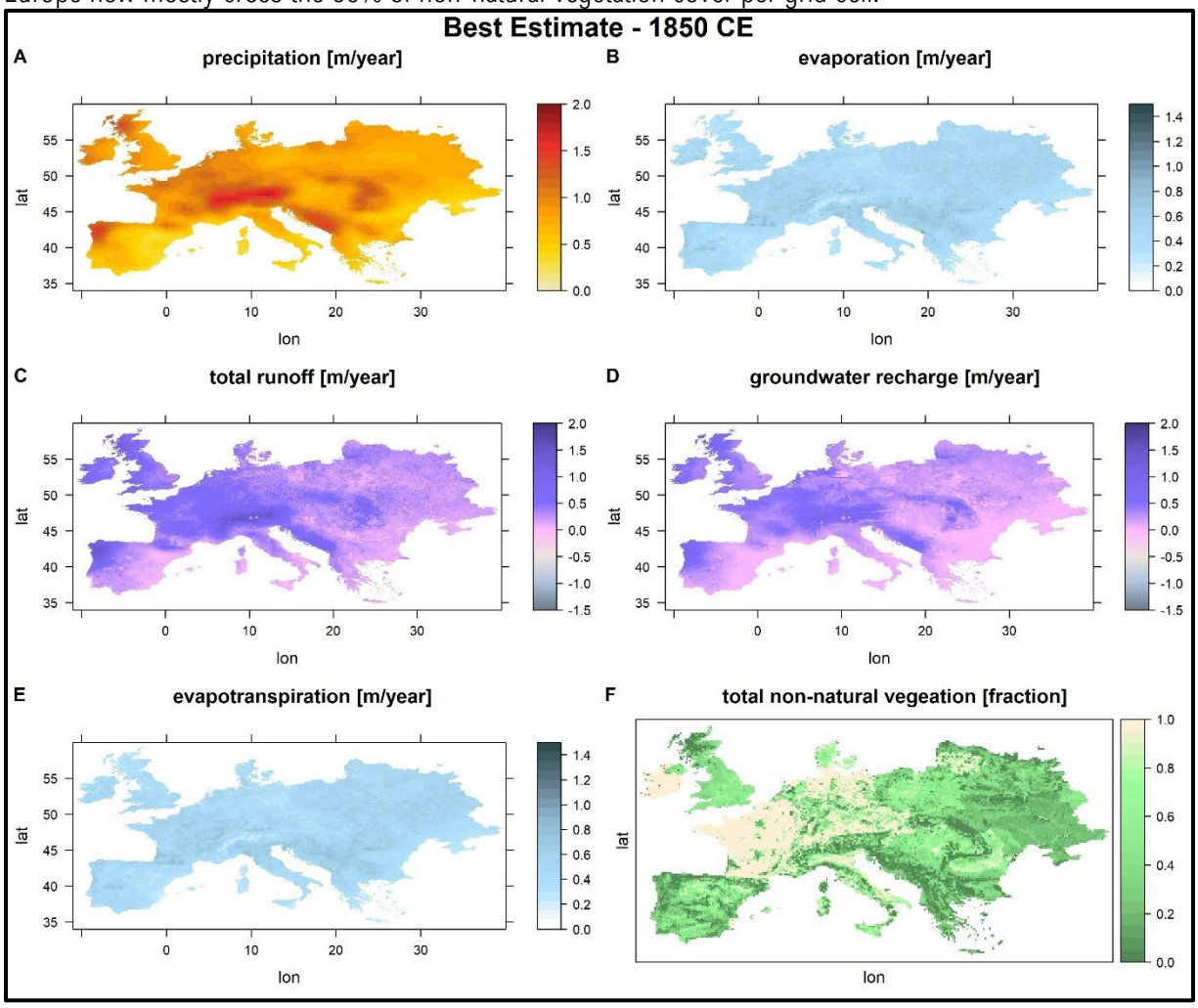

Figure 31, the hydrological variables for time slice 5, 1850 CE, in the best estimate scenario

LUC in Ireland is visible on the hydrological variables, but not as extensively as drastic change in LUC. To further analyze this, the different types of vegetation are showed in figure 32. Here it becomes clear that most of the Irish land area transformed from natural vegetation and rangeland to pasture, which might not have that much of an impact compared to a more rainfed crop dominated non-natural vegetation fraction type. The land areas that changed to irrigated crops in Italy and Iberia show an increased value for evaporation and a decreased value for runoff, indicating opposite effects from rainfed crop and pasture.

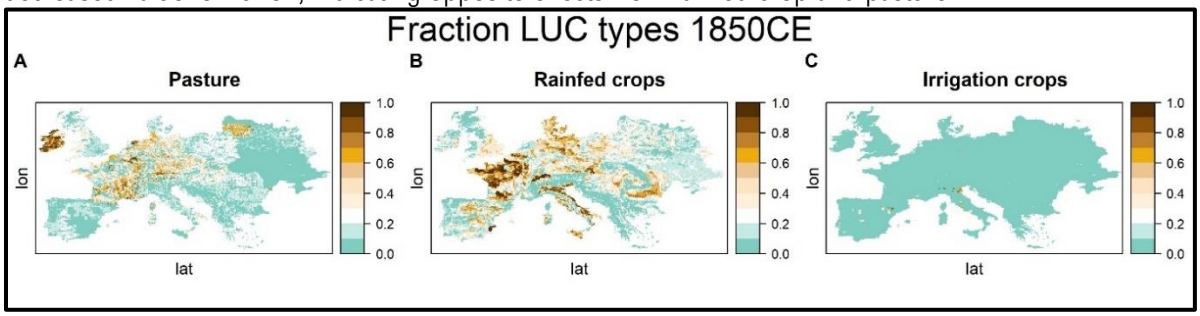

Figure 32, map of the different vegetation types for 1850 CE best estimate with A. fraction of pasture, B. fraction of rainfed crops and C. fraction of irrigated crops

Lastly, figure 33 shows year 1950 CE, which remarkably shows a decrease and increase in non-natural vegetation. In eastern Europe there is spatially large increase in non-natural vegetation, compared to previous time slices. In western Europe the fraction of total non-natural vegetation decreases slightly, and thus gains natural vegetation. The increase of non-natural vegetation in eastern Europe is mostly rainfed crops dominated. This results in an increased runoff with values around 0.2 m/year, similar to the increased runoff that western Europe experienced between 100 CE and 1300 CE.

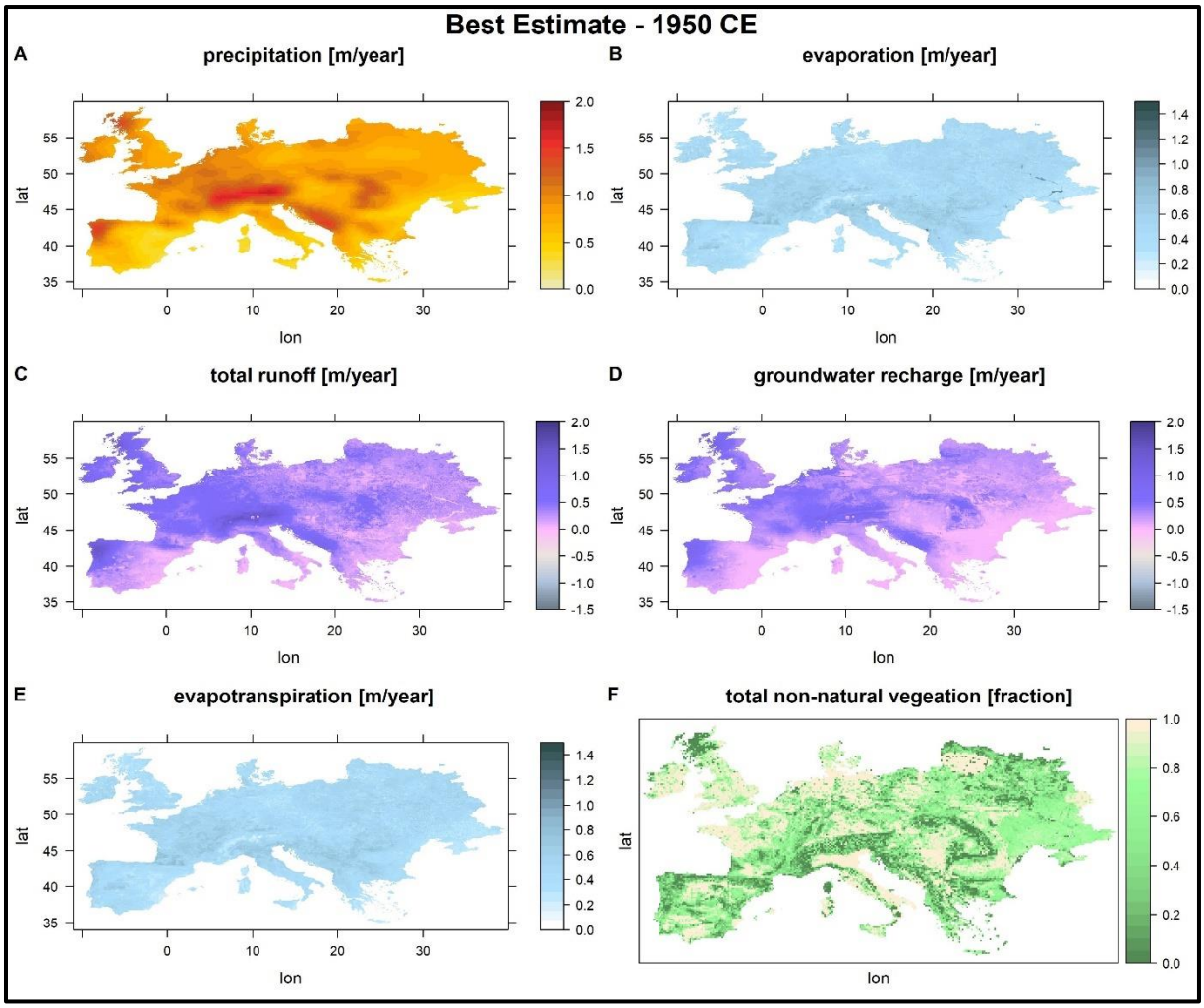

Figure 33, the hydrological variables for time slice 6, 1950 CE, in the best estimate scenario

The LUC, figure 34, shows an increase in rainfed crops in eastern Europe, and a sharp shift from rainfed to irrigated crops in northern Italy and patches of land in Iberia, explaining the reverse effects in evaporation and runoff.

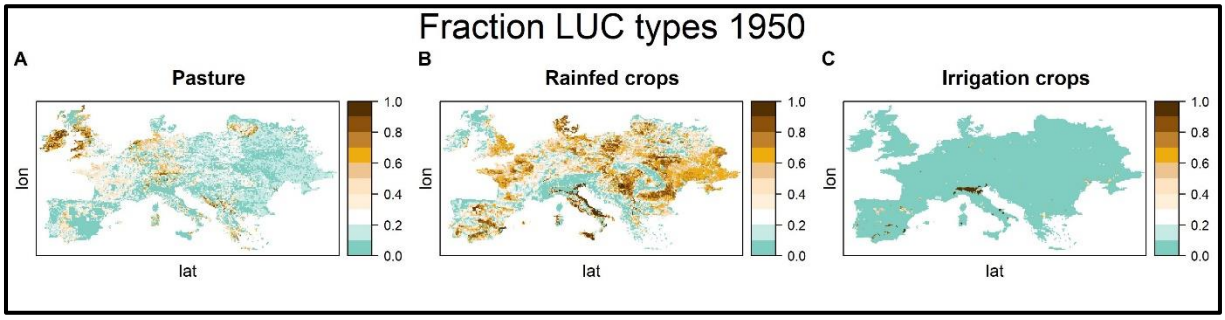

Figure 34, the LUC vegetation types in fraction for time slice 6 (1950 CE), with A. pasture, B. rainfed crops and C. irrigated crops

4.3.1 Evaporation

Looking more closely at the changes within the evaporation variable, figure 35 shows the difference in evaporation between two consecutive time slices, and with the last map, the difference between the first (10.000 BCE) and last (1950 CE). The range of change of the first four difference maps are equal, with an increase or decrease of about 0.1 m/year. But it changes completely for the figures 35E and F. For the evaporation variable a few lighter spots, a decrease in evaporation, can be seen in figure 35A. Overall that map contains a minimal value of -0.206 and maximum value of 0.166. However, spatial differences are mostly between 0.0 and -0.1 m/year. A general decrease in evaporation can be seen over the course of 1.000 BCE to 100 CE, but small spatial patches in central Europe do indicate local increases of evaporation. In figure 35D an increase in evaporation can be seen locally, due to the cultivation of irrigated crops. In that similar time difference, a decrease in evaporation for eastern Europe can be seen, which then knows a large increase of rainfed crops. With the last time slice, the river Dniepr appears, creating both very large increases and decreases in evaporation due to its reservoirs. Although eastern Europe still shows decreasing values compared to 1850 CE, most of central and western Europe shows an increase, which is due to the gain of natural vegetation in 1950 CE. The maximum value in figure 35F is 0.868 m/year, mostly in the riverine area of the Dniepr, and the minimum value is -0.272 m/year.

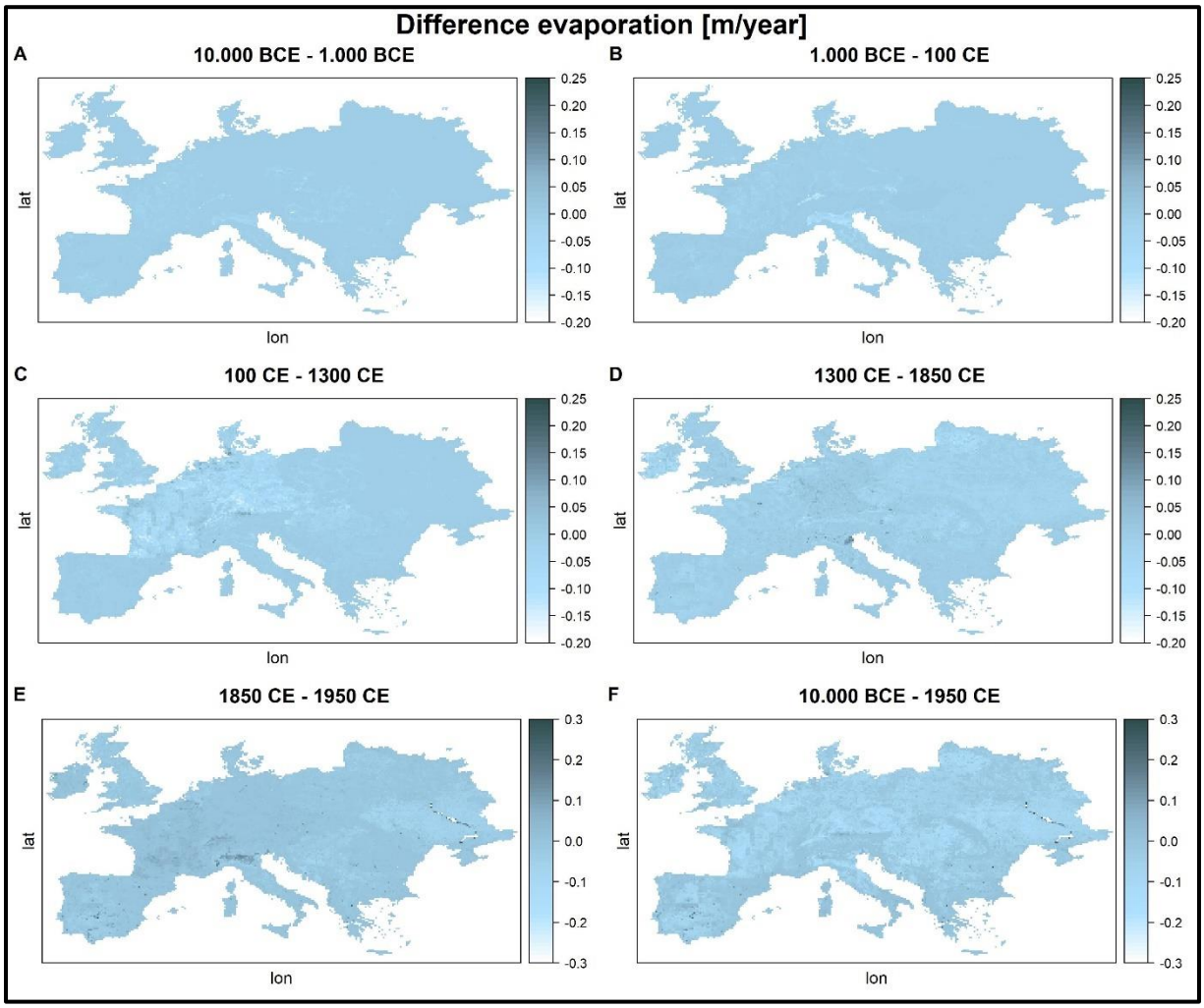

Figure 35, the maps of evaporation that show the differences between A. 10.000 BCE and 1.000 BCE, B. 1.000 BCE and 100 CE, C, 100 CE and 1300 CE, D. 1300 CE and 1850, E. 1850 CE and 1950 CE, and at last F. 10.000 BCE and 1950 CE.

To further analyze how evaporation correlates with LUC, a raster correlation analysis is performed, that correlates each cell for Pearson's method. In figure 36 the results can be seen, in which each map represents the cell to cell correlation coefficient between the LUC and evaporation for each time slice. The correlation strength is presented with the darkening of colors, as classified by 1 to 0.91 = very strong, 0.90 to 0.71 = strong, 0.70 to 0.51 = moderate, 0.50 to 0.31= weak, 0.30 to 0.01 = very weak. The sign indicates a positive (red) or negative (blue) correlation. In figure 36A there still is limited non-natural vegetation fraction, thus not many cells correlate. What stands out from the correlation matrix is that the negative correlation coefficients in eastern Europe becomes the strongest in the last two time slices, which is explained by the general increase of non-natural vegetation in that area, specifically rainfed crops.

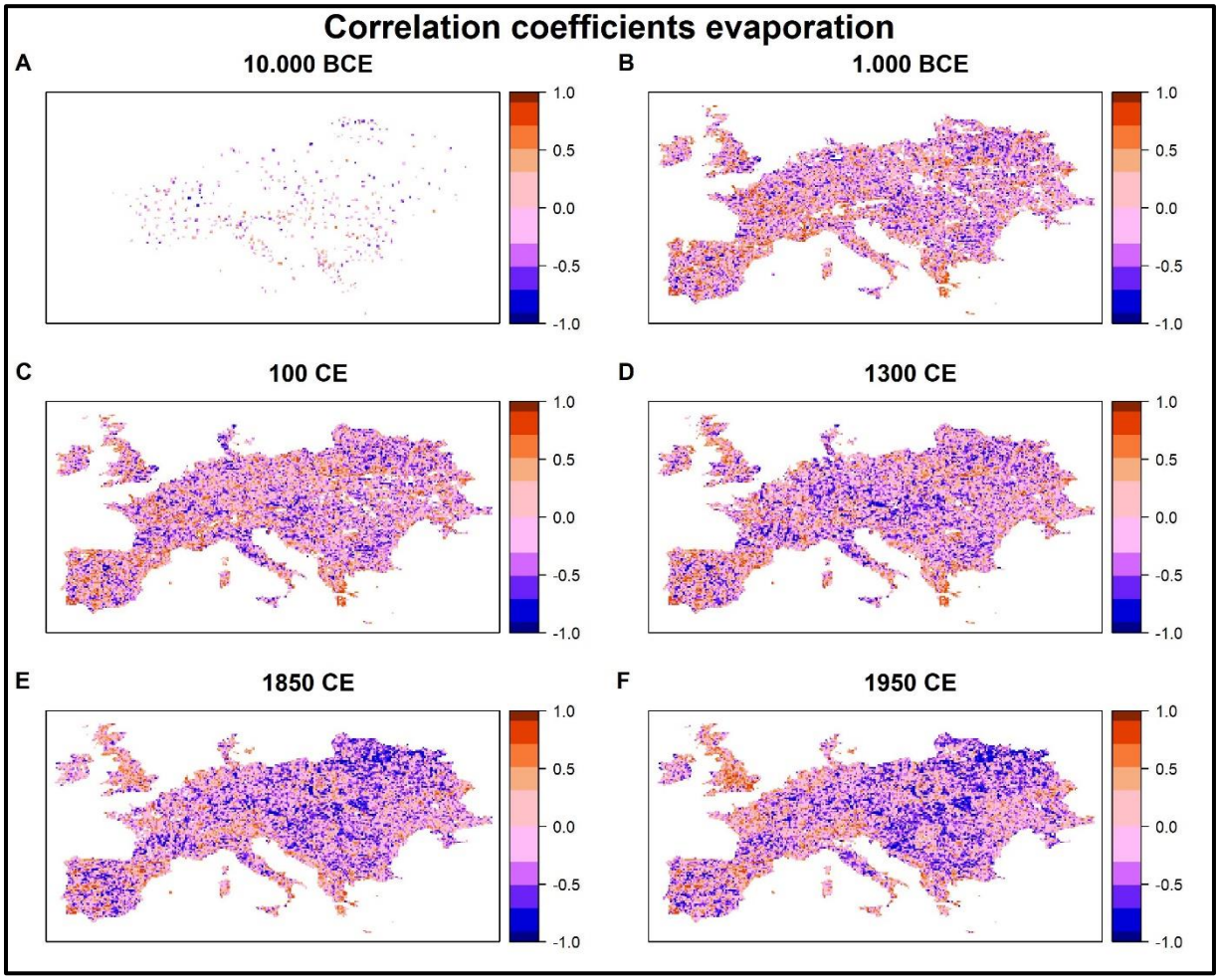

Figure 36, raster correlation with Pearson's correlation coefficients for A. LUC and evaporation at time slice 10.000 BCE, B. at 1.000 BCE, C. 100 CE, D. 1300 CE, E. 1850 CE and F 1950. The color ramp indicates the strength of the correlation coefficient as the 1 - 0.91 is very strong, 0.9 to 0.71 is strong, 0,70 to 0.51 is moderate, and 0.50 to 0.31 is weak, and at last 0.30 to 0.01 very weak. The sign indicates a positive or negative correlation coefficient.

To look more closely into with which vegetation types evaporation shows the strongest correlations, figure 37 shows the distinction between rainfed crops, irrigated crops and pasture. The land areas with more fraction of a specific vegetation type will give stronger correlation coefficients, thus only the very strong correlation coefficients are considered. Only time slices 1850 and 1950 are considered, as they entail the most vegetation covers and types of covers. The relation between irrigated crops and evaporation are generally very strong to strong, and mostly positive. Furthermore, rainfed crops show a more strong negative relations to evaporation, than pasture. On the contrary, pasture has more positive relationship with evaporation in the Dniepr region, Britsh isles, and the western European coast.

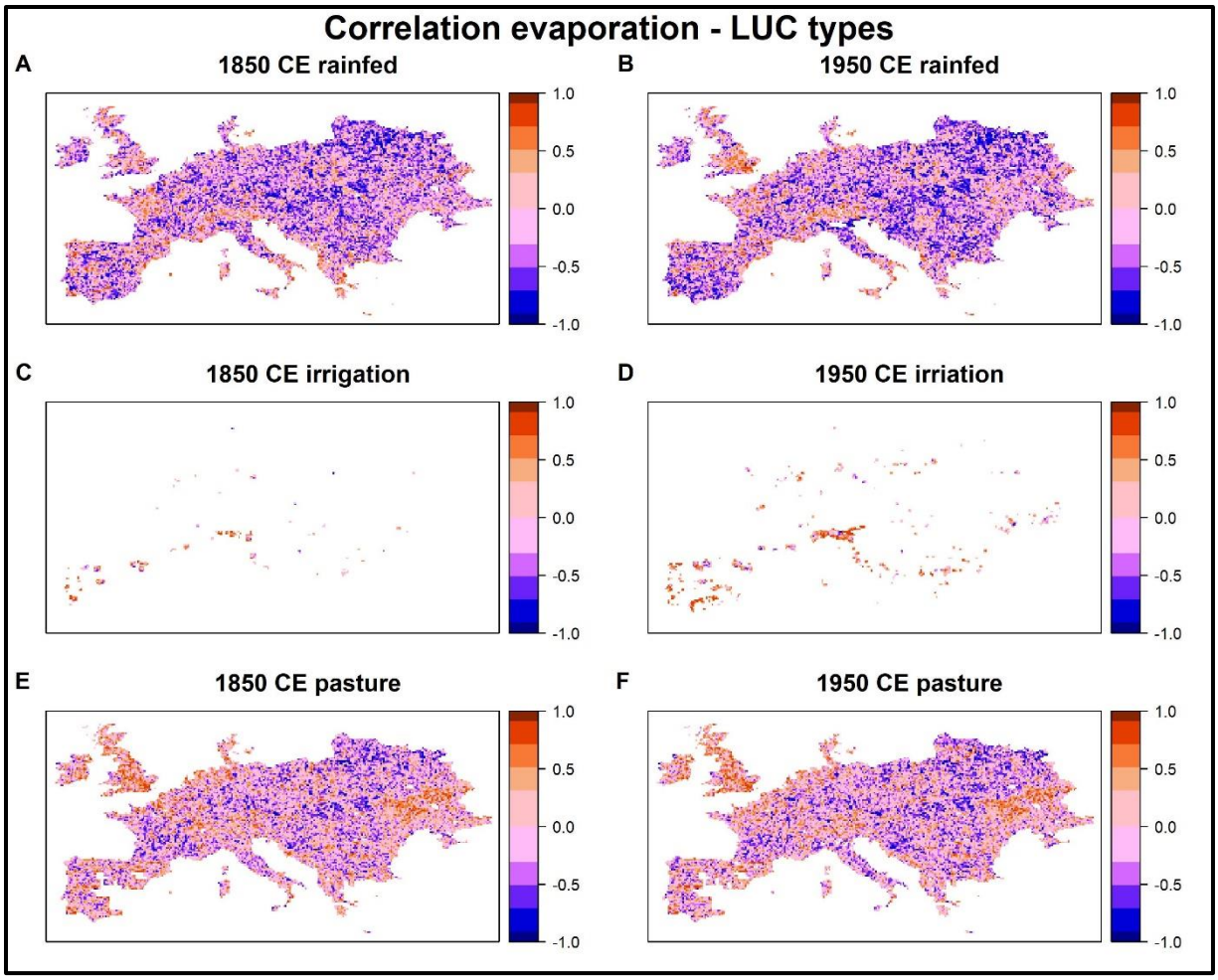

Figure 37, raster correlation with Pearson's correlation coefficients for A. rainfed crops LUC and evaporation at time slice 1850 CE, and B. at 1950 CE. The irrigated crops for C. 1850 CE and D. 1950 CE, in relation to evaporation, for pasture E. at 1850 CE and F. in 1950. The color ramp indicates the strength of the correlation coefficient as the 1 - 0.91 is very strong, 0.9 to 0.71 is strong, 0,70 to 0.51 is moderate, and 0.50 to 0.31 is weak, and at last 0.30 to 0.01 very weak. The sign indicates a positive or negative correlation coefficient.

4.3.2 Runoff

For the variable runoff the differentiation map, figure 38, shows between 10.000 BCE and 1.000 BCE a regional heavy increase of 0.2 m/year, but also decreases of about 0.1m/year in for example Italy, Spain and Albania. Figure 38A has a minimum value of -0.167 and a maximum value of 0.219, seen the relatively low increase of vegetation between these time slices, the runoff increases quite a lot. Between 1.000 BCE and 100 CE the minimum value is -0.052 and 0.176, but the changes are visible on a larger scale. The regions like Italy and central Europe, that have previously been discussed on their LUC, show increased runoff values, whereas the Hungarian region shows a minor decrease. This decrease will be further analyzed in the raster correlation analysis. Figure 38C has a minimum value of -0.369 and maximum value of 0.253, and for D minimum -0.272 and maximum 0.368, indicating that between 100 CE and 1300 CE there generally is more negative peaks in runoff, and between 1300 CE and 1850 higher positive peaks in runoff. This is visually also visible for the maps. In time slice 6 the river Dniepr undergoes some changes in de model, due to the inclusion of reservoirs. The Dniepr river has many reservoirs and dams, that also make this map have a minimum value of -0.868. In 1950 CE, when natural vegetation returns in western Europe, the runoff also decreases for western Europe. The overall change, figure 38F, actually reveals that the net change in runoff for west European decreased, but in that calculation the higher increases of runoff are not included. As could be seen in the river analysis and vegetation maps, it were the time slices 4 and 5 that were significantly different from the others, as natural vegetation returns in 1950. The map of figure 38F thus shows a net increase, but does not consider the higher values of runoff that have occurred in the meantime.

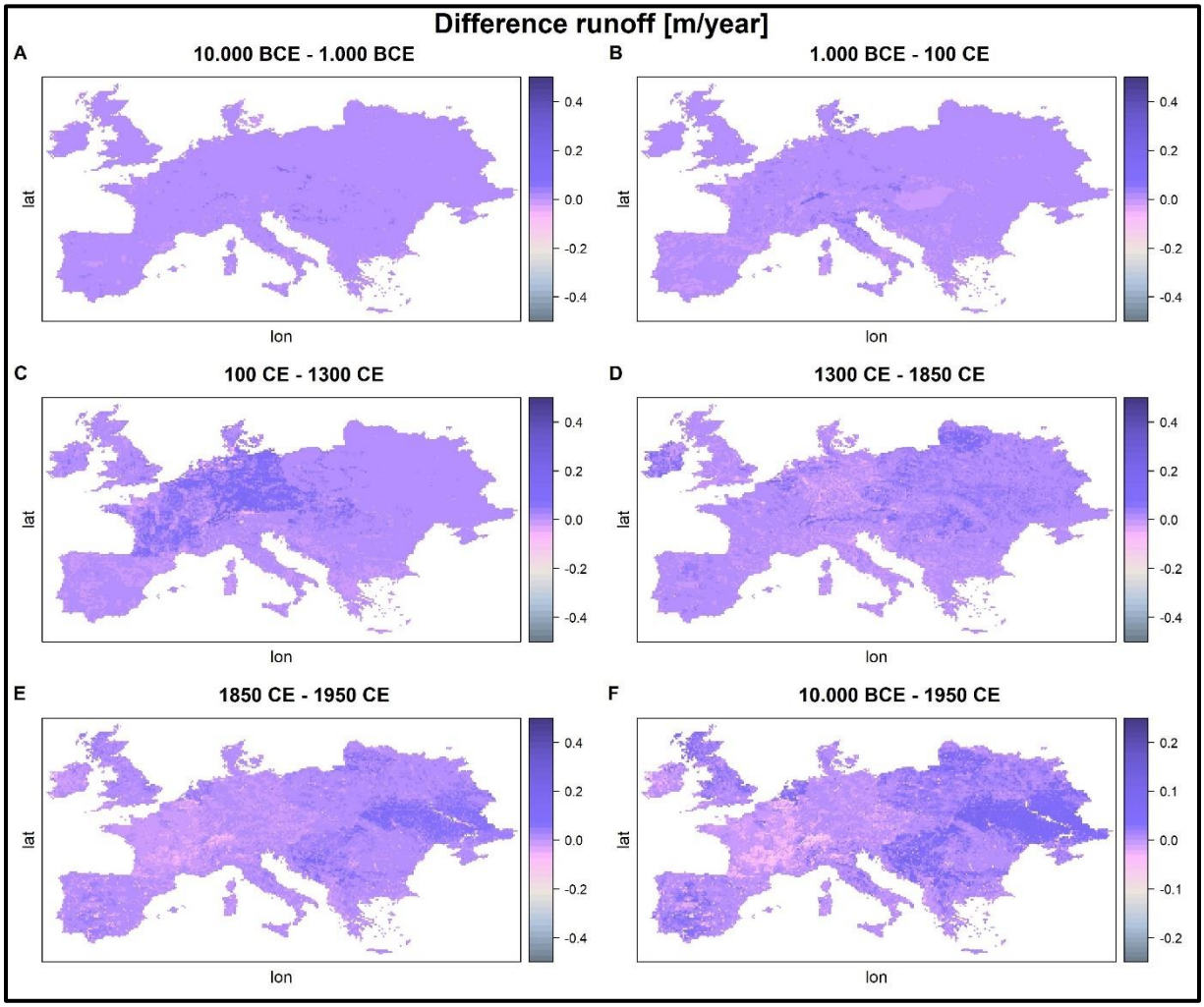

Figure 38, the maps of runoff that show the differences between A. 10.000 BCE and 1.000 BCE, B. 1.000 BCE and 100 CE, C, 100 CE and 1300 CE, D. 1300 CE and 1850, E. 1850 CE and 1950 CE, and at last F. 10.000 BCE and 1950 CE.

Now the variable runoff is compared on each time slice to the LUC with a Pearson's correlation method to see the relationship between total runoff and LUC. Again only the very strong correlations are considered. The very strong correlation grids are widely dispersed over the map, but the very strong positive correlation in middle Iberia and middle Ireland in 1.000 BCE to 1300 CE stand out. In 1850 and 1950 the eastern European region has many very strong relations, whereas 1.000 BCE and 100 CE are more negatively correlated in both eastern and western Europe. If we compare the correlation rasters of figure 39 from runoff with the correlation rasters figure 36 of evaporation, there is a slight inversed result visible between the variables. Where 1850 and 1950 are strongly red for runoff, the evaporation rasters shows strong blue in the same regions, which is in line with the water balance when precipitation is equal.

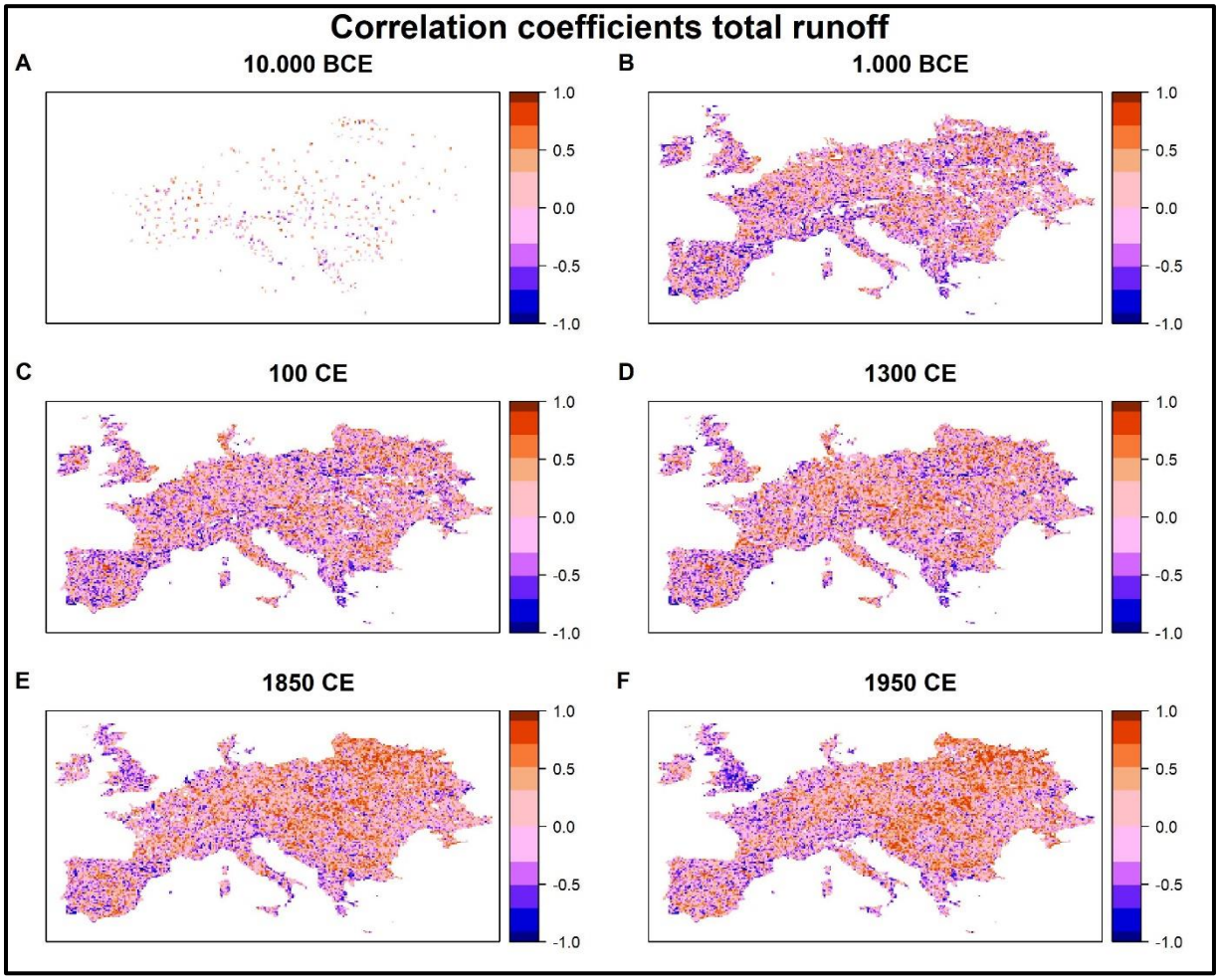

Figure 39, raster correlation with Pearson's correlation coefficients for A. LUC and runoff at time slice 10.000 BCE, B. at 1.000 BCE, C. 100 CE, D. 1300 CE, E. 1850 CE and F 1950. The color ramp indicates the strength of the correlation coefficient as the 1 - 0.91 is very strong, 0.9 to 0.71 is strong, 0,70 to 0.51 is moderate, and 0.50 to 0.31 is weak, and at last 0.30 to 0.01 very weak. The sign indicates a positive or negative correlation coefficient.

For different LUC types, the correlation has also been performed to see the differences in correlation for pasture, irrigated crops and rainfed crops. Especially the net decrease in runoff for western Europe needs some close attention. Figure 40 shows a very strong negative correlation with irrigated crop in the northern Italian regions where irrigated crops are cultivated. The correlation signs are again the inverse of the evaporation figure, which can best be seen figure 40C and D. Again, the Dniepr region has a very strong negative relation between runoff and pasture land cover, as well as England. Both these regions have high fractions of pasture, and high runoff values in m/year. For western Europe, small patches of very strong correlations were found, both positive and negative.

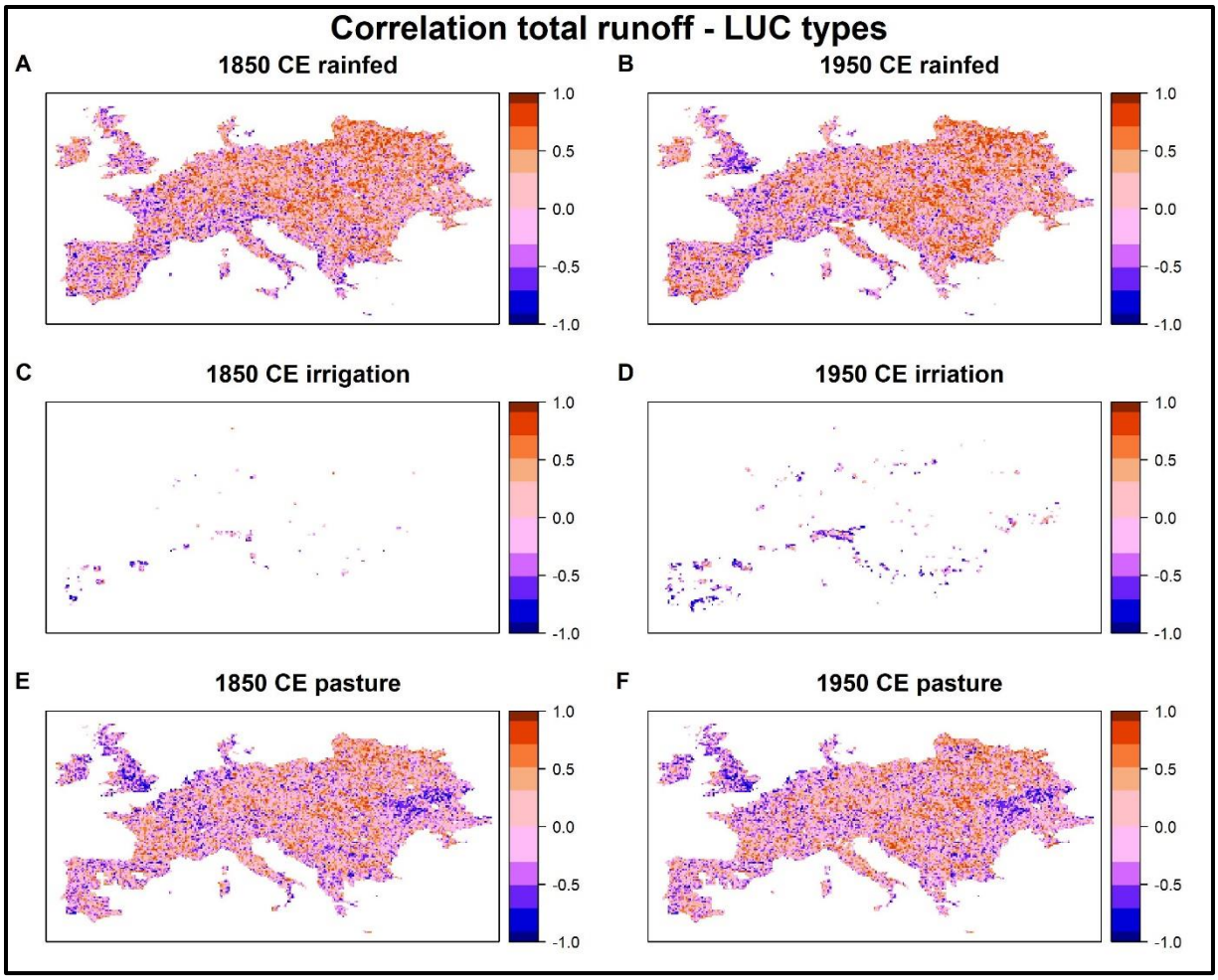

Figure 40, raster correlation with Pearson's correlation coefficients for A. rainfed crops LUC and runoff at time slice 1850 CE, and B. at 1950 CE. The irrigated crops for C. 1850 CE and D. 1950 CE, in relation to evaporation, for pasture E. at 1850 CE and F. in 1950. The color ramp indicates the strength of the correlation coefficient as the 1 - 0.91 is very strong, 0.9 to 0.71 is strong, 0,70 to 0.51 is moderate, and 0.50 to 0.31 is weak, and at last 0.30 to 0.01 very weak. The sign indicates a positive or negative correlation coefficient.

4.3.3 Groundwater recharge

In figure 41 the variable groundwater recharge and the differences between the time slices can be seen. As already discussed in the general spatial analysis per time slice, the groundwater recharge component has showed ambiguous results. The groundwater recharge is what remains of the precipitation that has fallen on a grid cell, and is then either evaporated or runs off, but with negative values groundwater is not recharged. Despite being the sum of evaporation and precipitation, the response of this component to LUC could become clear from this research with the distinction on vegetation types. Furthermore, human irrigation from groundwater sources has not been included in this research, thus irrigated crops will not show a result representative for of actual irrigated areas. A general overview of the differences with each time slice indicates a large ground water increase in northern Italy (41B) central Europe (41C) and eastern Europe (41E). Standing out, are the decreases in groundwater recharge, in Hungary (41B) and Switzerland (41E) for example. This could be explain by the topography of those regions, as the mountainous area has a less stable groundwater component, and groundwater flows to lower areas. In terms of LUC a supposed decrease of groundwater recharge in non-natural vegetation fraction was not visible.

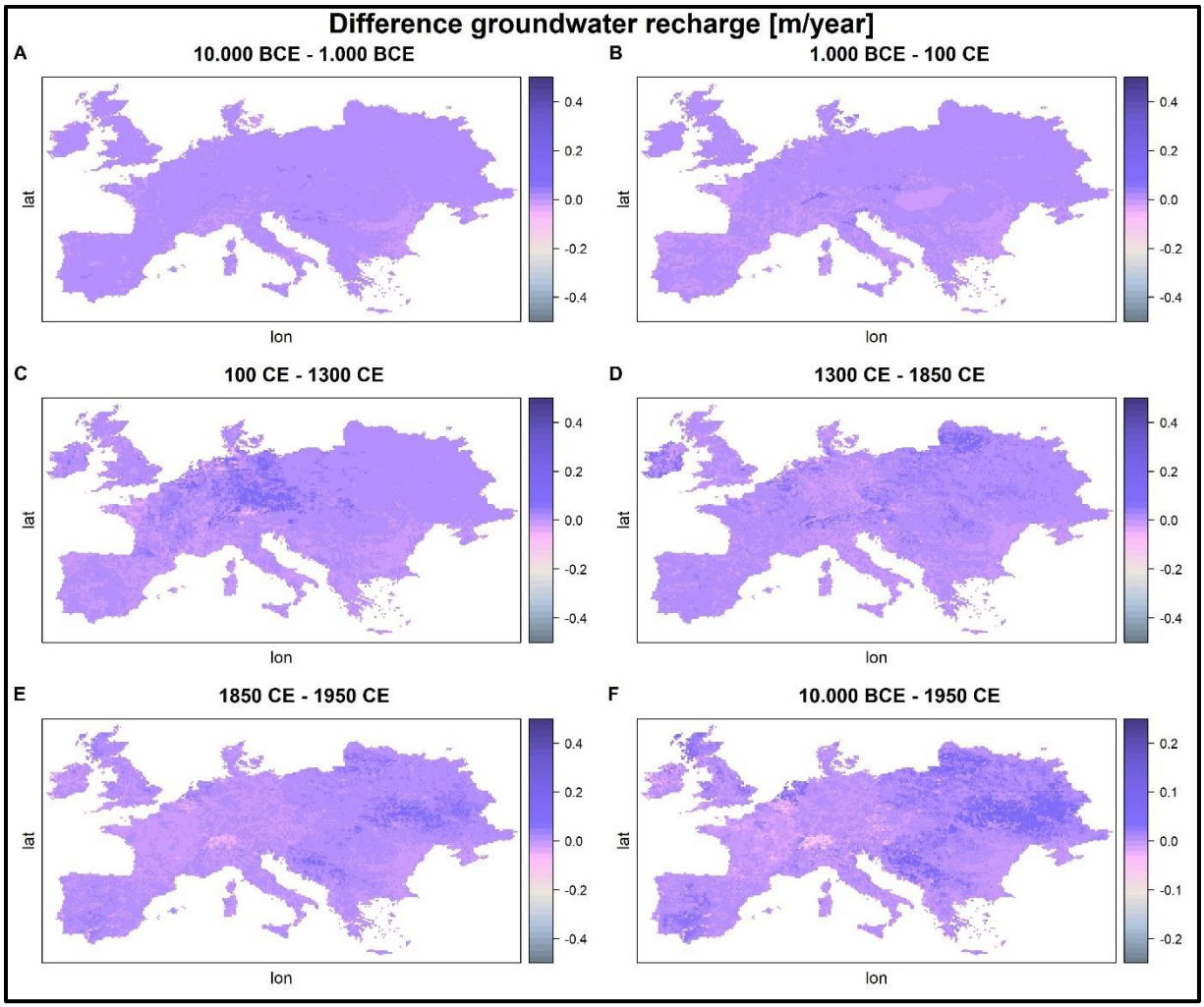

Figure 41, the maps of groundwater recharge that show the differences between A. 10.000 BCE and 1.000 BCE, B. 1.000 BCE and 100 CE, C, 100 CE and 1300 CE, D. 1300 CE and 1850, E. 1850 CE and 1950 CE, and at last F. 10.000 BCE and 1950 CE.

Raster correlation shows a shift from negative correlations in figure 42 A to C to positive correlation is D to F for eastern and central Europe. The very strong negative correlation in eastern Europe confirms the relatively low changes in both LUC and groundwater recharge, compared to other regions. The very strong correlation that then occurs in 1850 and 1950 indicate the high responsiveness to LUC. In central Europe some patches

of high positive correlation are found, but generally the correlations are not as strong in western and central Europe.

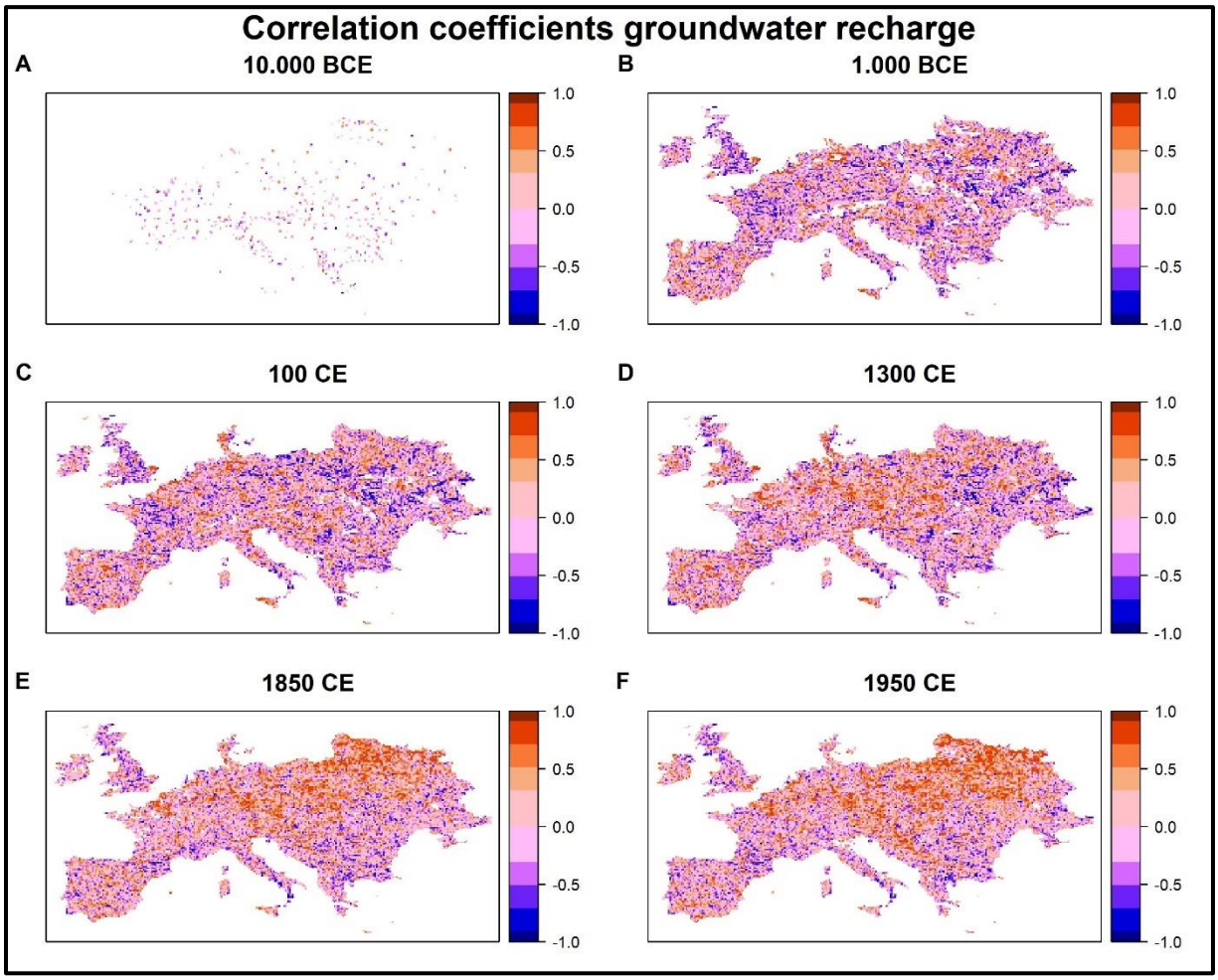

Figure 42, raster correlation with Pearson's correlation coefficients for A. LUC and groundwater recharge at time slice 10.000 BCE, B. at 1.000 BCE, C. 100 CE, D. 1300 CE, E. 1850 CE and F 1950. The color ramp indicates the strength of the correlation coefficient as the 1 - 0.91 is very strong, 0.9 to 0.71 is strong, 0,70 to 0.51 is moderate, and 0.50 to 0.31 is weak, and at last 0.30 to 0.01 very weak. The sign indicates a positive or negative correlation coefficient.

With LUC types for 1850 and 1950 it becomes clear that rainfed crops have a particular very strong positive correlation with groundwater recharge in eastern Europe. The pasture area of the western tip in France shows a strong negative correlation, and overall very strong negative correlation are associated with irrigated crops. The western tip of France, figure 43, was converted from pasture area to rainfed crops, indicating that groundwater recharge does strongly correlates to transformation of pasture area to rainfed crops.

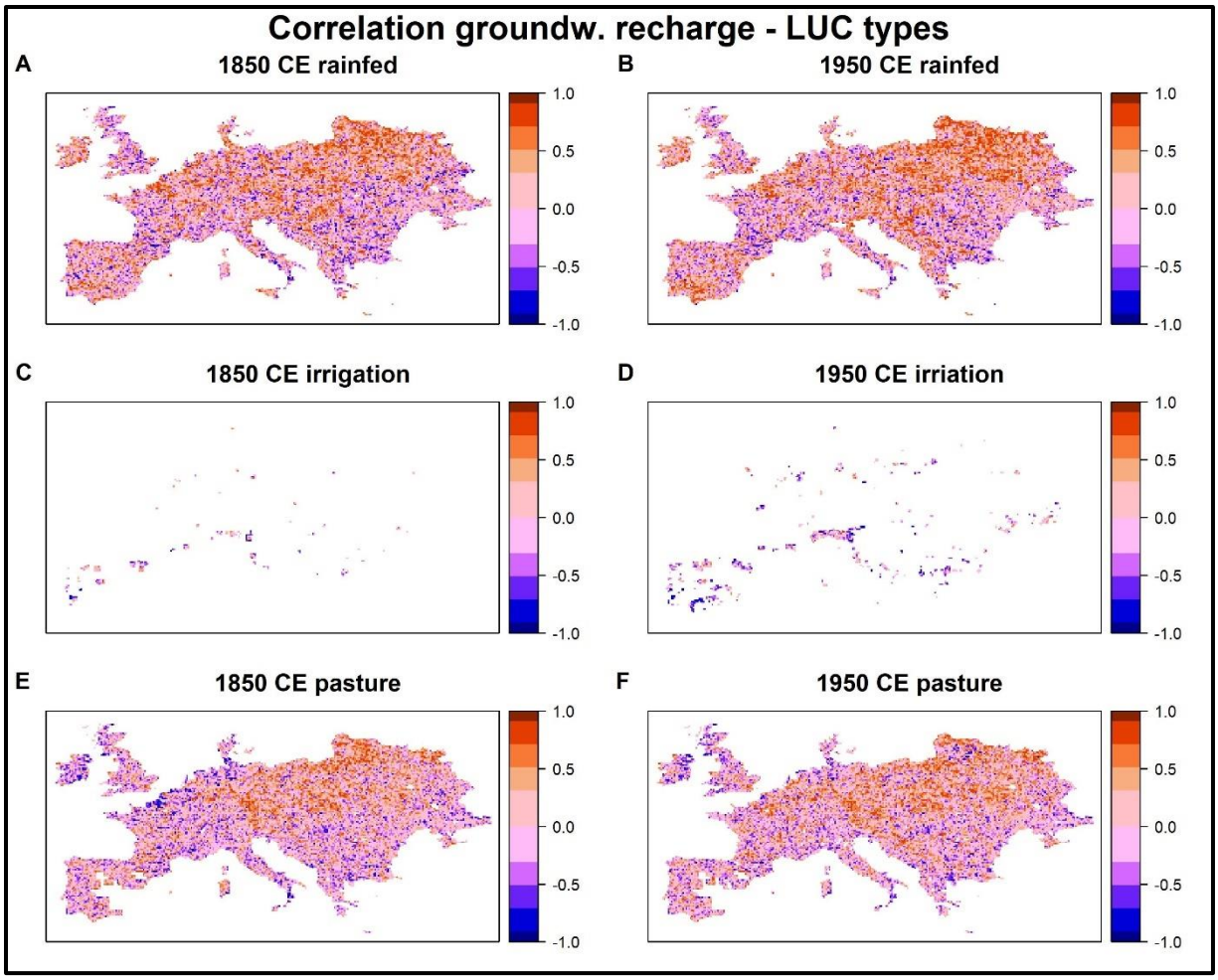

Figure 43, raster correlation with Pearson's correlation coefficients for A. rainfed crops LUC and groundwater recharge at time slice 1850 CE, and B. at 1950 CE. The irrigated crops for C. 1850 CE and D. 1950 CE, in relation to evaporation, for pasture E. at 1850 CE and F. in 1950. The color ramp indicates the strength of the correlation coefficient as the 1 - 0.91 is very strong, 0.9 to 0.71 is strong, 0,70 to 0.51 is moderate, and 0.50 to 0.31 is weak, and at last 0.30 to 0.01 very weak. The sign indicates a positive or negative correlation coefficient.

4.3.4 Evapotranspiration

In figure 44 the variable evapotranspiration and the differences between the time slices can be seen. To see how much of the evapotranspiration is the result of evaporation, the maps were compared. More heavy decreasing, than increasing values were found as the minimum value for 44A was -0.207, and the maximum value 0.087, for 44B the minimum -0.169 and maximum 0.056, for 44C the minimum -0.254 and maximum value 0.191, and 44D the minimum value of -0.218 and maximum value of 0.179 were found. This is not the case in figure 44E, where the minimum value is -0.225 and the maximum is 0.304. Overall the evapotranspiration shows results similar to evaporation, except for the Dniepr river. The evaporation from the reservoir surface is not visible for the time slice 1950. Other than that, no big differences were indicated between evapotranspiration and evaporation, thus further analysis have not been presented.

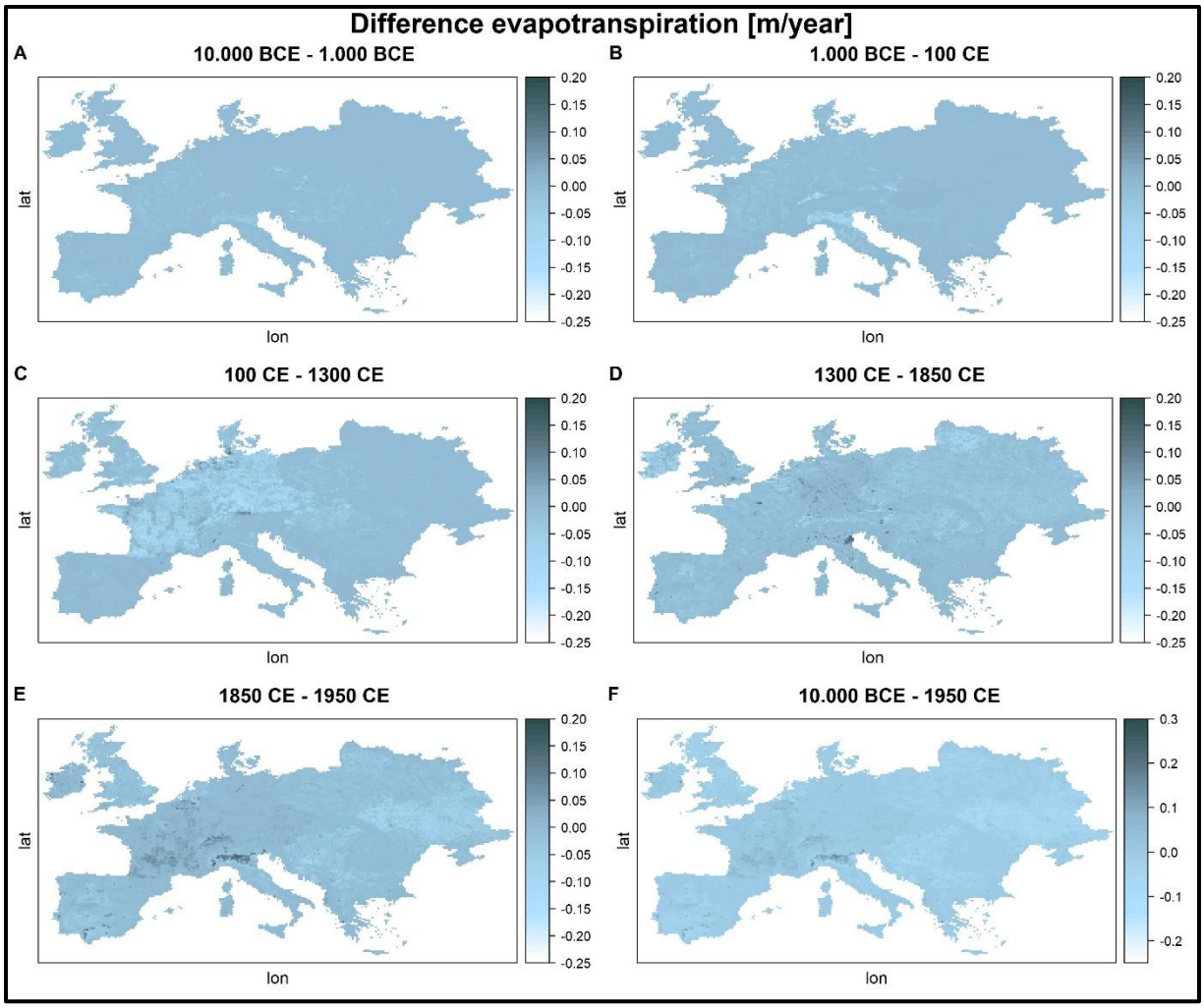

Figure 44, the maps of evapotranspiration that show the differences between A. 10.000 BCE and 1.000 BCE, B. 1.000 BCE and 100 CE, C, 100 CE and 1300 CE, D. 1300 CE and 1850, E. 1850 CE and 1950 CE, and at last F. 10.000 BCE and 1950 CE.

5. Conclusion

From the results can be concluded that the hydrological variables from PCR GLOBWB did respond differently to the LUC among the time slices. Overall the agricultural practices increased over the time slices, with 1950 for western Europe as an exception, as there non-natural vegetation fraction decreases.

The river analysis showed four very different fluxes for the river, but also knew very different LUC pattern. The Dniepr river data points knew a sharp increase in LUC and with the inclusion of reservoirs in the last time slice, showed significant differences. Evaporation significantly increased under LUC, whereas runoff decreased, opposing the hypothesis that LUC would decrease evaporation and increase runoff. This may be explained by the development of reservoirs in the Dniepr river. The discharge significantly differs between 1850 and 1950, and both are significantly different from the preceding time slices.

The Danube river has a gradual increase in LUC over the time slices, and from the results a slight gradual decrease in evaporation is visible, confirming the hypothesis. However only 10.000 BCE and 1950 CE were found to have a significant decrease in evaporation. For runoff between 10.000 BCE and 1950 CE a significant increase was found, confirming the hypothesis.

The Po river showed a gradual decrease up to 1300 CE in evaporation, and increase in runoff and groundwater recharge. But after time slice 4 the evaporation increased, and the runoff and groundwater recharge decreased again to values close to the level of time slice 1 for 1950. In this analysis the first real differences between the lower, best and upper scenario are visible. The significant difference between time slices for the variables highly differ between the estimation scenarios. Furthermore, for the Po river it becomes evident that with the introduction of irrigated crops the hydrological variables return to the value levels of a complete natural vegetation land cover. As irrigation itself, as a water demand, is not included in the model, the extraction of groundwater cannot be seen, but the increase in evaporation and decrease in runoff as a result to the irrigated crops is visible.

The Rhine showed, similar to the Po, a small decrease in the last time slice. The variables groundwater recharge and discharge did not find a significant difference between the time slices, but for evaporation 10.000 BCE, 1.000 BCE and 100 CE significantly differ from 1300 CE, 1850 CE and 1950 CE. The runoff significantly increased between 10.000 BCE and 1950 CE in the lower estimate. Also for the Rhine differences between the estimation scenario were noticed, emphasizing that the amount of LUC can make the difference is a statistical difference between time slices for hydrology.

Overall the cell based study on the effects of LUC on hydrology show that the loss of non-natural vegetation does impact the variables of the water balance. Most importantly, the type of vegetation influences the extent of change in hydrological properties. With the increase of rainfed crops, the response in hydrological values in evaporation, runoff and groundwater discharge was higher than with pasture. For irrigated crops, the opposite happened, and a decrease in runoff and groundwater recharge and increase in evaporation could be seen. As groundwater is a more slow responding water balance component, the results concerning groundwater recharge in reaction to a land use change are highly uncertain in this study, that is further discussed in the following chapter. Groundwater recharge showed very strong negative correlation is west France when pasture was converted to rainfed crops. This indicates a strong relation between the LUC type pasture and groundwater recharge, which indicates has pasture has better percolating qualities than rainfed crops. As for evaporation, the most decrease was found up to 1950, where an increase was shown with the introduction of more natural vegetation in western Europe. The evaporation variable showed very high values

for irrigated crops in Italy and Iberia, and for the Dniepr river reservoirs. An almost equal amount of increase or decrease was seen in evaporation and evapotranspiration, apart from the time slices where irrigated crops were introduced in Italy and the river Dniepr reservoirs. Lastly, runoff was found to be correlating more strongly negatively with the decrease in pasture and thus an increase in rainfed crop in regions such as the Biritsh isles, western Europe and the Dniepr area. Runoff was found positively correlated to rainfed crops in north eastern Europe. Overall where evaporation was found mostly with very strong negative correlations with LUC, runoff was found with strong positive correlations, and vice versa.

6. Discussion

To discuss the findings of this study, firstly the internal assumptions and flaws must be considered. There after the results are examined in relation to the literature. At last some points of attention are given for follow-up research.

In the preparation of the model, there was no distinction between natural vegetation types, according to MIRCA, as planned. From the literature it became clear that the type of vegetation, with for example differing root systems or leaf area index, give different responses in term of evaporation, evapotranspiration, runoff and groundwater recharge. In this research an equal distinction of tall and natural vegetation was made, which were set to a similar set of parameters in terms of root depth, kc and more parameters that can be found in appendix section 1. Also, the equal division of natural vegetation between short and tall vegetation in each grid cell is not representative for the various biomes of Europe. Furthermore, mistakenly a climate input file was used in the model, that lead to only 25 years of valuable data to be analyzed. With observations for only 25 years, the credibility of climate analysis is severely jeopardized. The general rule of thumb is having a minimum years of 30 climatic years, and seen the many outliers that the model found the results would have benefitted from the initial aim of 100 climate years. With more climatic cycles, extreme droughts or wet periods would have not been considered outliers, and the data would have been less skewed. Especially the variable groundwater recharge showed ambiguous results, and might have shown more conclusive results for longer running periods. At last, the human water demand was not included in this model, meaning the irrigation, livestock, industrial and domestic water use was not incorporated. This on the one hand gave the ability to solely identify the response of hydrological properties to LUC. On the other hand, human water use, especially looking at irrigated crops, has a great impact on the water balance, that now is not fully incorporated. The variables now showed values that resembled the levels of natural vegetation in 10.000 BCE, and it would be interesting to if this is still the case when irrigation demand is enabled.

Looking critically at the results, the input data must also be evaluated. The LUC fractions from the HYDE 3.2 database are extremely uncertain for the earliest years, hence the use of estimation scenarios. The absence of any natural vegetation in a very large region of western Europe and Ireland for the latest time slices (1300 and 1850) contributes to a lot of change in evaporation and runoff that could clearly be seen for the results. The decrease of non-natural vegetation in 1950 for these areas is explained by Klein Goldewijk et al. (2011) as the result of extreme human population growth that lead to decreased pasture and rainfed crop lands. In that way, HYDE considers the intensification of agriculture, where to the fraction of non-natural vegetation decreases. The theory sections acknowledges the intensification of LUC, rather than land area, but in the model the intensification of LUC has not been accounted for. This could have been accounted for by including the irrigation. Furthermore it must be noted that HYDE does not provide deforestation rates, and thus human activities such as logging and cultivation shifts are not presented. Within this research the land area fraction that was not assigned to non-natural vegetation was considered natural vegetation, but it should be considered that the HYDE data base approaches LUC in the light of agriculture activities, not deforestation.

If this topic would be studied further, the most important aspect to be included are the biomes and crop types. This more accurately presents the local response of the water balance to a specific LUC. Secondly, the inclusion of irrigation, especially when studying the most recent decades, as an indicator of intensification, is a facet to be included to obtain valid results from the water balance. This would especially make the groundwater recharge component in the water balance more interesting as a focal study aspect.

Despite the simplification of the hydrology cycle, and other flaws in this research, the results do show the impact LUC has on hydrology. Strong correlations between LUC and changes in evaporation, runoff, groundwater recharge and on some occasion discharge, indicate that in Europe the hydrological cycle has changed significantly since humans started altering the surface of the earth by transforming natural vegetation to agricultural productive land. Along with many other studies conducted on the simulation of hydrology in response to LUC, that found results jeopardizing ecosystem health, water quantity and quality and desertification (Wijesekara et al., 2014; Teklat et al., 2019; Rojas et al., 2012), this should be an incentive for future studies to include LUC in interactive climate models to better understand the implications it has. When considering water management, vegetation and specifically agricultural practices are a key concept that must be considered when mitigating climate change (Alfieri, 2017; Falloon & Betts, 2010). As the results showed decreasing values as soon as the decrease of fraction of non-natural vegetation occurred, shows potential for the afforestation and diversification to mitigate hydrological impacts of agricultural systems.

Acknowledgements

I want to thank dr. Kees Klein Goldewijk for his supervision, feedback and support during the master thesis process. Also, I want to thank dr. Rens van Beek for the instructions and feedback I received while and after working with PCR GLOBWB.

References

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao, Rome, 300(9), D05109.

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., ... & Feyen, L. (2017). Global projections of river flood risk in a warmer world. Earth's Future, 5(2), 171-182.

Armit, I., Swindles, G. T., Becker, K., Plunkett, G., & Blaauw, M. (2014). Rapid climate change did not cause population collapse at the end of the European Bronze Age. Proceedings of the National Academy of Sciences, 111(48), 17045-17049.

Arneth, A., F. Denton, F. Agus, A. Elbehri, K. Erb, B. Osman Elasha, M. Rahimi, M. Rounsevell, A. Spence, R. Valentini, 2019: Framing and Context. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D.C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press.

Astill, G. G., & Langdon, J. (Eds.). (1997). Medieval farming and technology: The impact of agricultural change in Northwest Europe (Vol. 1). Brill.

Bacci, M. L. (2017). A concise history of world population. John Wiley & Sons.

Bartosiewicz, L. (2013). Animals in Bronze Age Europe. The Oxford handbook of the European Bronze Age. Oxford University Press, Oxford, 328-347.

Bender, B. (1978). Gatherer-hunter to farmer: A social perspective. World archaeology, 10(2), 204-222.

Benito, G., Rico, M., Sánchez-Moya, Y., Sopeña, A., Thorndycraft, V. R., & Barriendos, M. (2010). The impact of late Holocene climatic variability and land use change on the flood hydrology of the Guadalentín River, southeast Spain. Global and Planetary Change, 70(1-4), 53-63.

Betts, R. A. (2006, December). Forcings and feedbacks by land ecosystem changes on climate change. In Journal de Physique IV (Proceedings) (Vol. 139, pp. 119-142). EDP sciences.

Bonan, G. B., Defries, R. S., Coe, M. T., & Ojima, D. S. (2012). Land use and climate. In Land change science (pp. 301-314). Springer, Dordrecht.

Bork, H. R., & Lang, A. (2003). Quantification of past soil erosion and land use/land cover changes in Germany. In Long term hillslope and fluvial system modelling (pp. 231-239). Springer, Berlin, Heidelberg.

Bork, H.-R., Bork, H., Dalchow, K., Faust, B., Piorr, H.-P., Schatz, T., 1998. Landschaftsentwicklung in Mitteleuropa. Gotha & Stuttgart, Klett-Perthes, 328pp.

Bosmans, J. H., van Beek, L. P., Sutanudjaja, E. H., & Bierkens, M. F. (2017). Hydrological impacts of global land cover change and human water use.

Brauch, H. G., Spring, Ú. O., Bennett, J., & Oswald, S. E. S. (Eds.). (2016). Addressing Global Environmental Challenges from a Peace Ecology Perspective. Springer International Publishing.

Briffa, K. R. (2000). Annual climate variability in the Holocene: interpreting the message of ancient trees. Quaternary Science Reviews, 19(1-5), 87-105.

Brorson, E. (1975). The earliest farming: demography as cause and consequence. In S. Polger (ed.), Population, Ecology and Social Evolution. The Hague: Mouton, pp. 53–78.

Büntgen, U., Tegel, W., Nicolussi, K., McCormick, M., Frank, D., Trouet, V., ... & Luterbacher, J. (2011). 2500 years of European climate variability and human susceptibility. Science, 331(6017), 578-582.

Calder, I. R. (1992, July). The hydrological impact of land-use change. In Proceedings of the conference on priorities for water resources allocation and management, natural resources and engineer advisers conference, Southampton (pp. 91-101).

Calder, I. R. (1998). Water use by forests, limits and controls. Tree physiology, 18(8-9), 625-631.

Carmean, W. H. (1957). The structure of forest soils.

Childe, V. G. (1935). New Light on the Most Ancient East. London: Routledge and Kegan Paul.

Cohen, M. N. (1977). The Food Crises in Prehistory: Overpopulation and the Origins of Agriculture. New Haven: Yale University Press.

Compagnucci, R., da Cunha, L., Hanaki, K., Howe, C., Mailu, G., Shiklomanov, I., Stakhiy, E., 2001. TAR Climate Change 2001: Impacts, Adaptation and Vulnerability. Chapter 4: Hydrology and Water Resources. In press.

Crabtree, P.J., (2010). Agricultural innovation and socio-economic change in early medieval Europe: evidence from Britain and France, World Archaeology, 42:1, 122-136, DOI: 10.1080/00438240903430373

Deevey, E. S. (1960). The human population. Scientific American, 203(3), 195–204.

DeFries, R., & Eshleman, K. N. (2004). Land-use change and hydrologic processes: a major focus for the future. Hydrological processes, 18(11), 2183-2186.

Diamond, J. (1997-a). Guns, Germs, and Steel: The Fates of Human Societies. New York: W. W. Norton, 480 pp.

Diamond, J. (1997-b). Location, location, location: the first farmers. Science, 278(5341), 1243-1244.

Dickinson, R. E. (1991). Global change and terrestrial hydrology—a review. Tellus A: Dynamic Meteorology and Oceanography, 43(4), 176-181.

Doughty, C. E., Wolf, A., & Malhi, Y. (2013). The legacy of the Pleistocene megafauna extinctions on nutrient availability in Amazonia. Nature Geoscience, 6(9), 761-764.

Ellis, E. C., Kaplan, J. O., Fuller, D. Q., Vavrus, S., Goldewijk, K. K., & Verburg, P. H. (2013). Used planet: A global history. Proceedings of the National Academy of Sciences, 110(20), 7978-7985.

Erdkamp, P. (1999). Agriculture, underemployment, and the cost of rural labour in the Roman world. The Classical Quarterly, 49(2), 556-572.

Falloon, P., & Betts, R. (2010). Climate impacts on European agriculture and water management in the context of adaptation and mitigation—the importance of an integrated approach. Science of the total environment, 408(23), 5667-5687.

Gammage, B., & Gammage, B. (2011). The biggest estate on earth: how Aborigines made Australia (pp. 1-4). Sydney: Allen & Unwin.

Garg, V., Aggarwal, S. P., Gupta, P. K., Nikam, B. R., Thakur, P. K., Srivastav, S. K., & Kumar, A. S. (2017). Assessment of land use land cover change impact on hydrological regime of a basin. Environmental Earth Sciences, 76(18), 635.

Garnsey, P., & Saller, R. (2014). The Roman Empire: economy, society and culture. Univ of California Press.

Geraghty, R. M. (2007). The Impact of Globalization in the Roman Empire, 200 bc—ad 100. The Journal of Economic History, 67(4), 1036-1061.

Gottfried, R. S. (2010). Black death. Simon and Schuster.

Govindasamy, B., Duffy, P. B., & Caldeira, K. (2001). Land use changes and Northern Hemisphere cooling. Geophysical Research Letters, 28(2), 291-294.

Guo, L. B., & Gifford, R. M. (2002). Soil carbon stocks and land use change: a meta-analysis. Global change biology, 8(4), 345-360.

Hall, J., Arheimer, B., Aronica, G. T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Burlando, P., Castellarin, A., Chirico, G. B., Claps, P., Fiala, K., Gaál, L., Gorbachova, L., Gül, A., Hannaford, J., Kiss, A., Kjeldsen, T., Kohnová, S., Koskela, J. J., Macdonald, N., Mavrova-Guirguinova, M., Ledvinka, O., Mediero, L., Merz, B., Merz, R., Molnar, P., Montanari, A., Osuch, M., Parajka, J., Perdigão, R. A. P., Radevski, I., Renard, B., Rogger, M., Salinas, J. L., Sauquet, E., Šraj, M., Szolgay, J., Viglione, A., Volpi, E., Wilson, D., Zaimi, K., and Blöschl, G.: A European Flood Database: facilitating comprehensive flood research beyond administrative boundaries, Proc. IAHS, 370, 89–95, https://doi.org/10.5194/piahs-370-89-2015, 2015.

Harbor, J. M. (1994). A practical method for estimating the impact of land-use change on surface runoff, groundwater recharge and wetland hydrology. Journal of the American Planning Association, 60(1), 95-108.

Harris, J. A., Hobbs, R. J., Higgs, E., & Aronson, J. (2006). Ecological restoration and global climate change. Restoration Ecology, 14(2), 170-176.

Houghton, R. A. (1994). The worldwide extent of land-use change. BioScience, 44(5), 305-313.

Houghton, R. A., & Goodale, C. L. (2004). Effects of land-use change on the carbon balance of terrestrial ecosystems. Ecosystems and land use change, 153, 85-98.

Huck, S. W. (2015). Statistical misconceptions: Classic edition. Routledge. Taylor & Francis

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: review and synthesis. Journal of Hydrology, 319(1-4), 83-95.

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: review and synthesis. Journal of Hydrology, 319(1-4), 83-95.

IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Jacob, J., Disnar, J. R., Arnaud, F., Gauthier, E., Billaud, Y., Chapron, E., & Bardoux, G. (2009). Impacts of new agricultural practices on soil erosion during the Bronze Age in the French Prealps. The Holocene, 19(2), 241-249.

Jenny, J. P., Koirala, S., Gregory-Eaves, I., Francus, P., Niemann, C., Ahrens, B., ... & Zolitschka, B. (2019). Human and climate global-scale imprint on sediment transfer during the Holocene. Proceedings of the National Academy of Sciences, 116(46), 22972-22976.

Klein Goldewijk, K., & Verburg, P. H. (2013). Uncertainties in global-scale reconstructions of historical land use: an illustration using the HYDE data set. Landscape ecology, 28(5), 861-877.

Klein Goldewijk, K., Beusen, A., & Janssen, P. (2010). Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1. The Holocene, 20(4), 565-573.

Klein Goldewijk, K., Beusen, A., van Drecht, G., & de Vos, M. (2011) The HYDE 3.1 spatially explicit database of human induced global land use change over the past 12,000 years, Global Ecology & Biogeography, 20(1): 73-86.

Klein Goldewijk, K., Beusen, A., Van Drecht, G., & De Vos, M. (2011). The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. Global Ecology and Biogeography, 20(1), 73-86.

Kling, G. W., Adams, H. E., Bettez, N. D., Bowden, W. B., Crump, B. C., Giblin, A. E., ... & Shaver, G. R. (2014). Land–water interactions. Alaska's changing arctic: Ecological consequences for tundra, streams, and lakes, 143-172.

Kron, G. (2013). Agriculture, Roman Empire. The Encyclopedia of Ancient History.

Levis, C., Costa, F. R., Bongers, F., Peña-Claros, M., Clement, C. R., Junqueira, A. B., ... & Castilho, C. V. (2017). Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. Science, 355(6328), 925-931.

Li, K. Y., Coe, M. T., Ramankutty, N., & De Jong, R. (2007). Modeling the hydrological impact of land-use change in West Africa. Journal of hydrology, 337(3-4), 258-268.

Liu, J., Zhang, Q., Singh, V. P., Song, C., Zhang, Y., Sun, P., & Gu, X. (2018). Hydrological effects of climate variability and vegetation dynamics on annual fluvial water balance in global large river basins. Hydrology & Earth System Sciences, 22(7).

Mahoney, E., & Nardo, D. (2016). The Black Death: Bubonic Plague Attacks Europe. Greenhaven Publishing LLC.

Marsicek, J., Shuman, B.N, Bartlein, P.J, Shafer, S.L. & Brewer, S., 2018. Reconciling divergent trends and millennial variations in Holocene temperatures. Nature, 554, 92-96. doi: 10.1038/nature25464

Martin, P. S. (1967). Prehistoric overkill. In P. S. Martin and H. E. Wright (eds) Pleistocene Extinctions. New Haven: Yale University Press, pp. 75–120.

Martín-Retortillo, M., & Pinilla, V. (2015). Patterns and causes of the growth of European agricultural production, 1950 to 2005. Agricultural History Review, 63(1), 132-159.

McConnell, D. J. (1992). The forest-garden farms of Kandy, Sri Lanka (No. 3). Food & Agriculture Org.

McCormick, M., Büntgen, U., Cane, M. A., Cook, E. R., Harper, K., Huybers, P., ... & Nicolussi, K. (2012). Climate change during and after the Roman Empire: reconstructing the past from scientific and historical evidence. Journal of Interdisciplinary History, 43(2), 169-220.

McDermott, F., Frisia, S., Huang, Y., Longinelli, A., Spiro, B., Heaton, T. H., ... & van der Borg, K. (1999). Holocene climate variability in Europe: evidence from δ18O, textural and extension-rate variations in three speleothems. Quaternary Science Reviews, 18(8-9), 1021-1038.

National Oceanic and Atmospheric Administration (NOAA), sd. Global Historical Climatology Network Monthly - Version 4. Retrieved from https://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/global-historical-climatology-network-monthly-version-4

Neris, J., Jiménez, C., Fuentes, J., Morillas, G., & Tejedor, M. (2012). Vegetation and land-use effects on soil properties and water infiltration of Andisols in Tenerife (Canary Islands, Spain). Catena, 98, 55-62.

Notebaert, B., Verstraeten, G., Ward, P., Renssen, H., & Van Rompaey, A. (2011). Modeling the sensitivity of sediment and water runoff dynamics to Holocene climate and land use changes at the catchment scale. Geomorphology, 126(1-2), 18-31.

Nurtaev, B. S., & Nurtaev, L. (2016). Long Term Trends in Climate Variability of Caucasus Region. JOURNAL OF THE GEORGIAN GEOPHYSICAL SOCIETY, 19(22).

Paul, B. K., & Rashid, H. (2017). Climatic Hazards in Coastal Bangladesh.

Pearson, K. L. (1997). Nutrition and the early-medieval diet. Speculum, 72(1), 1-32.

Pielke, R. A. (2005). Land use and climate change. Science, 310(5754), 1625-1626.

Ray, D. K., Gerber, J. S., MacDonald, G. K., & West, P. C. (2015). Climate variation explains a third of global crop yield variability. Nature communications, 6(1), 1-9.

Roberts, N. (1989). The Holocene: An Environmental History. Oxford: Basil Blackwell.

Rojas, R., Feyen, L., Bianchi, A., & Dosio, A. (2012). Assessment of future flood hazard in Europe using a large ensemble of bias-corrected regional climate simulations. Journal of Geophysical Research: Atmospheres, 117(D17).

Roosevelt, A. C. (2013). The Amazon and the Anthropocene: 13,000 years of human influence in a tropical rainforest. Anthropocene, 4, 69-87.

Rösch, M. (2008). New aspects of agriculture and diet of the early medieval period in central Europe: waterlogged plant material from sites in south-western Germany. Vegetation history and archaeobotany, 17(1), 225-238.

Rösch, M. (2013). Land use and food production in Central Europe from the Neolithic to the medieval period: Change of landscape, soils and agricultural systems according to archaeobotanical data. Economic Archaeology: From Structure to Performance in European Archaeology, 237, 109-128.

Ruddiman, W. F. (2013). The anthropocene. Annual Review of Earth and Planetary Sciences, 41, 45-68.

Russo, T. A., & Lall, U. (2017). Depletion and response of deep groundwater to climate-induced pumping variability. Nature Geoscience, 10(2), 105-108.

Shukla, P.R., Skea, J., Slade, R., van Diemen, R., Haughey, E., Malley, J., Pathak, M., Portugal Pereira, J., (eds.) Technical Summary, 2019. In: Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M, Belkacemi, J. Malley, (eds.)]. In press.

Smith P., M. Bustamante, H. Ahammad, H. Clark, H. Dong, E.A. Elsiddig, H. Haberl, R. Harper, J. House, M. Jafari, O. Masera, C. Mbow, N.H. Ravindranath, C.W. Rice, C. Robledo Abad, A. Romanovskaya, F. Sperling, and F. Tubiello, 2014: Agriculture, Forestry and Other Land Use (AFOLU). In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Sutanudjaja, E. H., Van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H., Drost, N., ... & Karssenberg, D. (2018). PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geoscientific Model Development, 11(6), 2429-2453.

Teklay, A., Dile, Y. T., Setegn, S. G., Demissie, S. S., & Asfaw, D. H. (2019). Evaluation of static and dynamic land use data for watershed hydrologic process simulation: A case study in Gummara watershed, Ethiopia. Catena, 172, 65-75.

Thompson, F. M. L. (1968). The second agricultural revolution, 1815-1880. The Economic History Review, 21(1), 62-77.

Thompson, S. E., Harman, C. J., Heine, P., & Katul, G. G. (2010). Vegetation-infiltration relationships across climatic and soil type gradients. Journal of Geophysical Research: Biogeosciences, 115(G2).

Trenberth, K. E., Smith, L., Qian, T., Dai, A., & Fasullo, J. (2007). Estimates of the global water budget and its annual cycle using observational and model data. Journal of Hydrometeorology, 8(4), 758-769.

Twine, T. E., Kucharik, C. J., & Foley, J. A. (2004). Effects of land cover change on the energy and water balance of the Mississippi River basin. Journal of Hydrometeorology, 5(4), 640-655.

United Nation University, sd. Evapotranspiration in forest and fields. Retrieved from http://archive.unu.edu/unupress/unupbooks/80635e/80635E0d.htm

United Nations of Roma Victrix (UNRV), sd-a. *Economy*. Retrieved from https://www.unrv.com/economy.php

United Nations of Roma Victrix (UNRV), sd-b. *Roman Aqueducts*. Retrieved from https://www.unrv.com/culture/roman-aqueducts.php

Van Beek, L.P.H. and M.F.P. Bierkens (2008), The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Report Department of Physical Geography, Utrecht University, Utrecht, The Netherlands, http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf

Veldkamp, A., & Lambin, E. F. (2001). Predicting land-use change. Elsevier: Agriculture, Ecosystems & Environment. Vol. 85 (1–3), 1-6.

Vörösmarty, C. J., & Sahagian, D. (2000). Anthropogenic disturbance of the terrestrial water cycle. Bioscience, 50(9), 753-765.

Vörösmarty, C., Hinzman, L., Peterson, B., Bromwich, D., Hamilton, L., Morison, J., ... & Webb, R. (2001). The Hydrologic Cycle and its Role in Arctic and Global Environmental Change.

W. W. Verstraeten, B. Muys, J. Feyen, F. Veroustraete, M. Minnaert, et al.. Comparative analysis of the actual evapotranspiration of Flemish forest and cropland, using the soil water balance model WAVE. Hydrology and Earth System Sciences Discussions, European Geosciences Union, 2005, 9 (3), pp.225-241. ffhal-00304821f

Weisdorf, J. L. (2005). From foraging to farming: explaining the Neolithic Revolution. Journal of Economic surveys, 19(4), 561-586.

Weisdorf, J. L. (2005). From foraging to farming: explaining the Neolithic Revolution. Journal of Economic surveys, 19(4), 561-586.

Wijesekara, G. N., Farjad, B., Gupta, A., Qiao, Y., Delaney, P., & Marceau, D. J. (2014). A comprehensive land-use/hydrological modeling system for scenario simulations in the Elbow River watershed, Alberta, Canada. Environmental management, 53(2), 357-381.

Wilkinson, G. E., & Aina, P. O. (1976). Infiltration of water into two Nigerian soils under secondary forest and subsequent arable cropping. Geoderma, 15(1), 51-59.

Willi, A., & Kolb, A. (2014). Land division and water management in the west of the Roman Empire.

Wirtz, K. W., & Lemmen, C. (2003). A global dynamic model for the Neolithic transition. Climatic Change, 59(3), 333-367.

Wood, P. J., Hannah, D. M., & Sadler, J. P. (Eds.). (2008). Hydroecology and ecohydrology: past, present and future. John Wiley & Sons.

Zhang, L., Dawes, W. R., & Walker, G. R. (2001). Response of mean annual evapotranspiration to vegetation changes at catchment scale. Water resources research, 37(3), 701-708.