Universität für Bodenkultur Wien
University of Natural Resources
and Applied Life Sciences, Vienna

Lincoln
University
Te Whare Wānaka o Aoraki
CHRISTCHURCH • NEW ZEALAND

- Master's Thesis-

"Principal Component Analysis on Thermal Infrared Time-Series of two pre-Alpine catchments in Lower Austria, Ybbs and Traisen"

"Hauptkomponentenanalyse einer Zeitreihe thermischer Infrarotbilder für zwei voralpine Wassereinzugsgebiete in Niederösterreich, Ybbs und Traisen"

Claude Meisch

May 2015

University of Natural Resources and Life Sciences, Vienna
Department of Water – Atmosphere – Environment
Institute of Water Management, Hydrology and Hydraulic Engineering

Lincoln University Christchurch, New Zealand
Centre for Advanced Computational Solutions

First Supervisor: Univ.Prof. Dr. Karsten Schulz
Second Supervisor: Prof. Dr. Don Kulasiri
Assistant Supervisor: Benjamin Müller (M.Sc.)

Submitted by: Claude Meisch
claudemeisch@yahoo.com
Lobenhauerngasse 3/11
1170 Vienna, AUSTRIA
Matriculation no.: 1241210

Natural Resources Management and Ecological Engineering (Dipl.-Ing.)

Date of submission: 05 May 2015

Table of Contents

List of Figures

List of Tables

Abstract

Distributed hydrological models provide important contributions in understanding human impacts on water resources by analysing spatially variable hydrological behaviour. However parameterization of spatially distributed properties related to physical/hydrological behaviour exhibits some major difficulties in the models. A potential solution is the identification of heterogeneity of water related physical processes by displaying variances of thermal properties of surface and subsurface characteristics. Following the approach developed by Müller et al. (2014), Principal Component Analysis (PCA) has been applied to a time-series of Thermal Infrared (TIR) remote sensing data of the Ybbs and Traisen catchment. The spatiotemporal distributions of thermal properties with regard to the water/energy balance could be extracted within principal components. The first three principal components explain a cumulative proportion of 89% resp. 86% of the total variance inherent in the time-series of 12 resp. 11 Landsat 5 TM TIR-images. The remaining variance was attributed to transient effects, such as background noise or atmospheric disturbances. The relation between patterns of thermal variability in the first components and the most dominant landscape elements, exhibiting controls on thermal properties, is drawn. It was further assessed how sensitive the principal components are to the quality of the dataset. Spatial and temporal dimensions as well as shadowing effects affecting mountainous environment are discussed. PCA on TIR time-series was found useful in the deduction of patterns representing the variability of thermal properties at catchment scale. PCA represents a rigorous method in the deduction of physical meaningful hydrological patterns in the establishment of credible parameters.

Kurzfassung

Diskretisierte hydrologische Modelle liefern wichtige Beiträge für das Verständnis menschlicher Einflüsse auf die Wasserressourcen durch die Analyse von räumlich verteiltem Verhalten hydrologischer Variablen. Jedoch ist die Parametrisierung von räumlich verteilten physikalischen und hydrologischen Eigenschaften eine der größten Herausforderung in der Entwickelung von solchen Modellen. Eine mögliche Lösung wird in der Identifikation der heterogenen Verteilung der wasserbezogenen physikalischen Prozesse gesehen, in dem Varianzen der thermischen Eigenschaften von ober- und unterirdischen Merkmalen im Wassereinzugsgebiet dargestellt werden. Im Anschluss an den von Müller et al. (2014) entwickelten Ansatz, wird eine Zeitreihe von TIR-Fernerkundungsdaten mittels Hauptkomponentenanalyse für die Einzugsgebiete der Ybbs und Traisen untersucht. Es wird gezeigt, dass die raumzeitliche Verteilung der thermischen Eigenschaften mit Bezug auf den Wasser- und Energiehaushalt mittels Hauptkomponenten-analyse extrahiert werden kann. Die ersten drei Hauptkomponenten erklären einen kumulativen Anteil von 89% bzw. 86% der Gesamtvarianz der thermischen Eigenschaften in den 12 bzw. 11 Landsat 5 TM TIR-Bildern. Die verbleibende Differenz wurde vorübergehenden Effekten, wie Hintergrundrauschen oder atmosphärischen Störungen zugeschrieben. Den Zusammenhang zwischen den hervorgebrachten Mustern thermischer Variabilität und den dominierenden Landschaftselementen wird aufgezeigt. Zudem wurde die Empfindlichkeit der Methode gegenüber der Qualität des Datensatzes getestet. Die räumliche und zeitliche Verteilung der thermischen Ausprägung im Wassereinzugsgebiet sowie Schatten-effekte in den bergigen Regionen wird diskutiert. Die Hauptkomponentenanalyse stellte sich bei der Ableitung der Muster, welche die Verteilung direkter interner thermischer Eigenschaften darstellen, als hilfreich heraus. Hauptkomponentenanalyse von TIR-Bilder stellt eine gründliche Methode dar, welche durch die Ableitung von physikalisch aussagekräftigen Mustern, die Glaubwürdigkeit hydrologischer Parameter erhöht.

1 Introduction

It is a well-accepted fact among scientists that human activities influence the hydrological cycle, for instance by altering the chemical constitution and as such the radiative properties of the atmosphere or by degrading environmental systems through pollution and land use transformations (Beighley et al., 2005; Kirchner, 2006). However the science of hydrology was long characterized by the development of predictive lumped models focussing on the reproduction of real events without the implementation of intrinsic functioning of hydrological processes. Since then, hydrological modelling has been evolved towards a more descriptive approach by developing distributed models which now also consider the heterogeneity of hydrological processes (Kirchner, 2006). Models are hence no longer meant to only produce knowledge about quantity and quality of the water but also about the spatial occurrence of environmental degradation and how to tackle these problems (Refsgaard, 1997). However, catchments are complex, dynamic and heterogeneous systems and the definition of hydrological processes represents one of the most difficult tasks in the derivation and evaluation of human impacts on the environment (Beven, 2001). Spatial variability in particular is a characteristic and intrinsic feature of natural hydrological systems and "*the characteristics of that variability often have a substantial influence on the behaviour of the system"* (Grayson et al., 1997). Hence, knowledge about interior water fluxes and state variables becomes the key aspect in integrated water resources management as the impetus for "*getting the right answers for the right reasons*" (Kirchner, 2006) is more than ever of actuality.

Due to the complexity of these heterogenic systems, there exist substantial difficulties in the observation and the implementation of terrain data into distributed models. In order to describe the hydrological processes as accurately as possible, distributed model parameters are defined for every grid cell (Samaniego et al., 2010). These parameters act as indicators that can be inferred from measured data but are often optimised through calibration (Chang, 2012). While point measurements provide hydrological information for the parameterization at small scales, substantial problems arise when describing spatial variability at larger scales (Abbott & Refsgaard, 1996). When model parameters are estimated for a distributed model over a large catchment area, the ratio between unknown model parameters and the resolution of the model runs up. Hence, the linked computational effort grows considerably with a finer resolution and, as a consequence, the parameters risk to become ambiguous. This ambiguity effect during the parameterization process is called over-parameterization. Thus, the estimation of parameters, which calibrate the model to match the observed behaviour, one of the main challenges in scientific hydrological research.

Consequently there have been several calls for the collection of new data in hydrological science allowing for a better hydrological understanding and an improvement in the predictive capability of hydrological models (Kirchner, 2006; Grayson et al., 2002; Beven, 2001; Rosso, 1994; Beven et al., 1993).

Above all, the advancement of remote sensing data has brought substantial changes to Hydrological science. One of the most important aspects of remote sensing, is the capability of measuring information on a wide range of spatial scales, as opposed to point information (Rango & Shalaby, 1998). Such datasets provide information about the spatial and temporal variability of landscape elements that influence the hydrological processes (Grayson et al., 2002). The direct deduction of interior physical values is still in the early stages of development. Several approaches can be found in the literature, however, these are mostly published with respect to satellite sensors and bandwidths (Alcântara, 2013; Anderson et al., 2008; Sánchez et al., 2008). Nonetheless it is possible to deduce information about the spatial and temporal variability of controlling hydrological variables and conditions with remote sensing (Refsgaard, 1997). By putting the focus on the characterization of patterns rather than on the direct intrinsic physical properties, the spatial variability can be accounted for in hydrological models and the coupled parameterization process can be improved (Grayson et al., 2002).

By implementing Thermal InfraRed (TIR) remote sensing in studies of landscape functioning there is the potential to "*observe the states and dynamics of energy fluxes across and between landscapes, from patch to the regional levels*" (Quattrochi & Luvall, 1999). As TIR data is based on the physical principle of emitted radiation to determine thermal characteristics of the earth's surface the derivation of patterns can be related to hydrological process heterogeneity with regard to the energy balance; and as such it can be used as surrogate for parameterization. Quattrochi et al. (1999) explain that knowledge about "*energy balance characteristics (i.e., fluxes and redistribution of thermal energy within and across landscape elements) is an implicit and important aspect of landscape dynamics and landscape functioning".* Thus, thermal infrared data provide means of the distribution of all landscape elements influencing energy balance characteristics. The extraction of thermal energy variation leads to an understanding of how these complex ecosystem state variables are distributed and connected across a catchment. Risser et al (1984) have shown that the understanding of flows and transfers among spatial components are especially important in understanding the interlinkages of these landscape elements. Thus, TIR data needs to be processed into surrogate patterns, in order to be further implemented in the parameterization process of hydrological models accounting for the heterogeneity of landscape characteristics and their interlinkage.

One approach was shown by Müller et al. (2014), who extracted patterns of thermal dynamics with regard to internal energy fluxes by analysing time-series of TIR images with a principal component analysis. These patterns of energy balance variability imply information about the water balance characteristics related to landscape functioning and their contribution to the thermal energy response (e.g. latent energy fluxes controlled by soil moisture content). Consequently, information about the connections and relations between the landscape elements is inherent in the nominal principal component values producing the patterns of thermal variability. Through a further analysis they were able to relate the patterns to major landscape characteristics. Allowing them to extract continuous pixel based information of distinct hydrological behaviour with regard to the energy balance. These can further be classified into functional units or may serve as input in the parameterization process. Hence the approach should ultimately lead to an improvement of the credibility of hydrological models by providing a rigorous methodological framework for establishing parameters.

Therefore this framework will be applied in the scope of this thesis. The performance of the methodology will be tested to a wider range of catchment characteristics. A time-series of TIR Landsat 5 images of two pre-Alpine catchments, will be analysed using principal component analysis (PCA). The focus will be laid on the more pronounced topographical characteristic, which allows for an evaluation of the applicability of PCA in mountainous regions. The outputs will be analysed to assess the dominant landscape elements causing these patterns of spatial temperature variability. Before, an analysis of the differences between the catchments will help establishing a scientific statement on the inherent constraints.

1.1 Problem Definition and Relevance

Over-parameterization means that the optimization and calibration process of the hydrological model evolves into such a high computational task, that the values of the results of the estimation software become numerically unrealistic. In the worst case the values don't converge at all, which stops the parameterization process (Doherty, 2003). This may further evolve into parameters that can no longer be identified. This situation is called equifinality (Beven, 2001). This means that the set of parameters is decisive for the models performance, rather than the singular parameter being process defining at the cell. This results in different models displaying equally acceptable representations of hydrological behaviour due to high parameterization which comprises the predictive validity of a model (Kumar, 2009). To this fact the distributed hydrological models risk to become less reliant in producing new scientific insights on the intrinsic functioning of the water-balance system.

One solution to this problem is to couple multiple variables controlling hydrological behaviour into several homogenous units and reduce the spatial complexity to groups of cells having analogical controlling variables, so-called Hydrological Response Units (HRU's) (Flügel, 1995). Thus, they reduce the complexity in the parameterization process by assigning sets of parameters to spatially condensed units of identical hydrological behaviour (Flügel, 1995). Consequently, each HRU gets attributed one parameter independent of its location in the spatial domain of the model, equal HRU's are assigned the same parameter values. The difficulty arises from HRU's being defined as categorical classifications with explicit geographic delineations, this results in losing information about the interconnections in the parameters that should in reality be bound by local attributes (Doherty, 2003). Examples are the difficulties in estimating realistic variables displaying spatiotemporal variations for non-point source pollution or change within the catchment. Consequently, the need emerges to reduce complexity in the deduction of spatial heterogeneity, all by filling knowledge gaps of interconnectivity (Schulz et al., 2006).

Remotely sensed imagery might become a source of distributed data serving as surrogate patterns for the above mentioned issue of over-parameterization (Beven, 2012). Up to date, the main purpose of remotely sensed data in hydrological models is the estimation of precipitation, land cover types and vegetation parameters, or the derivation of information allowing vegetation and soil classification (Beven, 2012). Not all the capacities within the available data are explored due to methodological obstacles, as for example unknown meteorological conditions with regard to atmospheric correction and incoherence's in the datasets. Thus there is a lot of potential knowledge waiting to be exhausted from the available data (Clark, 2014).

One of these datasets waiting for full exploration is TIR remote sensing. Even though TIR data are widely available, the time-series analysis of TIR data for example, is still a novel field, due to the low availability of temporally coherent datasets. An approach was introduced by Müller et al. (2014), who analyse TIR time series with principal components. The Hypothesis is that when orthogonally transforming a time-series of TIR data into principal components it becomes possible to extract dominant patterns of thermal variability that can be related to controls on sensible and latent energy fluxes, such as texture, vegetation and elevation. Thus these patterns might reveal interesting heterogenic structures of hydrological behaviour. As such PCA enables the deduction of distributions of physically-based nominal surrogate patterns on behalf of which hydrological behaviour might be numerically inferred. Thus they propose a new approach for the resolution of the over-parameterization problem. However this approach needs more testing under different landscape conditions. This thesis will give a supplementary analysis of TIR time-series for two catchments in Lower Austria, following the approach of Müller et al. (2014), in order to identify potential strengths and weaknesses of the method.

1.2 Research Question and Research Objectives

This thesis presents the results of a course of research designed to address the above mentioned problems by answering this question:

> What knowledge and information is inherent in the structures and spatial patterns represented by principal components of TIR time-series for two catchments in Lower Austria?

To address these questions, the following research objectives were considered in this study:

General: Identify the information content and constraints of analysing spatial variability of physical patterns by analysing time-series of remotely sensed TIR images for two catchments.

Specific: Process TIR remotely sensed images with principal component analysis to identify spatial patterns of physical thermal characteristics.

Identify surface and subsurface characteristics of the spatial patterns represented in principal components of time-series of TIR images

Analyse the differences in the information represented in the spatial patterns of the principal components between the catchments

Analyse the influence of the seasonality and quantity of the thermal data on the spatial structures of the principal components.

1.3 Thesis Outline

The thesis comprises a general introduction, a theoretical background and a methods and materials part, before dealing with the applied methodology, the obtained results and the conclusion.

The theoretical background is meant to introduce the reader into the technical context of the thesis and give an overview of the theoretical aspect of the literature review. Thus Chapter 2 deals with the basic principles of remote sensing focussing on the physical properties of TIR remote sensing. This allows screening the strength and weaknesses of TIR-data and finding reasons for the little exploitation of TIR-data in science. In the next step, the thermal dynamics of land surfaces will be exploited before underlining the possibilities and potentials evolving from the extraction of variations of thermal energy fluxes within time-series of TIR-data.

The methods and materials describe the study area, the database and its manipulation and give the reader the necessary mathematical background of the principal component analysis. As such Chapter 3 gives detailed information about the study areas and displays major surface and subsurface characteristics of the catchments. The choice for the database will be elucidated and the further preparation of the data will be illustrated. Then the mathematical approach of the PCA-transform will be clarified.

The results part shows the information drawn from the PCA. Chapter 4 presents the results of the pattern extraction from time-series with principal component analysis displaying thermal variability. The different weights of the images for the different components will be shown before getting to the analysis of the impact of the input data for the pattern extraction. In a further step, the focus will be put on finding the sources explaining the thermal variability.

Chapter 5 is the discussion part, considering the identification of the major surface and subsurface characteristics causing thermal variability and the analysis of the difference between the catchments as well as the impact of the quantity and seasonality of the initial input data on the PCA. The constraints encountered in the processing of the time-series will be illustrated.

In the sixth chapter the conclusion of the thesis will be drawn. Here the significance of the here presented approach in the light of future developments in hydrological modelling will be explained. This will be done by screening some of the possibilities inherent to the method and by giving recommendations for further research.

2 Theoretical Background

2.1 Definition and Principles of Remote Sensing

In general terms, remote sensing is the technique of acquiring information about an object without actually being in contact with it (Jorgensen & Fath, 2008). Accurately speaking, we will understand remote sensing as the capturing information from a distance (rather than in situ) about the Earth's terrestrial and aquatic ecosystems and atmosphere, by sampling reflected and emitted radiation (Kitchin & Thrift, 2008).

Jensen (2007) defines remote sensing as:

> *„the non-contact recording of information from the ultraviolet, visible, infrared, and microwave regions of the electromagnetic spectrum by means of instruments such as cameras, scanners, lasers, linear arrays, and/or area arrays on platforms such as aircraft or spacecraft, and the analysis of acquired information by means of visual and digital image processing".*

The results of remote sensing are images with an abundance of information allowing us to "*determine the composition and nature of the Earth's surface and atmosphere from local to global scales*" (Kitchin & Thrift, 2008). Through the repetitive nature of the remote sensing, these images hold information about changes at different points in time. Thus remote sensing is a powerful tool to assess information that would otherwise be difficult to obtain within the same scale of time and space. Consequently, remote sensing becomes "*a multi-faceted, multi-disciplinary endeavour to acquire information from targets*" (Clark & Rilee, 2010).

As the foundation of remote sensing lies within the measurement of reflected and emitted electromagnetic radiation from the Earth's surface so this radiation is measured and categorized on a logarithmic scale of its wavelength (Jorgensen & Fath, 2008). Remote sensing uses this physical principle of radiation to determine characteristics of surface depending on the emitted or reflected wavelength (Kitchin & Thrift, 2008). These wavelengths are in the following categories: Ultraviolet, Visible, meaning visible to the human eye (VIR). In near-Infrared (NIR), where reflected radiance is measured. In the mid-IR, where reflected and emitted radiance can be measured and in the far-IR, where only emitted energy is measured in the form of thermal energy, here referred to as TIR (Clark & Rilee, 2010). The reason of multispectral remote sensing is that different types of materials can be distinguished on the basis of differences in their spectral signatures (Schowengerdt, 2007). The remote sensors are categorized into active and passive sensors. Active sensors generate their own signal which is sent to the Earth's surface and the reflection is measured, e.g. RADAR or LIDAR. Passive sensors measure emittance and reflection from the Earth's surface of solar energy (Kitchin & Thrift, 2008).

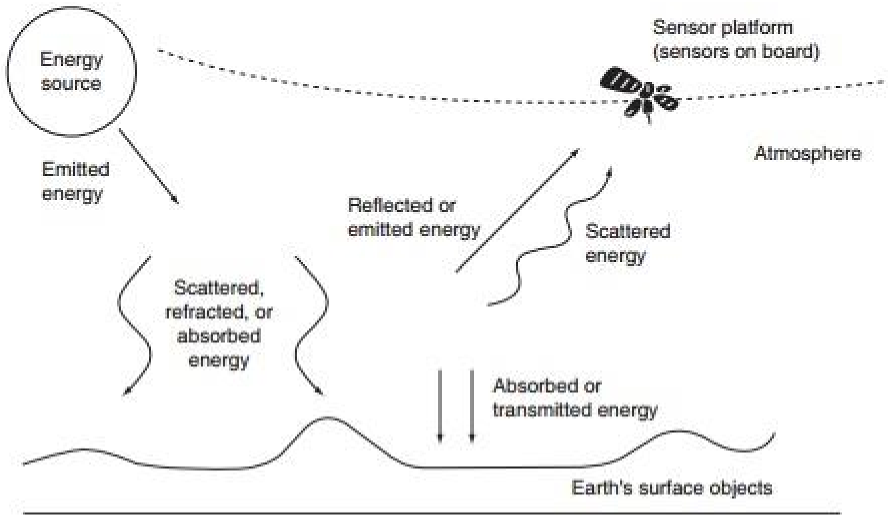

Figure 1: A typical remote-sensing system. Components and Energy Pathways (from Clark & Rilee, 2010)

Figure 1 shows the basic components of a remote sensing systems: an energy/light source, an energy sink (Earth's surface objects) and a sensor platform capturing the reflected and emitted radiance onto a recording medium (Schowengerdt, 2007). With respect to analog arial photography where a lens is focusing the light on photographic film, digital sensors record photons for a specific range of wavelength (Kitchin & Thrift, 2008), and these are stored as digital numbers representing brightness values as an array of pixels.

Several characteristics distinguish the qualities and detail of a remotely sensed image, traditionally referred to as the four types of resolution (Jensen, 2007). First of all, there is spatial resolution, which defines the size of the pixel (smallest discrete picture element) if it were projected onto the earth's surface.

Spectral resolution refers to the range or bandwidth of wavelength that are detected in an image. This also refers to band placement, which defines the portion of electromagnetic wavelength captured for a specific domain, for example 9.5 to 12 µm for TIR.

Radiometric resolution refers to the bandwidth of potential brightness values that can be stored in a pixel. Sensors with higher radiometric resolution capture more distinct differences in the brightness values, as opposed to lower radiometric resolution.

Temporal resolution defines the repeat frequency, which is how frequently a sensor takes image of the same location. Another element of the temporal resolution is the date of the acquisition, it refers to timing issues in order to remotely monitor effects or ecological events, e.g. floods.

Every object that has a temperature greater than absolute zero, emits TIR energy. So all the features in a typical landscape emit TIR radiation, even though this radiation is not visible to the human eye. To understand these physical properties and the functioning of TIR remote sensing, the following chapter will explain the most important physical properties these functions.

2.2 Thermal Infrared Remote Sensing

Satellite sensors that can detect the TIR radiation emitted by the earth's surface features are able to show information about the thermal properties of these materials. Like sensors measuring reflections, these sensors are passive systems measuring radiant emission, relying on the solar radiation as an energy source.

When matter is in random motion, we can measure the energy of its particles in the form of kinetic heat (often also referred to as real, or true, heat). All objects are in motion at the molecular scale when having a higher temperature than absolute zero. When these particles in motion collide, they emit electromagnetic radiation, which is measured in watts. The amount, or sum, of electromagnetic radiation exiting an object is its radiant temperature. The radiant temperature is correlated to the kinetic temperature. Sensors that are responsive in the thermal domain are able to measure the TIR radiation. In contrast to other processes of heat/energy transfer (like conduction or convection where the energy is passed through matter by collision with, or physically moving, the heated particles) radiation is an energy transfer process that can take place in a vacuum through electromagnetic waves (Kuenzer & Dech, 2013). By implication TIR remotely sensed images are able to monitor the thermal radiation of objects on the earth's surface. These images are able to display radiant temperatures of objects at the resolution of the measuring sensor. While the most commonly used derivations of TIR imagery are land and sea surface temperature, the thermal data can be used in manifold other ways than just the derivation of these standard products. As Kuenzer & Dech (2013) writes:

> *"These data enable the assessment of thermal anomalies (forest fires, coal fires, thermal pollution, energy leaks in buildings, inflamed areas in thermal medical imagery), the analysis of moisture conditions, or even the monitoring of machine performance in industrial applications, and – depending on sensor and resolution – the assessment of thermal dynamics at different scales."*

2.3 The TIR Domain and Atmospheric Windows

No strict or physical definition of the TIR domain exists (Quattrochi & Luvall, 1999). According to Kuenzer & Dech (2013), the TIR wavelength extends from "*about 3 to 14 µm*". Also needing consideration arte the atmospheric windows, which are wavelength domains of atmospheric transmittance (Alcântara, 2013). Atmospheric attenuation happens when radiation is passing though the atmosphere to the sensor and is interacting with particles and gasses in the atmosphere. This causes scattering of the radiation and so energy is redirected and/or absorbed at molecular scale (Chang, 2012). The main molecules responsible for absorption of electromagnetic radiation are H_20, O_2, O_3 and CO. In Figure 2 below, we see the atmospheric windows of TIR at 3-5 µm and 8-14 µm. Within the domain of 8-14 µm, only Ozone interferes with the radiometric sensors, and so this band is used by most sensors. Within the 3-5 µm window the reflected sunlight can still be interfering with the radiation and needs to be taken into account when analyzing daytime TIR images, recorded at 3-5 µm.

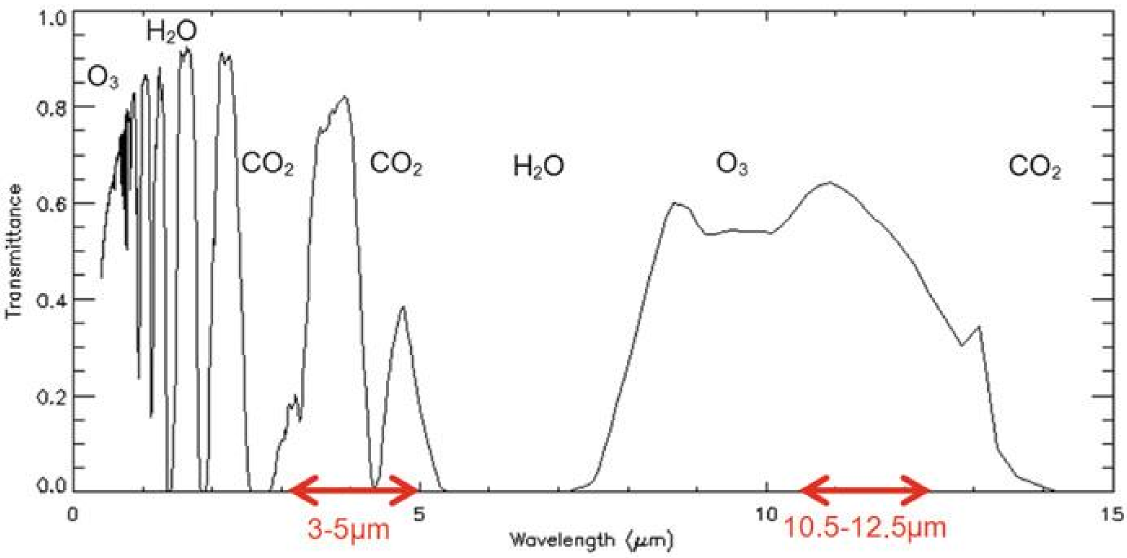

Figure 2: The TIR wavelength domain, typical adsorption domains induced by gases and water and atmospheric windows. (after Richter in Kuenzer & Dech, 2013)

Even though different authors define the TIR domain differently, the main characteristic, which is common to all definitions, is that while multispectral remote sensing in the visible (VIS), near infrared (NIR) records reflected radiation, TIR remote sensing records emitted radiation.

2.3.1 Planck's Law

Planck's law describes the electromagnetic radiation emitted by a blackbody (perfect emitter) at a given wavelength as a function of the blackbody's kinetic temperature (Planck, 1900). A blackbody is a hypothetical, non-existant ideal radiator, which totally absorbs all the energy falling upon at every wavelength and re-emits the maximum radiation, whilst not reflecting or transporting the incident energy. The spectral radiance of a body M_λ describes the amount of energy that is emitted as radiation in terms of power (watts) per unit area of the body, per unit solid angle and per unit frequency. Hence, we can calculate the radiance emitted by a blackbody (M_λ) simply by specifying a certain wavelength:

$$M_\lambda = \frac{2\pi hc^2}{\lambda^5 \left(e^{hc/\lambda kT} - 1\right)}$$

With:

M_λ =spectral radiant exitance [W m^{-2} μm^{-1}]

h = Planck's constant [6.626 x 10^{-34} J s]

c = speed of light [2.9979246 x 10^8 m s^{-1}]

k = Boltzmann constant [1.3806 x 10^{-23} J K^{-1}]

T = kinetic temperature [K]

Λ = wavelength [µm]

How the temperature and energy radiated from a blackbody correlate, as well as the wavelength the blackbody emits the most, can be described by the Stefan Boltzmann's law and Wien's law.

2.3.2 Stefan-Boltzmann Law

The Stefan-Boltzmann law states that the total spectral radiant flux exiting (M_{radBB}) a blackbody is proportional to the fourth power of its kinetic temperature (T) (Kuenzer & Dech, 2013).

With:

$$T_{radBB} = \sigma T^4$$

T_{radBB}= radiant flux exiting Blackbody [W/m^2]

T = absolute kinetic temperature [T]

σ = Stefan-Boltzmann constant [5.6697×10^{-8} W m^{-2} K^{-4}]

The equation shows that the higher the temperature of the radiating blackbody the greater the total amount of radiation it emits.

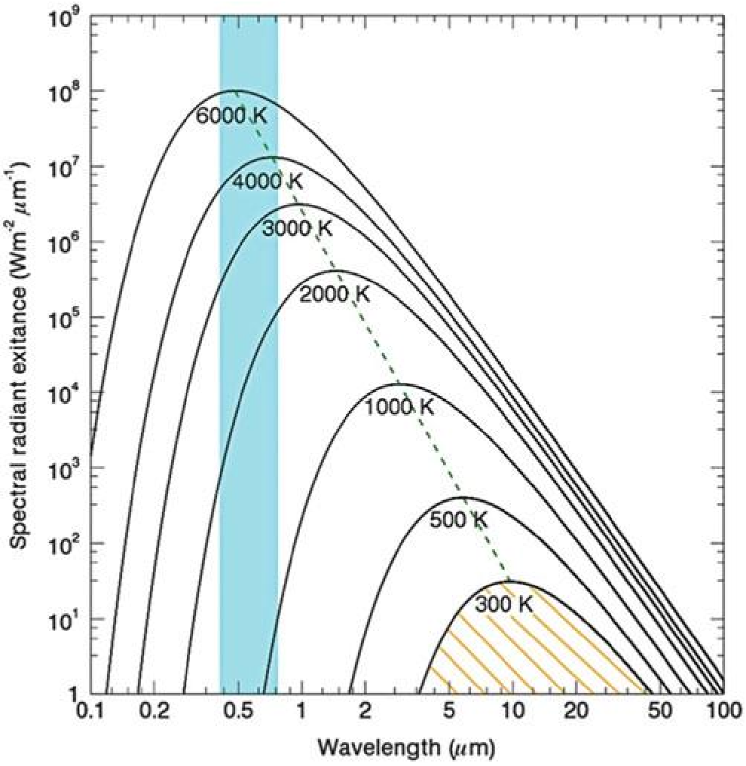

Figure 3: Blackbody radiation curves for different temperatures (after Planck's Law). The yellow shows the Stefan-Boltzmann law, green line is Wien's Law. The blue bar shows the VIS wavelength domain.

In *Figure 3*, the yellow marked area under the 300 K curve, which is the integration of all the area under the blackbody's radiation curve, shows the total spectral exiting radiation at 300K.

2.3.3 Wien's Displacement Law

Wien describes that an object at a constant kinetic temperature has a characteristic peak of radiant power (Clark, 2014). The relationship between the true kinetic temperature of a blackbody (T) in Kelvin and its peak spectral exitance (dominant wavelength) is described by the Wien's displacement law:

$$\lambda_{max} = \frac{A}{T}$$

With:

λ_{max}= dominant wavelength [µm]

A = Wien's constant [2898 µm K]

T = absolute kinetic temperature [K]

Illustrated by the green line in *Figure 3*, we see the maximum exitance λ_{max} moves to shorter wavelength with increasing temperature. The higher the temperature the shorter the wavelength of the peak (Clark, 2014). As a conclusion, we understand that while the sun has an average kinetic temperature of 5.778 K (5.505 °C), the sun has its peak emission in the visible domain of the wavelength spectrum (Kuenzer & Dech, 2013). The earth with its peak temperature of 300 K, is much colder; thus its peak emission is in the TIR domain.

2.3.4 The relevance of emissivity

Through the chapters above we see that the kinetic temperature of an object and its emitted radiant temperature, are linked. So we can use remotely installed radiometers to measure radiant temperatures, and by these means get knowledge about the linked kinetic temperature, which can be seen as the idea standing behind TIR remote sensing. Nevertheless, the linked relationship between kinetic and radiant heat is not perfectly true, and the remotely measured radiant temperature is always less than the true temperature of an object. This is due to a physical property called emissivity. Emissivity is defined as:

> "the ratio of the energy radiated from a material's surface to that radiated from a blackbody (perfect emitter) at the same wavelength and temperature and under the same viewing conditions" (Jensen, 2007).

As no object exists that is a perfect emitter, the result appears that the emissivity of an object and hence the remotely sensed radiant temperature is influenced by a number of factors (but is not dependent on kinetic temperature):

Firstly there is the effect of colour. Darker coloured objects absorb radiance better than lighter coloured objects which tend to reflect more of the incident energy.

Secondly the surface roughness has an effect on the emitted radiance. "*The greater the surface roughness of an object relative to the size of the incident wavelength, the greater the surface area of the object and potential for adsorption and re-emission of energy*" (Jensen, 2007).

Another factor having effects on the emissivity is the moisture content. The more moisture an object contains the greater will be its ability to absorb energy and become a good emitter. Wet soil particles have a high emissivity similar to water (Kuenzer & Dech, 2013).

Linked to this effect is compaction. The degree of soil compaction can effect emissivity as compaction affects soil moisture content (Kuenzer & Dech, 2013).

Knowledge of the effect of emissivity is crucial for thermal data analysis as we can have situations where two distinct materials have the exact same kinetic temperature (T_{kin}) but remotely we could observe very different radiance temperatures (T_{rad}) (Kuenzer & Dech, 2013). Depending on the situation and circumstances of the research prospective, a TIR image needs to be corrected for emissivity when true kinetic temperatures need to be calculated. This is especially crucial, within urban areas where the effect of emissivity is strong. As an example we could state the low emissivity of metal, hence industrial zones with high amount of metal buildings (aluminium, copper, and tin roofs) may appear much colder in radiance temperature than kinetic temperature. On the other extreme when remotely sensing kinetic temperatures of water surfaces, vegetation or under moist conditions, the emissivity effect is close to one, and thus a quite exact assessment of the kinetic temperature without emissivity correction is possible (Kuenzer & Dech, 2013).

2.4 Thermal dynamics of Land Surfaces

The high information content stored within TIR remote sensing makes this sources of information able to generate knowledge of high potential for hydrological model development. The information of surface heat/energy fluxes inherent in TIR data is integral for understanding landscape related processes for a high variety of conditions and scales in space and time. This allows us to monitor surface energy fluxes and by implementing TIR data in landscape studies; it provides measurements of basic input to the energy budget and its distribution in a landscape ecosystem. Quattrochi and Luvall (1999) note:

> *"Understanding how thermal energy is partitioned across a landscape, and the magnitude or variations in surface temperatures emanating from various landscape elements (e.g. forest, crops, pasture, water), is essential to defining the overall mechanisms that govern land-atmosphere interactions."*

As such we may quickly describe the thermal dynamics driving the landscape interactions which we may derive from Thermal Remote Sensing. The energetic dynamics may be defined as (after Quattrochi & Luvall, 1999):

- The coupling of extant energy balances with the environment
- The level of energy inputs (and, hence outputs)
- The kinds of energy transformations that occur
- The mix of energy outputs, which can be regarded as yields

TIR remote sensing data can hence be important, as it provides measurements of thermal energy fluxes. Thus, its usefulness is apparent for e.g. the parameterization of soil moisture or the better simulation of landscape energy exchanges over a variety of conditions on several scales of space and time.

Let's start with the most basic and broad equation of the energy flow through a natural system:

$$\text{Energy Input} = \text{Energy Output} + \text{Energy Storage Change}$$

This equation is valid for long time periods and for most of the natural systems. On shorter intervals, the energy balance differs significantly from equality. The reason is accounted for in energy accumulation or depletion in the system's energy storage. For this example, see land-atmosphere system, if the energy storage is increasing, it is probably an increase in soil or air temperature. Consequently, there is an important link, between the process (energy flow) and its response (temperature change). Quattrochis and Luvall call this a "*process-response system*" (1999), because it attempts to link causes with effects. In this case this would include the abilities to absorb, transmit and emit radiation and the ability of a system to conduct, convect and store energy (Oke, 1987). Consequently, with the use of TIR data we can infer soil moisture or evapotranspiration. To do so, we would need to write the general equation above into a more useful and specific form, by segmenting the net radiation (which represents the amount of energy available at the surface) into other energy balance and flux components. So we write

$$R_N = S(1-\alpha) + L_{wd} - \epsilon\sigma T_s^4$$

Where R_N being the net radiation, S is the total solar radiation available at the surface, α is the broad-band surface albedo, L_{wd} being the downwelling long-wave radiation, T_s being the surface temperature, ϵ being the emissivity and σ representing the Stefan-Boltzmann constant. The R_N on its turn can furthermore be explained by the equation below:

$$R_N - G = H + L$$

With G being the soil heat flux, H the sensible heat flux and LE the latent heat flux. LE can be seen as ET, nevertheless it is often referred to as LE, simply as ET is measured in terms of water in mm/day and LE is measured in W/m^2 (Schmugge, Kustas, Ritchie, Jackson, & Rango, 2002). It becomes clear that land surface temperature, can be understood as the result of "*the equilibrium thermodynamic state dictated by the energy balance between the atmosphere, surface and subsurface soil and the efficiency by which the surface transmits radiant energy into the atmosphere (surface emissivity)"* (Schmugge et al., 2002). As such, opposite to reflected solar radiation spatial temperature variations may also be caused by invisible processes and material properties below the surface, making the interpretation of TIR data a very complex discipline (Grayson & Blöschl, 2001).

Temperatures of all land surface materials change in a cycle of cooling and heating, but the response of the materials towards heating and cooling may differ considerably (Smith, 2012). For example darker materials absorb more incoming radiation/energy and heat up faster than reflective materials. Additionally if two materials absorb the same amount of temperature, the maximum temperatures may be differing (Kuenzer & Dech, 2013). All of these facts are results of the thermal capacities of materials. The properties of materials define the heat/energy required to induce a raise in temperature for example. As the land surface heats up energy may be transferred to cooler levels in the vertical soil column by conduction. On cooler days the reverse effect can be observed. These temperature exchanges may extend as deep as 30 cm below the surface (Smith, 2012). Additionally these heat transfer rates strongly depend on material properties such as bulk density or thermal conductivity. These effects are summed up under the term of "thermal inertia" controlling, diurnal and seasonal surface temperature variations. Thermal inertia is a function of density, thermal capacity and thermal conductivity. These properties are strongly dependant on soil characteristics such as composition, porosity and is as such often related to soil moisture content (Pratt & Ellyett, 1979; Lu et al., 2009). Consequently, with TIR remote sensing we might not only observe land surface functioning with regard to hydrological processes, but also spatial variability. If observations are available for consecutive time-steps, we can gather information about the temporal variability. The link between hydrological process and energy balance characteristics can be translated through the ET-term in the above mentioned equation. Which in turn is closely linked to the water balance, which is simply written as (Smith, 2012):

$$\frac{\Delta S}{\Delta t} = P - ET - Q$$

Where $\frac{\Delta S}{\Delta t}$ is the change in water storage in the soil, P is precipitation, ET is evapotranspiration and Q is runoff. From a hydrological point of view, it is hence possible to understand flows and transfers of water across and between landscapes by inferring them from these energy flux coefficients.

Several studies have been done to deduct evaporation, evapotranspiration or soil moisture conditions from these coefficients (Grayson et al., 1997; Kalma, McVicar, & McCabe, 2008; Pratt & Ellyett, 1979; Sánchez et al., 2008). However, difficulties arise as "*technological means as well as analytical rational are needed to gain more knowledge about energetic processes for a specific landscape*" (Quattrochi & Luvall, 1999). By this, they mean that direct point measurements are needed for validation. Additionally some meteorological constants need to be known for atmospheric corrections as well as several surface and subsurface characteristics in order to define the coefficients of the energy balance. Consequently, the focus of this thesis lies on the deduction of patterns representing direct internal states rather than inferring physically meaningful values from the radiation measured at the sensor.

3 Methods and Materials

3.1 Study Areas

Two Austrian study areas have been chosen to delineate patterns representing hydrological heterogeneity using principal component analysis. The catchments of the Rivers Ybbs and Traisen are part of the Danube basin in Lower Austria and drain the region towards the North into the Danube. The selection of the study area was done firstly on behalf of investigating the method under different landscape characteristics than those shown in the paper of Müller et al. (2014), i.e. elevation, soil types etc. Secondly the criteria of data availability played an important role in the process of choosing the appropriate study areas. As both catchments are being monitored and play roles in several research projects, datasets providing information on the landscape characteristics were available.

The Ybbs river catchment is characterized by warm temperate climatic conditions with an annual precipitation of up to 900mm and belongs to the northern limestone pre-Alps.

The Traisen river catchment is similarly summarized with a warm temperate climate and reaching from the northern limestone pre-Alps northwards toward the Danube.

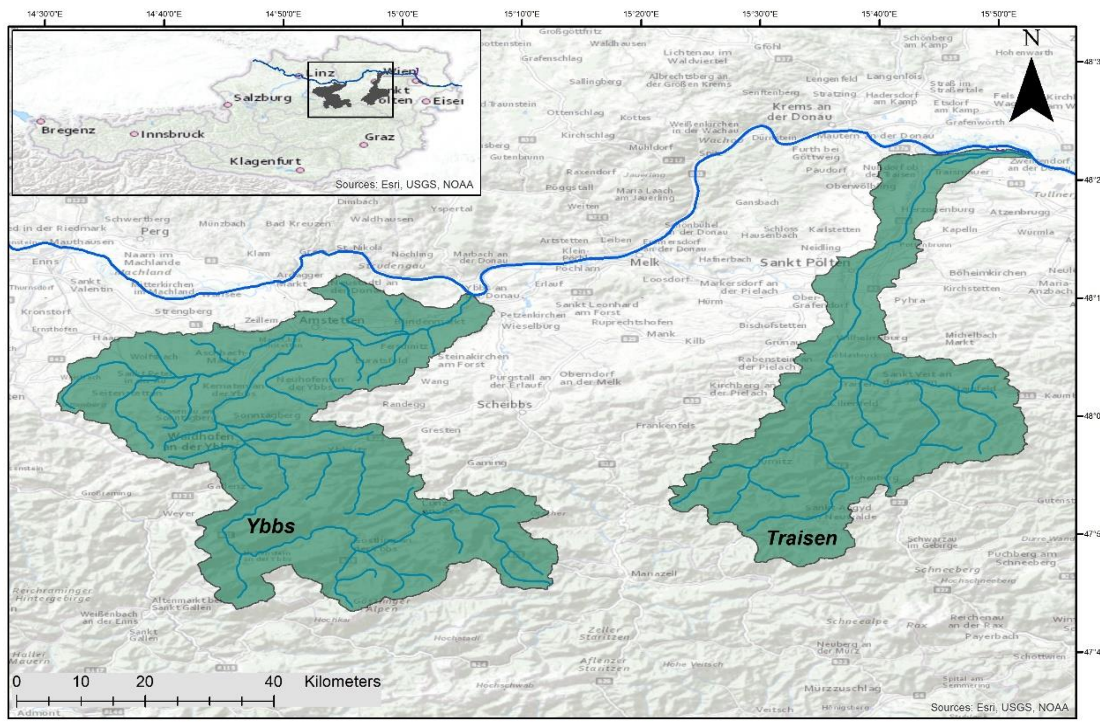

Figure 4: Ybbs and Traisen Catchments in Lower Austria, showing Ybbs and Traisen Rivers, its outlets into the Danube River and its location relative to the Austrian Territory.

3.1.1 Ybbs River Catchment

The upper Ybbs catchment is a mountainous region with high slopes and rocky underground (daNUbs, 2003). The lower Ybbs catchment is characterized by a large valley of pastures and floodplains. The main characteristics of the Ybbs river catchment are listed in the table below.

Table 1: Main Characteristics of the Ybbs river catchment

	Ybbs	
Catchment area	[Km²]	1370
Mean annual Precipitation	[mm/a]	862
Average Terrain Slope	[%]	23.71
Population Density	[inh/Km²]	68
Landuse Characteristics		Arable land, Pastures, Forest

In the following chapter the most important landscape defining the Ybbs catchment will be displayed and explained.

Elevation Characteristics

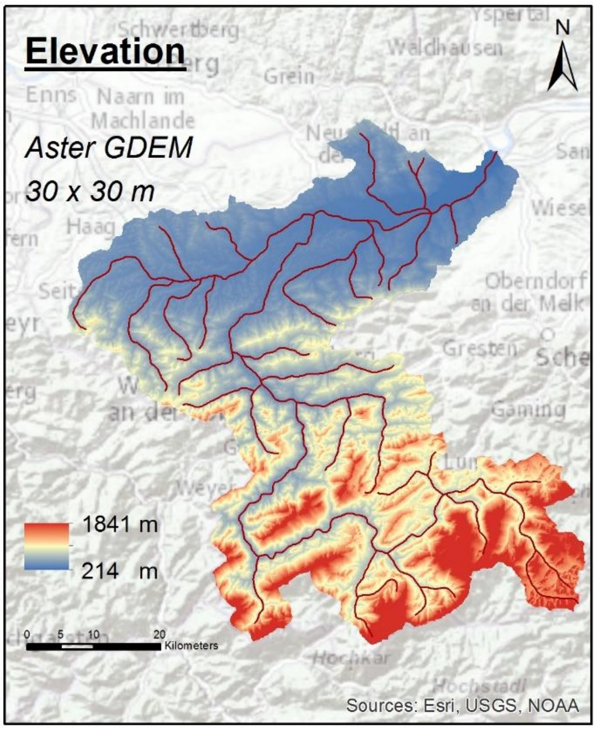

Figure 5: Digital Elevation Model of the Ybbs Catchment, showing the Ybbs River. Aster GDEM on a 30x30m grid

The height differential in the Ybbs river catchment ranges from 250m to 1900m above sea level. The average slope of the catchment is 23%. Nevertheless approximately one third of the catchment, see the downstream region, is flat or hilly with mean elevations below 500 m above sea level. The upstream region is characterized by narrow valleys and steep slopes with elevations higher than 500 m above sea level.

Landuse characteristics

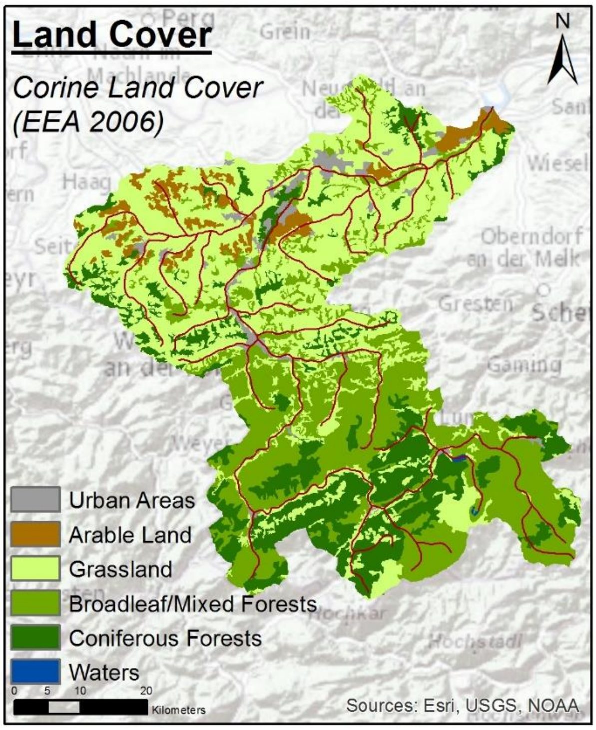

Figure 6 Landuse characteristics for the Ybbs Catchment

Landuse characteristics were derived from the Corine Land Cover maps issued by the European Environmental Agency. The grid resolution is 100x100 m. The dominant land cover types in the Ybbs catchment are pastures, forestry and arable lands (European Environment Agency, 2007).

Hydrogeological Characteristics

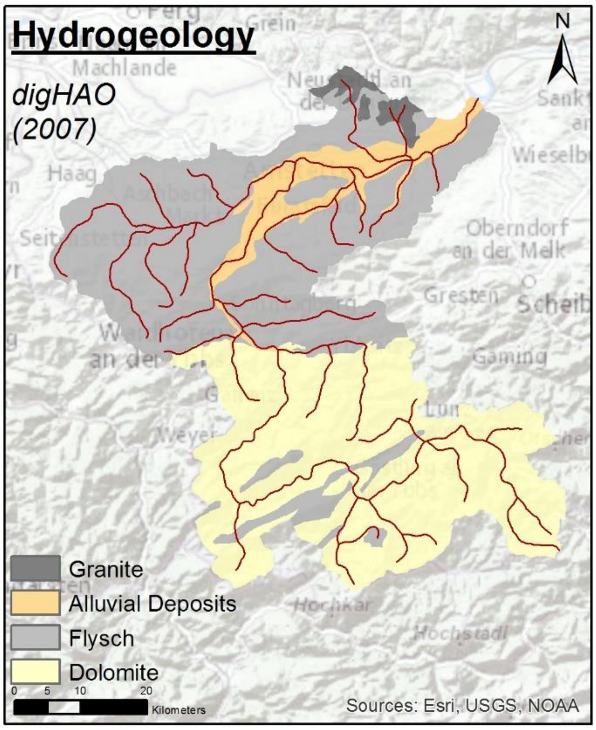

Figure 7: Hydrogeology for the Ybbs Catchment

The hydrogeological map of the hydrological Atlas of Austria (digHAO) was used to derive geological subsurface characteristics affecting hydrology. The three major geological characteristics are Flysch (46%), Dolomite (44%) and alluvial deposits (8%). Granite covers 2% of the total catchment area.

3.1.2 Traisen River Catchment

The upper Traisen catchment is a mountainous region with high slopes and rocky underground. The lower Traisen catchment is characterized by larger floodplains. In direct relation to the the high elevation range (200m – 1800m) climate changes north to south (from pannonian to alpine). Below are listed the main characteristics for the Traisen river catchment.

Table 2: Main Characteristics of the Traisen river catchment

		Traisen
Catchment area	[Km²]	911
Mean annual Precipitation	[mm/a]	600 - 1500
Average Terrain Slope	[%]	26.42
Population Density	[inh/Km²]	13
Landuse Characteristics		Arable land, pastures, Forest, Floodplains

Elevation Characteristics

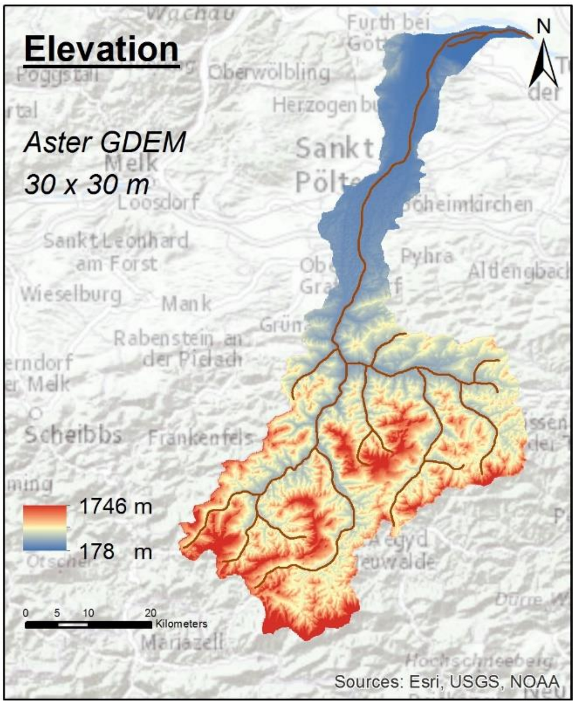

Figure 8 Digital Elevation Model of the Traisen Catchment, showing the Traisen River. ASTER GDEM on a 30x30m grid

We see the high altitudinal difference from the South being in the Pre-alpine environment to the lower North, where the Traisen River flows into the Danube. Remarkable for this catchment is the big expansion of the drainage area in the alpine environment, compared to the small drainage area in the valley of the catchment

Landuse Characteristics

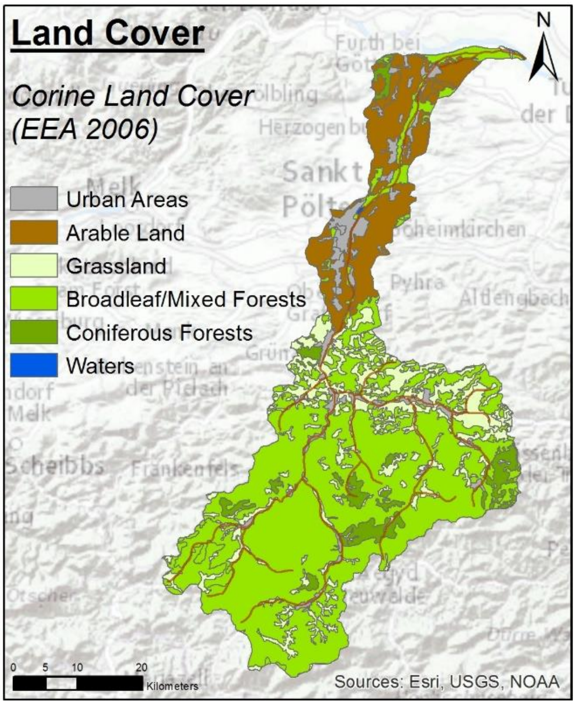

Figure 9: Landuse characteristics for the Traisen Catchment

We can see the dominance of forestry in the southern part of the catchment and the agriculturally intense northern areas of the basin. Land cover characteristics were derived from the Corine Land Cover maps (100x100m) made available by the European Environmental Agency (EEA, 2006).

Hydrogeological Chararcteristics

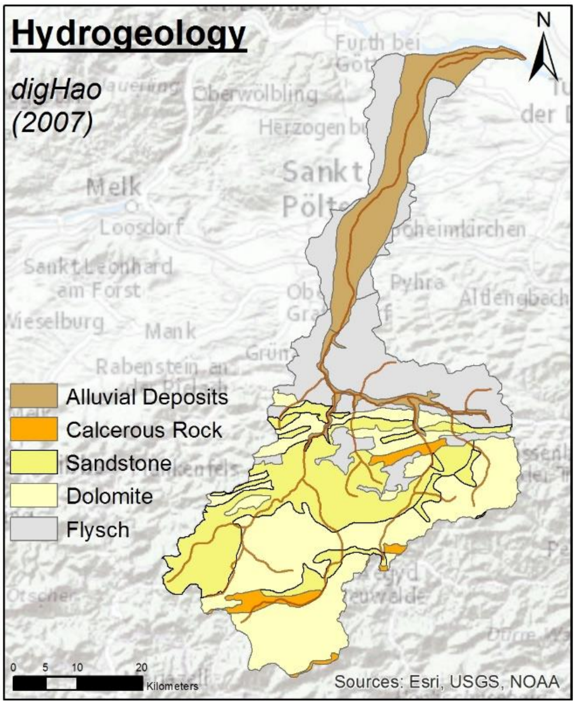

Figure 10: Hydrogeology for the Traisen River Catchment

The hydrogeology of the basin is classified into five major zones, these being namely Dolomite (31%), Flysch (27%), Sandstone (24%), Calcerous Rock (3%) and Alluvial Deposits (15%). The dataset was derived from the digital hydrological Atlas of Austria.

3.2 TIR Database

Tests using different satellite systems were completed. For example Landsat 7 only has usable products for a time range of 4 years. In 2003, a so called Scan Line Corrector (SLC), compensating the forward motion of Landsat 7 failed, and to date could not be repaired. The effect of this failure is that approximately 22 per cent are lost in an image and no-data gaps are as wide as 450 m. Hence Landsat 7 was rejected.

Aster is a satellite system of Japanese and US American cooperation on board the TERRA satellite, and is an efficient remote sensing system producing images of high spatial and spectral resolution with 14 bands from visible to thermal infrared. However both catchments were laying between overflying zones, which made the use of ASTER impossible as it would need to interlock two images produced on two different dates and so giving different temperature values, this would be scientifically untenable

As a consequence Landsat 5 was chosen as the satellite sensor system for this study. Landsat 5 was operated by NASA and USGS and launched in 1984. It was officially decommissioned in 2013 producing granules for a time span of 28 years, thus being the "*longest-operating Earth observation satellite*" (Betz, 2013). The satellite orbit is sun synchronous and mean solar overpass time is 9.45 am. It needs 99 minutes for one orbit (14 orbits a day). The temporal resolution is 16 days (16 days to scan the entire Earth) and granules are produced for 7 bands covering a spectral bandwidth from 0.45µm to 12.50µm. Each TIR image covers an area of approximately 185km east to west and 170 km north to south (Betz, 2013). Spatial resolution for the TIR band is 120 m which is resampled to 60 meter (since 2010 to 30 m) pixel size in the delivered data product (USGS, 2014). Landsat 5 operates at a height of 705 km. The TIR band of Landsat 5 operates at the atmospheric window, between 10.4 µm and 12.5 µm (USGS, 2014).

For the two different catchments, a total of 51 images were downloaded from the USGS Global Visualization Viewer (http://glovis.usgs.gov/) and the homepage of the USGS EarthExplorer (earthexplorer.usgs.gov), where Landsat 5 granules are freely available for a time period from 1984 to 2011. The pictures are composed of 23 images for the Ybbs river catchment and 28 images for the Traisen catchment. A further screening was done, in order to select only cloud-free images for the areas of the catchments of interest. In the end, a total of 11 images for the Ybbs catchment and 12 images for the Traisen catchments were retained.

3.3 Data Preparation

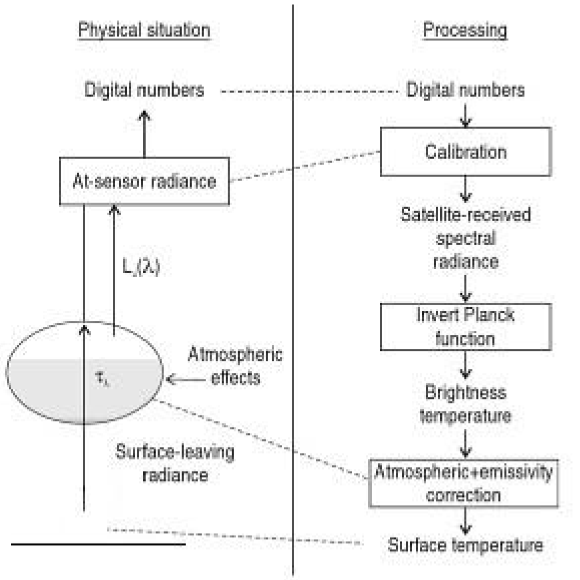

Figure 11: Left hand side shows the physical processes affecting the surface leaving spectral radiance, right-hand side explains the processing required to retrieve the surface kinematic temperature form the recorded digital numbers at the satellite-sensor (after Harris, 2013)

The factors that needed to be accounted for temporal coherence, are common to all remote sensing. These include calibration, georeferencing and, depending on the atmospheric window, atmospheric correction. While calibration means the transformation of sensor counts, i.e. binaries, into brightness temperature on behalf of sensor specific calibration coefficient considering deterioration effects, georeferencing is the process of associating locations with the map, to assure spatial coherence (Alcântara, 2013). Furthermore, atmospheric correction is the process of accounting atmospheric adsorption effects, to assure temporal coherence.

In order to discriminate between product artefacts and land surface change in the information provided by the TIR time-series of Landsat 5, the granules need to be further processed. First of all, the images were “mosaicked” north to south in R-studio, using mean temperatures at overlapping raster cells (www.rstudio.com; R Core Team, 2014). The data lacked an accurate georeferencing. To counteract this effect which might influence the following principal component analysis, an automated georeferencing tool was implemented in R-studio. The tool was part of the “landsat” package provided by Goslee (2011) and uses the root mean square error as a simple routine providing relative georeferencing to one manually georeferenced image. Further the results were checked visually for coherence. In the next step, using an iterative model in the model-builder implemented in ArcMap-Tool of the ArcGis for Desktop software package 10.2.2 (www.esri.com; Environmental Systems Research Institute, 2015) the TIR data was clipped to the relative catchment extents.

Additionally, the data was converted from Binaries to At-Sensor radiance and from radiance to brightness temperature in Kelvin (see Figure 11).

3.3.1 Conversion to At-Sensor spectral radiance

In order to calculate the image data of the Landsat 5 granules into a physical meaningful common radiometric scale, the first step is to calculate at-sensor spectral radiance. As discussed above before the granules reach the distribution media, they are rescaled and calibrated from the raw digital numbers transmitted from the satellite system to calibrated digital numbers in order to standardize the scaling for all scenes (Chander et al., 2009). This initial radiometric calibration is done using 32-bit floating point calculations to convert the unprocessed image digital numbers to units of absolute spectral radiance (Chander et al., 2009). For Landsat 5 they are then rescaled to 8-bit (values range from 0 to 255). When now converting the Digital Numbers to At-Sensor spectral radiance, knowledge of the lower and upper limit of the original rescaling factor is needed (Chander et al., 2009). The following equation is used to convert the calibrated digital numbers to at-sensor radiance:

$$L_\lambda = \left(\frac{(LMAX_\lambda - LMIN_\lambda)}{Q_{calmax} - Q_{calmin}}\right)(Q_{cal} - Q_{calmin}) + LMIN_\lambda$$

Where:

L_λ= Spectral radiance at the sensor [W / (m^2 sr µm)]

Q_{cal}= Quantized calibrated pixel value [DN]

Q_{calmin}= Minimum quantized calibrated pixel value corresponding to $LMIN_\lambda$ [DN]

Q_{calmax}= Maximum quantized calibrated pixel value corresponding to $LMAX_\lambda$ [DN]

$LMIN_\lambda$= Spectral At-Sensor radiance that is scaled to Q_{calmin} [W / (m^2 sr µm)]

$LMAX_\lambda$= Spectral At-Sensor radiance that is scaled to Q_{calmax} [W / (m^2 sr µm)]

(Chander et al., 2009)

3.3.2 Conversion to At-Sensor brightness temperature

In the next step all thermal band data was converted from at-sensor spectral radiance to effective at-sensor brightness temperature. The at-sensor brightness temperature is a one-to-one conversion from radiance to temperature, assuming that the earth's body is a black-body, hence a perfect emitter (spectral emissivity is 1)[1]. Due to the effects of emissivity, two different materials may have the same temperature but appear as different brightness temperatures in the TIR time-series. As emissivity is strongly linked with soil hydraulic properties (e.g. compaction), differences in temperature appearances are beneficial to assess units of the same hydrological behaviour. By using a reverse calculation of

[1] As a consequence, brightness temperature may appear higher than the actual real temperature, as the effect of emissivity is not accounted for.

emissivity to get absolute land surface temperatures for example these advantageous effects visualizing patterns of hydrological behaviour would reduce considerably. Hence a conversion to at-sensor brightness is the adequate approach here when accentuating patterns of hydrological behaviour.

The conversion from at-sensors spectral radiance to at-sensor brightness temperature is:

$$T = \frac{K2}{ln(\frac{K1}{L_\lambda} + 1)}$$

Where:

T= Effective At-Sensor brightness Temperature [K]

K2= Calibration constant 2 [K]

K1=Calibration constant 1 [W / (m^2 sr µm)]

L_λ= Spectral radiance at the sensor [W / (m^2 sr µm)]

(Chander et al., 2009)

3.3.3 Atmospheric and Emissivity Correction

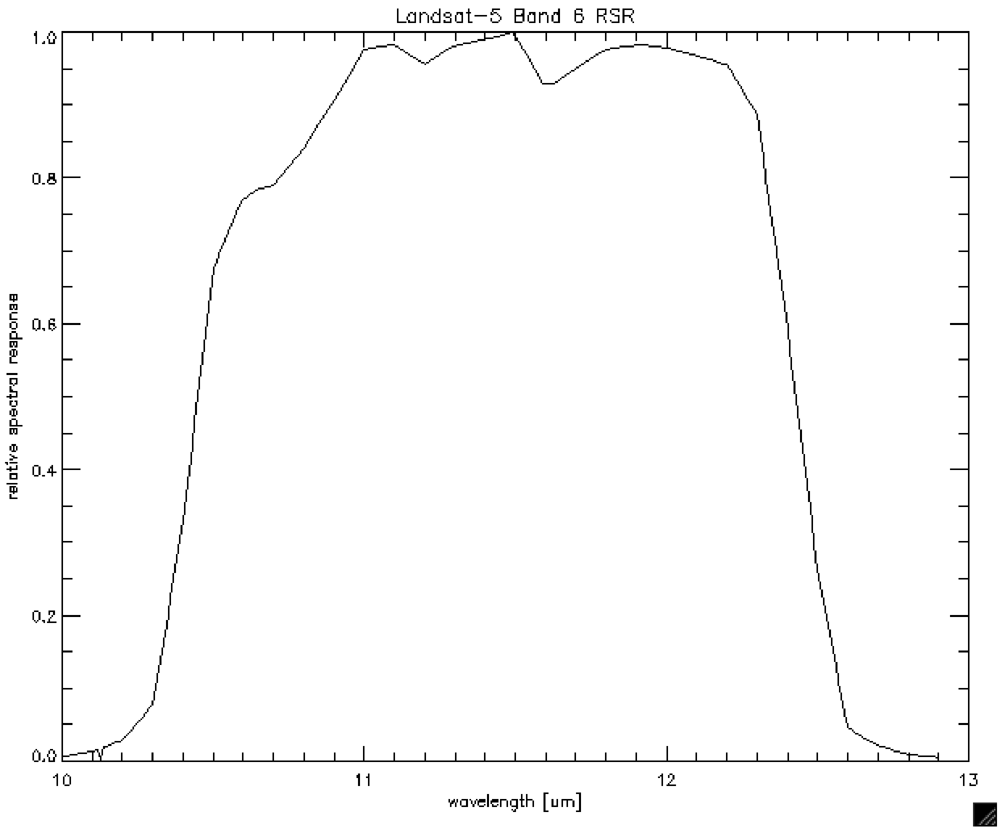

Figure 12: Landsat 5 Band 6 (TIR) Relative Spectral Response Curve (after Barsi, Barker, & Schott, 2003)

When applying atmospheric correction, parameters describing the atmospheric physical interception processes linked to water vapour content, need to be inserted into a radiative transfer equation. However accurate data coherent with the overpass times of the satellite are often unavailable in reality. This uncertain atmospheric contribution is one of the main problems for the remote sensing of surface temperature at infrared wavelength (French et al., 2008). These key difficulties still represent considerable uncertainty about the accuracy in the derived land surface temperatures (Kalma et al., 2008). Which is the reason why the potential of TIR time-series has not been fully realized yet. The peak emissions for mean terrestrial surface temperatures (~300K) occur in the 8-12.5 µm range of the wavelength (Cracknell, 1997). Figure 12 shows the relative spectral response at this wavelength, the same range where the atmosphere allows peak atmospheric transmission. That being the range of the wavelength where the atmosphere is the most transparent, see Chapter 1.1. Hence the question arises as to whether one should carry out atmospheric correction of the data or not (Cracknell, 1997). Qin et al (2001) state that it is rational to assume that the part of the intercepted radiance that is reflected by the atmosphere can be set equal with the radiance emitted of the atmosphere for clear days (see Figure 11 L(λ) and $\tau_{(\lambda)}$). There is still some atmospheric attenuation due to water vapor. So the magnitude of the atmospheric adsorption will depend on the water vapor content in the atmosphere. The at-sensor brightness temperature may vary from the actual land surface temperature by as much as 5°C (Cracknell, 1997). Cracknell further writes: "*While it would be naïve to regard the correction as a simple fixed offset over the whole scene, it is, nevertheless, reasonable to assume that the value of the correction varies relatively slow across the scene.*" Furthermore, both Landsat 5 and 7, each with a single thermal band provide no coherent possibility of correcting atmospheric effects unlike multi-thermal band instruments (Barsi et al. , 2005). Hence additional meteorological data become indispensable for the conversion from Top of Atmosphere temperature to surface-leaving temperature (Barsi et al., 2003). Furthermore have Müller et al. (2014) done tests in comparing PC's computed on extracted land surface temperature with initial binaries which showed almost identical results. They further state: "*a proper conversion to LST is in our opinion not fundamentally needed*". In conclusion due to the above stated facts, and to the high relative spectral response of Band 6 of Landsat 5 the atmospheric correction was omitted at this place. Note that clear and cloud-free days need to be selected for the calculation due to this omission.

Another interfering process when deriving land surface temperatures is the Emissivity. As stated in Chapter 2.3.4, Emissivity can be potentially very useful for distinguishing surface compositions. Salisbury et al. (1985) show for example that "*emissivity's of soil* *and vegetation are commonly distinct and do not rely upon plant chlorophyll content*". Hence by back-calculation of emissivity to land surface temperatures, a potential supplementary tool for distinguishing hydrological patterns would be lost and hence emissivity was retained in the dataset, calculation was done on brightness temperature data.

3.4 Principal Component Analysis

In order to identify patterns of similar functional behaviour the TIR time-series are analysed using Principal Component Analysis (PCA). PCA is a multivariate statistical model which focuses on the variables. PCA generates new sets of components, which represent alternative descriptions of the data. PCA transforms the original pixel vectors previously related to temperature values into new components, which potentially make features visible that were not identifiable in the original data set. Further to this PCA can reduce the number of dimensions in a dataset without considerably reducing the information content. When we think of the specific example of this work, the information of interest (e.g. Landcover, geology, elevation) of one time-step is integrated within the other time-step and hence we see redundancy in the dataset. In order to get rid of the redundancy and reduce "noise", principal component analysis transforms the original variables orthogonally into a new set of uncorrelated variables, called principal components. These are linear combinations of the original variables, whereas the idea of data reduction consists in trying to represent as much of the variation of the original variables through as less principal components as possible.

3.4.1 The PCA-transformation

The time-series of TIR remote sensing images will be understood as a multidimensional space, with as many dimensions as there are time-steps. For each pixel in the area under investigation a vector space with as many axes as there are time-steps is mapped. The measurements at different spatial locations (see pixels) are treated as variables and the time steps play the role of observations. Each pixel of each time-step within the image under investigation is plotted as a point in such a space with coordinates corresponding to the temperature values of the pixels. When applying PCA the "*most dominant controlling factors for the temporal dynamics*" (Müller et al., 2014) can be identified.

In order to do so the position of the points (representing a pixel) in the multi-temporal space can be described by vectors, whose components represent the individual thermal response of each time-step. Assuming that the points (pixel values) correlate, and that the assumption of the data not being completely independent holds true, so their axes of correlation are not orthogonal. The statistical test of dependence is covariance (Clark & Rilee, 2010). Small covariance indicates independence while the highest covariance indicates the most dependence. The task is now to find the strongest covariance relationship between the dates which derives the primary component axis. Each of these axes is created by a linear transformation-rotation and translation (Ng, 2008).

Given a dataset of values representing temperatures, at each pixel we have:

$$\{x^{(i)}, \ldots., x^{(n)}\} \text{ with } x^{(i)} \in R^n$$

We can transform and reduce the dataset to a k-dimensional dataset (k ≤ n). After organizing the dataset in a matrix we need to do some pre-processing. First we need to centre our data. This means that for each temperature image we subtract the means of the images of themselves, so that each of the dimension in the dataset has zero mean (Ng, 2008).

Set:

$$u = \frac{1}{m}\, x^{(i)}$$

We replace:

$$x^{(i)} \text{ with } x^{(i)} - u$$

In the next step we compute the variances of each of our time-steps and divide the centred data by the standard variation, so that our data set is represented in unit variance. This process is called scaling, or normalizing. It makes one scale of features of different scales. This process might not be necessary if we have only one scale within the dataset, as it is the case for this dataset. However, if this step is neglected the principal component analysis will tend to give more emphasis to those variables that have higher variances than to those variables that have very low variances. Then we speak of non-standardised-PCA. This normalisation prevents certain features to dominate the analysis because of their large numerical values. If this step is computed, we speak of standardised PCA. Thus we compute the normalization process:

$$\sigma_j^2 = \sum_{i=1}^{m} (x_j^{(i)})^2$$

And replace

$$x_j^{(i)} \text{ with } \frac{x_j^{(i)}}{\sigma_j}$$

Now as we have centred and scaled our data, in a next step we would find the axis of highest variation (Clark & Rilee, 2010). This could be translated as to find the axis or direction which the projected data to that direction are widely spaced out and vary as much as possible (Ng, 2008). PCA now consists of searching the direction of the vector u that maximises the variance when dataset is projected upon. Hence the need to compute a measure of covariance (if data was not scaled we speak of a measure of correlation), by calculating the covariance (or the correlation) matrix of the dataset. The covariance matrix is defined by:

$$\Sigma = \frac{1}{m} \sum_{i=1}^{m} x^{(i)} \; x^{(i)T}$$

By definition, the direction u where the dataset should be projected on is computed by calculating the eigenvector of the covariance matrix (Richards & Jia, 2006). Eigenvectors are defined as:

$$A\,u = \lambda\,u$$

With:

A being the covariance matrix Σ in this case.

Λ being the eigenvalue

μ being the eigenvector.

And as such the principle eigenvector is the one that maximises the covariance matrix (in the old projection), or eigenvalue (covariance matrix in the new projection) that gives us the best direction onto which to project our data (Richards & Jia, 2006). This is also known as the Karhunen-Loève transform or Hotelling transform (Jolliffe, 2002). Generally, if we want a k-dimensional subspace onto which we want to project our data, we choose the {u, … , u_k} to be the top k eigenvector of our covariance matrix Σ, and hence the top k eigenvectors corresponding to the top k eigenvalues, in descending order

(Jolliffe, 2002). These eigenvectors now give us the new basis onto which we can project our data. To do the projection we would simply represent the original dataset x $^{(i)}$ in the new {u $_{(1)}$… x $_{(k)}$} basis by multiplying:

$$y^{(i)} = (u_1^T x^{(i)}, u_2^T x^{(i)}, \dots, u_k^T x^i) \text{ with } y^{(i)} \in R^k$$

The figure below illustrates the orthogonal transformation in a 2-dimensional space. The initial data x^1 and x^2 are highly correlated. After the rotation, corresponding to the above explained Karhunen-Loève transform or Hotelling transform, the components y^1 and y^2 are uncorrelated, with y1 explaining most of the information or variance of the dataset.

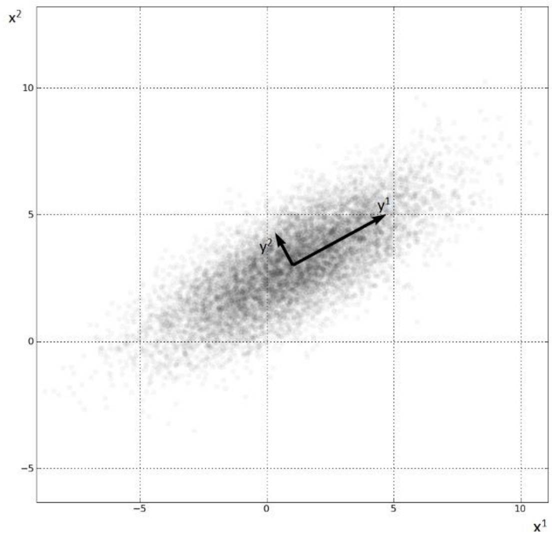

Figure 13: Illustration of PCA rotation. y^1 and y^2 correspond to the eigenvectors of the covariance matrix of the initial dataset and form the new basis onto which the data is projected. (Source: Wikimedia Commons)

4 Results

Thermal dynamics of land surfaces are controlled by their thermal properties which are strongly linked to the water balance through the latent heat fluxes that are inherent. To deduce functional units of similar hydrological behaviour, a principal component analysis was performed with a time-series of Landsat 5 TIR images for the catchments Ybbs and Traisen. For the Traisen catchment 12 cloud-free images were used as an input, whereas for the Ybbs catchment the PCA was calculated for 11 time-steps. The extent of the time-step was clipped using ArcMap (Environmental Systems Research Institute, 2015) to the extent of the catchment. The choice of changing or widening the extent was limited due to the overpass area of the satellite being close to the western edges of the Ybbs catchment. As such, to not further reduce the dataset, this work is computed on the spatial extent of the catchments.

Table 3 and 4 show the absolute and cumulative proportion of variance explained through the uncorrelated Principal Components. By definition the first principal component explains the most of the variance within the thermal time-series, whereas the second component explains the second most important uncorrelated variance, and so on. The principal component analysis was computed in Rstudio (R Core Team, 2014). The following components are centred in pre-processing and were further standardised (i.e. scaled) to give each image similar influence/weights on the components (Hirosawa et al., 1996).

Table 3: Importance of Components according to centred and standardized Principal Component Analysis of 11 images displaying thermal variability in Ybbs Catchment

Ybbs	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Proportion of Variance	0.792	0.074	0.024	0.021	0.017	0.016	0.014	0.012	0.012	0.010	0.009
Cumulative Proportion	0.792	0.866	0.890	0.911	0.928	0.944	0.958	0.970	0.982	0.991	1

In the Traisen catchment, the first three components explain 86 % of the thermal variability in the 11 images. For the Ybbs catchment 86% of the variability can already explained through PC1 and PC2.

Table 4: Importance of Components according to centred and standardized Principal Component Analysis of 12 images displaying thermal variablility in Traisen Catchment

Traisen	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12
Proportion of Variance	0.755	0.072	0.033	0.028	0.024	0.019	0.015	0.014	0.010	0.009	0.009	0.006
Cumulative Proportion	0.755	0.828	0.861	0.889	0.914	0.934	0.949	0.963	0.974	0.984	0.993	1

Figure 14 shows the images generated by the principal component analysis. After a visual inspection we see that for the first two principal components, some patterns become apparent that might explain hydrological behaviour and/or thermal inertia characteristics of soil and land cover types. In the first component a strong spatial variability is visualized with the river Ybbs that can be differentiated as well as urban areas and of course relief characteristics in the south. Furthermore, forested areas can be distinguished visually from pastures, riverbeds and urban areas. For the second principal component, a first visual estimation of the represented source of variation is a more difficult task. From south to north, the mountainous region can be distinguished as well as some patterns in an east-west corridor can be seen in the valley in the north of the catchment. The third and the fourth images show only very short gradients of variations. The third principal component shows some patterns in the urban areas of the catchment. Furthermore, these images show already some noise as can be seen by the representation of small clouds in the alpine environment. Upon further inspection of the proportions of variance in Table 3, we see that these images merely represent 2% of the total variance contained in the dataset.

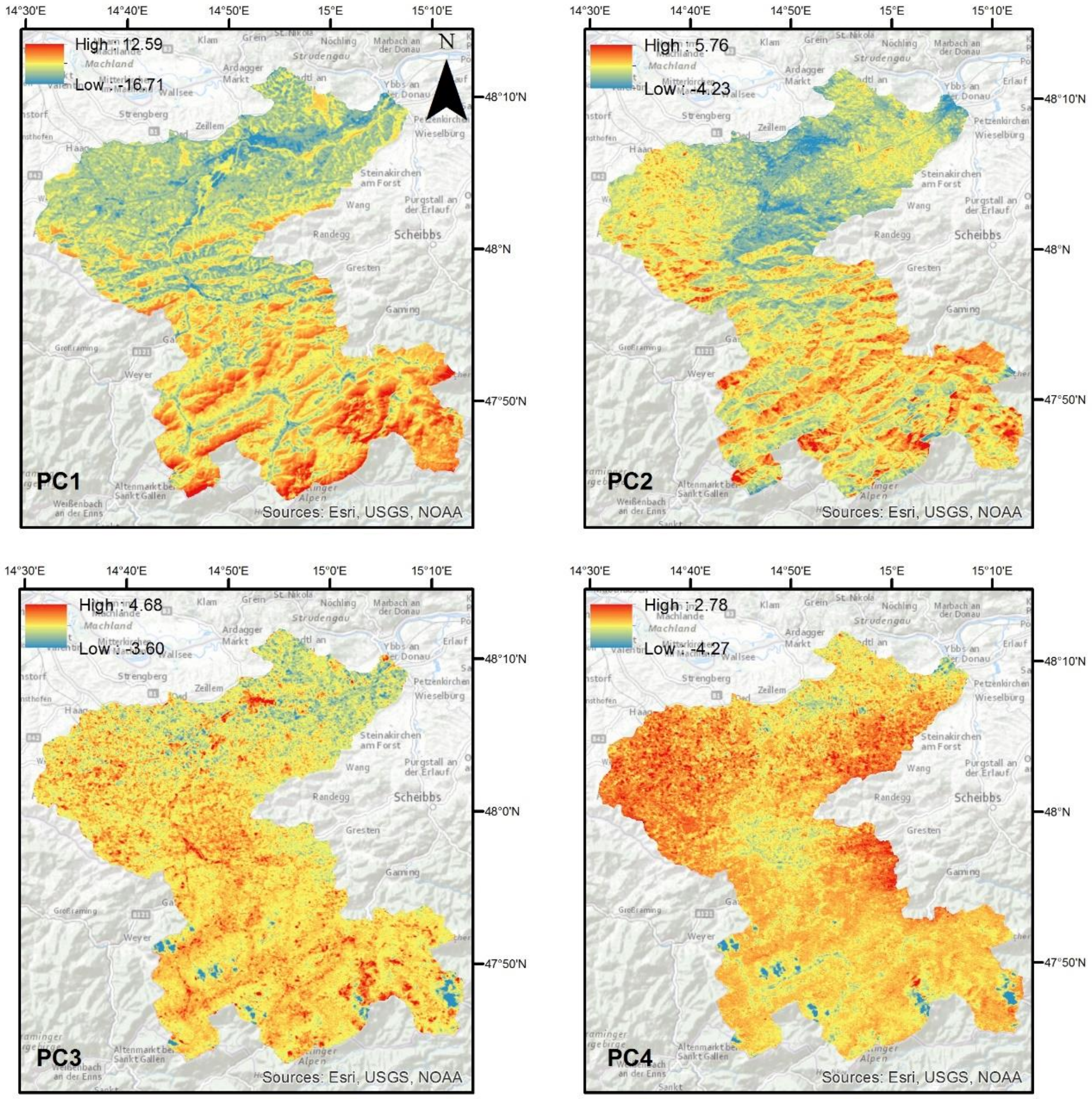

Figure 14: First 4 Principal Components for the Ybbs Catchment explaining a Total of 91% of the variance in temperature values

Figure 15 below shows the first four principal components for the Traisen catchment explaining the sources of uncorrelated variance in decreasing order. While the first principal component shows the variation that is explained by spatial distribution, with land cover types and elevation characteristics being represented. Furthermore, the riverbed is represented as well as urban areas. Component number 2 shows a distinct pattern, with a less stringent distinction between the valley and the mountainous areas. Moreover, characteristics potentially representing underlying hydrological behaviour might be represented. The components 3 and 4 each represent approximately 3% of total variance comprised in the dataset. The third component shows some patterns strikingly on the areas that are known to be urbanized regions, similar to the Ybbs catchment's principal component 3.

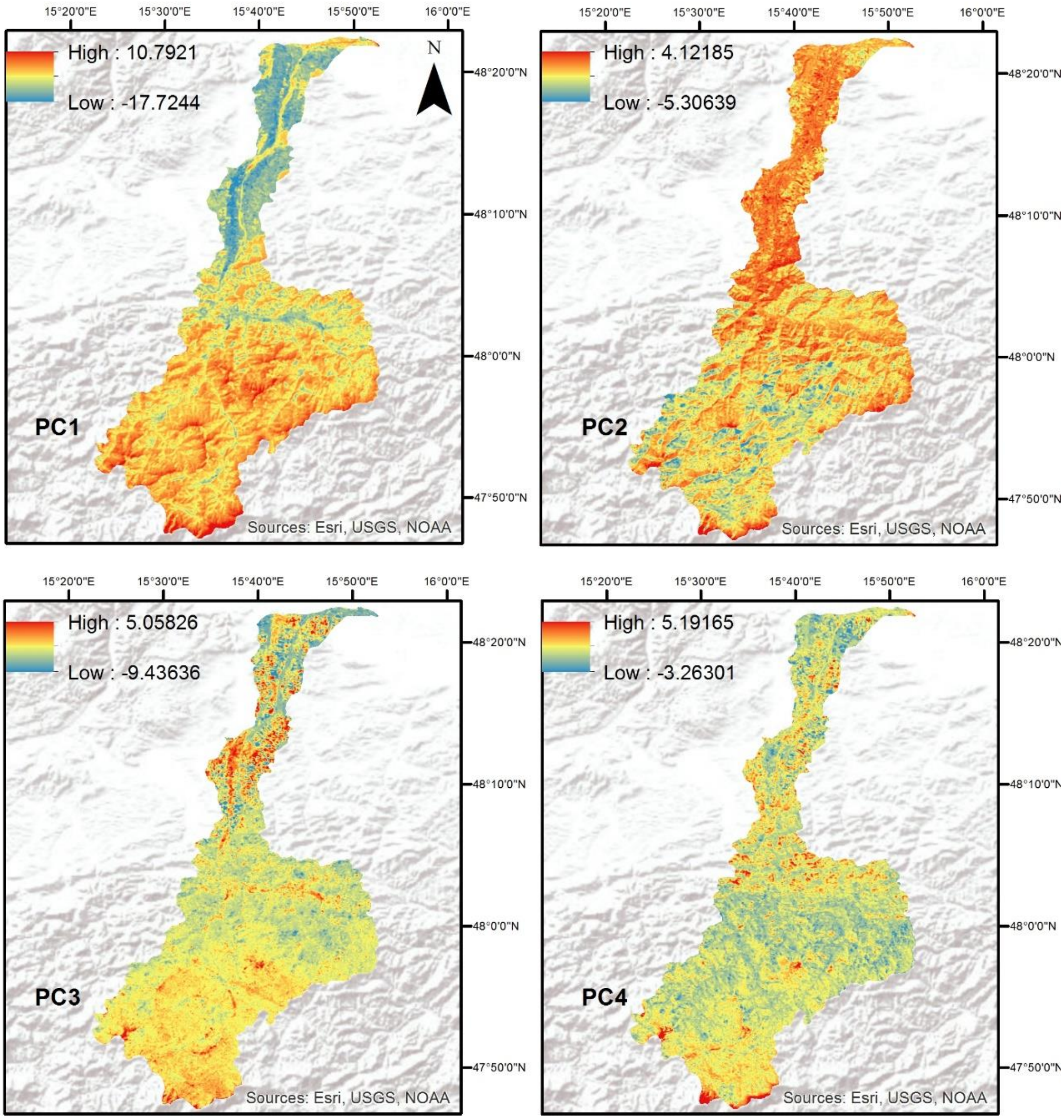

Figure 15: First 4 Principal Components for the Traisen Catchment explaining a Total of 88% of the variance in temperature values

To further deduct knowledge from the principal component analysis we should look at the component loadings, representing the weight/influence each of the time-steps has on the output produced by the PCA.

4.1 Component Loadings

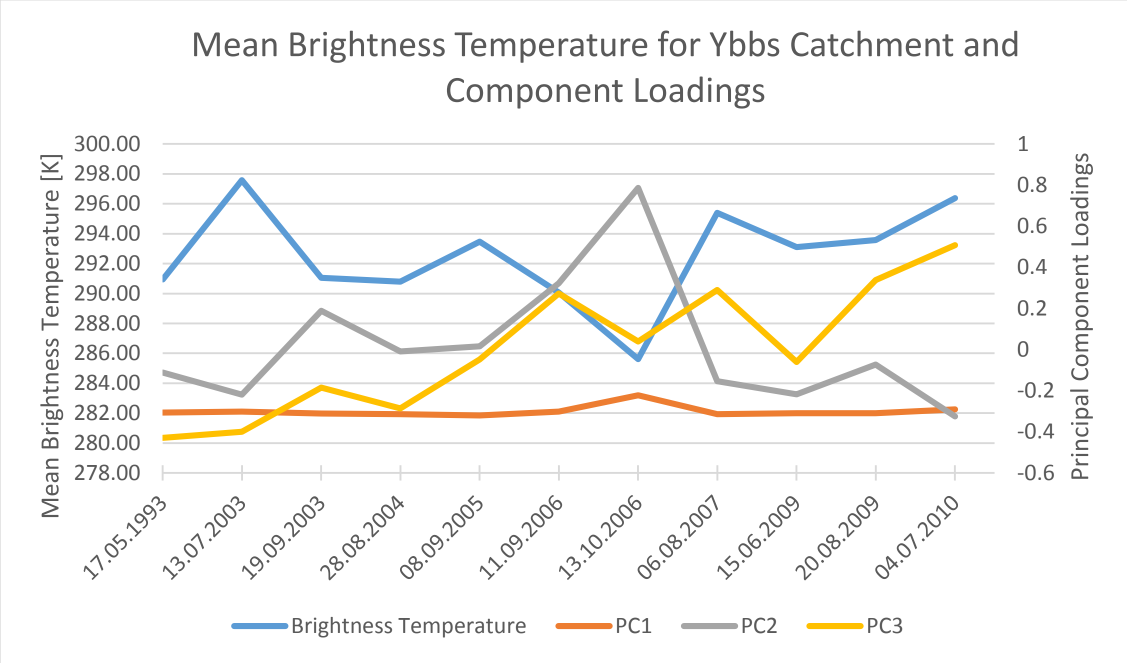

Figure 16: Brightness Temperature and Component Loadings (Y-axis) on the original TIR images (X-axis) of the first three principal components for the Ybbs Catchment. Mean Brightness Temperature for the catchment area was calculated for each acquisition date.

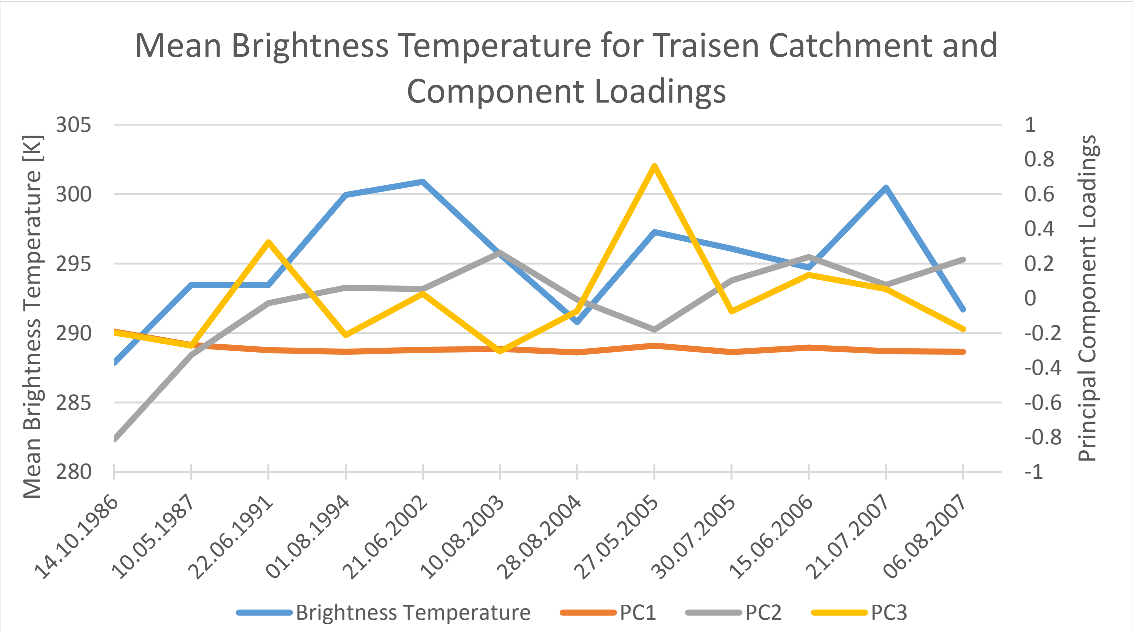

Figure 17: Brightness Temperature and Component Loadings (Y-axis) on the original TIR images (X-axis) of the first three principal components for the Traisen Catchment. Mean Brightness Temperature for the catchment area was calculated for each acquisition date.

Figure 16 and Figure 17 illustrate the correlation between each of the time-steps with the components being diagrammed. We see the influence of the time-steps to each of the components as well as the mean brightness temperature at each time-step. On the 13.10.2006 in Figure 16, we see a negative correlation with PC2 and PC3, which indicates that this time-step has a “latent” pattern in PC2 that is inverse to the one shown in PC3. Similarly, PC3 correlates positively with this time-step which indicates a similarity, even though it might be hidden or unapparent to some extent, of this time-step with the patterns in PC3. Analogical observations had been done in Figure 17, (i.e. time-step 27.05.2005). Another relationship that could be established is the one that becomes apparent by the loadings of Component 1. No seasonality can be observed, the loadings are consistent over the entire period. Furthermore, the relationship between principal component 1 and the major element explaining variability in temperature can be established, this being the one that occurs spatially. Nevertheless, before further analysing the patterns that are inherent in the components, the importance of a good distribution of the images for distinct conditions is shown in a sensitivity analysis, which is treated in detail in the following chapter. The idea was to analyse whether images with high influence on a component, show some special characteristics that might lead to a more profound interpretation of the components analysis.

4.2 Sensitivity Analysis

As shown in Chapter 4.1, specific images have specific influence on the different principal components. This raises the question what information within these images have these influences on the outcomes of the PCA and if more knowledge about the behaviour of PCA could lead to better results. This might point to a more efficient methodological approach, improving the representation of temperature variability in the first and more importantly in the consecutive components. To answer this question, a script was coded in R that automatically calculates principal components with lesser images and compares these with the initial principal component analysis calculated with 11, resp. 12 images. Absolute correlations of the means give a measure of similarity in information content between the consecutively calculated principal components ignoring images and the initial principal component.

Table 5: Mean absolute correlations and standard deviations between initial Principal Component Analysis and PCA with less input images. For each scenario, all possible variations of PCA's were calculated and mean absolute correlations as well as standard deviations with the initial PCA are shown for the first and second component. Traisen catchment

Traisen	Principal Component 1		Principal Component 2	
Scenario	Correlation	SD	Correlation	SD
1 Out	0.999	0.0003	0.968	0.0860
2 Out	0.997	0.0006	0.933	0.1230
3 Out	0.996	0.0009	0.892	0.1560
4 Out	0.993	0.0013	0.845	0.1860
5 Out	0.990	0.0019	0.790	0.2133
6 Out	0.987	0.0028	0.727	0.2384

Table 5 reveals a major behavioural aspect of the principal component analysis with time-series of granules: with decreasing input the information content of the first principal component remains nearly identical for all possible variations of initial time-steps. However, the information content and the patterns represented differ from the initial second principal component when other combinations are used for PCA. Analogical behaviour can be seen when comparing the effects of the inputs on the represented structures for the first and second component analysis for the Ybbs in the table below.

Table 6: Mean absolute correlations and standard deviations between initial Principal Component Analysis and PCA with less input images. For each scenario, all possible variations of PCA's were calculated and mean absolute correlations, as well as standard deviations with the initial PCA are shown. Ybbs catchment

Ybbs	Principal Component 1		Principal Component 2	
Mean Absolute	Correlation	SD	Correlation	SD
1 Out	0.999	0.0003	0.979	0.0521
2 Out	0.997	0.0007	0.953	0.0810
3 Out	0.995	0.0012	0.919	0.1128
4 Out	0.993	0.0020	0.876	0.1516
5 Out	0.990	0.0031	0.820	0.1917
6 Out	0.985	0.0049	0.750	0.2329

A principal component analysis was calculated with 9 TIR images and then compared with the initial principal component that was calculated on 12 images. Thus the algorithm consecutively dropped 3 images and calculated all possible variations of images. This allowed an examination on the behaviour of the first principal component towards its inputs. The idea behind this computation was the identification of images with stronger impacts on the patterns and structures represented in the first component. With further analysis, this would allow a better identification of the information content represented within PC1. Subsequently, we interpret that the first principal component shows nearly identical patterns when PCA is computed on 9 images or on 12. Comparing with Table 5, we see that even for a principal component calculated with only 6 initial TIR images, see half original Input, the information content stored in the first principal components correlates for 98% with the first principal component calculated on 12 images.

Figure 18 shows correlations and standard deviation between information stored in PC2, computed on 9 images, and the PC2, computed with the initial 12 images. When analysing Figure 18 we see that the second components of the PCA calculated without the first and second images of the initial image-stack merely correlate with about 50 % for the initial second component. When the second image of the initial image stack is used as an input for component analysis, the correlations rise to about 75%.

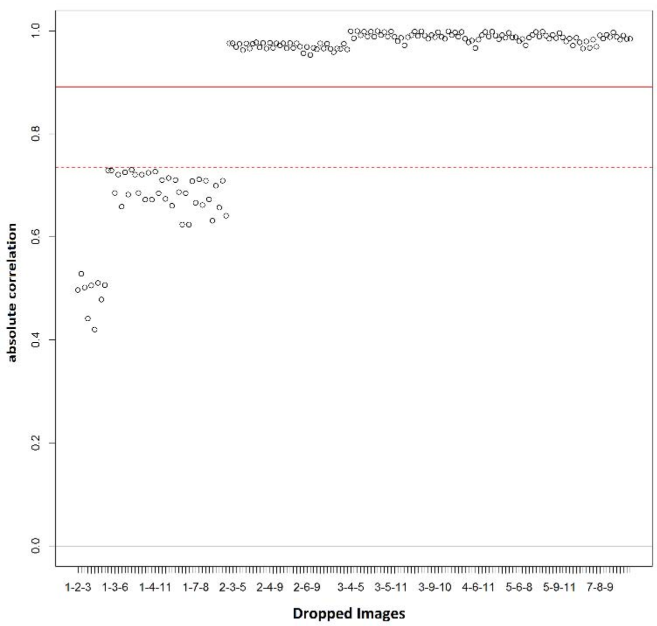

Figure 18 : Mean absolute correlations and standard deviations of all possible PCA's with 9 time-steps) as input for PC2 compared to Initial PCA (12 time-steps). Traisen.

As the further components only represented a proportion of variance of 2 to 3% of the total thermal variability, no further analysis was done on these components here.

As a consequence of these observations, it will be further investigated in the following chapters whether the first principal components might be a combination of the major surface characteristics that are spatially varying and controlling temperature variability across the landscape. This might be an explanation for the above seen behaviour in the sensitivity analysis. If this holds true, it might be stated that the second component would be the first "change" component, displaying change in temperature variability. As such the expectation of the second component displaying properties like thermal inertia and such deductible subsurface characteristics is held up here. To examine these expectations, the distributions of the patterns represented in the components were compared to the patterns represented in datasets of known surface and subsurface characteristics influencing hydrological behaviour.

4.3 Characterising Structure and Heterogeneity

Kernel densities were estimated using the density function provided by RStudio (R Core Team, 2014). This allows for a further interpretation of the patterns represented within the component values and whether these patterns fit the heterogeneity of a specific characteristic in the catchment, e.g. relationship with topography, land cover or geological variables. Furthermore the modal values of the distribution functions are displayed in the rug of the graphs allowing a better visualization and interpretation of clustered distributions. To calculate accurate kernel densities, the raster files displaying the specific surface and subsurface characteristics needed to be reprojected to match the resolution of the principal components. This was done using the reproject function of the raster package in R (Hijmans & van Etten, 2012; R Core Team, 2014). Several tests concerning the projection method were done. No significant difference could be observed in the resulting kernel densities except for the texture data. While all classified nominal values were interpolated using the nearest neighbour method, all continuous variables were reprojected with bilinear interpolation. All kernel densities were calculated on a 30 m grid raster. Except the probability distribution functions of the soil texture file were calculated on a 10m grid cell size as too much information would have been lost through the interpolation process.

4.3.1 Principal Component 1

Topography

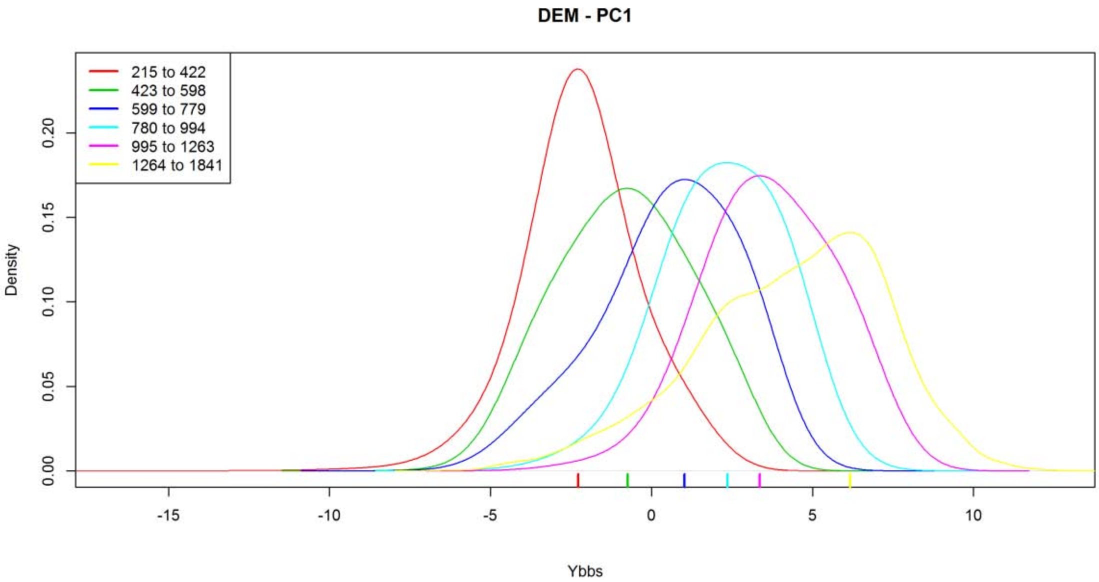

Figure 19: Density Plot for the Principal Component 1 and Digital Elevation Model for the Ybbs Catchment with mode displayed at the bottom. Bandwidth: 0.8.

Figure 19 shows the kernel density distribution of the first Principal Component for the Digital Elevation Model. The DEM was classified with k-means, a clustering algorithm that is aiming a partitioning of observations into k clusters with each observation being classified into the cluster that has the closest mean (MacKay et al., 2003). We see the distinct peaks of the probability distribution functions, meaning that the component values of the first principal component are representing the temperature variability influenced by the heterogeneity of the digital elevation model. The patterns represented in the component values result from the highest proportion of variance explained, that are being assumed to be the temperature variability caused by the heterogeneity of the catchment landscape. The influence of the digital elevation model on the landscape heterogeneity causing temperature changes is graphically visualized by the well spread distribution functions.

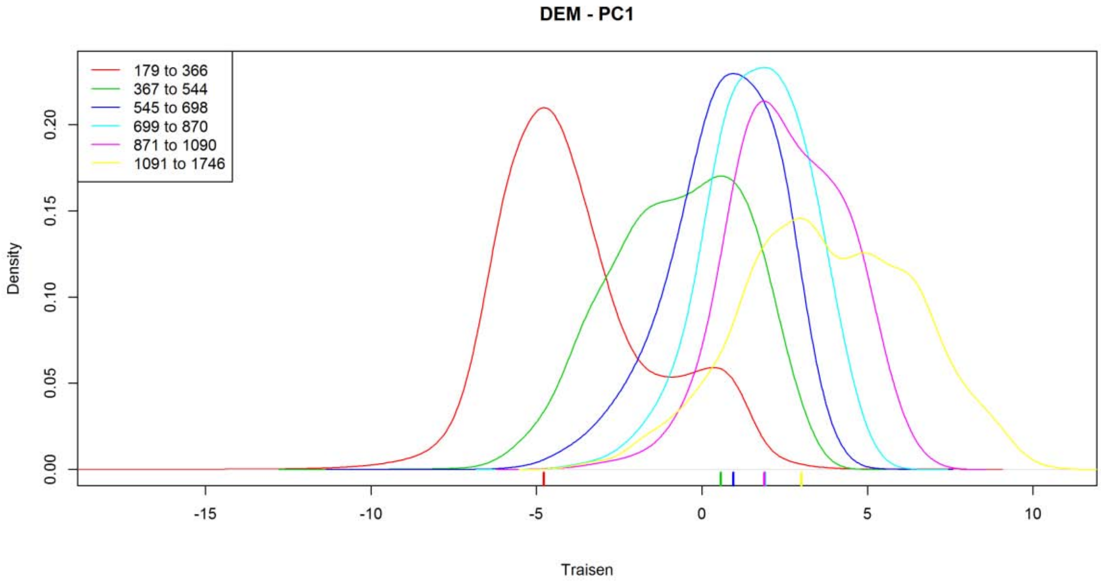

Figure 20: Density Plot for the Principal Component 1 and Digital Elevation Model for the Traisen Catchment with mode displayed at the bottom. Bandwidth: 0.8.

Figure 20 shows the density plots for the Traisen catchment. We see the distinction of the first class of the DEM which reflects the valley of the catchment. Within higher regions the distinction of topographical characteristics within the first principal component is less apparent, but yet significant enough to reveal the importance of elevation on the variability of temperature.

Land Cover Characteristics

Figure 21 shows the kernel densities for the Ybbs catchment according to their land cover characteristics. The Corine Land Cover Map was clipped to the extent of the catchment and classified into the 6 most dominant classes. We see the distinct peaks of the densities, showing the strong relationship between PC1 and land use characteristics.

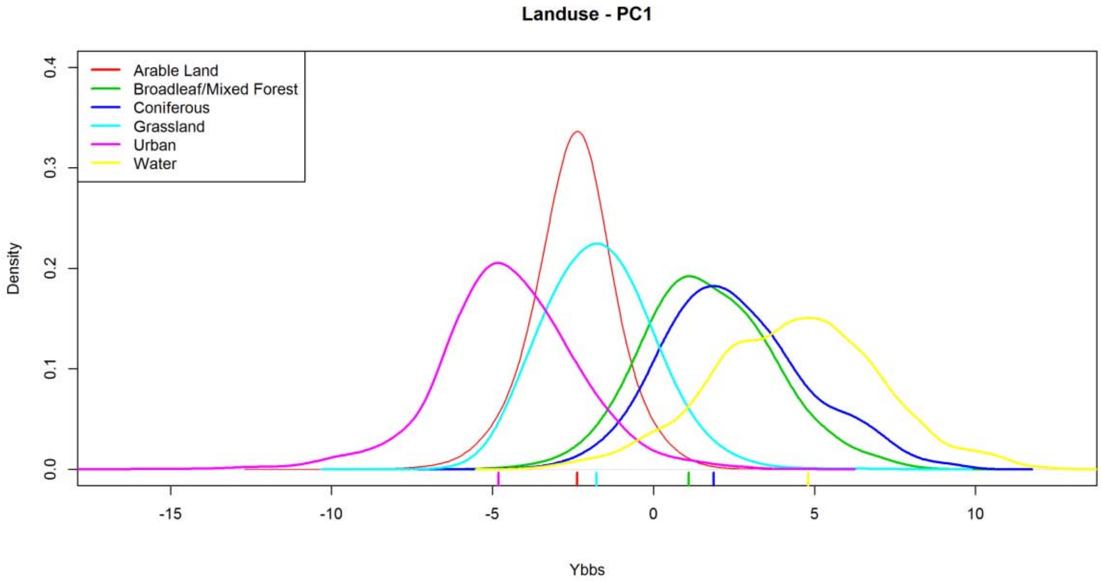

Figure 21: Density Distributions and corresponding modal values of Principal Component 1 for Landuse characteristics of the Ybbs catchment. Bandwidth= 0.4.

The strong influence of land cover characteristics on the temperature variability patterns becomes visible here: urban areas have a completely different behaviour than water related to temperature variability. The other patterns, in between these very distinct classes related to temperature change, are also well defined. Grassland and arable land are somewhat behaving similarly and the forested areas are less well distinguishable from each other. However the modal values indicate that land cover information is still extractable from the nominal values of the first component.

The same estimation of distribution functions was computed for the first principal component and the land cover classification of Traisen catchment, displayed in Figure 22.

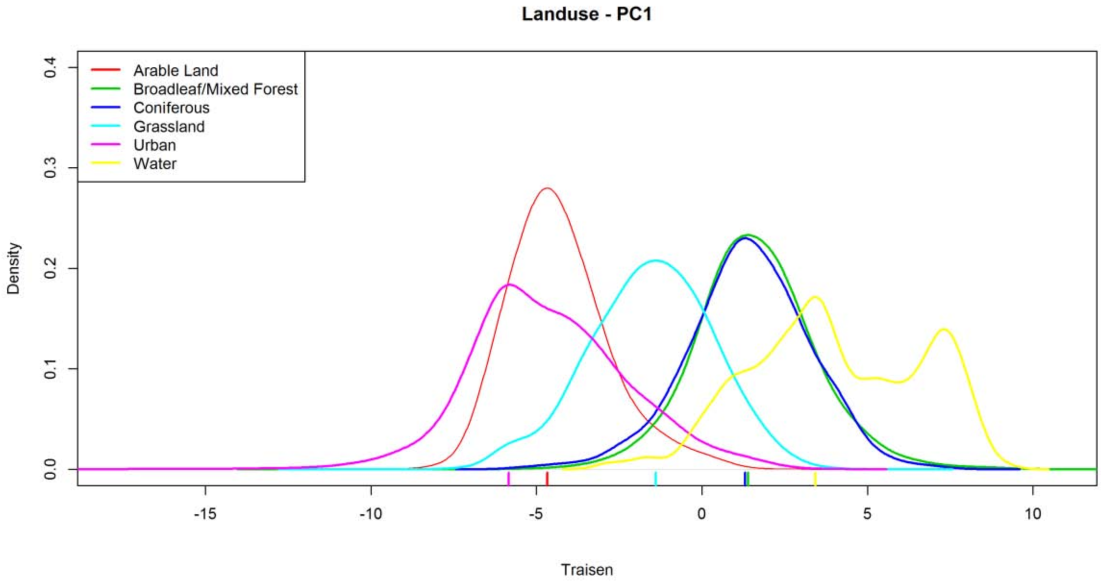

Figure 22: Density Distributions and corresponding modal values of Principal Component 1 for Landuse characteristics of the Ybbs catchment. Bandwidth= 0.4.

We can see in Figure 22, that for the Traisen catchment the distributions for the land cover characteristics are well distributed. Distributions for agriculture, urban areas and pastures are well distinguishable, whereas broadleaf forest and coniferous forests are clustered and strongly overlapping.

Soil Texture Characteristics

In the next step, it was tested whether soil hydraulic properties were identifiable. The dataset used for comparison was to a greater extent incomplete for both regions and only very dispersed point measurements were at hand. Consequently no patterns could be found due to these incomplete data. However a dataset with soil texture was at hand covering up to 44 % of the overall area of the Ybbs catchment and 30% of the Traisen catchment with soil texture characteristics information. This was the best dataset at hand for soil characteristics and can be seen in the ANNEXE II. Its interpretation should be done carefully due to the incompleteness of the dataset. However due to the assumed strong link between soils texture characteristics, its associated hydraulic properties and thermal properties of the soil (see Chapter 2.4), the inherent patterns are compared here.

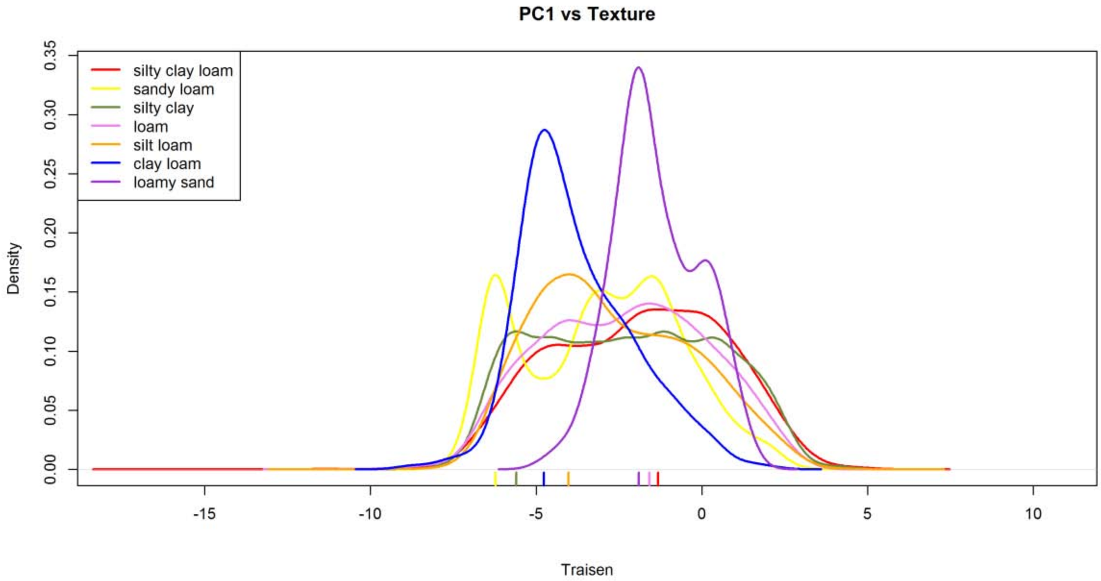

Figure 23: Density Distributions and corresponding modal values of Principal Component 1 for Soil texture characteristics of the Traisen catchment. Bandwidth= 0.4.

Figure 23 shows the patterns in the component values, which are not well representing texture characteristics. Silt loam and loam are the most dominant soil textures in the areas covered by the dataset in both catchments. Nevertheless, while clay loam and loamy sand are the most dominant patterns identifiable in the first component, the rest of the classifications could not be determined nor classified using the first principal component.

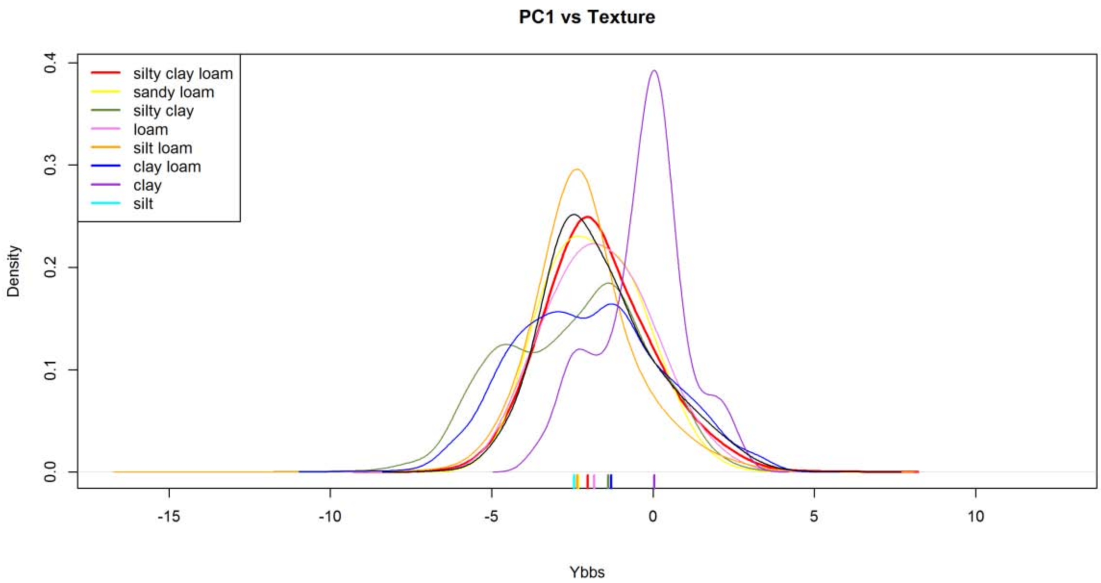

Figure 24: Density Distributions and corresponding modal values of Principal Component 1 for Soil texture characteristics of the Ybbs catchment. Bandwidth= 0.4.

The component values show similar poor representations of soil characteristics for the Ybbs catchment. While silt loam and loam represent the majority of the textures observable in the catchment, none of this information could be identified in the allocation of the component values.

Hydrogeology

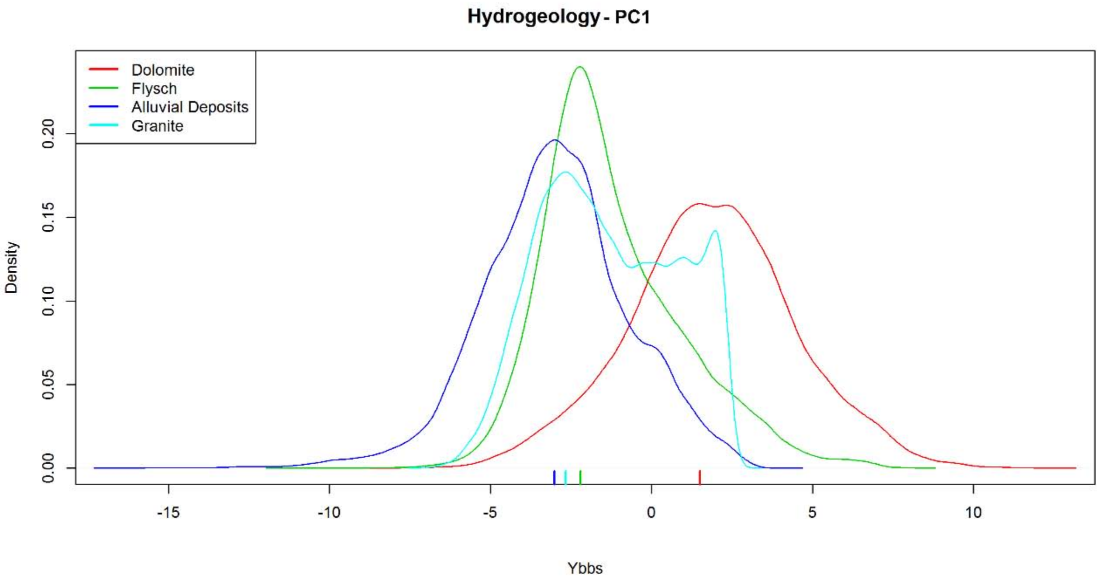

Figure 25: Density Distributions and corresponding modal values of Principal Component 1 for Hydrogeological characteristics of the Ybbs catchment. Bandwidth= 0.4.

Figure 25 shows the conjugated characteristics of hydrogeology and the first principal component for the Ybbs catchment. We see the relation of Flysch and alluvial deposits through the bounded distributions and the strong distinction towards dolomite. Noticeably, granite shows two peaks.

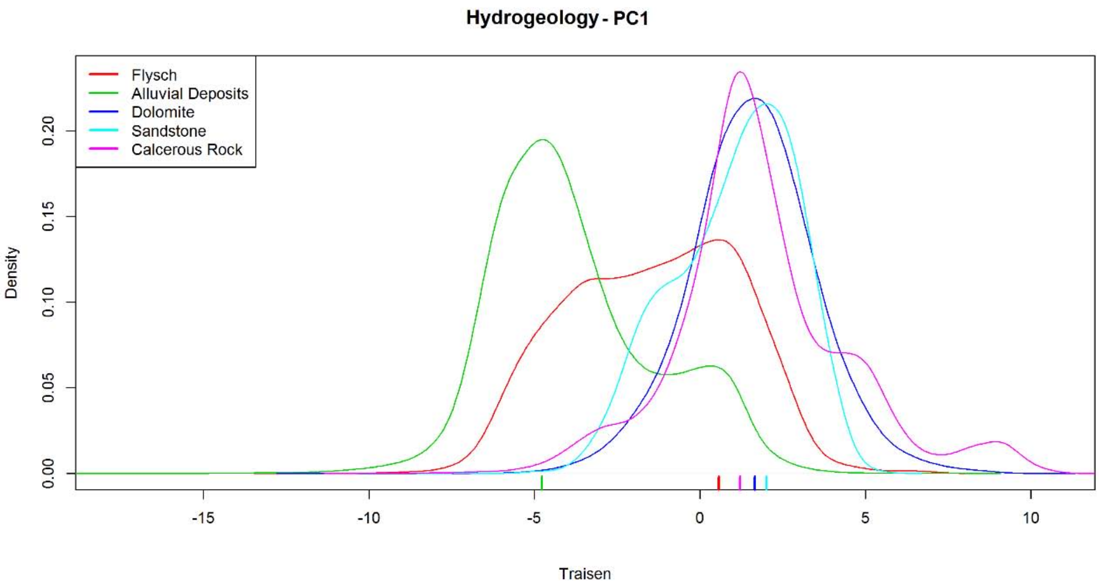

Figure 26: Density Distributions and corresponding modal values of Principal Component 1 for Hydrogeological characteristics of the Ybbs catchment. Bandwidth= 0.4.

Figure 26 shows the density distributions for the component values of PC1 of the Traisen catchment and its hydrogeological characteristics. The distinction of the classes is a difficult task in this case, as only alluvial deposits can be significantly differentiated from the other classes. A strong tendency in representing valley and mountainous areas is observable. When visually inspecting the hydrogeological map, it is outstanding that the classes of dolomite, sandstone and calcareous rocks are the subsurface characteristics of the Pre-Alpine environment. While Flysch shows some distinguishable behaviour in the component values, the rest of the hydrogeological characteristics are not distinctly represented in the first principal component.

Slope

Suspicious about the influence of the mountainous regions of the catchments on the patterns represented in the component values, the slope in degree rise was calculated from the digital elevation model using the slope-tool in ArcMap. It was suspected that some agglomeration patterns might be represented in the patterns.

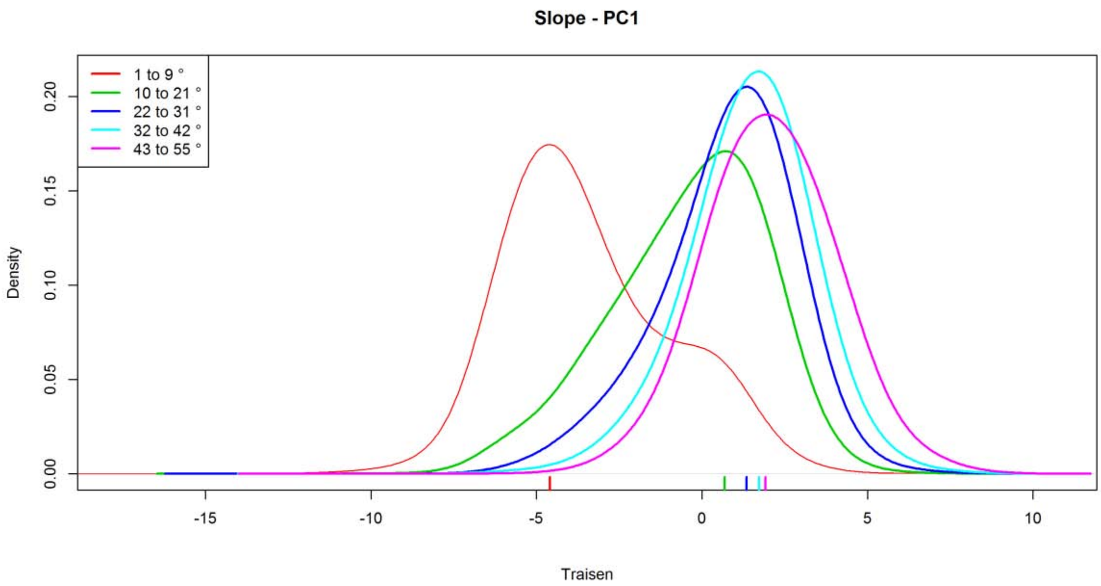

Figure 27: Density Distributions and corresponding modal values of Principal Component 1 and slope in degree rise over run of the Traisen catchment. Bandwidth= 0.4.

The classifications of the continuous slope variables were done with k-means in Rstudio using the "classINT" package (Bivand, 2015) and allowed for a comparison of the first principal components and slopes. For the Traisen catchment, we see the strong distinction of the valley to the Pre-Alpine environment in the component values. The valley shows vastly different hydrological behaviour than the Pre-Alpine environment. It can be noticed that throughout all the probability distribution functions that were done to this point this behaviour in the spatial distribution of the component values holds true for the first principal components.

In Figure 28 we see that the distinctions are similarly difficult for the first component of the Ybbs catchment. Here we have a bigger valley with more degradations in relief, resulting in a less abrupt rise in elevations. This is to some extent reflected in the first principal component, with the highest peak being a reflection of the flattest part of the valley in the north of the catchment. The modes of the other distribution might still represent a significant allocation of the slopes to the principal components. Nevertheless, when the slopes rise in the mountainous regions, a differentiation can no longer be done.

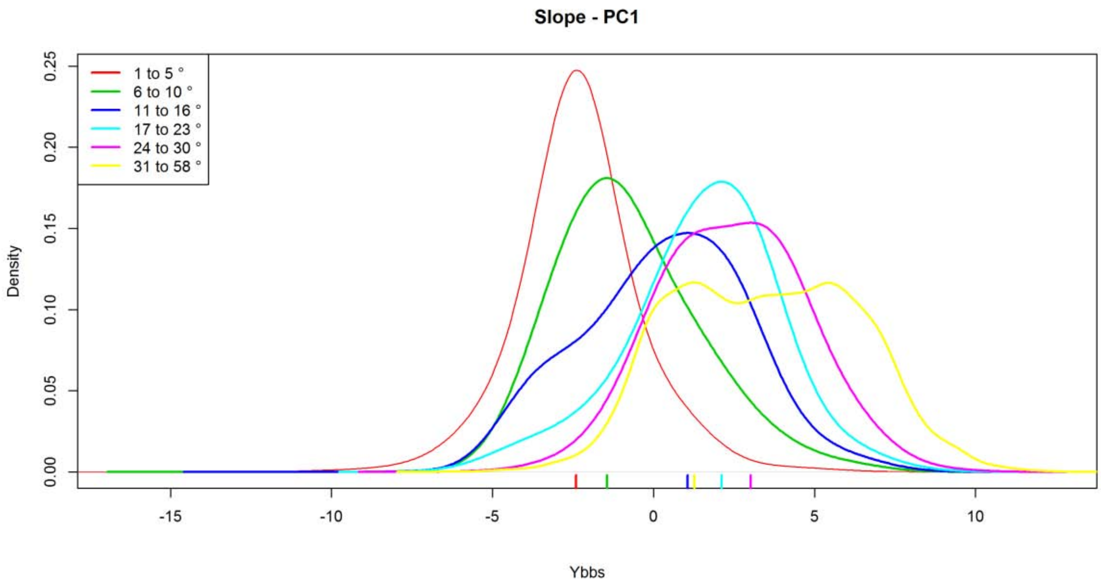

Figure 28: Density Distributions and corresponding modal values of Principal Component 1 and slope in degree rise over run of the Ybbs catchment. Bandwidth= 0.4.

Aspect

To see whether patterns related to aspect and effects of shadowing could be found in the first principle components, aspect was calculated within ArcMap from the digital elevation model at hand.

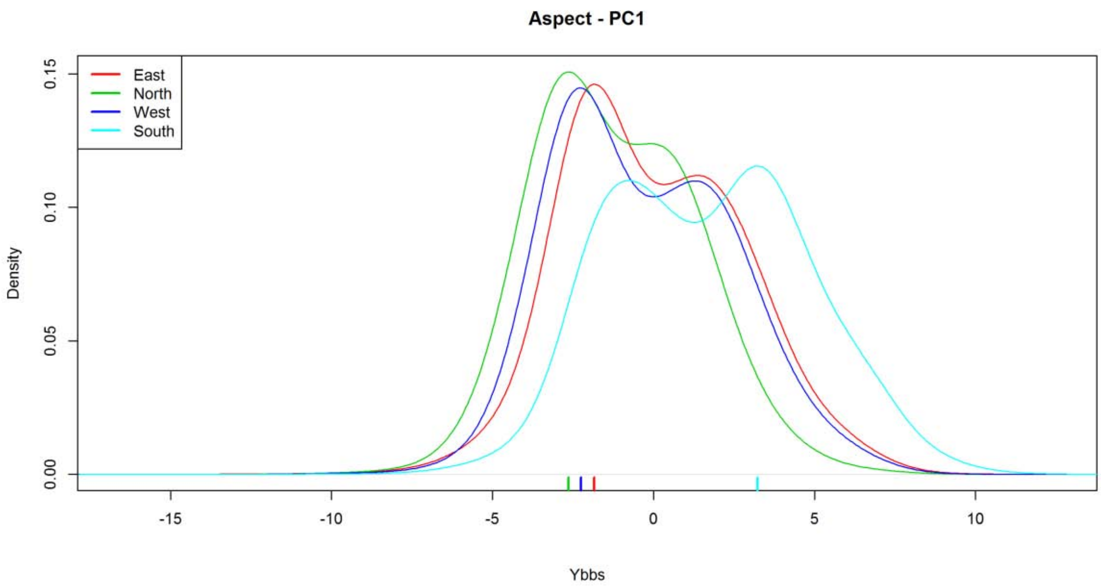

Figure 29: Density Distributions and corresponding modal values of Principal Component 1 and Aspect for the Ybbs catchment. Bandwidth= 0.4.

We can see in Figure 29 that there is a strong differentiation in the values of the first principal component for North-, and then South-facing areas of the Ybbs catchment.

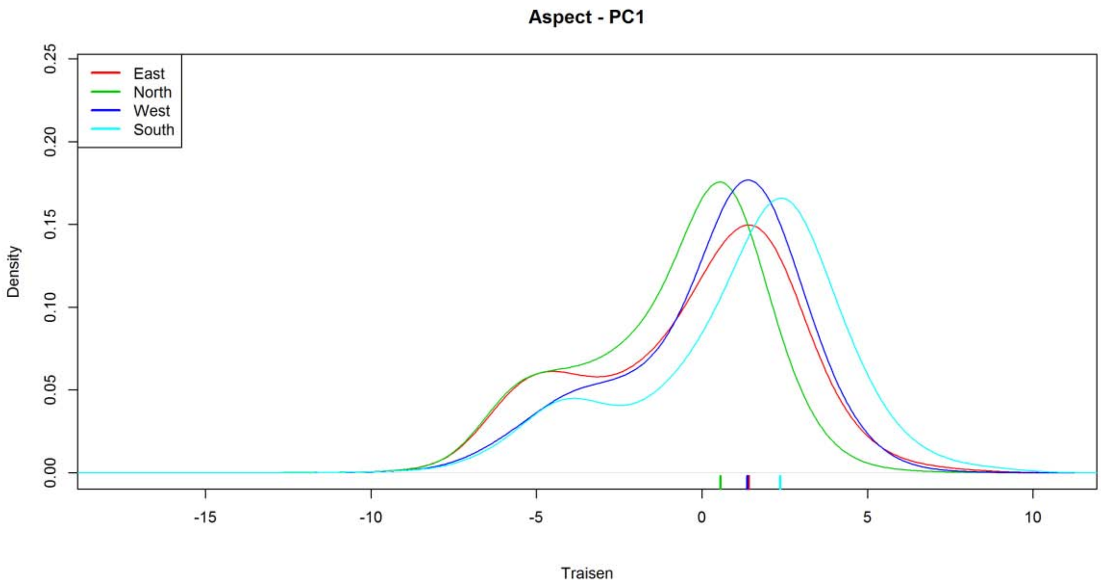

Figure 30: Density Distributions and corresponding modal values of Principal Component 1 and aspect for the Traisen catchment. Bandwidth= 0.4.

The comparison of the patterns for the calculated aspects and the first principal component of the Traisen catchment, shows similar results to those of the Ybbs catchment. While north and south facing parts of the catchment are clearly identifiable in the values of the first components, it becomes apparent that the east and westward facing parts of the catchment are not distinctly represented in the component.

Intersections

Further the experiment was done to separately calculate probability densities for land cover heterogeneity in the first PC in the valley and the alpine regions. Results can be seen in Figure 32, 33, 34 and 35. The distinction of the flat and mountainous region was done on behalf of the major changes in geology of the catchments, this being from Flysch to Dolomite.

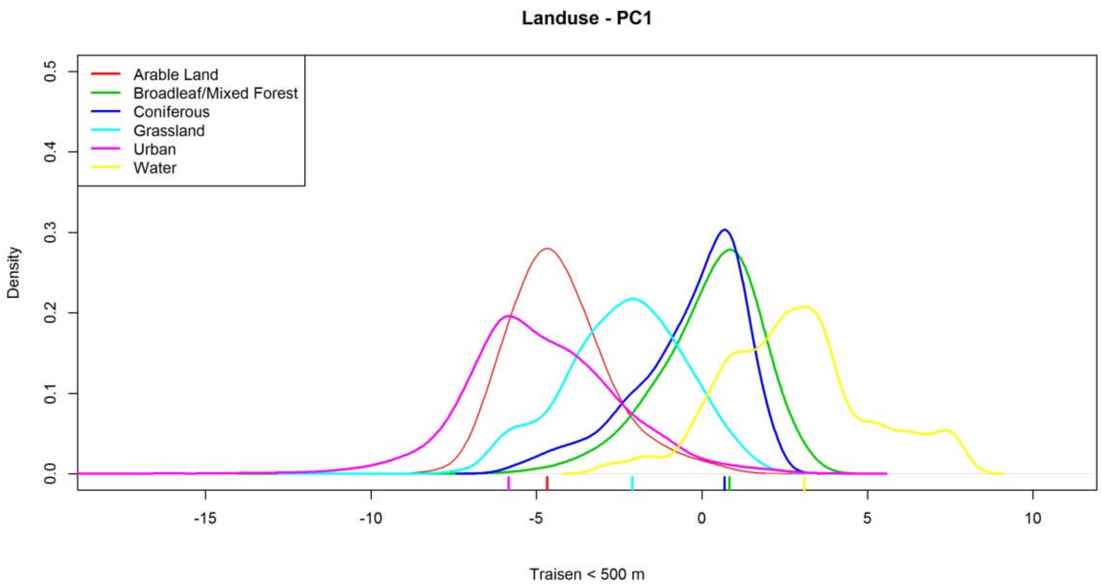

Figure 31: Density Distributions and corresponding modal values of Principal Component 1 and Land cover characteristics for the valley of the Traisen catchment. Bandwidth= 0.4.

The probability distributions of Figure 31 show the good representation of all land cover classes for the flat area of the Traisen catchment. The main constraint is that no differentiation can be done between mixed/broadleaf forest and coniferous forest.

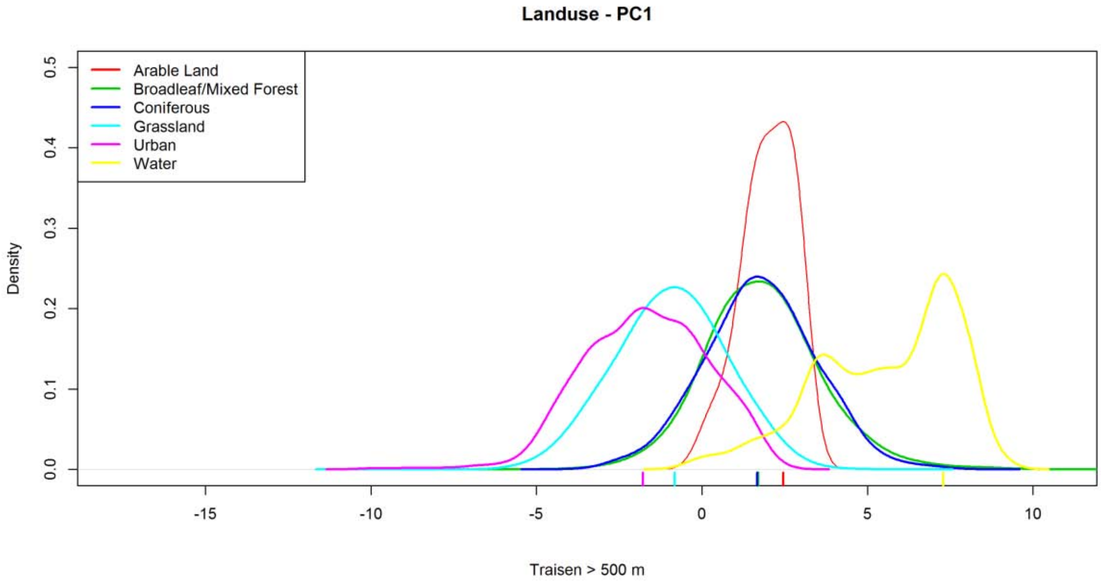

Figure 32: Density Distributions and corresponding modal values of Principal Component 1 and Land cover characteristics for the mountainous environment of the Traisen catchment. Bandwidth= 0.4.

The most striking point to be addressed in this visualization of the pattern similarity is that no distinction can be made between arable land and forests. Furthermore, the land cover class "arable land" shifted from negative component values in the valley to positive component values in the mountainous environment.

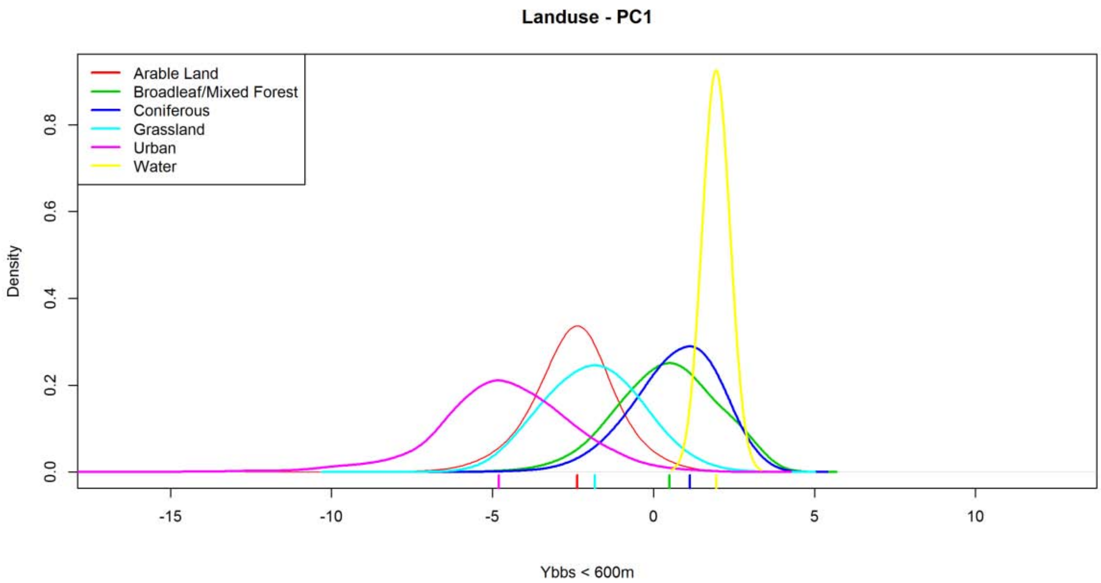

Figure 33: Density Distributions and corresponding modal values of Principal Component 1 and Land cover characteristics for the valley of the Ybbs catchment. Bandwidth= 0.4.

The similarity in the patterns for the flat environment of the Ybbs catchment and the first component of the temperature variability analysis is striking. While all classes are well distinguishable, the distribution of the classes from negative values for urban and arable land over grassland towards positive component values for less varying areas like forests and water shows that the patterns of the valley are strongly linked to land cover related changes. The first component of the mountainous area of the Ybbs catchment still shows some good variability in patterns. However, the patterns are less distinct as can be seen in Figure 34.

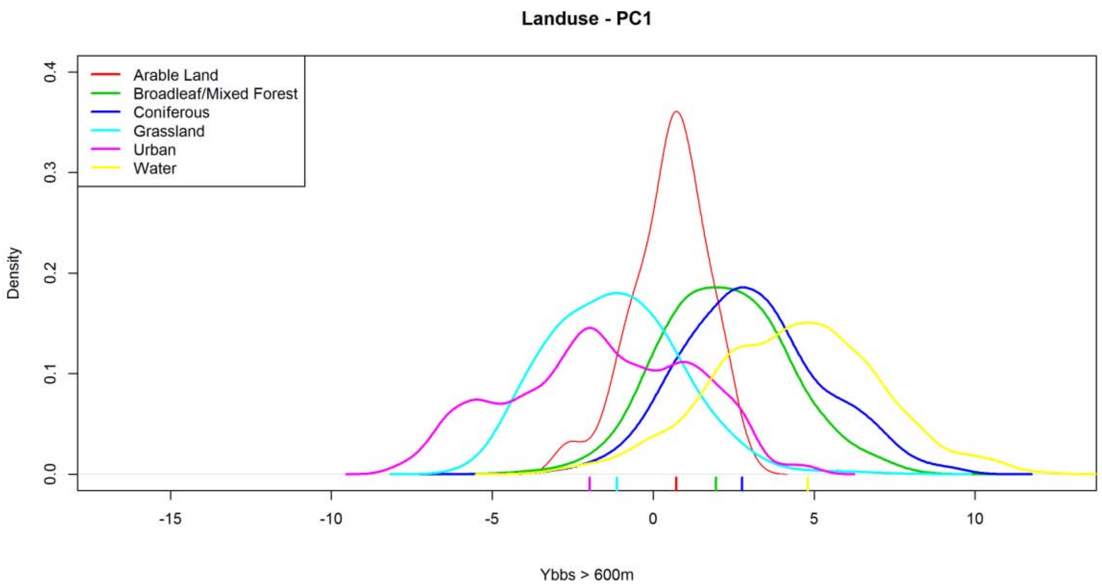

Figure 34: Density Distributions and corresponding modal values of Principal Component 1 and Land cover characteristics for the mountainous environment of the Ybbs catchment. Bandwidth= 0.4.

While the distribution of the classes is still well represented in the component in question, the class of the urban area is less well distinguishable in the mountainous region, which is a small class in comparison to the other classes in the mountainous area and the urban space of the valley. Strikingly, the arable land is again differently represented in the patterns of the mountainous environment of the component compared to the patterns in the flat areas.

4.3.2 Principal Component 2

The distributions were identically computed for the second principal component of the time-series analysis of the catchments in question. The best resemblance in the distribution of patterns could be found for major subsurface characteristics, as no other distributions showed a potential qualification for hydrological pattern differentiation. It should be added that the gradient of the component values of PC2 is considerably smaller than the gradient of the component showing highest proportion of variance. As a consequence, a well distinct distribution of probability functions is a difficult task. The first salient difference between the two second principal components is that the ranges in the component values are inverse to each other. As the sign of the eigenvectors is arbitrarily assigned there is no affect on the structures represented in the components (Soliman, Brown, & Heck, 2011).

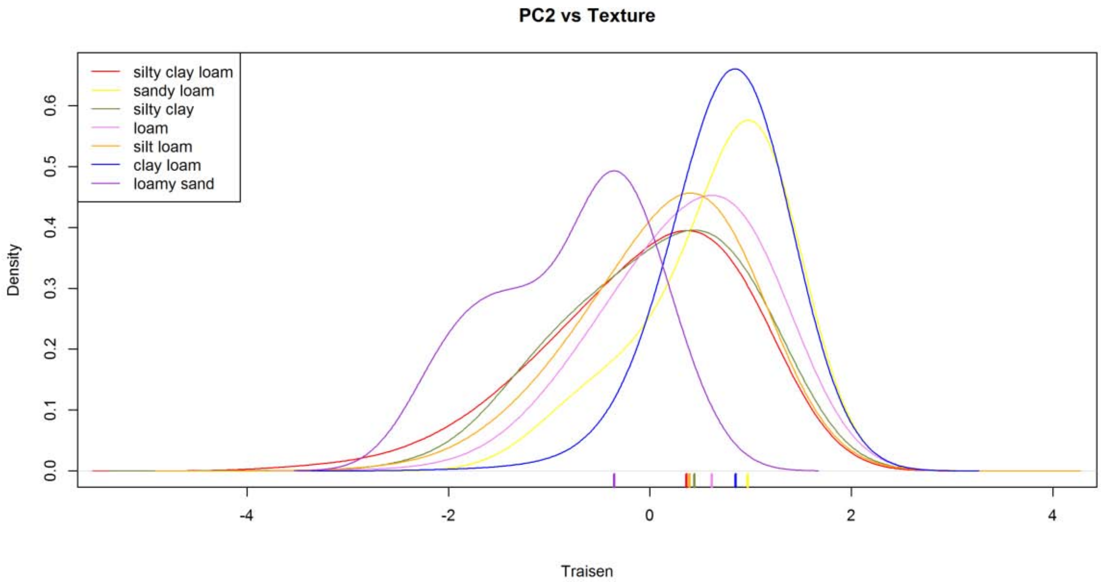

Figure 35: Density Distributions and corresponding modal values of Principal Component 2 for soil texture characteristics of the Traisen catchment. Bandwidth= 0.4.

For illustration purposes, the densities of the texture classification being represented in the distributions of the values of PC2 are shown, as it was assumed that texture would show a high resemblance in the patterns due to the thermal characteristics emerging from soil texture and their hydraulic capacity. Nevertheless, no information could be deduced from PC2 allowing an extrapolation towards texture characteristics. The distributions of the Ybbs catchments with regard to pattern similarity can be seen in Figure 36.

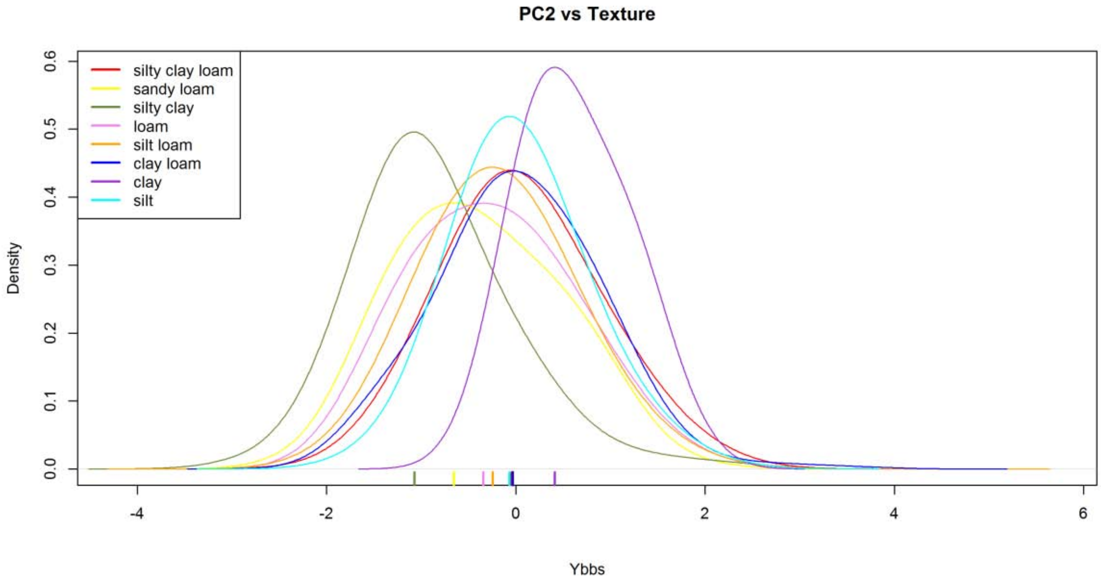

Figure 36: Density Distributions and corresponding modal values of Principal Component 2 for Soil texture characteristics of the Ybbs catchment. Bandwidth= 0.4.

We see that mainly clay, silt and silty clay are distinguishable. While clay is only represented in the dataset for 0.03 % of the covered area, silt and silty clay are merely representing 3 and 5% of the dataset respectively. As for the dominant soil texture classes, that being silt loam, loam and silty clay loam, no recognition can be found. Even though, Zaheer and Iqbal (2014) have recently displayed the potential of deducing percentages of silt and clay in the soil from TIR imagery, this could not be proven for PCA with TIR time-series. According to their study, a comparison with the content of organic matter in the soil could as well have been fruitful, but no data to validate this assumption could be found.

Next to this, the structures represented in PC2 showed coherence with the patterns of the hydrogeological distributions. As illustrated by Figure 37, the probability distribution functions and their corresponding modal values, would allow an extrapolation from the hydrogeological characteristics to continuous pixel based values based on the structures represented in PC2. We see that the component values of the second component represent the features of hydrogeology.

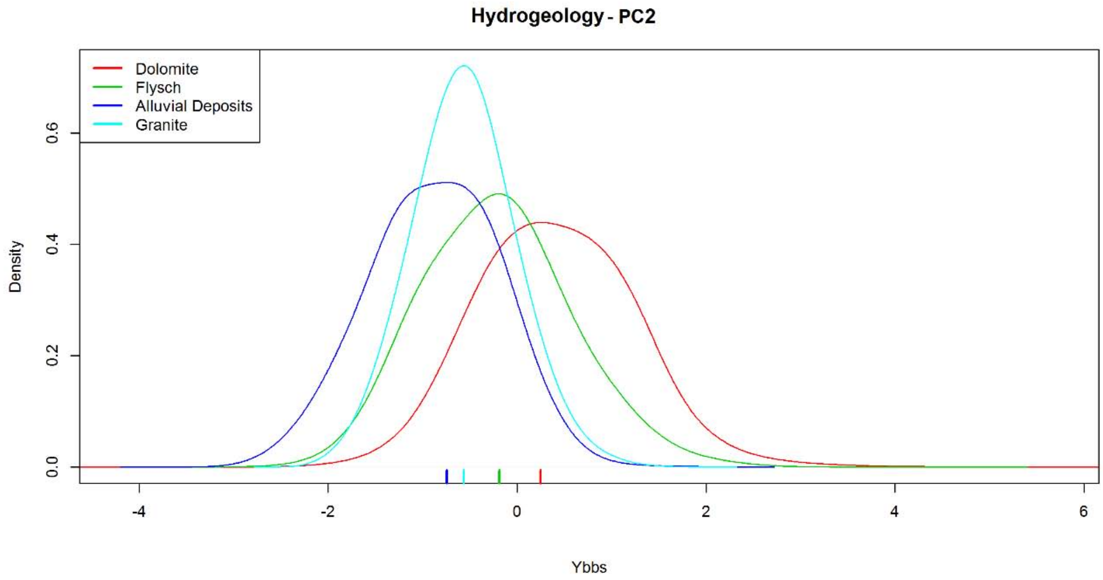

Figure 37: Density Distributions and corresponding modal values of Principal Component 2 for hydrogeological characteristics of the Ybbs catchment. Bandwidth= 0.4.

Figure 38 further displays the potential for an extrapolation of the hydrogeological characteristics from component values of PC2. As the hydrogeological data is to some extent interpolated and hence loses some of its physical relationship with reality, PC2 could potentially become the more accurate indicator for hydrogeology with further development of TIR remote sensing techniques. However, field measurements would be needed to validate the assumption.

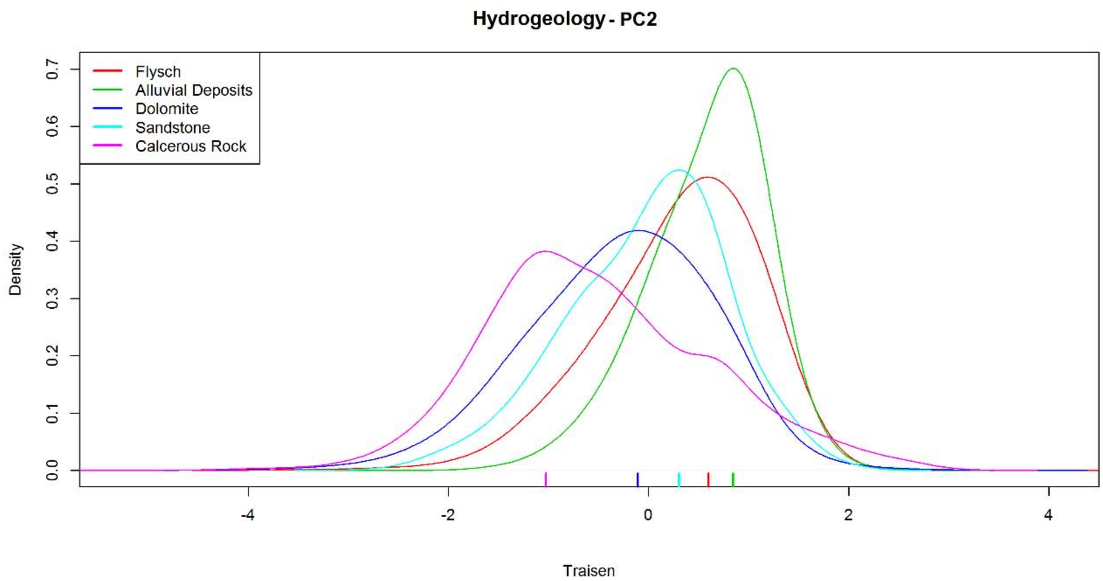

Figure 38: Density Distributions and corresponding modal values of Principal Component 2 for hydrogeological characteristics of the Traisen catchment. Bandwidth= 0.4.

The visual inspection of the second component's distribution showed some similarities with characteristics as aspect or slope. Hence, to validate these first visually assumed similarities, densities for aspect and slope were calculated for PC2. As it is assumed, that PC2 is the first component actually showing change in the spatial variability patterns, it was reasoned that some subsurface characteristics like agglomeration or soil water content, dependant on shadowing or/and inclination could be represented in the distribution of the component's values.

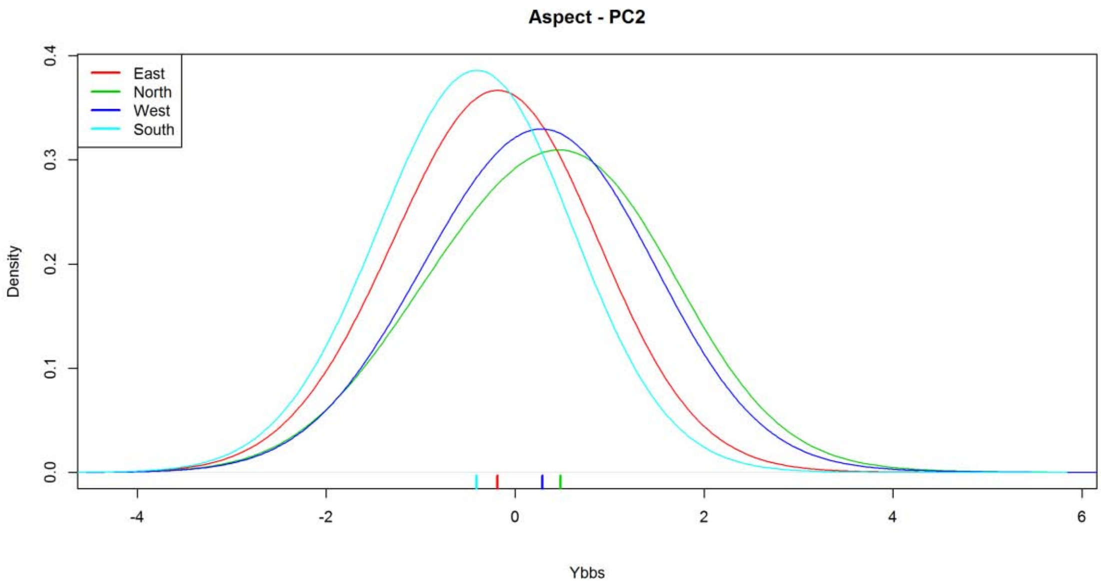

Figure 39: Density Distributions and corresponding modal values of Principal Component 2 and aspect for the Ybbs catchment. Bandwidth= 0.4.

Figure 39 shows the distributions of aspect within the structures of the component values. The strong distinction in component values for north and south inclined mountain ridges is remarkable. Furthermore, east and west facing land surfaces are distinctly represented. Consequently, we can assume that some information about the aspect has an effect on change on temperature variance, as has been confirmed by Zaheer & Iqbal (2014). Figure 40 shows similar patterns in the representation of aspect in PC2. The main difference is represented in the adverse component values.

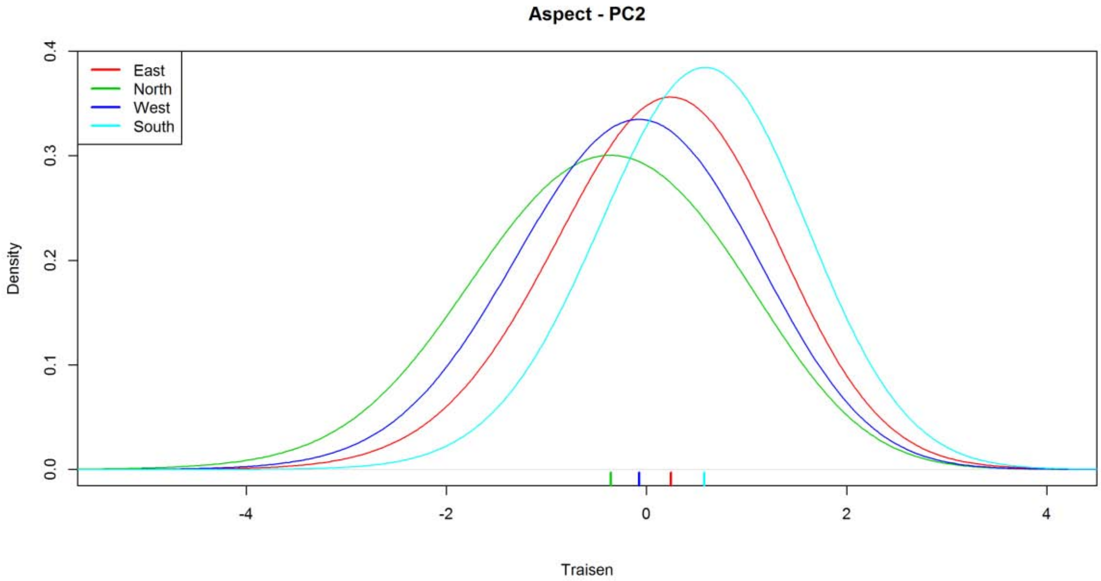

Figure 40: Density Distributions and corresponding modal values of Principal Component 2 and aspect for the Traisen catchment. Bandwidth= 0.4.

Though expected, the patterns of the slopes are not represented in the second principal component. The main distribution identifiable is the one representing the valley next to the clustered classes representing the mountainous environment as can be seen in ANNEXE III.

4.3.3 Principal Component 3

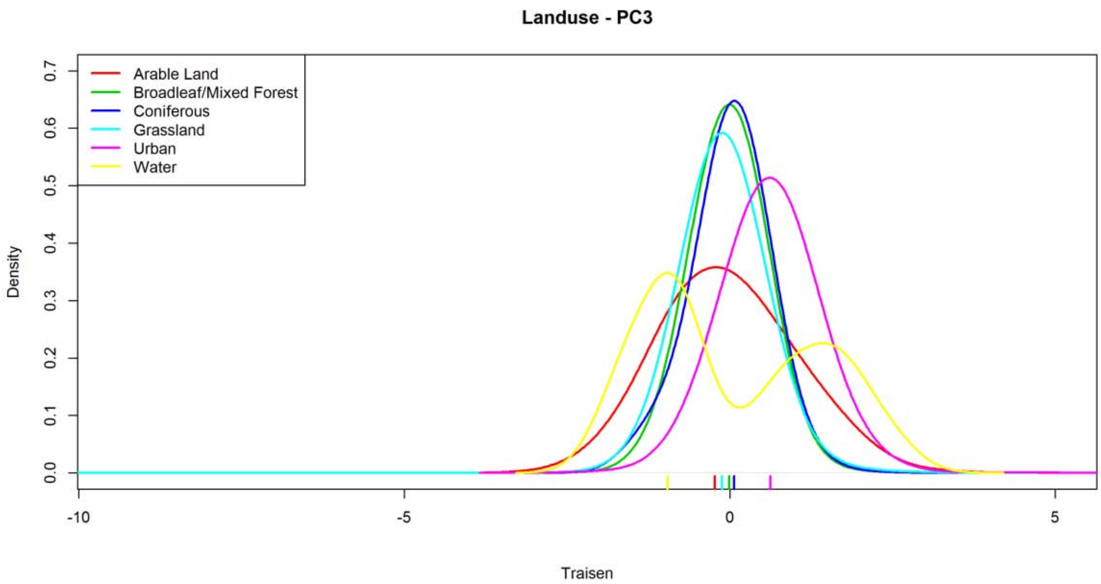

Figure 41: Density Distributions and corresponding modal values of Principal Component 3 for Landuse characteristics of the Traisen catchment. Bandwidth= 0.4.

The slight consistence of patterns in the distribution of the component values of PC3 could be found with land cover characteristics. The most notable effect that can visually be recognized is that the patterns show a different behaviour for urban areas. As such urban areas are somehow distinctly represented in the third principal component.

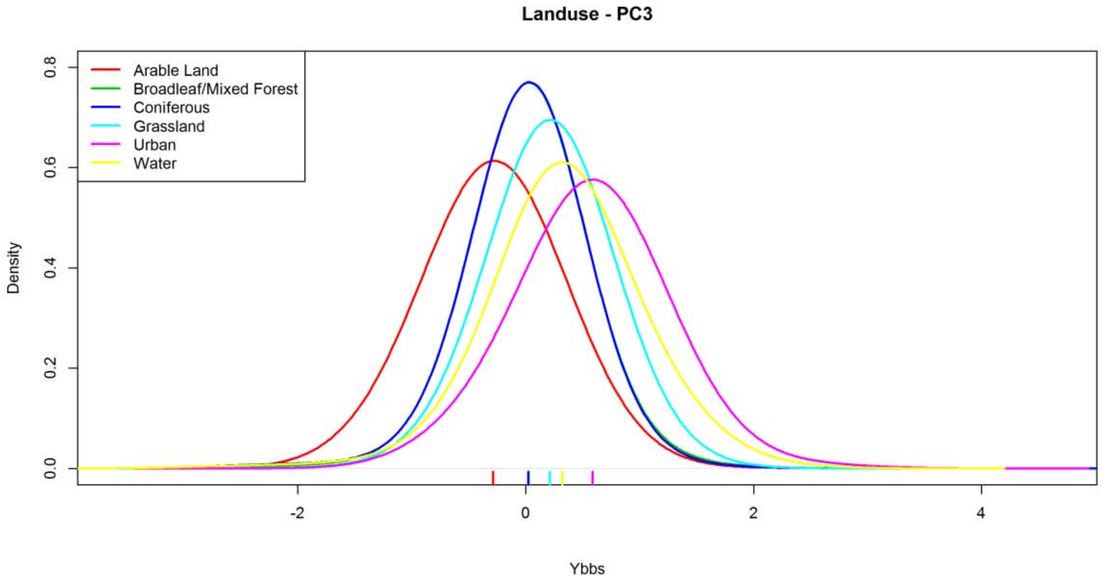

Figure 42: Density Distributions and corresponding modal values of Principal Component 3 for landuse characteristics of the Traisen catchment. Bandwidth= 0.4.

The rest of the components account for less than 2% of total variance and no structures in the distributions of the component values could be identified. As of this reason, the rest of the components were considered to be representing noise in the dataset and were as such not further analysed. As a strong relationship between thermal inertia characteristics of the catchment soil and the second principal component was expected but the availability of the data wouldn’t allow for a good distribution of the thermal images for distinct meteorological conditions, this could not be shown in the distributions.

5 Discussion

By processing TIR imagery with principal components it was possible to identify spatial patterns by relating them to the most dominant landscape elements known to be affecting the thermal characteristics at the surface and subsurface. As such the first component showed a mix of the major landscape characteristics in the catchments having an effect on temperature variability. These being the land cover characteristics and topography. It was further shown that relief properties like aspect are inherent in the patterns of PC1 and that the nominal values are differently related to landscape characteristics in mountainous areas as for flat areas.

The results further show the patterns in the second component having similarities with the distributions of hydrogeological characteristics.

The third component showed patterns revealing a “latent” influence of land cover characteristics, see urban areas. The deciding reason for this behaviour in PC3 is a matter of assumptions as no data for validation are available. However it can be discussed that some atmospheric attenuation could have been causing the signals in the urban areas. Or some change in temperature variations in the cities might have caused these patterns. It was shown that in the first principal component the influence of relief and shadowing was not to be neglected when analysing patterns of thermal variability.

The rest of the components displayed less than 2% of the total variance and were, after no structures could be visually identified, considered as noise.

The differences in the components between the catchments could mainly be revealed in the comparison of their patterns with their elevation properties. While the topography was immanent in the component values of the first component of the Ybbs, the distinction of elevation characteristics was less apparent for the first component of the Traisen catchment. This could be explained through the different geomorphological characteristics of the two catchments. The Ybbs catchment has a broad plain in the north potentially influencing the variances in temperatures distinctly from the Traisen catchment, characterised by a narrow floodplain towards the drainage in the north.

5.1 Sensitivity Analysis

The sensitivity analysis of the quantity of the input images towards the information content on the output, revealed the importance of well distributed images for the second component. The first component represented the same patterns when comparing the ones of computations with no more than 4 time steps, with the ones computed with the initial full dataset. It could be pointed out that the first principal component stays unaffected of the temporal thermal variance, i.e. seasonal changes, that are immanent in the time-series. Consulting the component loadings to this effect, it was shown that these stay consistent over the entire period and display no seasonality. To that effect it was revealed that the first component display patterns of thermal variability that occur spatially, rather than temporally. Consequently, the second component was the first component influenced by the seasonal distribution of the time-series.

This effect can also be seen in fluctuations of the component loadings in contrary to the equally distributed loadings for PC1. As such the second component was considered as being the first "change" component, displaying temporal thermal variability. We further see, that the second components of the PCA calculated without the first and second images of the initial image-stack merely correlate with about 50 % for the initial second component. When the second image of the initial image stack is used as an input for component analysis, the correlations rise to about 75%. Nevertheless, the correlations of the information within PC2 are still lying below their standard deviation. As a result, an analysis of the first and second images needed to be done to further be able to correctly interpret these results. This showed that while the first image represents the coldest day of the whole dataset, the second image shows cold conditions on a small temperature gradient. Intuitively we understand the importance of seasonally well distributed images on the second component. The second component shows to be very sensible for variability in between images. When recapitulating that PC2 has the ability to explain the second highest proportion of variance in temperature, we might be able to identify the second component as a combination of major surface and subsurface characteristics being at the source for causing variability in between the time-steps. We can deduct that the influence of thermal variability within the landscape is equally represented within every TIR-stack composition having fewer time-steps than the initial time-series. As such a combination of patterns causing thermal variability between the major landscape elements, can be deducted with very few TIR images. Thus increasing the dataset would not increase the information content represented in the patterns of PC1.

Furthermore, when analysing the mathematical background, we understand that through scaling the images before analysing them through the components, we expand or compress the gradients of the images to unit variance and hence account each image the same importance with regard to variance within the components (Jolliffe, 2002). If we had standardized the dataset, the images with high variances (summer) would have had a higher influence on the outcome of the analysis than images with lower variances (winter). Contrary, when scaling the dataset to unit variance, the influence of cold thermal images within the dataset rises, as their correlation with the rest of the images is low. This means that when only a few "cold" images are at hand, scaling is recommended to ensure the influence of these images on the outcome and allow a more precise distinction of patterns in the second component.

Furthermore, the comparison of the heterogeneity of known landscape characteristics with the pattern inherent in the components was made.

5.2 Probability Density Functions

5.2.1 DEM

The influence of topography on temperature variability patterns could be very well represented. However it should be mentioned, that one major difference between the Ybbs and the Traisen river catchment was that for the Ybbs catchment the nominal values of the first principal component are representing distinct and well distributed classes of elevations, while the component values of the Traisen catchment only showed the main distinction between the valley and the mountainous areas. This effect was tracked in all the comparison of patterns. This is to be explained through the characteristic relief, with the Pre-alps playing a distinct role in hydrological behaviour in the upper catchment area. As a consequence, we see the clustered distributions of characteristic interrelated land cover types in Pre-Alpine environment, as for example forests in the mountainous areas and arable land in the valley.

5.2.2 Land Cover Characteristics

The densities comparing land cover characteristics with patterns of thermal variability showed some effects needing consideration. So for example is Grassland and arable land somewhat behaving similarly for the Ybbs catchment, which might be reasoned by the time-series as the majority of the images are in summer and spring due to low cloud coverage, with crop growth being strong on arable lands, behaving comparably to grasslands when related to temperature variability. The influence of gardens and roof tiles within urban areas might explain the close positioning of the distribution function of the urban areas to the classes of grass- and arable lands. Forested areas are less well distinguishable from each other, which might be affected by the amount of conifers in mixed forests or the few images representing broadleaf forests under winter conditions. Evergreen and broadleaf in summer and spring might show similar behaviour in temperature variability with both having closed canopies. Complementary no knowledge about the amount of conifers within the mixed forests located in the catchments under investigation could be gained. However it is known the amount of conifers does not exceed 25% of the canopy closure for the class of mixed forests in Corine Land Cover maps (European Environment Agency, 2007). Further distinguishable are the water areas in the components, which shows the good emission properties of water with respect to thermal radiance and the low variability in temperature.

5.2.3 Hydrogeology

The most striking point here is the double peak of Granite in the density functions. When comparing the number of pixels of the distributions and the hydrogeological distribution of Granite, we discover, that the group is small compared to the other three classes and covers a mere 3% of the catchment. Furthermore the extent of the granite is exactly covered by a broadleaf forest which might indicate some problems with the representation of the granite class in the hydrogeological dataset or in the patterns of thermal variability.

5.2.4 Aspect and Slope

Zang and Kuenzer (2007) demonstrated the high influence of aspect and slope on thermal data of land TIR imagery. They showed that at the overpass time of Landsat 5 (+-10am) "*the same object/surface can differ by up to 10°C*" due to aspect. As such it was tested whether the influence of the overpass-time of the satellite, solar altitude and aspect was comprised in the first principal component and if some shadowing affects could be identified in the distributions of the component values.

While the overpass time of the Landsat 5 satellite system is in the morning, the expectation was that the east-exposed surfaces would show up stronger in the signals than they actually do in this case (see Zhang & Kuenzer, 2007). No heating or cooling patterns for eastern or westwards facing surfaces could be found in the first component's values. Nevertheless, the southern-facing areas are behaving remarkably different than the northern-facing areas. So this might represent the effect of shadowing. One reason for this relief induced behaviour might be the importance of the sun-sensor-object geometries, as the alignment of the Pre-Alpine mountains and the overpass route of the satellite being from north to south. Hence, some patterns of apparent similar temperature variability might emerge from these shadowing effects caused by imperfections in the alignment of the triangle between the sensor, the earth's surface and the sun.

Similar behaviour was observed for the Ybbs catchment: while the satellite is overpassing the catchments from south to north at approximately 10 am and the mountain ridge is mainly aligned to face south, respectively north, the patterns represented here might similarly be caused by some varying sun-sensor-object geometries. In further research it should thus be experimented whether an "hillshade-effect" is represented in the components, this being a function of slope, aspect and sun-sensor-object geometry (Hais et al., 2009).

5.2.5 Intersections

The above mentioned problems raised questions on the influence of the alpine region on the calculated densities. Thus it was necessary to test whether behaviour of component values changed between valley and mountainous regions. The results show, that differences in the patterns between the valley and mountainous areas could be observed. So was for example no differentiation between coniferous and mixed forest possible within the environment of the valley. The influence of conifers on the temperature patterns of the mixed forest class is assumed to be the reason. The problem of the imperfectly distributed images throughout the seasonality of the year potentially causes a similar temperature variability for all forest classes. As the closed canopies and the related micro-climate remain similar. More strikingly however was that patterns in the first component of the mountainous areas showed no differentiation between arable land and forested areas. The land cover class "arable land" shifted from negative component values in the valley to positive component values in the mountainous environment. While the majority of arable land is in the valley, the remaining area of arable land being at the foot of the alpine region is extremely small and neighbouring a vineyard which is classified as evergreen grassland. As such it might be possible that through this small amount of pixels and the resolution of the land cover data, the pattern of arable land is not representable.

However the Traisen catchment showed similar behaviour when comparing different regions of the components with patterns of different land uses: arable land was again distinctly represented. The reasons for this distinct representation in patterns might be numerous - it could be different stages of ripeness, combined with different cultivation methods, i.e. different harvesting times or simply different subsurface characteristics inherent in the mountainous soils causing the shift of the arable land's distribution function. The effect might be caused by different saturations of the grassland. The soil layer of the grassland in the mountainous areas might be less thick and dispersed with bare rock, which affects the temperature variability and makes it behave differently between the environments.

5.2.6 Aspect in the Second Components

For the second components, patterns were similarly being related to issues of Aspect. As explained above the correlating results might be reasoned from shadowing, i.e. heating effects and sun-sensor-object geometry. Consequently, we can assume that while the main spatially varying patterns remain land cover and topography (see PC1), the main change in this spatial variability (see PC2) is a result of a mixture of hydrogeological and elevational characteristics. In effect, flow patterns, soil water content and agglomeration patterns could potentially be represented in the configuration of the component values. However, these assumptions would need to be confirmed with field measurements.

5.2.7 Third Principal Component

In the comparison between the third components values and the heterogeneity of the land cover characteristics, correlations between urban areas and component values were found. This leads to the assumptions, that the third principle component might translate into specific characteristics affecting the signal strength received at the satellite. As an example, one might state the possibility of ozone and other greenhouse gases interfering with the signals being represented here. Similarly PC3 might potentially represent patterns of air water content or relative humidity which is known for intercepting and absorbing the emitted wavelength and therefore influencing the temperature variability received at the satellite sensor. A further explanation for these patterns, might be that throughout the time span of the selected series of images, the several land use characteristics are subject to temperature change. As an example; an explanation for the patterns in PC3 revealing urban areas, could be that there has been a change in the temperature variation in the cities. Thus showing that the cities have become warmer or temperature variation decreased throughout the investigated time range (e.g. due to increased surface sealing) (Kuenzer & Dech, 2013). Nevertheless, to derive a real trend, the images would need to be atmospherically corrected and the cities would need to be remotely sensed each year at the same date. Unfortunately, this is not possible due to restrictions of many causes in the availability of the data needed. Eventually, these assumptions would further need to be validated through comparison with actual field measurements.

The performance in the deduction of patterns of land surface functioning using PCA on TIR time series with regard to energy balance shows some limitations which have to be considered when interpreting the results. The main points of criticism will be shown in the next passage.

5.3 Cloud-Coverage in the TIR Datasets

The first limitation in the use of the method is that clouded images cannot be processed without further modification. Clouds appear as soon as in the second component, disturbing the extraction of patterns. By masking them and interpolating temperature values for these areas, it might be possible to solve the above mentioned issue. This could result in a higher availability, and a larger, better distributed dataset. However this might in return revoke the extraction of the interpretation of the patterns considering the interpolation process of temperature values and as such thermal variability and/or landscape functioning.

Secondly there is a lack of well distributed images due to cloud coverage. The database has considerably less images for winter and autumn than for summer and spring. As shown in Chapter 4.2 the method is depending on seasonally well distributed images, especially for the extraction of subtle patterns in the second component. If more cloud-free winter and autumn images were available, the information content in the components following PC1 would be of higher significance with more pronounced patterns allowing a better determination of the thermal properties that are displayed. At the current state of technology available, i.e. with overflight times of up to 16 days, there is a difficulty to gather images that show the desired meteorological conditions (cloud-free) and are well distributed in the seasons. To extract the full potential of this method, for example by displaying thermal inertia of the soils and as such extract knowledge about soil texture characteristics or percentage of organic matter, one would need to apply a principal component analysis on a diurnal remotely sensed TIR time-series. Remotely sensing a catchment day and night would result in an additional gain of information. Further analysing the consequent temperature variability of the diurnal time-steps would potentially establish connections between thermal variability and intrinsic soil properties, as shown by Soliman et al. (2011).

5.4 Image Size and Resolution

The resolution, the scene size and the resulting swath that the satellite is monitoring at each time-step needs consideration. The size of the units displaying thermal variance is limited to the resolution of the initial TIR data. Moreover a catchment lying in between two orbits cannot be analysed without further manipulation of the data. It is obvious, that at each overpass the sensor is sensing the thermal characteristics under different meteorological conditions. This is limiting the principal component analysis to the field of view of the satellite system. Though for analysing areas larger than the field of view of the satellite several PCA's are needed, and need distinct consideration as they are potentially analysing different conditions, i.e. images of different time-steps and/or cloud-coverage.

5.5 Interpretation of Principal Components

Validation of the patterns is problematic. To identify the drivers of thermal variance displayed in the principal components, data for validation is needed. Hence it was only possible to compare the patterns in the components with the ones of the major landscape elements at hand. If it could be further shown that the thermal variances displayed in the components are connected to other surface and subsurface characteristics i.e. organic matter content, bulk density, porosity, the component analysis of TIR time series might become a substantial part as Input, in hydrological modelling. As such complete datasets explaining hydraulic soil properties and/or hydrothermal physical properties are needed for comparison with the outcomes of the PCA. Thus more research and monitoring is needed on the pattern-process relationship of the PC's as well as on their explanatory power with regard to surface energy characteristics.

Accordingly, the orthogonal transformation-process that is part of the principal component analysis, makes an interpretation of the patterns inherent in the principal components a difficult and intricate task. Through the Karhunen-Loève- or Hotelling-transform, the deduction of direct internal physical principles becomes impossible (Page et al., 2012). The PCA shows patterns of thermal variability of TIR time series, but through the reduction of dimensionality and reprojection towards a new space, the interpretation of the patterns as a linear combination of all landscape characteristics needs to be handled with care (Jolliffe, 2002). As such the interpretation of the heterogeneity represented in the components becomes difficult as the patterns represent spatial and temporal variance of rotated, and as such unidentifiable, physical thermal properties. While on one hand this seems as a constraint it is on the other one of the strength of the approach. Through the rotational transformation, PCA allows the extraction of patterns of landscape functioning with all existing TIR datasets and is not sensor-specific. Complementary through the rotational transform of time series, maps can be produced incorporating all dominant landscape characteristics affecting thermal properties.

Through the constraint of orthogonality, an interpretation of the consecutive components might pose a problem. As such, rotation of the consecutive components is influenced by the maximized variance of the first component, with its position not always being unambiguous in the new projection. Thus small changes in the first components might have considerable effects on the interpretation of the consecutive components. Spinning this thought further, we encompass the importance of well distributed images on the distinguishability of the patterns within the first component for good interpretation of patterns in the second component. One way to aid interpretation is the manual rotation of the components as is done with factor loadings in factor analysis (Jolliffe, 2002). However also factor analysis knows several drawbacks. PCA for example is consecutively maximizing variance. In factor analysis which is proportions of variances are more evenly redistributed throughout the factor components. This compromises the further interpretation of factor components.

6 Conclusion

In this work, the approach proposed by Müller et al (2014) was applied to two mountainous catchments. In contrast to the Attert catchment analysed by Müller et al. (2014), the here revealed patterns of thermal variability show strong influences of topographic characteristics. Energy fluxes controlled by landscape characteristics are shown to be affected by topography. The differences in the components values, when comparing the mountainous regions with the flat areas of the catchments, have been displayed. Furthermore, patterns representing shadowing effects are shown to play an important role in the distribution of temperature variability throughout the here analysed catchments. Consequently, the topographically induced temperature variability also represented the most dominant differences between the catchment's components. These ambiguities need further research to be evaluated. The question arises whether these patterns are effective differences in temperatures caused by topography or if a bias, caused by sun-sensor-object geometries of the remote sensing system, is being represented. These features show the necessity to expand research by providing in situ measurement on the influence of shadowing on hydrological processes or by applying this methodological framework to catchments under known meteorological conditions and environmental stresses.

Nonetheless, unlike any other approach extracting energy balance characteristics at catchment scale, the PCA on TIR time-series is able to extract patterns displaying high proportions of thermal variability, all while reducing "noise" in the dataset. The patterns incorporate all temperature-dominating landscape characteristics. Thus it has its main strength in a general, sensor-unspecific approach able to extract land surface functioning with regard to the energy balance, see thermal variability. The approach proposed by Müller et al. (2014) provides a practical tool that identifies the functional behaviour of catchments, following the calls for an increased use of surrogate patterns in future hydrological models (Western et al., 2001). As Grayson et al. (2001) write: "*These* [the surrogate patterns] *provide rigorous tests of the 'behavioural' nature of simulations of catchment response […]"*.

Over the last decades there was little exploitation for TIR datasets. The determination of the quantitative precision in the deduction of TIR remote sensing based hydrological fluxes is difficult through the lack of in situ validation datasets. Consequently there was a particular chance for data errors in deducing absolute variables from TIR images (i.e.

bias in atmospheric corrections, unknown emissivities of land surface elements) (Kuenzer & Dech, 2013). Here however potential data errors are efficiently extracted and can even be visualized through the rotational transform into principal components. Thus contraire to conventional deductions of catchment dynamics which require extensive field work or are sensor-specific, the here illustrated method becomes generally applicable on all available TIR-datasets. As such TIR remote data should be reconsidered with focus on the underlying patterns representing thermal variability. These patterns are useful for many applications in hydrological research, such as validation, classification and parameterization.

Simulations of hydrological quantities at the outlet of the catchment might not necessarily be enhanced with this new input in hydrological models. However the observed patterns of catchment dynamics can be used to improve the understanding of processes within and between landscape elements (Stisen et al., 2011). The use of the underlying patterns (as represented by the first PC for example) is intrinsically valuable for assessment of distributed modelling approaches (Stisen et al., 2011). One example for the future use could be the validation of simulated interior fluxes. The principal components might thus be used as surrogate patterns for a comparison between simulated and observed hydrological heterogeneity with regard to energy controls (Grayson et al., 2002).

Another potential use for the components might be the deduction of a classification scheme based on the principal components. As shown in the results and discussion, the component values are to be understood as parameters for a combination of landscape characteristics influencing landscape functioning with regard to energy balance. Thus a classification scheme might be applied to the component values to order and group grid cells of similar thermal behaviour into several functional units. These functional units might be used in upcoming hydrological modelling approaches. Further this time-series analysis allows for a distinction between seasonal changes in thermal variability. For example, PC's of TIR time-series could be analysed into distinct functional units representing functional behaviour for different seasons. Expanding the understanding of the effects of different meteorological conditions and seasonal changes on hydrological behaviour, it would allow a more accurate prediction of the effects of climate change on the catchment behaviour.

Furthermore becomes the potential of the method apparent when the question is raised of how to scale up descriptions of hydrological responses from small (where they are developed) to large scales. The here extracted patterns could potentially validate such issues of scaling and parameterization as they provide a measure of heterogeneity of the combined variables exerting controls on the energy balance. PCA on TIR time series allows for a derivation of direct indicators of hydrological response. As such they could be implemented in a parameterization approach like the disaggregation-aggregation model proposed by Viney and Sivapalan (2004) for example. The approach might provide a way of linking the catchment-scale variables with catchment-scale responses all while retaining some "essence" of small scale physics (Viney & Sivapalan, 2004).Or how Müller et al (2014) state: *"The strongest impact of the approach presented is expected when the derived component values from the PCA analysis will be incorporated into model parameter regionalization schemes (e.g. the multi-scale parameter regionalization (MPR) scheme presented by Samaniego et al., 2010)".* In other words, the PC's can be used as "proxy" or "soft" data in the deduction of for example connectivity functions. As the patterns provide continuous pixel based values, rather than nominal data (see HRU's), by representing thermal variability of the landscape, the patterns in the PC's could be used as "*natural skeleton*" onto which hydrological response is projected (Müller et al., 2014). Consequently they reduce the ambiguity in the calibration process and thus potentially increase the predictive value of hydrological models. It can therefore be shown, that PCA on TIR time-series provides a rigorous methodology that might be minimising the problems what Kirchner (2006) described as "*right results for wrong reasons*".

Research might further be expanded to other datasets such as active and passive radar or the near-infrared wavelengths. Considering near-infrared datasets it might become possible to extract distinct catchment hydrological characteristics. As it is shown by Zaheer & Iqbal, (2014), near-infrared Landsat5 TM data, has information about soil properties inherent. They approximated soil organic matter content through remote sensing with a combination of near-infrared and far-infrared wavelengths. When considering radar data, it is assumed to enhance the information content stored in the principal component's patterns, for example through a higher temporal resolution. As such better representations of specific soil hydraulic conditions is expected through a direct comparison between patterns of temperature variability and meteorological conditions. Thus the extraction of e.g., soil moisture content might become feasible. When analysing high-resolution active radar sensor data with principal components, one might be able to extract thermal variability in deeper soil layers as these radar systems are ground-penetrating. As such it might become possible to show spatial and temporal temperature variability adding a vertical representation of variances in soil depths. For example has been shown that soil moisture content can be estimated using active radar systems for a depth of up to 1.3m (Gacitúa, et al., 2012; Lunt et al., 2005) as well as freezing and thawing processes from spaceborne passive radar (Mironov & Muzalevsky, 2013). The here presented method would hence further introduce a temporal dimension to the above mentioned monitoring processes by displaying these processes within a time-series analysis using principal components.

All in all, this method provides a measure of connectivity between landscape elements with regard to internal energetic fluxes. Thus the approach encourages the development of the hydrological science through the implicit knowledge of the distribution of hydrological process with regard to the distribution of thermal characteristics. The state of the art of hydrological modelling is advanced in providing a new idea, new data and new experimental work. Through knowledge about interior water fluxes by deducing patterns of thermal variability, new data for parameterization based on direct physical principles rather than computational optimization, is provided. If implemented in hydrological models, a more profound understanding of the spatial distribution of hydrological behaviour is provided. The derivation and evaluation of human impacts, see knowledge about how and where degradation will affect water resources, can be deduced from knowledge about spatial variability (Grayson et al., 1997). Thus given the predicted climate change and the observed intensification of environmental degradation, the here provided information about the distribution of spatial variability of landscape elements having controls on the energy balance, is enhancing integrated water resources management through an improved assessment of interior water fluxes.

7 References

Abbott, M. B., & Refsgaard, J. C. (Eds.). (1996). *Water Science and Technology Library*. *Distributed Hydrological Modelling*. Dordrecht: Springer Netherlands.

Alcântara, E. (2013). *Remote sensing: Techniques, applications and technologies*. Hauppauge, New York: Nova Science Publishers Inc.

Anderson, M. C., Kustas, W. P., Norman, J. M., STARKS, P., & AGAM, N. (2008). A thermal-based remote sensing technique for routine mapping of land-surface carbon, water and energy fluxes from field to regional scales. *Remote Sensing of Environment, 112*(12), 4227–4241. doi:10.1016/j.rse.2008.07.009

Barsi, J. A., Barker, J. L., & Schott, J. R. (©2003). *IGARSS 2003: Learning from Earth's shapes and colors : 2003 IEEE International Geoscience and Remote Sensing Symposium : proceedings : Centre de Congrès Pierre Baudis, Toulouse, France, 21-25 July, 2003*. Piscataway, N.J.: IEEE.

Barsi, J. A., Schott, J. R., Palluconi, F. D., Hook, S. J., & Butler, J. J. (2005). In : *SPIE Proceedings, Optics & Photonics 2005* (pp. 58820E). SPIE.

Beighley, R. E., Dunne, T., & Melack, J. M. (2005). Understanding and modeling basin hydrology: interpreting the hydrogeological signature. *Hydrological Processes, 19*(7), 1333–1353. doi:10.1002/hyp.5567

Betz, L. (2013). *Landsat 5 Sets Guinness World Record For 'Longest Operating Earth Observation Satellite'*. Retrieved from http://www.nasa.gov/mission_pages/landsat/news/landsat5-guinness.html#.VNNsG52G_uw

Beven, K. J. (2012). *Rainfall-runoff modelling: The primer* (2nd ed). Hoboken: Wiley.

Beven, K. J., & Moore, I. D. (1993). *Terrain analysis and distributed modelling in hydrology*. *Advances in hydrological processes*. Chichester, England, New York: Wiley & Sons.

Beven, K. (2001). How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences, 5*(1), 1–12. Retrieved from http://www.hydrol-earth-syst-sci.net/5/1/2001/hess-5-1-2001.pdf

Bivand, R. (2015). ClassINT. Norway.

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment, 113*(5), 893–903. doi:10.1016/j.rse.2009.01.007

Chang, N.-B. (2012). *Multiscale hydrologic remote sensing: Perspectives and applications*. Boca Raton: Taylor & Francis.

Clark, P. E. (2014). *Remote sensing tools for exploration: Observing and interpreting the electromagnetic spectrum*. [Place of publication not identified]: Springer.

Clark, P. E., & Rilee, M. L. (2010). *Remote sensing tools for exploration: Observing and interpreting the electromagnetic spectrum*. New York: Springer.

Cracknell, A. P. (1997). *The advanced very high resolution radiometer (AVHRR)*. London: Taylor & Francis. Retrieved from http://www.loc.gov/catdir/enhancements/fy0745/97171558-d.html daNUbs. (2003). *Water balance calculations for the case study regions in Austria, Hungary and Romania: Report.*

Doherty, J. (2003). Ground Water Model Calibration Using Pilot Points and Regularization. *Ground Water, 41*(2), 170–177. doi:10.1111/j.1745-6584.2003.tb02580.x

Environmental Systems Research Institute. (2015). ArcGis for Desktop. Redlands, CA: ESRI. Retrieved from www.esri.com

European Environment Agency. (2007). *CLC2006 technical guidelines*. *Technical report: 17/2007*. Luxembourg: Publications Office.

Flügel, W.-A. (1995). Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany. *Hydrological Processes, 9*(3-4), 423–436. doi:10.1002/hyp.3360090313

French, A., Schmugge, T., Ritchie, J., Hsu, A., Jacob, F., & Ogawa, K. (2008). Detecting land cover change at the Jornada Experimental Range, New Mexico with ASTER emissivities. *Remote Sensing of Environment*, *112*(4), 1730–1748. doi:10.1016/j.rse.2007.08.020

Gacitúa, G., Tamstorf, M. P., Kristiansen, S. M., & Uribe, J. A. (2012). Estimations of moisture content in the active layer in an Arctic ecosystem by using ground-penetrating radar profiling. *Journal of Applied Geophysics, 79,* 100–106. doi:10.1016/j.jappgeo.2011.12.003

Goslee, S. C. (2011). Analyzing Remote Sensing Data in R: The landsat Package. Journal of Statistical Software, 43(4), 1-25. URL http://www.jstatsoft.org/v43/i04/.: The landsat Package. *Journal of Statistical Software, 43*(4), 1–25. Retrieved from http://www.jstatsoft.org/v43/i04/.

Grayson, R., & Blöschl, G. (2001). *Spatial patterns in catchment hydrology: Observations and modelling*. Cambridge, U.K., New York: Cambridge University Press.

Grayson, R. B., Blöschl, G., Western, A. W., & McMahon, T. A. (2002). Advances in the use of observed spatial patterns of catchment hydrological response. *Advances in Water Resources, 25*(8-12), 1313–1334. doi:10.1016/S0309-1708(02)00060-X

Grayson, R. B., Western, A. W., Chiew, Francis H. S., & Blöschl, G. (1997). Preferred states in spatial soil moisture patterns: Local and nonlocal controls. *Water Resources Research, 33*(12), 2897–2908. doi:10.1029/97WR02174

Hais, M., & Kučera, T. (2009). The influence of topography on the forest surface temperature retrieved from Landsat TM, ETM + and ASTER thermal channels. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*(6), 585–591. doi:10.1016/j.isprsjprs.2009.04.003

Hijmans, R. J. & van Etten, J. (2012). *raster: Geographic analysis and modeling with raster data*. R package version: 2.0-12. Retrieved from http://CRAN.R-project.org/package=raster

Hirosawa, Y., Marsh, S. E., & Kliman, D. H. (1996). Application of standardized principal component analysis to land-cover characterization using multitemporal

AVHRR data. *Remote Sensing of Environment*, *58*(3), 267–281. doi:10.1016/S0034-4257(96)00068-5

Jensen, J. R. (2007). *Remote sensing of the environment: An earth resource perspective* (2nd ed). *Prentice Hall series in geographic information science*. Upper Saddle River, NJ: Pearson Prentice Hall.

Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed). *Springer series in statistics*. New York: Springer.

Jorgensen, S. E., & Fath, B. (2008). *Encyclopedia of Ecology, Five-Volume Set*. Oxford: Elsevier Science.

Kalma, J. D., McVicar, T. R., & McCabe, M. F. (2008). Estimating Land Surface Evaporation: A Review of Methods Using Remotely Sensed Surface Temperature Data. *Surveys in Geophysics, 29*(4-5), 421–469. doi:10.1007/s10712-008-9037-z

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research, 42*(3).

Kitchin, R., & Thrift, N. J. (2008). *International encyclopedia of human geography*. Amsterdam: Elsevier.

Kuenzer, C., & Dech, S. (2013). *Thermal Infrared Remote Sensing* (Vol. 17). Dordrecht: Springer Netherlands.

Kumar, R. (2009). *Distributed Hydologic Model Parameterization. Application in a Mesoscale River Basin* (Dissertation). Friedrich-Schiller-Universität, Jena.

Lu, S., Ju, Z., Ren, T., & Horton, R. (2009). A general approach to estimate soil water content from thermal inertia. *Agricultural and Forest Meteorology, 149*(10), 1693–1698. doi:10.1016/j.agrformet.2009.05.011

Lunt, I. A., Hubbard, S. S., & Rubin, Y. (2005). Soil moisture content estimation using ground-penetrating radar reflection data. *Journal of Hydrology*, *307*(1-4), 254–269. doi:10.1016/j.jhydrol.2004.10.014

MacKay, David J. C. (2003). *Information theory, inference, and learning algorithms*. Cambridge, UK, New York: Cambridge University Press.

Mironov, V. L., & Muzalevsky, K. V. (2013). Spaceborne radar monitoring of soil freezing/thawing processes in the Arctic tundra. *Russian Physics Journal, 55*(8), 899–902. doi:10.1007/s11182-013-9898-6

Müller, B., Bernhardt, M., & Schulz, K. (2014). Identification of catchment functional units by time series of thermal remote sensing images. *Hydrology and Earth System Sciences Discussions*, *11*(6), 7019–7052. doi:10.5194/hessd-11-7019-2014

Ng, A. (2008). *Lecture 15: Machine Learning*. Retrieved from http://youtu.be/QGd06MTRMHs

Oke, T. R. (1987). *Boundary layer climates* (2nd ed). London, New York: Routledge.

Page, R. M., Lischeid, G., Epting, J., & Huggenberger, P. (2012). Principal component analysis of time series for identifying indicator variables for riverine groundwater extraction management. *Journal of Hydrology, 432-433,* 137–144. doi:10.1016/j.jhydrol.2012.02.025

Planck, M. (1900). Entropie und Temperatur strahlender Wa¨rme. *Ann Phys*, (306(4)), 719–737.

Pratt, D. A., & Ellyett, C. D. (1979). The thermal inertia approach to mapping of soil moisture and geology. *Remote Sensing of Environment, 8*(2), 151–168. doi:10.1016/0034-4257(79)90014-2

Qin, Z., Karnieli, A., & Berliner, P. (2001). A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. *International Journal of Remote Sensing, 22*(18), 3719–3746. doi:10.1080/01431160010006971

Quattrochi, D. A., & Luvall, J. C. (1999). Thermal infrared remote sensing for analysis of landscape ecological processes: methods and applications. *Landscape Ecology, 14*(6), 577–598. doi:10.1023/A%3A1008168910634

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. Retrieved from http://www.R-project.org/

Rango, A., & Shalaby, A. I. (1998). Operational applications of remote sensing in hydrology: success, prospects and problems. *Hydrological Sciences Journal, 43*(6), 947–968. doi:10.1080/02626669809492189

Richards, J. A., & Jia, X. (2006). *Remote sensing digital image analysis: An introduction* (4th ed). Berlin: Springer.

Risser, P. G., Karr, J. R., & Forman, R. T. (1984). Landscape Ecology: Directions and Approaches. *Illinois Natural History Survey Special Publication, 2*.

Rosso, R. (1994). *Advances in distributed hydrology: Selected papers from the international workshop : Bergamo, Italy, June 25-26, 1992*. Highlands Ranch, Colo.: Water Resources Publications.

Salisbury, J. W., & Eastes, J. W. (1985). The effect of particle size and porosity on spectral contrast in the mid-infrared. *Icarus, 64*(3), 586–588. doi:10.1016/0019-1035(85)90078-8

Samaniego, L., Kumar, R., & Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research, 46*(5), n/a. doi:10.1029/2008WR007327

Sánchez, J. M., Scavone, G., Caselles, V., Valor, E., Copertino, V. A., & Telesca, V. (2008). Monitoring daily evapotranspiration at a regional scale from Landsat-TM and ETM+ data: Application to the Basilicata region. *Journal of Hydrology, 351*(1-2), 58–70. doi:10.1016/j.jhydrol.2007.11.041

Schmugge, T. J., Kustas, W. P., Ritchie, J. C., Jackson, T. J., & Rango, A. (2002). Remote sensing in hydrology. *Advances in Water Resources, 25*(8-12), 1367–1385. doi:10.1016/S0309-1708(02)00065-9

Schowengerdt, R. A. (2007). *Remote sensing: Models and methods for image processing* (3rd ed). Burlington, MA: Academic Press.

Schulz, K., Seppelt, R., Zehe, E., Vogel, H. J., & Attinger, S. (2006). Importance of spatial structures in advancing hydrological sciences. *Water Resources Research, 42*(3), n/a. doi:10.1029/2005WR004301

Smith, R. B. (2012). *Introduction to: Remote Sensing of the Environment*. Retrieved from http://www.microimages.com/documentation/Tutorials/introrse.pdf

Soliman, A., Brown, R., & Heck, R. J. (2011). Separating near surface thermal inertia signals from a thermal time series by standardized principal component analysis.

International Journal of Applied Earth Observation and Geoinformation, 13(4), 607–615. doi:10.1016/j.jag.2011.03.004

Stisen, S., McCabe, M. F., Refsgaard, J. C., Lerer, S., & Butts, M. B. (2011). Model parameter analysis using remotely sensed pattern information in a multi-constraint framework. *Journal of Hydrology, 409*(1-2), 337–349. doi:10.1016/j.jhydrol.2011.08.030

USGS. (2014). *Band designations Landsat Satellites*. Retrieved from http://landsat.usgs.gov/band_designations_landsat_satellites.php

Viney, N. R., & Sivapalan, M. (2004). A framework for scaling of hydrologic conceptualizations based on a disaggregation–aggregation approach. *Hydrological Processes, 18*(8), 1395–1408. doi:10.1002/hyp.1419

Western, A. W., Blöschl, G., & Grayson, R. B. (2001). Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research, 37*(1), 83–97. doi:10.1029/2000WR900241

Zaheer, A., & Iqbal, J. (2014). Evaluation of Landsat TM5 Multispectral Data for Automated Mapping of Surface Soil Texture and Organic Matter in GIS. *European Journal of Remote Sensing*, (47), 557–573. Retrieved from http://server-geolab.agr.unifi.it/public/completed/2014_EuJRS_47_557_573_Ahmed.pdf

Zhang, J., & Kuenzer, C. (2007). Thermal surface characteristics of coal fires 1 results of in-situ measurements. *Journal of Applied Geophysics, 63*(3-4), 117–134. doi:10.1016/j.jappgeo.2007.08.002

ANNEXE I. TIR-Imagery and Principal Components

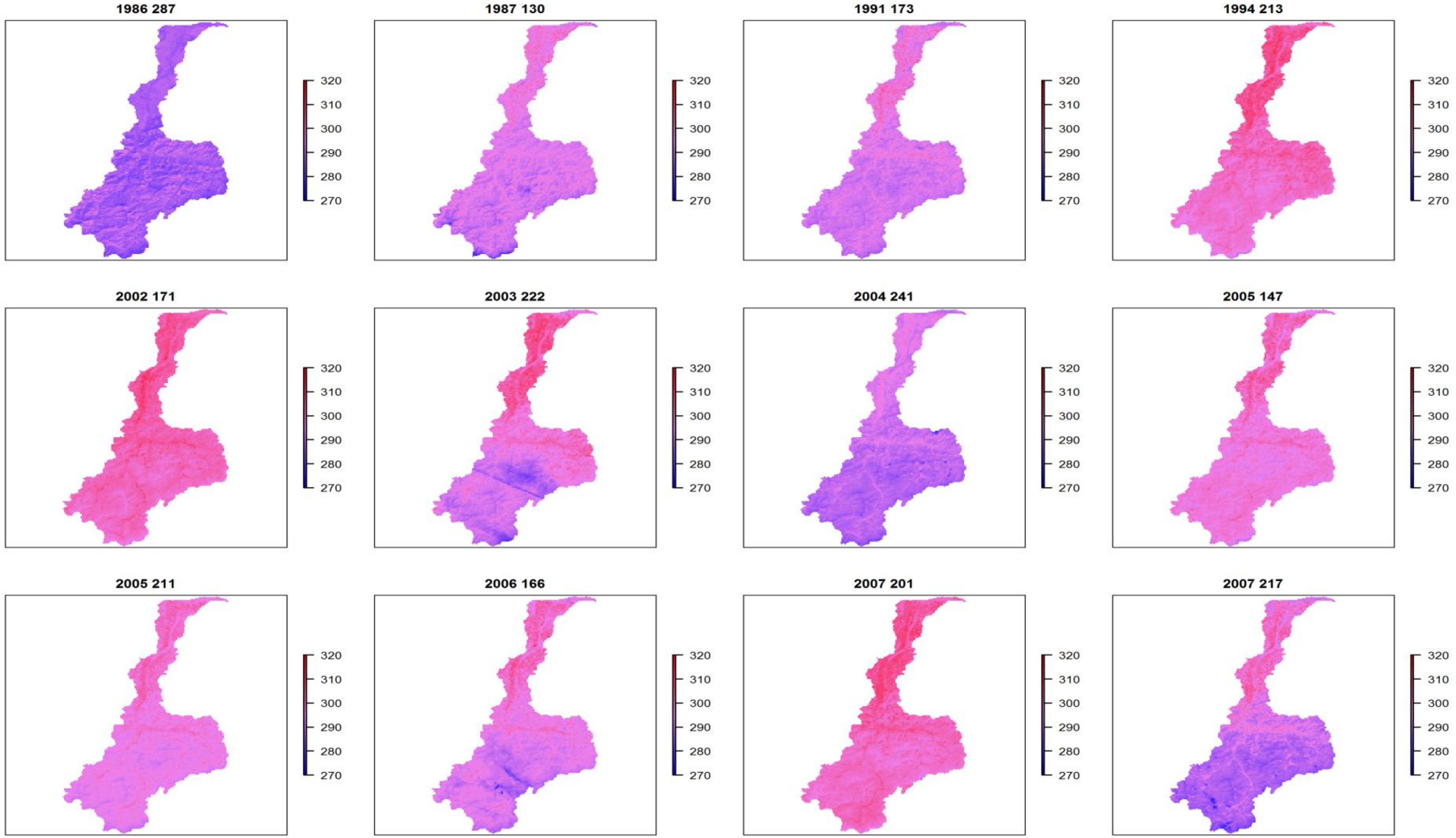

Figure: TIR Time-Series for Traisen River Catchment

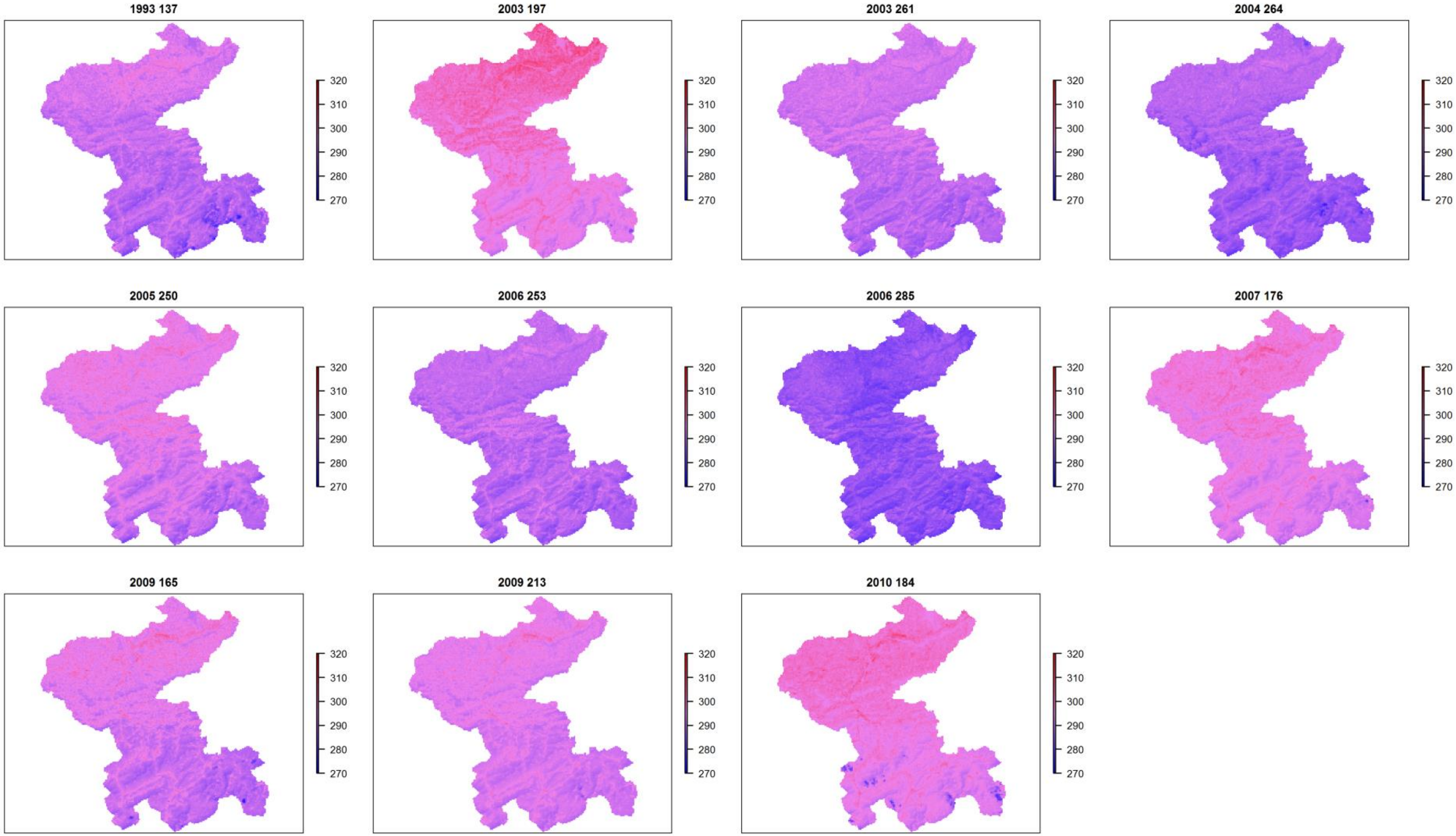

Figure: TIR Time-Series for Ybbs Catchment

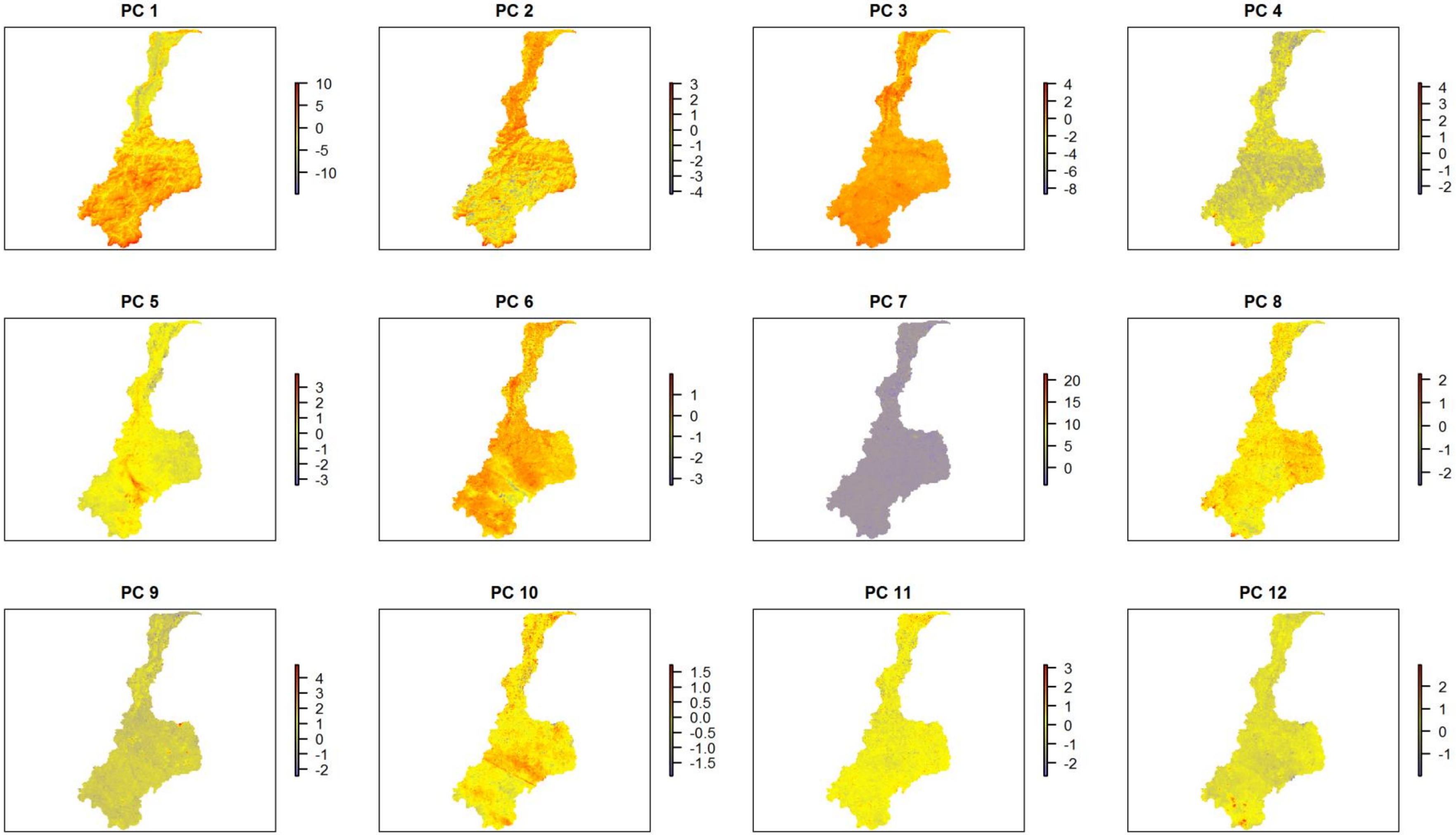

Figure: Principal Components of TIR Time-series Analysis for Traisen Catchment

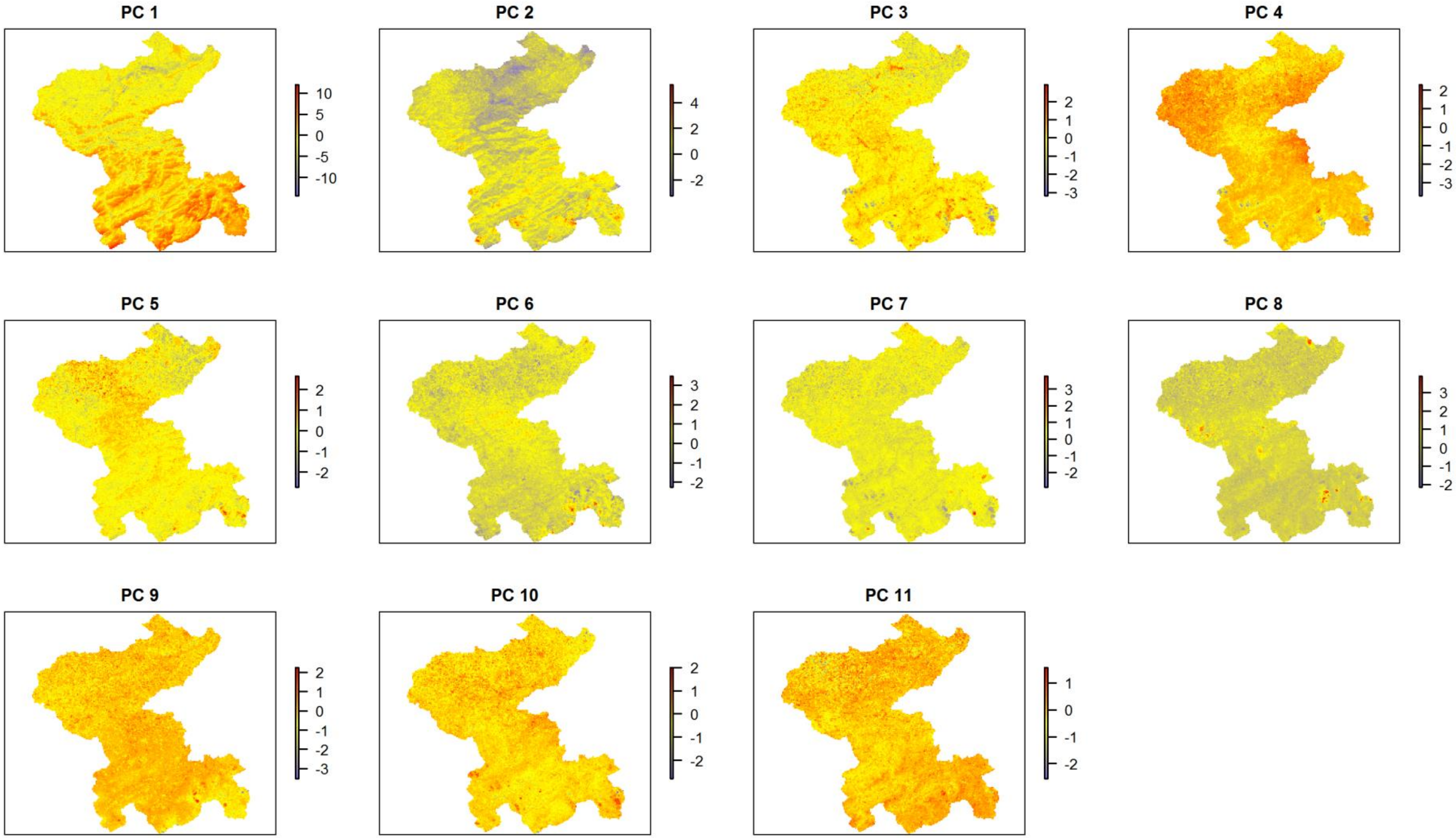

Figure: Principal Components of TIR Time-series Analysis for Ybbs Catchment

ANNEXE II. Soil Texture Characteristics

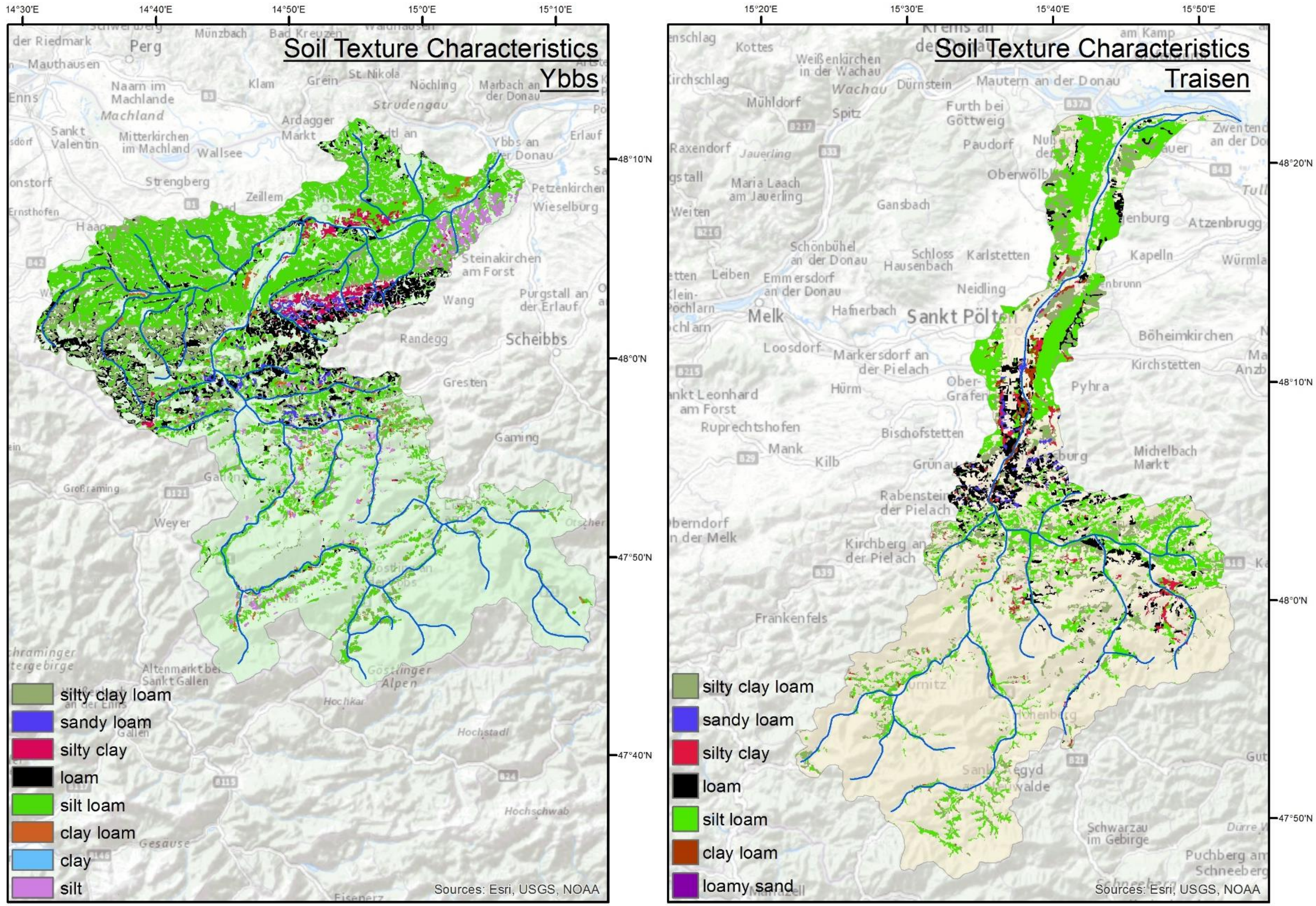

Figure: Soil Texture Characteristics for Ybbs and Traisen.

ANNEXE III. Probability Distribution Functions

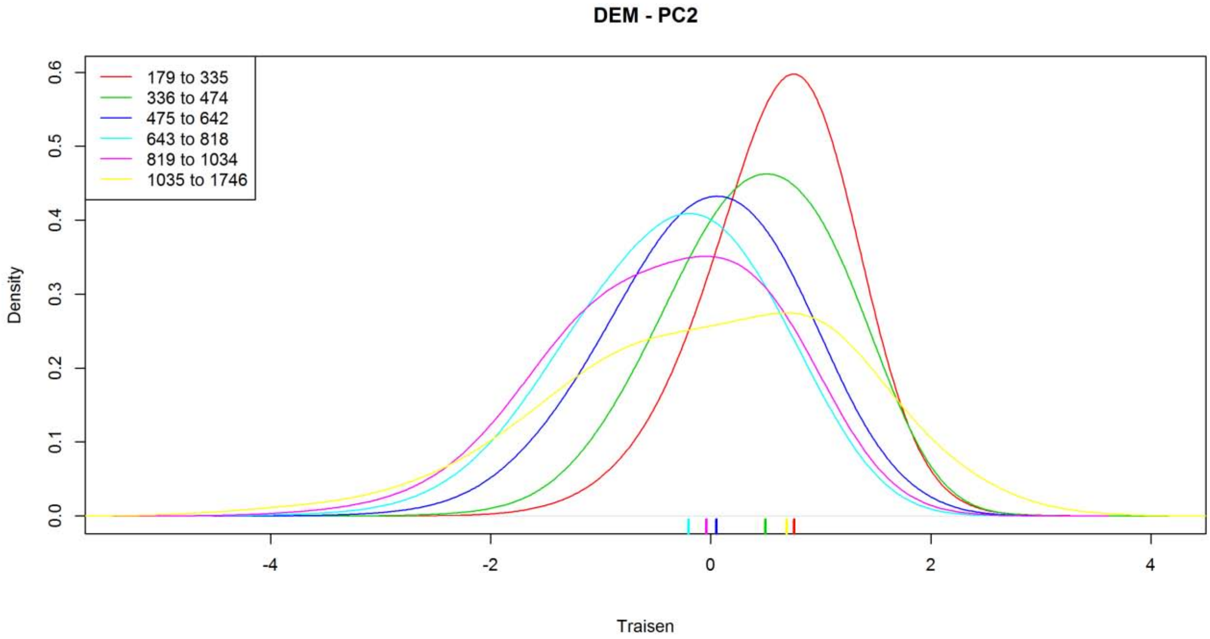

Figure: DEM and PC2. Traisen. Bandwith=0.8

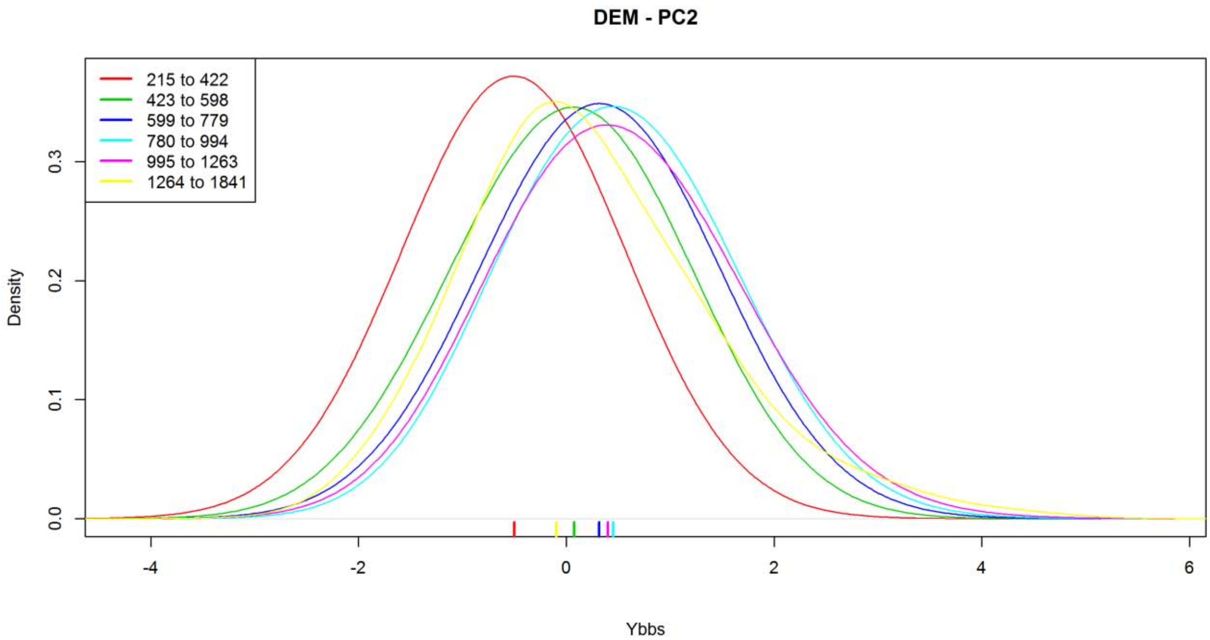

Figure: DEM and PC2. Ybbs. Bandwith=0.8

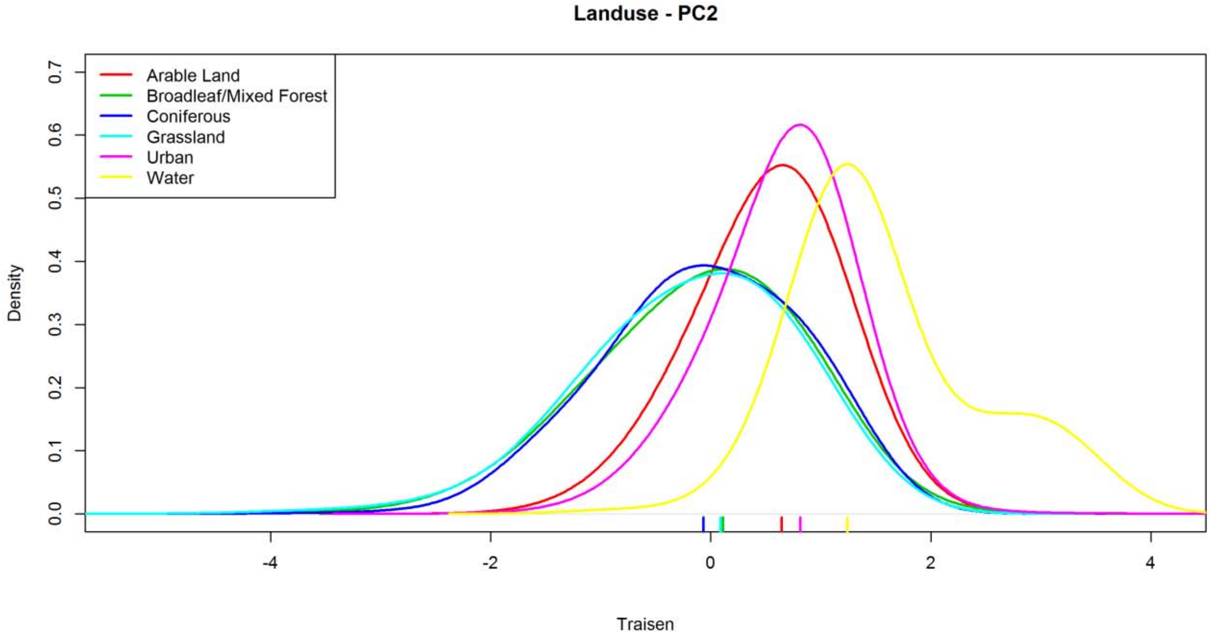

Figure: Landuse and PC2. Traisen. Bandwith=0.4

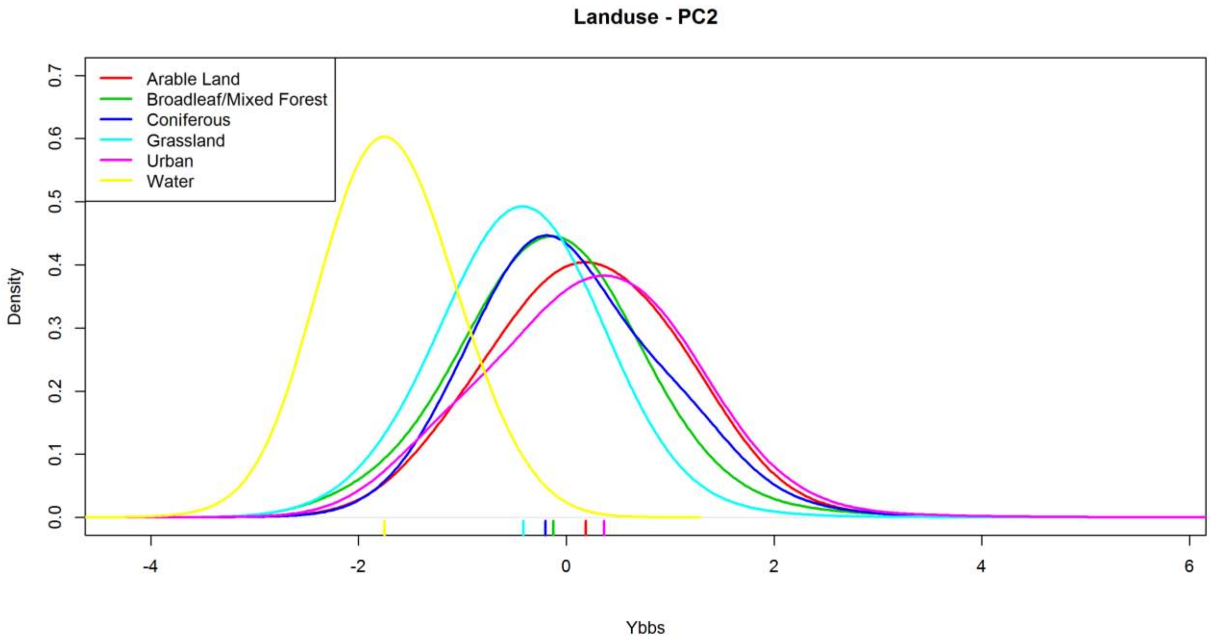

Figure: Landuse and PC2. Traisen. Bandwith=0.4

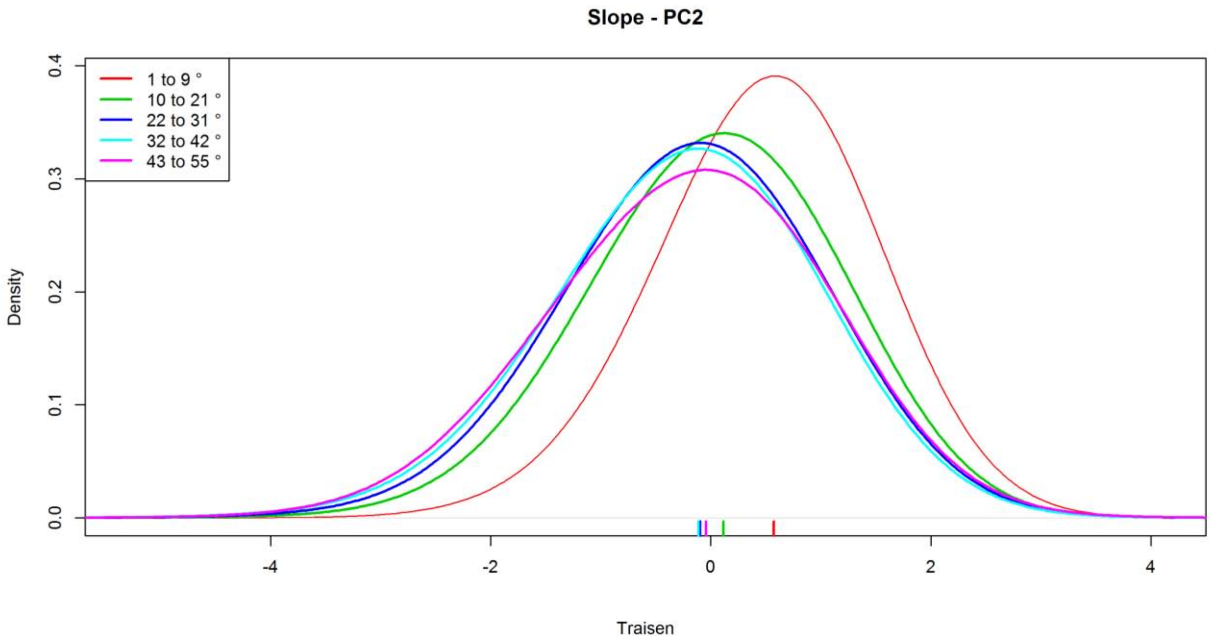

Figure: Slope and PC2. Traisen. Bandwith=0.4

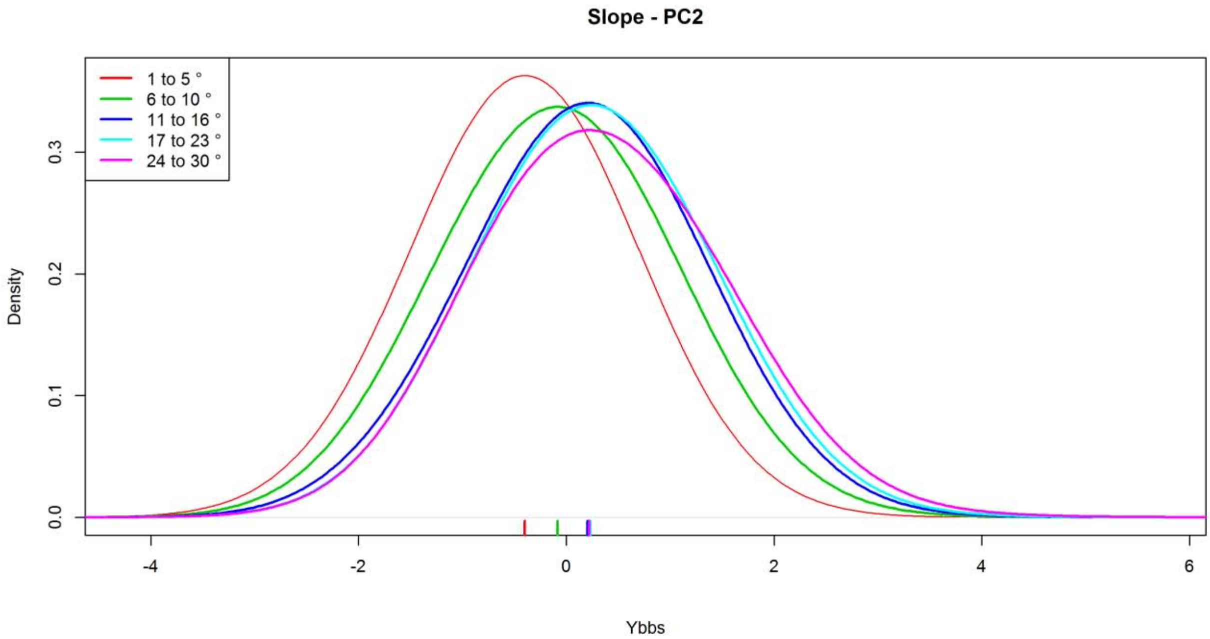

Figure: Slope and PC2. Ybbs. Bandwith=0.4

ANNEXE IV. R-scripts

Georeferencing

This script was coded in the “landsat”-Package to georeference the TIR imagery towards one initially manually georeferenced image.

```
#####Packages#####
require(raster)
require(sp)
require(rgdal)
require(landsat)

######################GEOREFERENCING######################

R1 <- raster(paste(refpath,file[1],sep="/"),band=1)
TARGET=as(R1, "SpatialGridDataFrame")
stats_DF=data.frame()
Name_DF=c()
stats_x=c()
stats_y=c()
for (i in 1:length(files)) {

  Ri <- raster(paste(refpath,file[i],sep="/"),band=1)
  Df_ri=as(Ri, "SpatialGridDataFrame")
  shift_georef=georef(TARGET, Df_ri, maxdist=50)
  shifted_Df_ri=geoshift(Df_ri, padx=10, pady=10, shiftx=shift_georef$shiftx,
shifty=shift_georef$shifty, nodata=NA)
  Name=paste(substr(file[i],1,16),"G.tif",sep="")
  Name_DF=append(Name_DF, Name)
  stats_x=append(stats_x, shift_georef$shiftx)
  stats_y=append(stats_y, shift_georef$shifty)

writeGDAL(shifted_Df_ri,fname=paste(Savepath2,Name,sep="/"),drivername="GTiff"
)
  print(Name)
}

plot(Df_ri, shifted_Df_ri)
stats_DF=data.frame(Name_DF, stats_x, stats_y)
write.table(stats_DF, "Directory/gref_stats.csv")
```

Raster Calculator

This script was coded to transform the raw TIR images from Binaries to Radiance values and from Radiance to Brightness Temperature.

```
#####Rastercalculator#####

refpath="Directory"
SavePath1="Directory/Radiance"
SavePath2="Directory/Kelvin"
files=list.files(refpath,pattern="M.tif")
files=files[which(apply(as.matrix(files),2,nchar)==16)]
nchar(files[1])

#####Needed R-Packages#####
require(raster)
require(sp)
require(rgdal)

######FUNCTIONS#######

funradiance=function (x) {(0.055376*(x)+1.18)}

FunToA=function (x) {1260.56/(log((607.76/x)+1))}

  ####First PART DN-to-Rad#####
```

```
for (ind in 1:length(files)) {
  R1=raster(paste(refpath,files[ind],sep="/"))
  Name=paste(substr(files[ind],1,11),"Rad.tif",sep="")
  Rrad=R1
  Rrad[]=(funradiance(Rrad[]))
  writeRaster(Rrad, filename=paste(SavePath1,Name,sep="/"), format="GTiff",
overwrite=T)
  print(Name)
}


  ####Second PART Rad-to-ToA######
  Name2=paste(substr(files[ind],1,12),"Kelv.tif",sep="")
  RToA=Rrad
  RToA[]=(FunToA(RToA[]))
  writeRaster(RToA, filename=paste(SavePath2,Name2,sep="/"), format="GTiff",
overwrite=T)
  print(Name2)
}
print("done")
```

Principal Component Analysis

```
##### Create PCA, stack Components, create Multilayer RasterFile #####

##### Packages #####
require(raster)
require(sp)
require(rgdal)

##### Sort Files #####
refpath="Directory/Kelvin "
savepath="Directory/PCA "
nchar(files[1])

##### Stack #####
 files=list.files(refpath,pattern="KelvG.tif")
 files=files[which(apply(as.matrix(files),2,nchar)==21)]
stackdasma=stack()

 for (i in 1:length(files)) {
 Ri <- raster(paste(refpath,files[i],sep="/"),band=1)# raster einlesen
 #Ri = projectRaster(Ri, R1, method='bilinear')
  stackdasma=stack(stackdasma,Ri)
  print(nlayers(stackdasma))
 }

##### PCA #####

## Ri=raster(paste(savepath,file[1],sep="/")) TIR imagery
nlayers(stackdasma) ##number of layers
DF=as.data.frame(stackdasma) ##data frame
DF=na.omit(DF) ##NA`s raus
PCA=prcomp(DF, scale=T, center=T) # PCA MAGIE
summary(PCA, loadings=F)

PCAval=predict(PCA) # PCA on Coordinate System
loadings_PCA=as.data.frame(PCA$rotation)
write.table(loadings_PCA, file=paste(savepath, "Loadings.csv",sep="/"))
rownames(PCAval) # ...
PCABlanko=mean(stackdasma, na.omit=T) # Blanc Raster with equal resolution and
extent as TIR imagery
PCABlanko[][as.numeric(rownames(PCAval))]=PCAval[,1]
plot(PCABlanko)

##### Stack Principal Components #####
PClist=list()
for (i in 1:nlayers(stackdasma)){
  PCABlanko[][as.numeric(rownames(PCAval))]=PCAval[,i]
```

```
  PClist[[i]]=PCABlanko
  print(i)
}

PCAStack=stack(PClist)

### create Multilayer RasterFile #####
plot(PCAStack)
summary(PCA, loadings=T)
writeRaster(PCAStack,
filename=paste(savepath,"PCA_stack_scaled_center",sep="/"), format="GTiff",
overwrite=T, bandorer="BIL")

### create Single-layer RasterFiles #####

d2=unstack(PCAStack)
outputnames <- paste(seq_along(d2))
Name="PCscaled"
for(i in seq_along(d2)){writeRaster(d2[[i]],
filename=paste(savepath,(paste(substr(Name,1,8),outputnames[i],sep="")),sep="/
"),format="GTiff", overwrite=T, bandorer="BIL")}


labels1=as.character(files)
#####Plot Loadings #####
plot(loadings_PCA$PC2, loadings_PCA$PC3, pch=19, col=4, xlim=c(-1, .5),
ylim=c(-.5, 1), xlab="Loadings PC3", ylab="Loadings PC2", main="DOY ~
Loadings", sub="Traisen")
text(loadings_PCA$PC2, loadings_PCA$PC3,
labels=paste(substr(row.names(loadings_PCA), 2, 8)), pos=2, srt=50)

plot(loadings_PCA$PC2, loadings_PCA$PC1, pch=19, col=4, xlim=c(-1, .5),
ylim=c(-.5, 1),  xlab="Loadings PC2", ylab="Loadings PC1",  main="DOY ~
Loadings", sub="Traisen")
text(loadings_PCA$PC2, loadings_PCA$PC1,
labels=paste(substr(row.names(loadings_PCA), 2, 8)), pos=2, srt=45)
```

Density Calculations

This exemplary R-script file was used to estimate the densitiy estimations for the textural classes. However, it can be used for all densitiy estimations. See CD-Rom for all scripts used.

```
##### Needed Packages #####
require(raster)
require(sp)
require(rgdal)

#####Input######
Input=raster("directory.tif")

PC1=raster("Directory_PC1.tif")
projectRaster(PC1, Input, filename="Directory_PC1.tif", overwrite=T,
method='bilinear')
PC1=raster("Directory.tif")

OR

PC1=raster("Directory/PC1.tif")

projectRaster(Input, PC1, filename="Directory/Input.tif", overwrite=T,
method='bilinear')
Input=raster("Directory/Input.tif")

####DataFrame###

DF=as.data.frame(PC1)
head(DF)
```

```
DF=cbind(DF, (as.data.frame(Tex)))
head(DF)

####Selection of Attributes (HGEO)####

silty_clay_loam=(DF$PC1_Traisen_repr_Tex[which(DF$NUM==5)])
sandy_loam=(DF$PC1_Traisen_repr_Tex[which(DF$NUM==9)])
silty_clay=(DF$PC1_Traisen_repr_Tex[which(DF$NUM==2)])
loam=(DF$PC1_Traisen_repr_Tex[which(DF$NUM==7)])
silt_loam=(DF$PC1_Traisen_repr_Tex[which(DF$NUM==8)])
clay_loam=(DF$PC1_Traisen_repr_Tex[which(DF$NUM==4)])
loamy_sand=(DF$PC1_Traisen_repr_Tex[which(DF$NUM==11)])

D1=density(silty_clay_loam,bw=0.4, na.rm=T) #die bandwidth (bw) muss t du f?r
alle gleich einstellen. Bisschen probieren, was gut aussieht
D2=density(sandy_loam,bw=0.4, na.rm=T)
D3=density(silty_clay,bw=0.4, na.rm=T)
D4=density(loam,bw=0.4, na.rm=T)
D5=density(silt_loam,bw=0.4, na.rm=T)
D6=density(clay_loam,bw=0.4, na.rm=T)
D7=density(loamy_sand,bw=0.4, na.rm=T)

OR ######Selection of Attributes (DEM with K-means Intervals)####

######classes k means ##########

require(classInt)
classes=as.list(c(classIntervals(na.omit(DF$DEM_Ybbs_reproj_PCscaled1), n=6,
style="kmeans")))
classes$brks


DEM_214to449=DF$PCscaled1[which(DF$DEM_Ybbs_reproj_PCscaled1>=classes$brks[1]
& DF$DEM_Ybbs_reproj_PCscaled1<classes$brks[2])]
DEM_450to667=DF$PCscaled1[which(DF$DEM_Ybbs_reproj_PCscaled1>=classes$brks[2]
& DF$DEM_Ybbs_reproj_PCscaled1<classes$brks[3])]
DEM_668to908=DF$PCscaled1[which(DF$DEM_Ybbs_reproj_PCscaled1>=classes$brks[3]
& DF$DEM_Ybbs_reproj_PCscaled1<classes$brks[4])]
DEM_909to1211=DF$PCscaled1[which(DF$DEM_Ybbs_reproj_PCscaled1>=classes$brks[4]
& DF$DEM_Ybbs_reproj_PCscaled1<classes$brks[5])]
DEM_1212to1841=DF$PCscaled1[which(DF$DEM_Ybbs_reproj_PCscaled1>=classes$brks[5
] & DF$DEM_Ybbs_reproj_PCscaled1<classes$brks[6])]#TEST
DEM_1100to1746=DF$PCscaled1[which(DF$DEM_Ybbs_reproj_PCscaled1>=classes$brks[6
] & DF$DEM_Ybbs_reproj_PCscaled1<=classes$brks[7])]

OR ######Selection of Attributes (Corine Land Cover) ####

Stadt=(DF$PC2[which(DF$clc2000_Ybbs_reproj_pc2==7 |
DF$clc2000_Ybbs_reproj_pc2==8 | DF$clc2000_Ybbs_reproj_pc2==10 |
DF$clc2000_Ybbs_reproj_pc2==13 | DF$clc2000_Ybbs_reproj_pc2==14 |
DF$clc2000_Ybbs_reproj_pc2==18)])
Ackerbau=(DF$PC2[which(DF$clc2000_Ybbs_reproj_pc2==3 |
DF$clc2000_Ybbs_reproj_pc2==4 | DF$clc2000_Ybbs_reproj_pc2==15)])
Meadow=(DF$PC2[which(DF$clc2000_Ybbs_reproj_pc2==16)])
Wald=(DF$PC2[which(DF$clc2000_Ybbs_reproj_pc2==2 |
DF$clc2000_Ybbs_reproj_pc2==5 | DF$clc2000_Ybbs_reproj_pc2==9)])
Gebüsch=(DF$PC2[which(DF$clc2000_Ybbs_reproj_pc2==12 |
DF$clc2000_Ybbs_reproj_pc2==19)])
#Wasser=(DF$PC2[which(DF$clc2000_Ybbs_recl_reproj==4)])
Wasser=(DF$PC2[which(DF$clc2000_Ybbs_reproj_pc2==5 |
DF$clc2000_Ybbs_reproj_pc2==11)])

OR ########Selection of Attributes (CLC for mountaineous and Flat
Terrain)#####

Ackerland=DF$PCscaled1[which((DF$DEM_Ybbs_r_PC1>=200 & DF$DEM_Ybbs_r_PC1<=600)
& (DF$Ybbs_clc_pro==2 | DF$Ybbs_clc_pro==4))]
```

```
LaubundMischwald=DF$PCscaled1[which((DF$DEM_Ybbs_r_PC1>=200 &
DF$DEM_Ybbs_r_PC1<=600) & (DF$Ybbs_clc_pro==3 | DF$Ybbs_clc_pro==7))]
Grunland=DF$PCscaled1[which((DF$DEM_Ybbs_r_PC1>=200 & DF$DEM_Ybbs_r_PC1<=600)
& (DF$Ybbs_clc_pro==9 | DF$Ybbs_clc_pro==8 | DF$Ybbs_clc_pro==12 |
DF$Ybbs_clc_pro==13))]
Nadelwald=DF$PCscaled1[which((DF$DEM_Ybbs_r_PC1>=200 & DF$DEM_Ybbs_r_PC1<=600)
& (DF$Ybbs_clc_pro==5))]
Fels=DF$PCscaled1[which((DF$DEM_Ybbs_r_PC1>=200 & DF$DEM_Ybbs_r_PC1<=600) &
(DF$Ybbs_clc_pro==14 | DF$Ybbs_clc_pro==18))]
Wasser=DF$PCscaled1[which((DF$DEM_Ybbs_r_PC1>=200 & DF$DEM_Ybbs_r_PC1<=600) &
(DF$Ybbs_clc_pro==1 | DF$Ybbs_clc_pro==15 | DF$Ybbs_clc_pro==17 |
DF$Ybbs_clc_pro==16))]
Urban=DF$PCscaled1[which((DF$DEM_Ybbs_r_PC1>=200 & DF$DEM_Ybbs_r_PC1<=600) &
(DF$Ybbs_clc_pro==6 | DF$Ybbs_clc_pro==10 | DF$Ybbs_clc_pro==11))]


#########DENSITY Calculations######
D1=density("ATTRIBUTE",bw=0.8, na.rm=T) #die bandwidth (bw) muss t du f?r alle
gleich einstellen. Bisschen probieren, was gut aussieht
D2=density("ATTRIBUTE",bw=0.8, na.rm=T)
D3=density("ATTRIBUTE",bw=0.8, na.rm=T)
D4=density("ATTRIBUTE",bw=0.8, na.rm=T)
D5=density("ATTRIBUTE",bw=0.8, na.rm=T)
D6=density("ATTRIBUTE",bw=0.4, na.rm=T)

#####Plotting####
png(filename="D:/MASTERARBEIT/El
Texto/Figures/ANALYSE/Texture_PC1_Traisen.png", antialias = "cleartype",
pointsize=10, width=3000, height=1600, res=300)
par(mfrow=c(1,1))

colors1 = c("red", "yellow", "darkolivegreen4", "violet", "orange", "blue",
"darkorchid3", "cyan", "black", "brown")

plot(D1, lwd=2, xlim=range(DF$PC1_Traisen_repr_Tex,
na.rm=T),ylim=c(0,max(c(max(D1$y),max(D2$y),max(D4$y), max(D5$y), max(D7$y),
max(D8$y), max(D9$y), max(D10$y)))),col=colors1[1],
     main="PC1 vs Texture",
     sub="Traisen",
     xlab="")
lines(D2,col=colors1[2], lwd=2)
lines(D3,col=colors1[3], lwd=2)
lines(D4,col=colors1[4], lwd=2)
lines(D5,col=colors1[5], lwd=2)
lines(D6,col=colors1[6], lwd=2)
lines(D7,col=colors1[7], lwd=2)

rug((D1$x[D1$y == max(D1$y)]), col=colors1[1], lwd=2)
rug((D2$x[D2$y == max(D2$y)]), col=colors1[2], lwd=2)
rug((D3$x[D3$y == max(D3$y)]), col=colors1[3], lwd=2)
rug((D4$x[D4$y == max(D4$y)]), col=colors1[4], lwd=2)
rug((D5$x[D5$y == max(D5$y)]), col=colors1[5], lwd=2)
rug((D6$x[D6$y == max(D6$y)]), col=colors1[6], lwd=2)
rug((D7$x[D7$y == max(D7$y)]), col=colors1[7], lwd=2)


classes=as.list(c("ATTRIBUTE1", "ATTRIBUTE2"…etc))
legend("topleft",legend=paste(classes[1:7]),col=colors1,lwd=2)
classes
dev.off()
```

Sensitivity Analysis

```
##### Sort Files #####
refpath="Directory"
savepath="Directory"

##### Built Stack #####
```

```
files=list.files(refpath,pattern="KelvG.tif")
files=files[which(apply(as.matrix(files),2,nchar)==21)]

stackdasma=stack()

for (i in 1:length(files)) {
  Ri <- raster(paste(refpath,files[i],sep="/"),band=1)
stackdasma=stack(stackdasma,Ri)
  print(nlayers(stackdasma))
}

DF=as.data.frame(stackdasma) ##data frame

DF=na.omit(DF) ##NA`s raus

PCA=prcomp(DF, scale=T, center=T)

summary(PCA, loadings=F)

ALL.mean_Traisen_pc1=c()
ALL.mean_Traisen_pc2=c()
ALL.sd_Traisen_pc1=C()
ALL.sd_Traisen_pc2=C()


################1 OUT###############################

cors1.1=c()
loadings=c()
loadings1=c()
Name_out=paste(seq_along(DF))

for (i in 1:11){
  #Spalte i wird ausgelassen
  PCA.i=prcomp(DF[,-i], scale=T, center=T)
  #Correlation berechnet
  cors1.1[i]=abs(cor(PCA$x[,1],PCA.i$x[,1]))
  loadings1=append(loadings1,loadings$PC1)
}
stats_DF=data.frame(Name_out, cors1.1)
write.table(stats_DF, "E:/Thesis/Landsat-
5/No_CloudCover/Traisen/Method_Analysis/One-Out_Pc1_cors.csv")

ALL.mean_Traisen_pc1=mean(cors1.1) ALL.sd_Traisen_pc1=sd(cors1.1)

####Plot####
png(filename="Directory", width=3000, height=3000, units="px", res=300)
plot(cors1.1, xlab="dropped images",ylab="absolute correlation",ylim=c(0,1))
abline(h=0)
abline(h=mean(cors1.1),col="red")
abline(h=mean(cors1.1)+c(1,-1)*sd(cors),lty=2,col="red")
dev.off()

cors1.2=c() #Initialisierung des Speichers fuer Ergebnisse
loadings=c()
loadings1=c()
Name_out=paste(seq_along(DF))

for (i in 1:11){
  #Spalte i wird ausgelassen
  PCA.i=prcomp(DF[,-i], scale=T, center=T)
  #Correlation berechnet
  cors1.2[i]=abs(cor(PCA$x[,2],PCA.i$x[,2]))
loadings=as.data.frame(PCA.i$rotation)
  loadings1=append(loadings1,loadings$PC1)
}
stats_DF=data.frame(Name_out, cors1.2)
write.table(stats_DF, "directory.csv")
```

```
ALL.mean_Traisen_pc2=mean(cors1.2)
ALL.sd_Traisen_pc2=sd(cors1.2)

png(filename="Directory", width=3000, height=3000, units="px", res=300)
plot(cors1.2, xlab="dropped images",ylab="absolute correlation",ylim=c(0,1))
abline(h=0)
abline(h=mean(cors1.2),col="red")
abline(h=mean(cors1.2)+c(1,-1)*sd(cors1.2),lty=2,col="red")
dev.off()

#################2 OUT#############################

cors2.1=list()
for (i in 1:11){
  for (j in (i+1):11){
    if(j<=11 & i!=j){
      PCA.ij=prcomp(DF[,-c(i,j)], scale=T, center=T)
      #Correlation berechnet
      cors2.1[[paste(i,j,sep="-")]]=abs(cor(PCA$x[,1],PCA.ij$x[,1]))
}}}

mean(unlist(cors2.1))
sd(unlist(cors2.1))
write.table(cors2.1, "directory.csv")

png(filename="directory.csv", width=3000, height=3000, units="px", res=300)
plot(unlist(cors2.1), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors2.1),names(cors2.1))
abline(h=0)
abline(h=mean(unlist(cors2.1)),col="red")
abline(h=mean(unlist(cors2.1))+c(1,-1)*sd(unlist(cors2.1)),lty=2,col="red")
dev.off()

cors2.2=list()
for (i in 1:11){
  for (j in (i+1):11){ #hier ist der Trick, dass du von i weiter laeufst.
Damit ist keine Kombi doppelt.
    if(j<=11 & i!=j){ #und damit du dann nicht ueber die Anzahl der Spalten
rauslaeufst, solltest du das hier eintragen.
      #Spalte i und j wird ausgelassen
      PCA.ij=prcomp(DF[,-c(i,j)], scale=T, center=T)
      #Correlation berechnet
      cors2.2[[paste(i,j,sep="-")]]=abs(cor(PCA$x[,2],PCA.ij$x[,2]))
    }}}

mean(unlist(cors2.2)) #mittlere Korrelation (gross ist gut)
sd(unlist(cors2.2)) #Standardabweichung (klein ist gut)
write.table(cors2.2, "E:/Thesis/Landsat-
5/No_CloudCover/Traisen/Method_Analysis/Two-Out_PC2_cors.csv")

png(filename="directory.png",width=3000, height=3000, units="px", res=300)
plot(unlist(cors2.2), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors2.2),names(cors2.2))
abline(h=0)
abline(h=mean(unlist(cors2.2)),col="red")
abline(h=mean(unlist(cors2.2))+c(1,-1)*sd(unlist(cors2.2)),lty=2,col="red")
dev.off()
####################3 OUT###########################

cors3.1=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11)
    if(j<=11 & i!=j & k<=11 & k!=j & k!=i){

      PCA.ijk=CA.ij=prcomp(DF[,-c(i,j,k)], scale=T, center=T)
```

```
      cors3.1[[paste(i,j,k,sep="-")]]=abs(cor(PCA$x[,1],PCA.ijk$x[,1]))
    }}}

mean(unlist(cors3.1))
sd(unlist(cors3.1))
write.table(cors3.1, "directory.csv")


png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors3.1), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors3.1),names(cors3.1))
abline(h=0)
abline(h=mean(unlist(cors3.1)),col="red")
abline(h=mean(unlist(cors3.1))+c(1,-1)*sd(unlist(cors3.1)),lty=2,col="red")
dev.off()

cors3.2=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11)
      if(j<=11 & i!=j & k<=11 & k!=j & k!=i){

        PCA.ijk=CA.ij=prcomp(DF[,-c(i,j,k)], scale=T, center=T)

        cors3.2[[paste(i,j,k,sep="-")]]=abs(cor(PCA$x[,2],PCA.ijk$x[,2]))
      }}}

mean(unlist(cors3.2))
sd(unlist(cors3.2))
write.table(cors3.2, "directory.csv")

png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors3.2), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors3.2),names(cors3.2))
abline(h=0)
abline(h=mean(unlist(cors3.2)),col="red")
abline(h=mean(unlist(cors3.2))+c(1,-1)*sd(unlist(cors3.2)),lty=2,col="red")
dev.off()
###############################4
OUT##############################################


cors4.1=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11){
      for (l in (k+1):11)
      if(j<=11 & i!=j & k<=11 & k!=j & k!=i & l<=11 & l!=j & l!=i & l!=k){

        PCA.ijkl=prcomp(DF[,-c(i,j,k,l)], scale=T, center=T)

        cors4.1[[paste(i,j,k,l,sep="-")]]=abs(cor(PCA$x[,1],PCA.ijkl$x[,1]))
      }}}}

mean(unlist(cors4.1))
sd(unlist(cors4.1))
write.table(cors4.1, "directory.csv")

png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors4.1), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors4.1),names(cors4.1))
abline(h=0)
abline(h=mean(unlist(cors4.1)),col="red")
abline(h=mean(unlist(cors4.1))+c(1,-1)*sd(unlist(cors4.1)),lty=2,col="red")
dev.off()
```

```
cors4.2=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11){
      for (l in (k+1):11)
        if(j<=11 & i!=j & k<=11 & k!=j & k!=i & l<=11 & l!=j & l!=i & l!=k){

          PCA.ijkl=prcomp(DF[,-c(i,j,k,l)], scale=T, center=T)

          cors4.2[[paste(i,j,k,l,sep="-")]]=abs(cor(PCA$x[,2],PCA.ijkl$x[,2]))
        }}}}

mean(unlist(cors4.2))
sd(unlist(cors4.2))
write.table(cors4.2, "directory.csv")


png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors4.2), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors4.2),names(cors4.2))
abline(h=0)
abline(h=mean(unlist(cors4.2)),col="red")
abline(h=mean(unlist(cors4.2))+c(1,-1)*sd(unlist(cors4.2)),lty=2,col="red")
dev.off()

###############################5 OUT#############################

cors5.1=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11){
      for (l in (k+1):11){
        for(m in (l+1):11)
        if(j<=11 & i!=j & k<=11 & k!=j & k!=i & l<=11 & l!=j & l!=i & l!=k &
m<=11 &m!=j & m!=i & m!=k & m!=l){

          PCA.ijklm=prcomp(DF[,-c(i,j,k,l, m)], scale=T, center=T)

          cors5.1[[paste(i,j,k,l,m,sep="-
")]]=abs(cor(PCA$x[,1],PCA.ijklm$x[,1]))
        }}}}}

mean(unlist(cors5.1))
sd(unlist(cors5.1))
write.table(cors5.1, "directory.csv")

png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors5.1), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors5.1),names(cors5.1))
abline(h=0)
abline(h=mean(unlist(cors5.1)),col="red")
abline(h=mean(unlist(cors5.1))+c(1,-1)*sd(unlist(cors5.1)),lty=2,col="red")
dev.off()

cors5.2=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11){
      for (l in (k+1):11){
        for(m in (l+1):11)
            if(j<=11 & i!=j & k<=11 & k!=j & k!=i & l<=11 & l!=j & l!=i & l!=k
& m<=11 &m!=j & m!=i & m!=k & m!=l){

              PCA.ijklm=prcomp(DF[,-c(i,j,k,l,m)], scale=T, center=T)

              cors5.2[[paste(i,j,k,l,m,sep="-
")]]=abs(cor(PCA$x[,2],PCA.ijklm$x[,2]))
```

```
            }}}}}

mean(unlist(cors5.2))
sd(unlist(cors5.2))
write.table(cors5.2, "E:/Thesis/Landsat-
5/No_CloudCover/Traisen/Method_Analysis/Five-Out_PC2_cors.csv")


png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors5.2), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors5.2),names(cors5.2))
abline(h=0)
abline(h=mean(unlist(cors5.2)),col="red")
abline(h=mean(unlist(cors5.2))+c(1,-1)*sd(unlist(cors5.2)),lty=2,col="red")
dev.off()

##############################6 OUT#############################

cors6.1=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11){
      for (l in (k+1):11){
        for(m in (l+1):11){
         for(n in (m+1):11)
            if(j<=11 & i!=j & k<=11 & k!=j & k!=i & l<=11 & l!=j & l!=i & l!=k
& m<=11 &m!=j & m!=i & m!=k & m!=l &
                 n<=11 & n!=j & n!=i & n!=k & n!=l & n!=m){

              PCA.ijklmn=prcomp(DF[,-c(i,j,k,l,m,n)], scale=T, center=T)

              cors6.1[[paste(i,j,k,l,m,n, sep="-
")]]=abs(cor(PCA$x[,1],PCA.ijklmn$x[,1]))
            }}}}}}

mean(unlist(cors6.1))
sd(unlist(cors6.1))
write.table(cors6.1, "directory.csv")

png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors6.1), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors6.1),names(cors6.1))
abline(h=0)
abline(h=mean(unlist(cors6.1)),col="red")
abline(h=mean(unlist(cors6.1))+c(1,-1)*sd(unlist(cors6.1)),lty=2,col="red")
dev.off()

cors6.2=list()
for (i in 1:11){
  for (j in (i+1):11){
    for (k in (j+1):11){
      for (l in (k+1):11){
        for(m in (l+1):11){
          for(n in (m+1):11)
            if(j<=11 & i!=j & k<=11 & k!=j & k!=i & l<=11 & l!=j & l!=i & l!=k
& m<=11 &m!=j & m!=i & m!=k & m!=l &
                 n<=11 & n!=j & n!=i & n!=k & n!=l & n!=m){

              PCA.ijklmn=prcomp(DF[,-c(i,j,k,l,m,n)], scale=T, center=T)

              cors6.2[[paste(i,j,k,l,m,n, sep="-
")]]=abs(cor(PCA$x[,2],PCA.ijklmn$x[,2]))
            }}}}}}

mean(unlist(cors6.2))
sd(unlist(cors6.2))
write.table(cors6.2, "directory.csv")
```

```
png(filename="directory.png", width=3000, height=3000, units="px", res=300)
plot(unlist(cors6.2), xlab="dropped images",ylab="absolute
correlation",ylim=c(0,1),xaxt="n")
axis(1,1:length(cors6.2),names(cors6.2))
abline(h=0)
abline(h=mean(unlist(cors6.2)),col="red")
abline(h=mean(unlist(cors6.2))+c(1,-1)*sd(unlist(cors6.2)),lty=2,col="red")
dev.off()

ALL.mean_Traisen_pc1=c(mean(cors1.1), mean(unlist(cors2.1)),
mean(unlist(cors3.1)), mean(unlist(cors4.1)), mean(unlist(cors5.1)),
mean(unlist(cors6.1)))
ALL.mean_Traisen_pc2=c(mean(cors1.2), mean(unlist(cors2.2)),
mean(unlist(cors3.2)), mean(unlist(cors4.2)), mean(unlist(cors5.2)),
mean(unlist(cors6.2)))
ALL.sd_Traisen_pc1=c(sd(cors1.1), sd(unlist(cors2.1)), sd(unlist(cors3.1)),
sd(unlist(cors4.1)), sd(unlist(cors5.1)), sd(unlist(cors6.1)))
ALL.sd_Traisen_pc2=c(sd(cors1.2), sd(unlist(cors2.2)), sd(unlist(cors3.2)),
sd(unlist(cors4.2)), sd(unlist(cors5.2)), sd(unlist(cors6.2)))

write.table(ALL.mean_Traisen_pc1, "directory.csv")
write.table(ALL.mean_Traisen_pc2, "directory.csv")
write.table(ALL.sd_Traisen_pc1, "directory.csv")
write.table(ALL.sd_Traisen_pc2, "directory.csv")
```

ANNEXE V. CD-ROM

This work and supplementary materials can be found in digital form on an attached CD-ROM:

- Master's Thesis (PDF)
- TIR-Datasets (TIFF)
- PCA-Maps (TIFF)
- PCA Loadings and Proportions of variance (XLSX)
- Landscape Characteristics (TIFF)
- Sensitivity Analysis - Quantitative Results (CSV)
- Probability Distribution Functions (PNG)
- R-scripts (R)

Personal Declaration

„Hiermit versichere ich, dass ich diese Masterarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich versichere alle Stellen der Arbeit, die wortwörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht und die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt zu haben."

Wien, den 5. Mai 2015 Claude Meisch