Bayesian Formulation of Uncertainty in Paleoclimate Reconstructions

A Tree Ring Width Case Study

Masterarbeit

im Fachgebiet Meteorologie

vorgelegt von

Christopher Danek

am 22. Dezember 2014 an der Universität Bonn

begutachtet von: Prof. Dr. A. Hense

und Dr. C. Ohlwein

Abstract

Paleoclimate modeling is important to understand the responses of the Earth system to forcing mechanisms (e.g. solar and volcanic). Inverse modeling of the past climate provides an estimation of variabilities prior to the observational period. We provide annual average temperature reconstructions based on tree ring width proxy for the 20th century (41 trees) and high–resolution gridded data (0.5°, Climate Research Unit, Jones and Harris [2013]). The proxy data is standardized with exponential functions and splines. Temporal and spatial cross validation is applied. Two transfer models are compared, the classical regression and the Bayesian conditional probability. Covariance estimation via *gLasso* method (Friedman et al. [2008], the penalty parameter reduces estimated coavariances to 80 %). Deterministic and stochastic reconstructions are verified via scoring rules and skill scores.
The Bayesian transfer model shows better prediction skills for stochastic reconstructions compared to the regression model and the climate reference. In the deterministic case, the regression yield better results. The formulation of the full conditional probability density may provides an additional information to the paleoclimate reconstruction problem and allows a better assessment of past climate uncertainty.

Zusammenfassung

Paläoklimatologische Modellierung ist wichtig um das physikalische Erdsystem und seine Rückkopplungen und Antriebe zu verstehen (z.B. solare oder vulkanische Antriebe). Inverse Modellierung des Paläoklimas liefert eine Schätzung der Variabilitäten vor Beginn der Beobachtungsera. Wir liefern jährlich gemittelte Temperaturrekonstruktionen basierend auf Baumring Dicke Proxy Daten für das 20. Jahrhundert (41 Bäume) und einen hochaufgelösten räumlichen Datensatz (0.5°, Climate Research Unit, Jones and Harris [2013]). Die Proxy Daten werden mit Exponential– und Splinefunktionen standartisiert. Zeitliche und räumliche Kreuzvalidierung wird verwendet. Zwei Transfermodelle werden hergeleitet, die klassische Regression und ein Bayesisches bedingtes Wahrscheinlichkeitsmodell. Kovarianzen werden mit der *gLasso* Methode geschätzt (Friedman et al. [2008], der Strafparameter reduziert die geschätzen Kovarianzen auf 80 %). Deterministische und stochastische Rekonstruktionen werden mit scoring rules und skill scores verifiziert.
Das Bayesische Transfermodell zeigt bessere Vorhersagen im Falle der stochastischen Rekonstruktionen vergleichen mit dem Regressionsmodell und der Klimareferenz. Im determinstischen Fall liefert das Regressionsmodell bessere Ergebnisse.

Die Fomulierung einer vollen bedingten Wahrscheinlichkeitsdichte liefert zusätzliche Informationen im paläoklimatologischen Rekonstruktionsproblem und erlaubt eine bessere Abschätzung der Unsicherheiten des vergangenen Klimas.

Contents

1 Introduction **1**

2 Background **5**

2.1 Dendroclimatological Research 5

2.2 Principles of Dendrochronology 9

2.2.1 Structure and Growth of a Tree Ring 9

2.2.2 Limiting Factors and Site Selection 10

2.2.3 Standardization 11

3 Methods & Data **15**

3.1 Notation 15

3.2 Model A: Multiple Linear Regression 18

3.3 Model B: Conditional Density of the Multivariate Normal Distribution 19

3.4 Data Preparation and Construction 23

3.4.1 Proxy Data 23

3.4.2 Environmental Data 24

3.4.3 Implementation 25

3.4.3.1 Cross Validation for Conditional Transfer Model B . 27

3.4.3.2 Estimate the Inverse of a Covariance Matrix via Lasso Regression 27

3.5 Verification 29

4 Results **33**

4.1 Model Comparison 36

4.2 Time Truncation 39

5 Discussion **45**

6 Conclusions **51**

Bibliography **53**

1 Introduction

The estimation of the variability of the climate system is a crucial task to accomplish in atmospheric sciences. With observational data of numerous environmental parameters of the climate spheres, climatologists are able to quantify processes that form our daily weather and year to year climate. Additionally, numerical models based on laws derived from theory of the dynamics of the climate system are used to obtain an added value to the limited availability of measurements.
However, by using measurements and statistical or numerical models, *uncertainties* occur at each step of the scientific process due to lack of information, wrong assumptions that have been made or due to the fact that there is only a single realization of the experiment 'climate' from which all the measurements are taken from. These uncertainties even increase, if there are gaps in the time series of the measurements or – even more problematic – if we want to analyze the past climate from pre–observational times, where no observational data are available.
Paleoclimatologists (paleo greek for "old" or "ancient") try to estimate the past state of the climate based on *proxy data* from which they assume that this data is affected by the climate itself. That is, if the climate conditions change, the proxy data should change as well (Fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC), Working Group I – The Physical Science Basis (WGI), Masson-Delmotte et al. [2013]).
There are many phenomena in the natural world from which we know that they react to changes of the surrounding meteorologic or climatologic conditions such as tilting sunflowers during the day or falling leaves from trees in mid latitude regions in autumn. Both objects the sunflowers and the leaves react to changes of the environment (to face the sun and a consequence of decreased amount of sun light respectively). Now if one wants to derive climate properties of the blooming time and autumn last year, a problem arise: the sunflowers and leaves are likely rotten.
Thus, natural objects are needed which preserve their reactions to changes of the surrounding conditions. The growth of a human being is preserved in the way, that one can measure size and weight as an parametrization for 'growth' during the lifetime of that animal in order to obtain a time series of 'growth' as a function of time. However, it is not reasonable to get an idea about the climate conditions during the lifetime of the human being because its 'growth' depends on factors such as the genetic endowments, nutrition and physical activity rather than climate.
So environmental phenomena which fulfill both of the desired properties, i.e. preserve reactions caused mainly through climate changes are needed. In paleoclimatology these are called *climate archives* or *climate indicators* and they preserve their reactions in form of climate proxies (Jones and Mann [2004]). Such a climate indicator is a tree ring, a layer of wood formed through ecophysiological processes

such as cell division, cell enlargement, and cell differentiation in a tree during an annual astronomic cyle or a growing season (Schweingruber [2007]). This tree ring carries several measureable climate proxies such as the width and density of the tree ring, wood anatomic properties, and stable isotopes of oxygen, carbon or hydrogen (Treydte et al. [2004]). From all of them we assume that their formation was affected mainly by the surrounding climate conditions (Fritts [1976]). In doing so, we are able to reconstruct information on climate prior to the instrumental era via a *response* or *transfer function* that connects proxy and environmental data. For tree ring proxy data, this field of atmospheric sciences is called dendroclimatology (dendron greek for "tree").

There are several more climate indicators like historical data, corals, pollen, ice cores, and ocean and lake sediments. All of them are frequently used by paleoclimatologists since their ability of reconstructing the past climate is well known (Masson-Delmotte et al. [2013]). For example, an ice core record carries the climate proxies fraction of melting ice, rate of accumulation of precipitation, concentrations of trace gases enclosed in the ice and isotope proportions from the time the ice was deposited (Jones and Mann [2004]). In a similar way, corals (Rimbu et al. [2001]) or lake sediments may used (Litt et al. [2012]).

Tree rings are advantageous for reconstructing past variations in climate from a paleoclimate perspective. For example, tree ring width can be easily measured and thus quantified compared to other climate proxies, and the obtained time series have a very high temporal resolution (annual or a growing season) and so the climatic information is exactly placed in time (Fritts [1976]).

The process of finding the transfer function between the environmental variable and the proxy data is also known as the *calibration* of the proxy data. The aim is to determine quantitatively in what manner the proxy data responds to the environment in means of its temporal or spatial variabilities and vice versa. This procedure belongs to the field of statistical modeling and is realized via its methods. Traditionally there are two directions of calibration: *classical* and *inverse* calibration – or forward and inverse modeling respectively. In classical calibration, the proxy variable is modelled on environmental data whereas in inverse calibration, the environmental variable is modelled on proxy data (Robertson et al. [1999]). Afterwards, proxy data from either non–observational era or times where observational data are available but that was not used in the calibration dataset (*training* data) is substituted in the derived statistical model. By that the model is *verified* against known observational or model–independent data either in the proxy data space (forward model) or environmental data space (invers model). Finally, the goodness of the statistical estimate of the past climate can be assessed (Fritts [1976]).

A crucial question that arises is how certain we can be about these reconstructions of the past climate. Model predictions are uncertain because of natural variations in climate, uncertain pathways of any forcing mechanisms as well as unknown feedback processes (Collins [2007]). Furthermore one may ask if it is possible to quantify these uncertainties? As a prominent example the IPCC deals with that problem by using qualitative rather than quantitative terms such as 'likely', 'unlikely' or 'virtually certain' representing certain degrees of uncertainty (Cubasch et al. [2013]). In order to quantify these uncertainties, *probabilities* are used. Basically, probabilites are numbers which measure how likely an outcome of an experiment will be. However, there are two fundamental perspectives on the meaning of probabilites in statistics: the *frequentist* and *bayesian* perspective. For the first, the probability of an event is defined as the relative frequency of occurences of that event in a large number of trials. In other words, the ratio of the occurences of the event to all opportunities, where the event could have been occured. So frequentists see derived parameters as some sort of true or fixed values Ellison [1996]. In contrast, bayesian statistics give an explicit formulation for the uncertainty of the derived parameters. The use of probabilities includes assessments of how likely it is that a parameter takes a certain value compared to another (Gelman et al. [2004]).

Aim of this Thesis

This is a Master thesis in the field of paleoclimate modeling and examines the ability of two different statistical transfer models to deal with uncertainties in paleoclimate reconstructions. Exemplarily for the dendroclimatologic proxy 'tree ring width', two different transfer functions are derived via 1) *linear regression* and 2) the formulation of *conditional probabilites* in order to reconstruct annual average air temperature in two meters height of the 20th century. The skill of these models in reconstructing past climates is assessed with and without an additional statistical noise.

In chapter 2, a review about paleo–climate modeling with tree ring proxies is given followed by a short overview about the principles of dendrochronology.
The two transfer functions are derived in chapter 3. Also, information about the used proxy and environmental data is given as well as a description of the implementation of the paleoclimate models and their verification. Finally, the results are shown (chapter 4) and discussed (chapter 5). Finally, a summary is given in chapter 6.

2 Background

This chapter provides a review about dendroclimatological research (sec. 2.1) and gives a short introduction in dendrochronological principles (sec. 2.2).

2.1 Dendroclimatological Research

Numerous dendroclimatologic studies revealed the relationship between tree ring proxies and past annual or seasonal average temperatures and precipitation patterns in many parts of the globe. In inverse modeling, i.e. if the environmental variable is predicted from the proxy data, transfer models are mostly derived by the principles of linear regression.
In forward modeling, growth models such as the "Vaganov–Shashkin" (Evans et al. [2006]) or "MAIDEN" (Misson [2004]) are used to generate artifical proxy data based on empirical derived relationships between the tree growth and environmental variables like solar radiaton, near–surface air temperature and soil water balance. Since forward modeling is beyond the scope of this study, it is not longer discussed here.
Figure 2.1 is taken from the recent 5th report of the IPCC (Masson-Delmotte et al. [2013]) and shows reconstructed Northern Hemisphere annual temperature anomalies during the last 2000 years with respect to the climate average of 1881–1980 (Masson-Delmotte et al. [2013]). Most of these anomalies are derived via inverse regression techniques. It becomes clear that these predictions are uncertain regarding

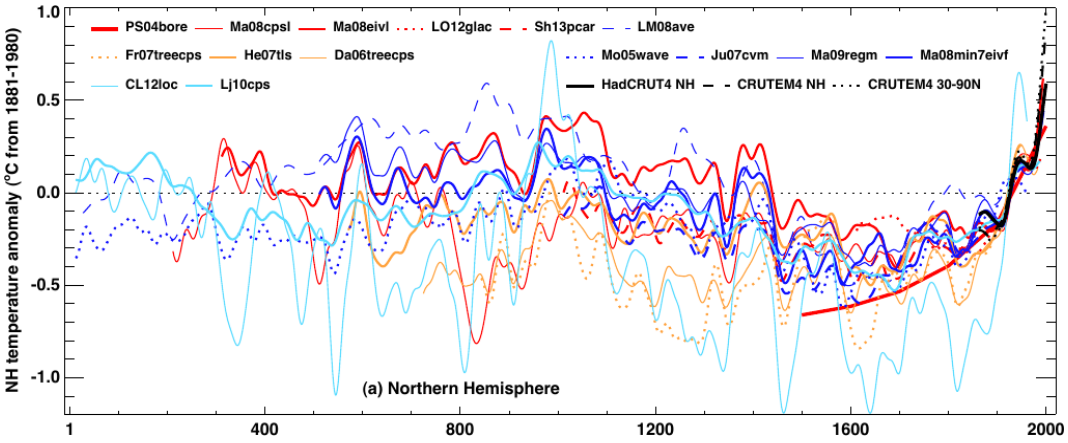

Figure 2.1: Illustration of several reconstructed Northern Hemisphere annual temperature anomalies during the last 2000 years with respect to the climate average of 1881–1980 (taken from Masson-Delmotte et al. [2013]).

differences in the variabilities of the time series as well as the sign of the anomalies, whether they are positive or negative. For example, the prediction by Christiansen and Ljungqvist [2012] (thin pale blue line in figure 2.1) fluctuates much stronger

than the one of Pollack [2004] (thick red line) from the year 1600 to 2000. The temporal resolution of the first study is annual, while the one of the second study is centennial. That is because different proxy data were used. Christiansen and Ljungqvist [2012] used multiple annually resolved proxies (e.g. tree ring densities and widths, varved lake sediments as well as documentaries), whereas Pollack [2004] used terrestrial boreholes from which past temperatures can be measured in certain depths. However, these are much less resolved yielding a smoothed time series. Nevertheless, both predictions show a positive trend in recent Northern Hemisphere temperature anomalies.
In contrast to the borehole reconstruction, the multi–proxy prediction of Christiansen and Ljungqvist [2012] dates further back in time. Also, the reconstruction obtained by Moberg et al. [2004] represents the last two millennia (dotted blue line in figure 2.1). Both used especially tree ring proxys to obtain long cliamte reconstructions.

Through the advantages of tree ring proxies, being annually resolved and able to reconstruct past climates of the last millenia, many dendroclimatological studies were conducted in the past centuries. Jacoby et al. [1996] used records of coniferous species across northern North America which were extended using archaeological and driftwood material via cross dating (see sec. 2.2). They point out that their derived reconstructions reflect decade to century scale climatic information but depend on region, site, species and standardization methods. Furthermore, they used tree ring records to validate climate models regarding predictions of climate forcings. Luckman [1996] show a relationship between suppressed tree ring widths during glacial events in the holocene corresponding to the early 18th, mid–19th and late 20th century as well as for glacial periods referred as “Little Ice Age” in the 15th and 16th century. Briffa et al. [1998a] regressed Tree ring densities against Northern Hemisphere Land and Marine Temperature and found that this represents a good estimator for Northern Hemisphere mean summer temperatures. Furthermore, signals from large explosive volcanic eruptions were reproduced which influenced the extents of Northern Hemisphere cooling during the past 600 years. Briffa et al. [1998b] showed that averaged tree ring width and tree ring density records show a strong relationship with summer temperatures averaged over the same areas (annual correlations between summer temperature (June to August) and tree ring width around 0.3 and 0.4). The ability of representing the mean temperature on continental or even Norhtern Hemisphere scale is emphasized.
Hughes and Gaumlich [1996] were able to reconstruct precipitation patterns and droughts in the Great Basin in western North America via Bristlecone pines records covering the period 6000 B.C. to A.D. 1979 (Bristlecone refers to different species

of pine trees growing in arid regions of western North America which are one of the oldest living trees). Stahle and Cleavelend [1996] showed a similar sensitivity of tree ring growth to rainfall amounts exemplarily on Baldcypress trees. Their March to June rainfall reconstructions represent 50 to 70% of the instrumentally recorded rainfall variance. They used the statistical method *Principle Component Analysis* or *Empirical Orthogonal FUnctions* in order to obtain the dominant patterns (*principal components*) of variability (von Storch and Zwiers [1999]). Besides, they found that growth patterns of Baldcypress trees reflect the large–scale Pacific–North America (PNA) circulation pattern. In this context, Lauterbacher et al. [1999] were able to reconstruct indices of the North Atlantic Oscillation (NAO) and Eurasian (EU) circulation back to 1675 with tree ring data and *canonical correlation analysis* (CCA). In contrast to EOFs, the correlation structure of a pair of variables is considered instead of the variability of only a single variable (von Storch and Zwiers [1999]).
Fritts et al. [1971] used a combination of the first principal components of annual temperature and precipitation data and regression to derive their paleoclimate transfer function. They found positive correlations between tree ring growth and average temperatures of the autumn of the year when the tree ring is formed, and the spring of the following year. Negative correlations with temperature were found especially for months of the growing season of the same year. They suggest that decreased ring width growth may be a consequence of increased consumption of available moisture and stored nutrients, induced by increased water loss and respiration if temperatues are high. Furthermore, they mention that the transfer functions may vary with site and taxon, because growth of one species at one site could be more or less related to the amounts of prior growth and refer to *autocorrelation* of the tree rings. That is, how the growth one year is inter–correlated with the growth of preceding or following years.
Cook et al. [1994] used principal components for spatial regression analysis. The term “spatial“ refers to a representation of proxy data from multiple sites that are incorporated in one transfer model. They found that records of late wood densities explained the variances of i) the montly mean surface air temperature from April to September (similar to findings of Schweingruber [1996]), ii) the Palmer draught severity index (a measure for the intensity of a draught) for June to August, and iii) the sea–level pressure from June to August.
Another approach are *pseudo* proxies to face the problem of a limited spatial representation and temporal resolution. Introduced by Mann and Rutherford [2002], pseudo proxies represent a subset of the output of a numerical global climate model which is ’degraded’ and pertubed by adding statistical noise to the data in order to represent the noise characteristics of real world proxy data. Afterwards, recon-

struction algorithms are applied to the pseudo proxies yielding reconstructions based on the climate, which was simulated by the global climate model (Smerdon [2012], Smerdon et al. [2011]).
Briffa et al. [1992] used a combination of regression and *power spectrum analysis* to derive their transfer function. The power spectrum refers to Fourier transform of the auto–covariance function of a variable and is useful for time series analyses since quasi–oscillatory components in a process become visible as peaks in its spectrum. The authors found significant timestable periods in the autocovariance functions of reconstructions based on tree ring width and density records from northern Sweden. These are on decadel rather than annual time scales of approximately 30 to 100 years. The authors conclude that the variance on these timescales might indicate a true cycle represented by near–periodic forcings external to the climate system or an internal atmospheric or ocean process influencing northern summer temperatures.

Most of these studies have in common, that the transfer functions were derived via regression analysis. However, recent findings suggest that these reconstructions tend to underestimate past climate variability. von Storch et al. [2004] set up a coupled atmosphere ocean global circulation model (AOGCM) for the past 1000 years to produce a reference for comparison with the regression based reconstructions estimated by Mann et al. [1999] or Mann et al. [1998]. Moreoever, they used pseudo proxies with an additional white noise. The climate variations produced by the numerical model were stronger than those of the regression based reconstructions. Similar results were obtained by Zorita et al. [2004], Zorita et al. [2003] and Rind et al. [2004].
In this context, Robertson et al. [1999] suggest to use a probabilistic approach for deriving statistical transfer functions in paleoclimate modeling. In doing so, some kind of prior belief or assumption about how a specific variable should be modified. A simple example is given by the authors: assume that all possible temperature values which may influence the formation of an annual tree ring of a particular species lie inside the range of 10 to 50 °C. This is true since all trees of that species are growing under these temperature conditions. Considering this as an iterative process, a better and better prior knowledge about the variable may be found. Bayesian statistics treats variables as random variables instead of deterministic quantities (Ellison [1996]). Thus, the possibilities of modeling a paleoclimate transfer function are broadened and probabilities may be used in order to estimate the uncertainty of a derived parameter. Moreover, a conditional probability may be derived. In words of inverse paleoclimate modeling, this quantity gives an estimation about how likely the occurence of a certain temperature is, given a certain tree ring width (see chapter 3).

Several Bayesian models have been developed in paleoclimate reconstruction studies and revealed the usefulness for these purposes (Korhola et al. [2002], Robertson et al. [1999], Haslett et al. [2006]).
Moreoever, comparisons with classical paleoclimate transfer concepts such as regression show that a Bayesian formulation may yield better results especially regarding the uncertainty of the model (e.g. Ohlwein and Wahl [2012] and Guiot et al. [2009] for reconstructions based on pollen abundances, Robertson et al. [1999] for carbon isotope indices derived from oaks). Röpnack et al. [2013] showed that a Bayesian formulation of a verification model is appropriate to compare an ensemble of numerical weather prediction models. Moreover, via averaging over a number of Bayesian models, more information can be obtained compared to the ensemble (Min et al. [2007]).
Buser et al. [2009] applied a Bayesian model to temperature changes as simulated by global and regional circulation model. They obtained a probabilistic estimate of climate change that can be seen as advantageous.
Often it is noticed that obtaining transfer functions from the Bayesian perspective needs a huge amount of computational efforts (e.g. Vasko et al. [2000] and Guiot et al. [2009]). Therefore, sampling algorithms such as *Markov chain Monte Carlo* (MCMC) are used to draw a large number of samples of probability distributions.

In this study, a Bayesian formulation of the paleoclimate reconstruction problem is derived and compared to the classical regression approach. Computationally intensive methods are omitted here explicitly and the model is constructed analytically. The following provides a short overview about the dendrochronological principles which are the basis for finding a link between environmental variables such as temperature and tree ring proxies.

2.2 Principles of Dendrochronology

This section provides a short overview of the principles of dendrochonology following Fritts [1976]. These principles aim to incorporate biological and statistical modeling. Some are more related to the system 'tree' whereas others are more general formulations and true for numerous biological systems.

2.2.1 Structure and Growth of a Tree Ring

Generally, a tree ring can be seen as the result of diameter growth of a tree. Figure 2.2 shows a schematic cross section through the stem of a tree. The *bark* is the outer

layer of a tree, which protects it and can be separated into outer and inner bark. Underneath the inner bark is the *cambium* which can be seperated in *phloem* and *xylem* (not shown in figure 2.2). If cells in the cambium divide, the derived cells un-

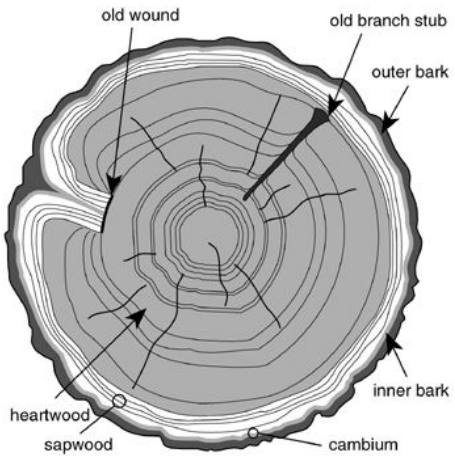

Figure 2.2: Schematic cross section of a tree (`http://pixgood.com/cambium-tree.html`, 21.12.2014)

dergo cell enlargement and differentiation which produces new xylem and phloem. It follows the *sapwood* followed by the *heartwood*. Without going into detail, the difference between both is that the process of *lignification* is finished in the heartwood, whereas continuing in the sapwood. Lignification means the thickening of cell walls by incorporation of *lignin*. In other words, the cell is transformed into wood. Finally, the center of the tree, the *pith*, is reached. According to Schweingruber [2007], radial growth of trees depends on age as well as site conditions. Concerning age related growth trends, the cell division capacity and velocity cell differentiation is slowing down with increasing age of a tree. In this context, *juvenile* wood refers wood structures of periods, when the growth of the tree is characterized by a progressive increase in cell dimensions. Whereas *adult* wood is represented through cells of more or less constant size. It has been found that radial growth of trees usually exhibit a decreasing trend with increasing age of the tree (Cook et al. [1995], Fritts [1976]).

2.2.2 Limiting Factors and Site Selection

From ecological sciences, the principle of *limiting factors* states that a tree cannot grow faster than allowed by its most limiting variable. The limiting factors for a tree may grouped in *external* and *internal* ones. Water, temperature, light, carbon

dioxide, oxygen and soil minerals are the external limiting factors. Internal limiting factors are the amount of nutrients, minerals, enzymes and water. Note, that is only a rough overview about limiting factors of a tree. Other possible influences such as an improved fertilization due to increased carbon dioxide and nitrogenous components, negative effects of acid rain, increased ultra–violet radiation (DeLuisi et al. [1994]), or increasing insect herbivory (Fleming and Volney [1995]) are issues of ongoing research. Hättenschwiler et al. [1996] showed in an experiment by simulating montane forest climate in containers, that radial stem growth increased with rising nitrogen concentrations. The wood density increased with rising carbon dioxide concentrations but decreased when more nitrogen was deposited. They conclude that rising atmospheric carbon dioxide will not lead to an enhanced radial growth of tree rings. In contrast nitrogen likely does.
By considering all possible limiting factors, the resulting growth of a tree can be seen as an *aggregated* growth. In doing so, the radial growth may be represented by a growth function G, which is defined as the sum of limiting factor related growth, yielding $G_t = A_t + C_t + D_t + \epsilon_t$, where t stands for a particular year. That is, the growth of the tree is a function of age– (A), climate– (C), and disturbances–related growth (C) as well as depending on unknown processes, indicated by ϵ (see the illustration of a wound in the stem of the tree in figure 2.2).
As can be seen, most of the limiting factors mentioned above are related to climate. Disturbances that can influence tree ring growth are manifold. Non–climatic noise may affect growth of the tree and can be caused by site–related effects such as wildfire, rockfalls, or anthropogenic influences and biological effects such as ageing or competition. Through the integrative character of the aggregated tree growth, biological and ecological impacts are also present in the measured parameter. As a consequence, the variability of the obtained measurement will depart from the variability associated with climate.
Note that correlations between tree ring proxy data and environmental data are not as high as normally expected in natural or engineering sciences. For example, correlation values of about 0.6 would not represent a strong statistical relationship in general but talking about proxy data such as tree ring width, values of $0.5 - 0.8$ are common (Robertson et al. [1999]).

2.2.3 Standardization

The mentioned disturbances and age–related growing trends during a lifetime of a tree result in variabilities which are likely not related to climate alone. These departures are often systematic and must be eliminated prior to any statistical analysis or

calibration purposes in dendroclimatology as well as other paleoclimate fields (Esper et al. [2003]). This detrending is called *standardization* or indexation and is assessed through modeling the actual measurements by deterministic functions.
In dendrochronological studies, *exponential* functions are often used to obtain standardized indices because they take into account the allready mentioned age–related growth trend. That is, the exponential function is constructed so that the influence of probably fast juvenile growth is removed. Furthermore, *spline* functions are used which can be seen as a generalized form of an exponential function. Splines are piecewise constructed polynomial functions. They are constructed around a certain number of *knots*.
In this study, a standardized proxy record is obtained by dividing the raw measurements by the modelled fit curve. Division through the fitted ring width values removes the trend in growth and scales the variance so that it is approximately the same throughout the entire length of the time series. The obtained chronology is unitless and has a mean value of one. Consequently, standardized location mean chronologies (average over all trees from one location) can be compared even if trees with very large and trees with very narrow rings are included in the dataset of the particular location. That means a standardization needs to be done before averaging over a specific number of trees, otherwise fast– and strong–growing trees would dominate the obtained signal (Fritts [1976]).
In contrast, standardization techniques imply a possible loss of low–frequency climate signals (Briffa et al. [2002], Briffa et al. [1996]) and variances on timescales above the age of the tree are certainly removed (Briffa et al. [1992]). The latter refers to the 'segment length curse' and is related to the fact that the maximum wavelength of recoverable climatic information is ordinarily related to the lengths of the individual series (Cook et al. [1995]). Note that an increasing number of knots for constructing the polynomial spline function yields in a better fit but also removes more variabilities from the time series and the climate induced signals.

In this thesis, four different standardization methods are applied on tree ring width data: a negative exponential fit (Fritts [1976]), a Hugershoff curve (Fang et al. [2010]), and two polynomial splines (cubic splines, constructed via 5 and 15 knots, Perkins and Swetman [1996]). Figure 2.3 illustrates exemplarily their effect on the raw tree ring width measurements. It becomes clear that the annual variability of the raw measurement is much higher than these of the fit curves. The red curve in figure 2.3 represents a negative exponential function of the form $f(t) = a\exp(-bt) + k$ and takes into account the juvenile growth trend mentioned before. The black curve represents the Hugershoff curve which can be seen as an extended exponential function of the form $f(t) = at^b\exp(-ct) + k$. Additionally, the Hugershoff curve is used to

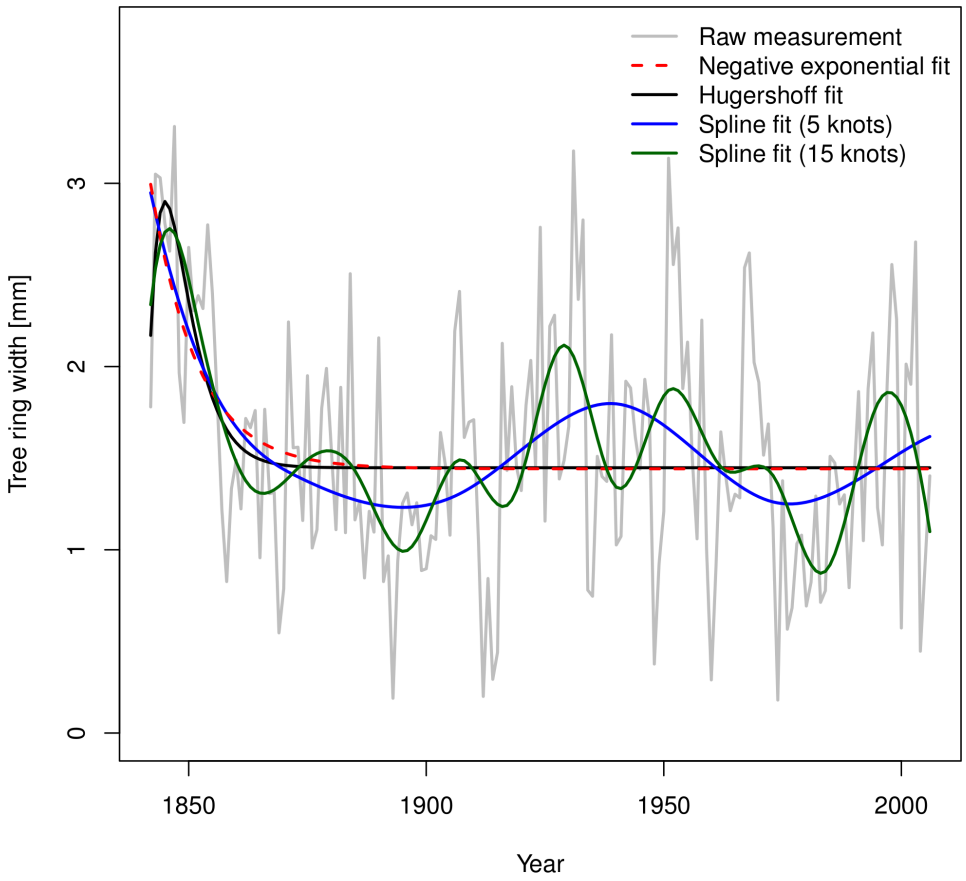

Figure 2.3: Example for fit curves of standardization process. Negative exponential function ($f(t) = a\exp(-bt) + k$, Fritts [1976]), Hugershoff curve ($f(t) = at^b\exp(-ct) + k$, Fang et al. [2010]) and polynomial spline functions with 5 and 15 knots, (Perkins and Swetman [1996]). For the exponential functions, the time t serves as the predictor.

remove age related growth trends in the beginning of the lifetime of a tree. This is indicated by the peak in the beginning of the time series. In both exponential functions the time t serves as the predictor to obtain the fit curve. The spline functions are shown via the blue (5 knots) and green (15 knots) curves in figure 2.3. Note that an increased number of knots yields in a fit curve of increasingly higher frequencies. Fritts [1976] notes that there is no theoretical basis for using splines for standardization of tree ring width data. However they seem most convenient compared with the negative exponential fit or the Hugershoff curve. Splines are more flexible than the other functions and can be used for complicated growth variations, and they can approximate the exponential function as well. However, an increasing number of knots used for the construction of the polynomial function also removes increasingly shorter wavelengths from the measurement chronology (Speer [2010]).

Another method of standardization is the *Regional Curve Standardization* (RCS). Although RCS has the potential to preserve long–timescale climate information (Briffa et al. [2013]) this standarization method is not used here due to the following limitations. 1) Exact dating of the tree ring width cores is necessary. That means not only the present (last) year needs to be included in the core but also the first growing year represented by the pith must be available. Often the first few years are not included in a core and the number of missing rings to the pith is estimated (Cook et al. [1995]). 2) A core of every age of a chronology is needed since all measurement series are aligned in the order of their cambial age. If a specific cambial age is missing, i.e. if a tree ring width record of a specific age is not avialable for a site, that age is underrepresented in the site chronology. In contrast, cambial ages which occur more often are overrepresented.

It is important to note that there is not a single method of standardization that fully removes age–related and non–climatic growth trends from every tree of every species from every site. Obtained index series need to be intently analyzed before continuing with statistical analyses.

3 Methods & Data

This section describes the statistical methods for deriving transfer functions between the proxy and the environmental data. A short introduction into multivariate statistics and notation is given (sec. 3.1), followed by a comparison of two different approaches of calibration. First, linear regression as an example from the Frequentist viewpoint (hereafter Model A, sec. 3.2) and second, the conditional probability density from the Bayesian viewpoint (hereafter Model B, sec. 3.3). Moreover, a description of the used data and its implementation is given (sec. 3.4) followed by model verification methods (sec. 3.5).

3.1 Notation

Let $\mathbf{X}$ represent a finite number of possible outcomes of an experiment (samples), drawn from the sample space $\mathbf{\Omega}$, which represents all possible outcomes of the experiment. The realizations $\mathbf{x}$ that $\mathbf{X}$ will take are not known in advance and depend on the events in $\mathbf{\Omega}$ that take place, why $\mathbf{X}$ is random. Furthermore let $\mathbf{X}$ be continuous, i.e. that it is impossible for $\mathbf{X}$ to take a distinct value. Instead, $\mathbf{X}$ takes an infinitesimal small interval of values where the distinct value is included (von Storch and Zwiers [1999]). Thus $\mathbf{X}$ can be seen as a multivariate continuous random variable definded as the set

$$\mathbf{X} = \left\{ (\mathbf{X}, f_{\mathbf{X}}(\mathbf{x})), \mathbf{x} \in \mathbf{\Omega}^n \right\}, \qquad \mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \,. \tag{3.1}$$

That is, $\mathbf{X}$ is a n dimensional quantity taken from the sample space $\mathbf{\Omega}$. The function $f_{\mathbf{X}}(\mathbf{x})$ stands for the *probability density function* (pdf) and is necessary to quantify the probabilities that the variable $\mathbf{X}$ takes the particular values $\mathbf{x}$. In fact, speaking about continuous random variables, the *probability* of an event is expressed as the integral of the pdf over the integral, that describes the event (von Storch and Zwiers [1999]). This integral is called the *cumulative distribution function* (cdf) of the continuous random variable and is defined as $F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{\mathbf{x}} f_{\mathbf{X}}(\mathbf{r})\, \mathrm{d}\mathbf{r}$. That is, the probability, that a continuous random variable takes particular values, is defined when this integral can be evaluated.
Generally, a pdf is definded via its *parameters* which are not known in advance and need to be *estimated* (von Storch and Zwiers [1999]). Assuming that the realizations of $\mathbf{X}$ are drawn from a *normal* or *Gaussian* distribution (second expression in

eq. 3.1), f is definded by two parameters: the *mean* $\boldsymbol{\mu}$ and the *variance* $\boldsymbol{\Sigma}$. The analytical definition of the pdf for the normal distribution is

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\,|\boldsymbol{\Sigma}|^{1/2}}\,\exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\,\boldsymbol{\Sigma}^{-1}\,(\mathbf{x}-\boldsymbol{\mu})\right), \tag{3.2}$$

with the dimension n of the random variable $\mathbf{X}$, $|\cdot|$ denoting the determinant, and $\cdot^T$ the transposed. From now on, the argument of the exponential term in the formulation of the pdf is called *Mahalanobis distance* of $\mathbf{X}$ and is defined as $\mathcal{D}^2(\mathbf{X}) = (\mathbf{X}-\boldsymbol{\mu})^T\,\boldsymbol{\Sigma}^{-1}\,(\mathbf{X}-\boldsymbol{\mu})$.
It becomes clear that knowing $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ is sufficient to calculate the full probability density function of the normal distribution. Moreover, the parameters of f can be estimated via the *likelihood function* l of $\mathbf{X}$ which represents certain values of f when its parameters are known. By claiming, that the likelihood should take its maximum value and taking the logarithmus ($l = \ln(f) \stackrel{!}{=} \max$), the parameters of f can be derived. This method is called maximum log–likelihood and may be found in Friederichs and Hense [2005] on page 16. The estimators for the two parameters can be defined as follows:

$$\boldsymbol{\mu} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\,, \qquad \boldsymbol{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i-\mu_i)(\mathbf{x}_i-\mu_i)^T\,, \tag{3.3}$$

with n the dimension of the ith realization of $\mathbf{X}$, $\mathbf{x}_i$, and its corresponding expectation value μ_i. Remarkably, the maximum log–likelihood estimators of the two parameters of the normal distribution equal the first two *moments* of $\mathbf{X}$, its expectation value or mean and variance.
The first moment stands for the mass of the data, i.e. which value most of the data take and is a vector of means in the multivariate case. The second moment stands for the variance of the data, i.e. how far the individual datapoints depart from their mean or the size of the spread of the data. The variance is responsible for the fluctuations of the time series shown in figure 2.1 in the introduction.
In our multivariate case, the variance quantifies the co–variabilites between all realizations of $\mathbf{X}$ and is therefore called *covariance*. This is not a single number but a quadratic matrix being symmetric as well as positive semi–definite (Wilks [2006]). This quantity is important for the statistical calibration process described in the introduction since it represents a link between all the realizations of $\mathbf{X}$ in a quantitative way. In the definition of the pdf, the Mahalanobis distance $\mathcal{D}^2(\mathbf{X})$ represents this distance of the individual datapoints and the expected value $\boldsymbol{\mu}$, weighted by their co–variabilities $\boldsymbol{\Sigma}$. The larger the distance is, the smaller the value of the probability density will be because of the minus sign in the exponential term. Note

that for the calculation of the Mahalanobis distance, the *inverse* of the covariances are needed. However, it is possible that the estimated covariance matrix of $\mathbf{X}$ is non–singular, i.e. its inverse does not exist. In this case it is impossible to evaluate the pdf. The problem of non–singularity of the covariance matrix is described in sec. 3.3.

Bayes Rule

The aforementioned probability, that the event $\mathbf{X}$ takes the particular value $\mathbf{x}$ may be extended to a *conditional* probability. Therefore, let $\mathbf{Y}$ be another multivariate continuous random variable that may represent unknown parameters of unobserved data. Then $[\mathbf{X}|\mathbf{Y}]$ means the probability that $\mathbf{X}$ takes the particular value $\mathbf{x}$ *given* that $\mathbf{Y}$ takes the value $\mathbf{y}$. Here, $[\cdot|\cdot]$ refers to a conditional probability density function.
In doing so, a measure for the *joint* probability, i.e. for the combined event $\mathbf{X}$ *and* $\mathbf{Y}$ can be defined as the product of the *marginal* or *prior* probability density of the variable which is the conditioning one (denoted by $[\cdot]$), and the conditional probability density yielding

$$[\mathbf{X}, \mathbf{Y}] = [\mathbf{X}|\mathbf{Y}] \cdot [\mathbf{Y}] = [\mathbf{Y}|\mathbf{X}] \cdot [\mathbf{X}] = [\mathbf{Y}, \mathbf{X}]\,. \tag{3.4}$$

The prior information can be seen as some knowledge that experimenters gain from already existing data before conducting the experiment (prior latin “before”). For example, this could be the expectation value or some frequency distribution of these data. Besides, a priori statements about the data, such as believes about what the data of interest should or could be like can be interpreted as prior probabilities (Ellison [1996]). Thus, the prior information needs to be given in advance and different priors may lead to different posteriors (Ohlwein and Wahl [2012]). This is obvious, if for example an uniform distribution is chosen as the prior distribution. In this case, all possible values are equally likely. In contrast, assuming the expectation value of the data in form of a normal distribution as the prior will shift the posterior into that direction.
By rearranging equation 3.4, the conditional probability of $\mathbf{X}$ given $\mathbf{Y}$ can be evaluated as

$$[\mathbf{X}|\mathbf{Y}] = \frac{[\mathbf{Y}, \mathbf{X}]}{[\mathbf{Y}]} = \frac{[\mathbf{Y}|\mathbf{X}] \cdot [\mathbf{X}]}{[\mathbf{Y}]}\,. \tag{3.5}$$

Vice versa it is possible to define the conditional probability of $\mathbf{Y}$ given $\mathbf{X}$ but is omitted here. Formulation 3.5 is also known as the *Bayes rule* and can be seen as a general form of a transfer function for the paleoclimate reconstruction problem

(Kumke et al. [2004]).
The conditional probability in the Bayes rule is also known as the *posterior* probability. As the name indicates is the posterior probability the result of the statistical calibration process since it measures the conditional probability of the two variables $\mathbf{X}$ and $\mathbf{Y}$ (post latin for "after" or "behind"). Hence, the right hand side of eq. 3.5 needs to be evaluated.
It is important to note that all these probabilities, whether they are prior, joint, or conditional, can be represented via probability density functions as introduced in eq. 3.2 and are functions of the parameters mean and variance by assuming a normal distribution.
In this study, two different approaches for paleoclimate transfer functions are examined. The first approach is to derive the relationship between the environmental variable annual average temperature in two meters height and the proxy variable annual tree ring width using linear regression (Model A, see sec. 3.2), which is the typical tool in paleoclimate calibration (see chapter 2). The second approach is related to the introduced Bayesian probability perspective and estimates the complete conditional probability of $\mathbf{X}$ given $\mathbf{Y}$, starting from equation 3.5 (Model B, see sec. 3.3).

3.2 Model A: Multiple Linear Regression

In earth sciences regression techniques are widely used methods for obtaining a quantitative relationship between variables. With regression, a variable that one wish to determine is estimated (predicted) via another variable or variables. These explaining variables are called *predictors*, the explained variable is the *response*. Analogue to the introduced notation lets consider $\mathbf{R}$ and $\mathbf{T}$ as two multivariate variables that represent the events 'annual tree ring width' and 'annual average temperature in two meters height' respectively. In a general formulation the regression model would be

$$\mathbf{T} = f(\mathbf{R}) + \boldsymbol{\epsilon} \quad \propto \quad \boldsymbol{\beta}\mathbf{R} + \boldsymbol{\epsilon} \quad = \quad \beta_0 + \sum_{i=1}^{p} \beta_i \mathbf{R}_i + \boldsymbol{\epsilon} \tag{3.6}$$

with a function f that links the two variables $\mathbf{R}$ (predictor) and $\mathbf{T}$ (response) and the regression coefficient matrix $\boldsymbol{\beta}$. The term 'linear' refers to the coefficients and how they enter the regression model (von Storch and Zwiers [1999]). That is, $\mathbf{T} = \beta_0 + \beta_1 \exp(\mathbf{R}_1)$ is a linear regression model but $\mathbf{T} = \beta_0 + \mathbf{R}_1^{\beta_1}$ is not. β_0 is called the intercept term and results from a constant $\mathbf{1}$ as predictor. $\boldsymbol{\epsilon}$ represents an additive noise to the model representing its variance.

Since we are dealing with physical quantities related to the climate system, the process f is unknown and very high–dimensioned (n). Thus one may reduce f to a finite set of p coefficients $\boldsymbol{\beta}$ so that $p \ll n$. This is possible under the assumption, that the additive noise is minimized so that $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon} = (\mathbf{T} - \boldsymbol{\beta}\mathbf{R})^T(\mathbf{T} - \boldsymbol{\beta}\mathbf{R}) \stackrel{!}{=} \min$. Following differentiation with respect to the unknown $\boldsymbol{\beta}$ gives an estimation for the regression coefficients. These derived regression coefficients are the statistical transfer model that measure the relationship between the two variables $\mathbf{R}$ and $\mathbf{T}$.
This method of minimizing the differences between predictors and response is called the method of *least squares* and has several important implications. First, reconstructed temperatures represented through linear combinations of the predictors and the derived regression coefficients have always an underestimatied variance, since the coefficients are derived under exactly that assumption (Robertson et al. [1999]). Second, this model can only be applied, if the variable to be modelled (in this case annual average temperature in two meters height) is drawn from a normal distribution. Third, the noise of the model is also assumed to be normally distributed, independent and of zero mean and constant variance ($\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon)$, Kumke et al. [2004]).
In other words, a regression predicts the mean of a normal distributed variable because the distances between the individual datapoints and the mean are minimized. This leads to an equality of the means of the predictors and the response variable over the calibration period, yielding in $\mathbf{T} \sim \mathcal{N}(\boldsymbol{\beta}\mathbf{R}, \boldsymbol{\epsilon})$.
Note that thinking about regression modeling in the sense of modeling causality between the response and the regressors would be inappropriate as it is always in statistical modeling (von Storch and Zwiers [1999]). Besides, it is possible to 'overfit' the regression model by using a large number of predictors. By that, the resulting transfer function fits very well for the calibration period but not necessarily for the prediction period (Wahl and Ammann [2007]).

3.3 Model B: Conditional Density of the Multivariate Normal Distribution

It has been stated that in order to define a probability density function of a continuous random variable drawn from the normal distribution, its first two moments are

needed. Now let $\mathbf{X} = (\mathbf{R}, \mathbf{T})^T$. According to eq. 3.3, the estimators of the moments of $\mathbf{X}$ may be defined as

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_R \\ \boldsymbol{\mu}_T \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} \mathbf{R}_i \\ \frac{1}{n}\sum_{i=1}^{n} \mathbf{T}_i \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{R}_i - \mu_{R_i})(\mathbf{T}_i - \mu_{T_i})^T . \tag{3.7}$$

Than the estimated covariance matrix $\boldsymbol{\Sigma}$ can be seen as a block matrix seperated in four submatrices, yielding

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{RR} & \boldsymbol{\Sigma}_{RT} \\ \boldsymbol{\Sigma}_{TR} & \boldsymbol{\Sigma}_{TT} \end{pmatrix}. \tag{3.8}$$

This block matrix represents a $(p+q)\times(p+q)$ matrix partitioned in four submatrices ($\boldsymbol{\Sigma}_{RR}$, $\boldsymbol{\Sigma}_{RT}$, $\boldsymbol{\Sigma}_{TR}$, $\boldsymbol{\Sigma}_{TT}$) with the dimensions $p \times p$, $p \times q$, $q \times p$, and $q \times q$. That is, $\boldsymbol{\Sigma}_{RR}$ and $\boldsymbol{\Sigma}_{TT}$ are of quadratic form and symmetric and represent the variances of two components of $\mathbf{R}$ and $\mathbf{T}$.
In contrast, $\boldsymbol{\Sigma}_{RT}$ and $\boldsymbol{\Sigma}_{TR}$ characterize the covariances of $\mathbf{R}$ and $\mathbf{T}$, that is how they depend on each other. Here, $\boldsymbol{\Sigma}_{RT}$ and $\boldsymbol{\Sigma}_{TR}$ are of quadratic form as well ($p = q$) but they are not covariance matrices on their own since they are not symmetric. However, because the full covariance matrix $\boldsymbol{\Sigma}$ is symmetric, $\boldsymbol{\Sigma}_{TR} = \boldsymbol{\Sigma}_{RT}{}^T$ (Wilks [2006]). Note that if $\mathbf{R}$ and $\mathbf{T}$ are independently distributed, $\boldsymbol{\Sigma}_{RT}$ and $\boldsymbol{\Sigma}_{TR}$ equal 0 (Morrison [1976]). In fact, the covariabilities represented by $\boldsymbol{\Sigma}_{RT}$ shall be used to derive the transfer function between temperature and tree ring width.
Therefore, recall the formulation of the conditional probability density function of the joint event of a continuous random variable expressed by the Bayes rule (eq. 3.5), $[\mathbf{T}|\mathbf{R}] = \frac{[\mathbf{R},\mathbf{T}]}{[\mathbf{R}]}$, and the variance of a multivariate random variable denoted by its Mahalanobis distance, $\mathcal{D}^2(\mathbf{X}) = (\mathbf{R} - \boldsymbol{\mu}_R)^T \boldsymbol{\Sigma}^{-1} (\mathbf{T} - \boldsymbol{\mu}_T)$. Since we assume that $\mathbf{X} = (\mathbf{R}, \mathbf{T})^T$ is a joint event from the multivariate normal distribution, its joint conditional probability density may be defined via a pdf according to the one introduced in eq. 3.2 over the joint sample space of $\mathbf{R}$ and $\mathbf{T}$: $\boldsymbol{\Omega}^{2n} = \boldsymbol{\Omega}_R^n \cup \boldsymbol{\Omega}_T^n$ with $\cup$ representing the union of both.
In doing so, the Mahalanobis distances of $[\mathbf{T}, \mathbf{R}]$ and $[\mathbf{R}]$ may be calculated via the inverse covariances $\boldsymbol{\Sigma}^{-1}$ (from eq. 3.8) and $\boldsymbol{\Sigma}_{RR}{}^{-1}$ respectively.
The question arises, whether these inverse covariance matrices exist or not. For that, let M be a quadratic block matrix defined as

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{3.9}$$

Furthermore, let the submatrix A be non–singular, i.e. that its inverse exist. Then

$$S = D - CA^{-1}B \tag{3.10}$$

is called the *Schur complement* of A in M. This identity was published by Issai Schur in 1917 (Horn and Zhang [2005]). There are three important followings of eq. 3.9 and 3.10:

$$\begin{aligned} &1)\, |S| \neq 0\,, \\ &2)\, M^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{pmatrix}, \\ &3)\, |M| = |A| \cdot |S|\,. \end{aligned} \tag{3.11}$$

The procedure to derive the inverse M^{-1} is based on Gaussian elimination and may be found in Horn and Zhang [2005] on page 17 or Frazer et al. [2007] on page 112. Note that the components A and D are interchangeable. So starting from a non–singular submatrix D, the Schur complement of D in M can be defined as $S = A - BD^{-1}C$.
Obviously, the inverse of the block matrix, M^{-1}, can be constructed by its components. For that, only the inverse matrices of A and S are necessary. Both exist, if the inverse of the component A is defined.
Now consider the estimated covariance matrix $\mathbf{\Sigma}$ from formulation 3.8 as M and its first component $\mathbf{\Sigma}_{RR}$ as A respectively. Consequently, $\mathbf{\Sigma}^{-1}$ exists if ${\mathbf{\Sigma}_{RR}}^{-1}$ is non–singular. Furthermore, the Schur complement of $\mathbf{\Sigma}_{RR}$ in $\mathbf{\Sigma}$ is defined as

$$\mathbf{S}_R = \mathbf{\Sigma}_{TT} - \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}\,, \tag{3.12}$$

and its inverse exists. However, the covariance matrix $\mathbf{\Sigma}$ sometimes is non–singular. In this case, the *lasso regression* is used in this study to estimate its inverse (see sec. 3.4.3.2 for further information).
Now let

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{RR} & \mathbf{C}_{RT} \\ \mathbf{C}_{TR} & \mathbf{C}_{TT} \end{pmatrix} = \mathbf{\Sigma}^{-1} = \begin{pmatrix} \mathbf{\Sigma}_{RR} & \mathbf{\Sigma}_{RT} \\ \mathbf{\Sigma}_{TR} & \mathbf{\Sigma}_{TT} \end{pmatrix}^{-1}. \tag{3.13}$$

According to the second formulation in eq. 3.11, the components of $\mathbf{C}$ can be defined as

$$\begin{aligned} \mathbf{C}_{RR} &= {\mathbf{\Sigma}_{RR}}^{-1} + {\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}{\mathbf{S}_R}^{-1}\mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1} \\ \mathbf{C}_{RT} &= -{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}{\mathbf{S}_R}^{-1} \\ \mathbf{C}_{TR} &= -{\mathbf{S}_R}^{-1}\mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1} \\ \mathbf{C}_{TT} &= {\mathbf{S}_R}^{-1}. \end{aligned} \tag{3.14}$$

Now, the definition of the inverse covariance matrix (eq. 3.14) can be used to derive the conditional probability density of $\mathbf{X} = (\mathbf{R}, \mathbf{T})^T$, i.e. the pdf of the temperature given fixed values of tree ring width.
From the definitions of the probability density function of a continuous multivariate normal distributed random variable (eq. 3.2) and the Bayes rule (eq. 3.5) it follows that

$$[\mathbf{T}|\mathbf{R}] = \frac{[\mathbf{R},\mathbf{T}]}{[\mathbf{R}]} = \frac{(2\pi)^{p/2}|\mathbf{\Sigma}_{RR}|^{1/2}}{(2\pi)^{(p+q)/2}|\mathbf{\Sigma}|^{1/2}} \frac{\exp\left(-\frac{1}{2}(\mathbf{X}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})\right)}{\exp\left(-\frac{1}{2}(\mathbf{R}-\boldsymbol{\mu}_R)^T{\mathbf{\Sigma}_{RR}}^{-1}(\mathbf{R}-\boldsymbol{\mu}_R)\right)}. \quad (3.15)$$

By using the definition of the determinante of the block matrix $\mathbf{\Sigma}$ (third formulation in eq. 3.11), the enumerator and denumerator of the first term in eq. 3.15 reduce to

$$\frac{1}{(2\pi)^{q/2}|\mathbf{\Sigma}_{TT} - \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}|^{1/2}} = \frac{1}{(2\pi)^{q/2}|\mathbf{S}_R|^{1/2}}\,.$$

For the second term, the Mahalanobis distance in the denumerator from the pdf $[\mathbf{R}]$ and the first of the two terms of the first component of $\mathbf{\Sigma}^{-1}$, $\mathbf{C}_{RR}$, (eq. 3.14) can be cancelled out by expanding the exponential terms.
Hence, the exponential term of 3.15 reduces to the Mahalanobis distance of the conditional density function $[\mathbf{T}|\mathbf{R}]$ and can be written as

$$\begin{aligned}\mathcal{D}^2(\mathbf{T}|\mathbf{R}) \quad = \quad & (\mathbf{R}-\boldsymbol{\mu}_R)^T{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}{\mathbf{S}_R}^{-1}\mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}(\mathbf{R}-\boldsymbol{\mu}_R) \\ & - (\mathbf{R}-\boldsymbol{\mu}_R)^T{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}{\mathbf{S}_R}^{-1} \\ & - (\mathbf{T}-\boldsymbol{\mu}_T)^T{\mathbf{S}_R}^{-1}\mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}(\mathbf{R}-\boldsymbol{\mu}_R) \\ & + (\mathbf{T}-\boldsymbol{\mu}_T)^T{\mathbf{S}_R}^{-1}(\mathbf{T}-\boldsymbol{\mu}_T)\,.\end{aligned} \quad (3.16)$$

The final expression for the conditional pdf yields by recalling that $\mathbf{\Sigma}_{TR} = {\mathbf{\Sigma}_{RT}}^T$ and ${\mathbf{\Sigma}_{RR}}^{-1^T} = {\mathbf{\Sigma}_{RR}}^{-1}$):

$$\begin{aligned}[\mathbf{T}|\mathbf{R}] = & \frac{1}{(2\pi)^{q/2}|\mathbf{\Sigma}_{TT} - \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}|^{1/2}} \\ & \exp\Big\{ -\frac{1}{2}[\mathbf{T}-\boldsymbol{\mu}_T - \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}(\mathbf{R}-\boldsymbol{\mu}_R)]^T(\mathbf{\Sigma}_{TT} - \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT})^{-1} \\ & [\mathbf{T}-\boldsymbol{\mu}_T - \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}(\mathbf{R}-\boldsymbol{\mu}_R)]\Big\}.\end{aligned} \quad (3.17)$$

Finally the two parameters of the conditional multivariate probability density function can be derived:

$$\boldsymbol{\mu}_{T|R} = \boldsymbol{\mu}_T + \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}(\mathbf{R}-\boldsymbol{\mu}_R)\,, \qquad \mathbf{\Sigma}_{T|R} = \mathbf{\Sigma}_{TT} - \mathbf{\Sigma}_{TR}{\mathbf{\Sigma}_{RR}}^{-1}\mathbf{\Sigma}_{RT}\,, \quad (3.18)$$

representing conditional mean and conditional covariance.
It becomes clear that the conditional covariance matrix equals the Schur complement of $\mathbf{\Sigma}_{RR}$ in $\mathbf{\Sigma}$. Furthermore, if temperature and tree ring width are independent from each other, the conditional parameters reduce to the normal mean and covariance because $\mathbf{\Sigma}_{TR}$ would equal 0 and the second terms on the right hand sides in eq. 3.18 would vanish.
As mentioned above, because of the symmetry of $\mathbf{\Sigma}$ one could start the other way around with the fomulation of the Schur complement $\mathbf{S}_T$ instead of $\mathbf{S}_R$ which would yield a formulation for the parameters of the conditional multivariate normal probability density function for tree ring width given fixed temperature values $[\mathbf{R}|\mathbf{T}]$.

3.4 Data Preparation and Construction

3.4.1 Proxy Data

In this study, European beeches (*Fagus sylvatica* L.) from three sites located in mid–west Germany were sampled, kindly provided by Dr. Burkhard Neuwirth. The earliest available tree ring was formed in 1837, the latest in 2011. As it is common in dendroclimatological studies, ring width values from a tree are obtained by several *cores* per tree. In this study mostly two cores are sampled per tree and the tree mean is obtained by averaging over the tree cores. The *location* or *site mean* is represented through the mean of trees from the location. Since the chronologies from different trees are of different length, missing data occur often in the early years of the chronologies. When site means are derived, missing values are stripped before the computation process. Especially for Model B, missing data in the estimated coavariance matrix are a big problem since they may affect its singularity. Because of that, the available time for the calibration process reduces to that timepoint, when tree ring values from all trees from a site are available. Table 3.1 provides an overview of the proxy data and their location is shown in figure 3.1. Note the small correlation coefficient between annual average temperature and tree ring width of location L1 and that all correlations are negative.
As described in sec. 2.2, tree ring width values need to be standardized before any statistical analyses. This is done via i) calculation of the tree mean by averaging tree ring width over the numbers of cores per tree and ii) standardizing every tree mean curve of a site. In this study, exponential functions as well as spline functions are used in order to remove age trends and possible non–climatic stand dynamics on

Table 3.1: Tree ring width proxy data used in this study. All proxy data are kindly provided by Dr. B. Neuwirth (2014).

ID	Trees	Lon, Lat	Site	Species	Calibration Years ($\boldsymbol{t}$)	n	Temperature Correlation**	ID of Reference
L1	15	8.98, 51.13	Kellerwald	FASY*	1901–2006	106	-0.01	dhk13
L2	15	6.49, 50.61	Eifel	FASY	1901–2009	109	-0.44	dre26
L3	11	8.98, 52.24	Wesergebirge	FASY	1907–2009	103	-0.32	drw03

* *Fagus sylvatica* L.
** Pearson correlation coefficient of annual average temperature and tree ring width series at the sample location of the calibration time $\boldsymbol{t}$.

Table 3.2: Standardization methods used in this study

No	Standardization	Construction	Implementation
1	Negative Exponential	$f(t) = a\exp(-bt) + k$	Non–linear Least Squares
2	Hugershoff Curve	$f(t) = at^b\exp(-ct) + k$	Non–linear Least Squares
3	Basis Spline	Cubic, 5 knots	Linear Least Squares
4	Basis Spline	Cubic, 15 knots	Linear Least Squares

interdecadal to subcentury length scales. Table 3.2 summarizes the four standardization methods used in this study.

3.4.2 Environmental Data

The temperature data is provided through a global monthly gridded dataset of the University of East Anglia Climatic Research Unit at a spatial resolution of $0.5° \times 0.5°$ from January 1901 to December 2009 (version CRU TS3.10, hereafter called CRU). Note that there are no values for sea surface temperature, hence the CRU temperature data is defined above land only. Further information may found in the documentation (Jones and Harris [2013]).
Annual temperatures are calculated through twelve months arithmetic mean (January to December). In contrast to tree ring data, much more places from which temperature values can be taken from are available. From these, only places within a certain correlation distance from the tree locations L1, L2, and L3 were chosen. This is done through calculation of the Pearson correlation coefficient ρ between the temperature time series of the tree locations and all other locations of the total available time. If the correlation exceeds a threshold ρ_{thr}, this location is chosen or rejected otherwise. From the chosen places within the correlation distance, 30 lo-

cations for temperature values were chosen randomly. By that, possible large–scale climatologic effects on tree ring width are likely included in the environmental data and not only local climate signals are considered. Here, ρ_{thr} is set to 0.8. Figure 3.1 shows the correlation distance and randomly chosen points from where the temperature vales are taken from exemplarily. Additionally, the three locations from where tree ring width values were sampled are shown.

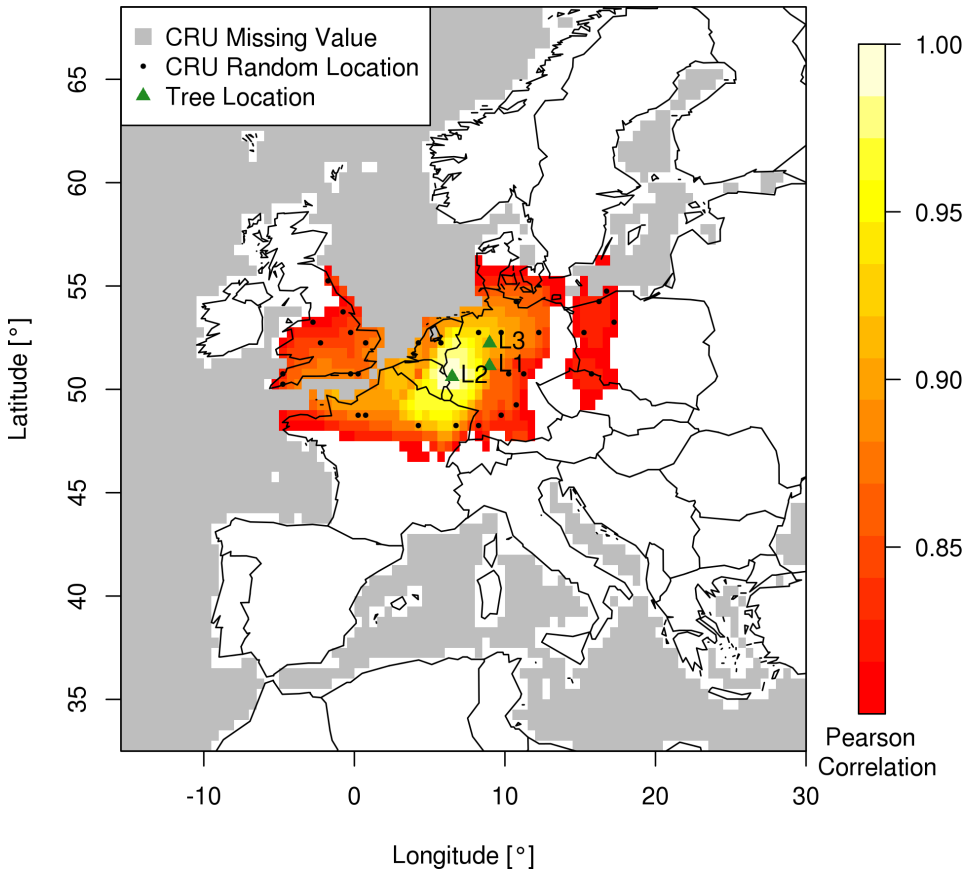

Figure 3.1: Correlation distance of CRU temperature values from 1901 – 2009 of all places and tree ring width sample location B (correlation threshold $\rho_{thr} = 0.8$). 30 randomly chosen places within the correlation distance were chosen for each tree ring width location and standardization method. Tree ring width data kindly provided by Dr. Burkhard Neuwirth, temperature data provided by University of East Anglia Climatic Research Unit (Jones and Harris [2013]).

3.4.3 Implementation

For the calibration process, two sets of data are needed – a training and a verification set – and subsamples from the available data may only be included in one of these

two. This is accomplished via the method of *cross validation*. Therefore, the timeline from year 1 to year n is divided in smaller time–windows of length d of approximately a decade (here $d = 11$). Additionally, a lag is used to shift the time–window of about 1 year. In summary, via the cross validation the relationship between tree ring width and temperature is examined decade–wise for every decade from year 1 to year n yielding the following notation:

$$\begin{aligned} t_i &= \text{year } i \\ \boldsymbol{t} &= (t_1, \ldots, t_i, \ldots, t_n) \\ \boldsymbol{t}_j &= (t_{i-l}, \ldots, t_i, \ldots, t_{i+l}) \,(\text{training data}) \\ \tilde{\boldsymbol{t}}_j &= \boldsymbol{t} \setminus \boldsymbol{t}_j \,(\text{verification data}) \end{aligned} \quad \text{with} \quad \begin{aligned} & i \in [1, n] \\ & d = 11 \,\text{years}\,, l = \frac{d}{2} \\ & j \in [1, n - d + 1]\,, \end{aligned}$$

where "\" denotes without. So in order to model the temperature of year t_i, the training decade $\boldsymbol{t}_j$ is excluded from the whole timeline of available data $\boldsymbol{t}$, resulting in the verification dataset $\tilde{\boldsymbol{t}}_j$. The length of $\boldsymbol{t}_j$ equals the length of the time–window d that is shifted, in this case 11 years. Hence, $l = 5$ years are omitted on both sides from t_i (rounded). Consequently, decade–wise cross validation leads to vectors $\mathbf{T}_j$ of length d. Via $\mathrm{T}_i = \mathbf{P}\mathbf{T}_j$, the reconstructed vector $\mathbf{T}_j$ is transformed so that $\mathbf{P}$ either picks its central located value or calculate its mean. In the first case, $\mathbf{P}$ is a vector of zeros of length d with a one on its central position. In the latter case, $\mathbf{P}$ is a vector of length d with $\frac{1}{d}$ at every position.
Additionally, both models run in a deterministic as well as in a stochastic version. By that, the jth model residuals $\boldsymbol{\epsilon}_j$ of the linear regression model A and the derived variabilities in form of the conditional covariance matrix $\boldsymbol{\Sigma}_{T|R_j}$ of the conditional transfer model B are neglected in the deterministic case and considered in the stochastic version. Figure 3.2 summarizes the described model implementation.

	Deterministic	Stochastic
A: Linear Regression	$\mathbf{T}_j = \boldsymbol{\beta}_j \mathbf{R}_j$	$\mathbf{T}_j = \boldsymbol{\beta}_j \mathbf{R}_j + \mathcal{N}(\mathbf{0}, \boldsymbol{\epsilon}_j)$
	$\mathrm{T}_i = \mathbf{P}\mathbf{T}_j$	
B: Conditional Probability	$\mathbf{T}_j = \boldsymbol{\mu}_{T\|R_j}$	$\mathbf{T}_j = \boldsymbol{\mu}_{T\|R_j} + \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{T\|R_j})$

Figure 3.2: Schematic summary of the calibration process of the two models A and B.

3.4.3.1 Cross Validation for Conditional Transfer Model B

For the conditional probability formulation of the Bayesian transfer model, covariances are given in matrix form (see eq. 3.7 and 3.8). Therefore, let $\mathbf{R} = (R_{i,j}) = R_{i+j-1}$ is a 2–dimensional datamatrix with the indices for timepoints $i \in [1, d]$ and $j \in [1, n-d+1]$ (analogue for $\mathbf{T}$). Then, $\mathbf{X} = (\mathbf{R}, \mathbf{T})^T$ is a $2d \times (n-d+1)$ dimensioned datamatrix. Figure 3.3 visualizes this procedure for $\mathbf{R}$.

$$\mathbf{R}_j = \begin{pmatrix} \overline{R}_1 & \dots & \overline{R}_{i-d+1} & \dots & \overline{R}_{i-l} & \dots & \overline{R}_i & \dots & \overline{R}_{n-d+1} \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & \dots & \vdots & \dots & \boxed{\overline{R}_i} & \dots & \vdots & \dots & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \overline{R}_d & \dots & \overline{R}_i & \dots & \overline{R}_{i+l} & \dots & \overline{R}_{i+d-1} & \dots & \overline{R}_n \\ \underbrace{\quad}_{\tilde{\boldsymbol{t}}_j} & & & & \underbrace{\quad}_{\boldsymbol{t}_j} & & & & \underbrace{\quad}_{\tilde{\boldsymbol{t}}_j} \end{pmatrix}$$

Figure 3.3: Exemplary cross validation procedure for the datamatrix $\mathbf{R}$ in the conditional transfer model B at the jth iteration. The red column indicates the removed training data $\boldsymbol{t}_j$, the blue field represents lag years and is also removed during the calibration. The remaining data stands for the verification data $\tilde{\boldsymbol{t}}_j$. The box around year i in training decade marks the year that is reconstructed, bar over variable symbolizes average (site mean in case of tree ring width data, average of 30 random locations within correlation distance for temperature data).

Again, for every calibration step, the training decade $\boldsymbol{t}_j$ is removed (indicated by the red column in figure 3.3). Additionally, all columns are removed that include one of the years within the training decade $\boldsymbol{t}_j$ (indicated by the blue field in figure 3.3). The verification data $\tilde{\boldsymbol{t}}_j$ results in the union of the remaining parts left- and righthand of the removed data.
Furthermore, all entries in the 2–dimensional datamatrix are averaged values: the site mean in case of tree ring width data and the mean of 30 random locations within the correlation distance for temperature data.

3.4.3.2 Estimate the Inverse of a Covariance Matrix via Lasso Regression

For the conditional model B, the inverse of the covariance matrix $\boldsymbol{\Sigma}$ of $\mathbf{X}$ is needed. Its inverse can be calculated, if the determinant is different from zero, $|\boldsymbol{\Sigma}| \neq 0$.
As it is a common problem in statistical analyses, this is not always the case in this study why the *gLasso* method published by Friedman et al. [2008] is used to

obtain an estimaton of the inverse of a non–singular matrix. This method is based on the lasso regression, which is a variante of regression as described in sec. 3.2. The idea is to estimate a sparse matrix that maximizes the log–likelihood function of multivariate normal distributed samples. This matrix is already maximized partially with respect to the mean of these samples. In other words, the off–diagonal elements of a sparse matrix which depend on the estimated moments of the data are coerced to smaller values via an increasing penalty parameter ρ_L in order to obtain a singular matrix. If ρ_L equals one, the gLasso estimator is an diagonal matrix with zeros on its off–diagonal elements. If ρ_L equals zero, nothing happens and the glasso estimator equals the analytical estimator. Further information can be found in Friedman et al. [2008], Röpnack et al. [2013], and Zerenner [2011]. We use the R package "glasso: Graphical lasso- estimation of Gaussian graphical models" (Friedman et al. [2008]) which is available in the repositories.

The conditional parameters of the transfer model B are defined via the covariances $\mathbf{\Sigma}_{RT}$ and the inverse of the variances of the proxy data, ${\mathbf{\Sigma}_{RR}}^{-1}$ (eq. 3.18). Via the gLasso method, penalized estimators for them are calculated. By that, the penalty parameter ρ_L is chosen so that the differences between the norm of the analytical covariance estimator and the gLasso estimator ($\mathbf{\Sigma}_L$) are set to a certain percentage threshold γ_{thr}, so that $||\mathbf{\Sigma}_L||_F - \gamma_{thr} \cdot ||\mathbf{\Sigma}||_F \stackrel{!}{=} \min$. Here, $||\cdot||_F$ denotes the Frobenius norm which is defined as $||\mathbf{\Sigma}||_F = \sqrt{\sum_{i,j=1}^{n} |\Sigma_{ij}|^2}$ with n the dimension of the matrix. Prior to that, the diagonal elements of the analytical matrix are set to zero, since the gLasso penalty method aims at the off–diagonal elements. In this study γ_{thr} is set to 80 %, i.e. the gLasso estimator of the covariances represent 80 % of the covariances derived via the analytical estimator (eq. 3.7).

Figure 3.4 shows exemplarily the influence of the gLasso method on the estimated covariance matrix $\mathbf{\Sigma}_{RT}$ and the estimated inverse covariances ${\mathbf{\Sigma}_{RR}}^{-1}$ as derived in sec. 3.3 (note that these are submatrices of the full covariance matrix). The first row in figure 3.4 shows the covariances $\mathbf{\Sigma}_{RT}$ before (a) and after (b) applying gLasso method. Each row and column of the matrix represent one year of the training decade $\boldsymbol{t}_j$ indicated by the numbers of the diagonal elements. Since $d = 11$, each of the shown submatrices is dimensioned 11×11.

The lower off–diagonal elements represent the covariabilities between the tree ring width and temperature, with preceding temperature realizations. That is, the tree ring width of year i mostly depends on the temperatures of the years $i-1$ and $i-2$ (the two lower off–diagonals). The upper off-diagonal elements represent the opposite case: covariabilities of tree ring width values of year i and temperatures of the following years. By comparing with the gLasso estimator (b), it becomes clear that the pattern of the analytical estimator is preserved but smoothed.

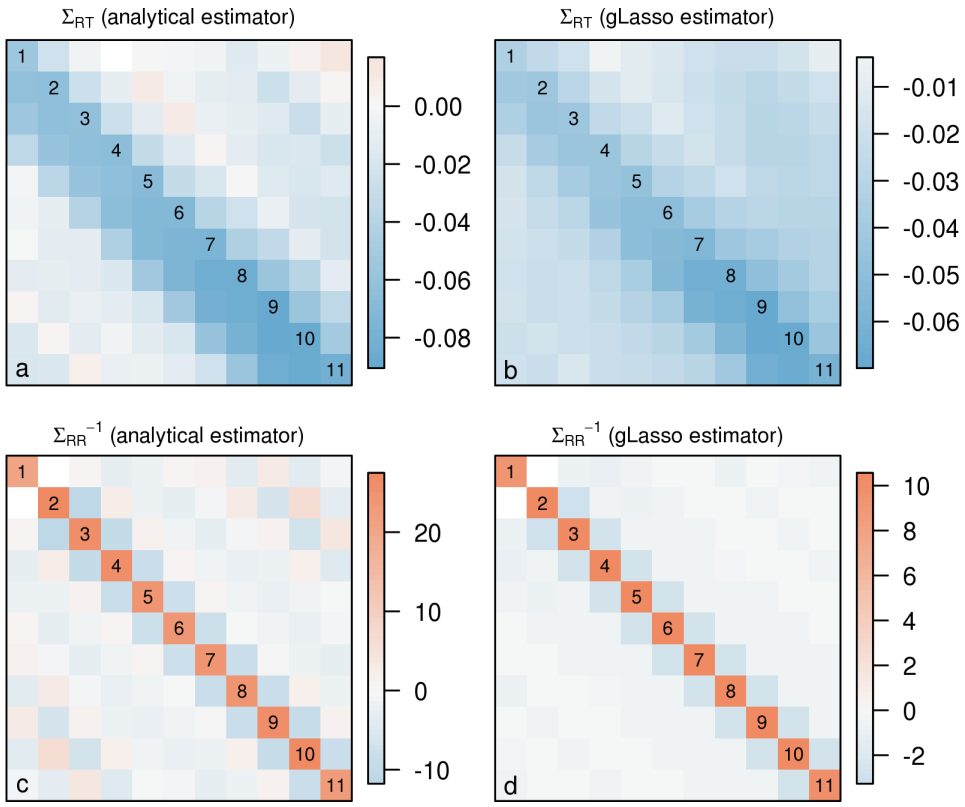

Figure 3.4: Visualization of the gLasso method to obtain an estimator for the covariances and its inverse exemplarily for the submatrices $\mathbf{\Sigma}_{RT}$ and $\mathbf{\Sigma}_{RR}{}^{-1}$ before (a and c) and after applying gLasso (b and d).

The second row of figure 3.4 shows the effect of the gLasso method applied on the analytical estimator of the inverse variances of the tree ring width data. As before, the basic pattern of the matrix is preserved. Furthermore, the sparsity is increased and the values from the off–diagonal elements in (c) are reduced to zero. For the conditional transfer model B, the gLasso estimators are used.
Further analyses of the covariances of the environmental and proxy data is given in chapter 5.

3.5 Verification

In order to estimate the ability or skill of a model to reconstruct past temperatures, both models are verified against a reference which is represented by the average climate as a primitive third model, and against each other. The average climate of the verification time $\tilde{\boldsymbol{t}}_j$ is used as reference, calculated through the mean of the randomly chosen locations from where the temperature values are taken from.

In statistical modeling, *scores* or *skill scores* are common tools to assess the quality of the model outcome. A score must take into account the character of the model, whether it is for example deterministic or stochastic. In this study, the *root mean square error* (RMSE) and the *mean absolute error* (MAE) are used as measures for the predictive skill of the deterministic models and the *continuous ranked probability score* (CRPS) for the stochastic. These scores have in common, that they measure the differences between the model outcome and an acutal observation or reference. The RMSE measures their quadratic difference and the MAE their absolute difference. They are defined as

$$\mathrm{RMSE}_j = \sqrt{\frac{1}{n-d-2}\sum_{i=1}^{d}(\mathbf{O}_{ji} - \mathbf{T}_{ji})^2} \tag{3.19}$$

and

$$\mathrm{MAE}_j = \frac{1}{n-d-2}\sum_{i=1}^{d}\left|\mathbf{O}_{ji} - \mathbf{T}_{ji}\right|, \tag{3.20}$$

with $\mathbf{O}_j$ being the vector of observations or the reference of the same time as the reconstruction ($\boldsymbol{t}_j$). The RMSE is a quadratic score. Because the the errors between observation and model are squared before averaging, large errors are emphasized. In contrast, the MAE is a linear score and all errors are weighted equally.
The CRPS applies to probabilistic forecasts that take the form of predictive cumulative distribution functions. It can be seen as a generalized formulation of the absolute error scores given above and for a deterministic system the CRPS reduces to the MAE (Hersbach [2000], Gneiting and Raftery [2007]). A simplified form of the CRPS for scalar quantities may be defined as

$$\mathrm{CRPS}_j = \sigma_j \left[z_j \cdot \Big(2\,F_Z(z_j) - 1\Big) + 2\,f_Z(z_j) - \frac{1}{\sqrt{\pi}} \right], \tag{3.21}$$

where $z_j = \frac{(x_j - \mu_j)}{\sigma_j}$, and $f_Z(z_j)$ and $F_Z(z_j)$ are the probability density and cumulative distribution functions of z respectively. The parameters for the probability density and cumulative distribution functions of the normal distributed variable z_j are provided via the model output $\mathbf{T}_j$ and its mean (μ_j) and variance (σ_j) respectively. The data x stands for the observation of the calibration time $\boldsymbol{t}_j$.
This version of the CRPS score is proposed by Gneiting and Raftery [2007] (their eq. 20) but with positive instead of negative sign so that all scores are positive oriented in this study (larger or equal to zero and the smaller the better).
From the definitions of the scores it can be seen that these are vectors since they are calculated for every verification time point j. In order to compare the scores of

different models in this study, a skill score is used to decide which reconstruction is better. For that, ratios between scores of two different models are derived via

$$\text{Skill score} = 1 - \left(\frac{\frac{1}{n-d-1} \sum_{j}^{n-d-1} \text{Score}_{A,j}}{\frac{1}{n-d-1} \sum_{j}^{n-d-1} \text{Score}_{B,j}} \right) . \tag{3.22}$$

Hence, the means of the scores of the models A and B are compared within a single number. The skill score is near zero, when both models have more or less the same skill. If the skill score is larger than zero, model A is likely to have a better (smaller) score than B. When the skill score has a negative sign, model A is not as good as B.

4 Results

In this chapter, the climate reconstructions derived from the two statistical transfer models are shown. Results from deterministic and stochastic as well as from decade–midyear and decade–mean reconstructions are distinguished.
First, the time series and scores of the reconstructed temperatures are shown exemplarily for one location and standardization method in order to get an idea about the general pattern of the model predictions. In sec. 4.1 and 4.2, the models are compared via skill scores against the climatic mean which serves as a reference, and with each other.

Temperature Reconstructions

Figure 4.1 shows temperature series of the deterministic (a) and stochastic (b) reconstructions exemplarily for tree location L2 and spline (5 knots) as standardization method. Note, that the timeline of the reconstruction is shorter than the one of the available proxy and environmental data due to the cross validation method. Because both models predict means in the deterministic formulation, only small variances are visible in the deterministic time series. Compared to the mean of the observations, both models are able to predict it as indicated by the gray, black and blue triangles at the left axis of the figure. Departures from annually observed temperatures are high for both models (of about ± 1 K). Furthermore, there are almost no differences between the model means, whether the reconstructed year i is the midyear (triangles on the left) of the reconstructed decade j or its mean (triangles on the right). In the stochastic version, both models show more fluctuations compared to (a). Again, the observation mean is modelled but differences to annual values are high (of about ± 1 K). For both models it is shown that the time series is smoother if the mean of the decade is chosen as the reconstructed temperature compared to the midyear (for deterministic and stochastic simulations).

Example RMSE Scores

For the same location and standardization method, the RMSE score is shown in figure 4.2. Firstly, the RMSE has small values of ca. 0.05 to 0.35 indicating a good prediction of the climate reference of both models. The reconstructions based on the mean of the decade j have generally a higher score (mean of ca. 0.2) than these reconstructed through the midyear of the decade (mean of ca. 0.05). Furthermore, the scores are of approximately the same size for both models A and B, so none of them seems to be better in predicting the climate reference if they are deterministic.

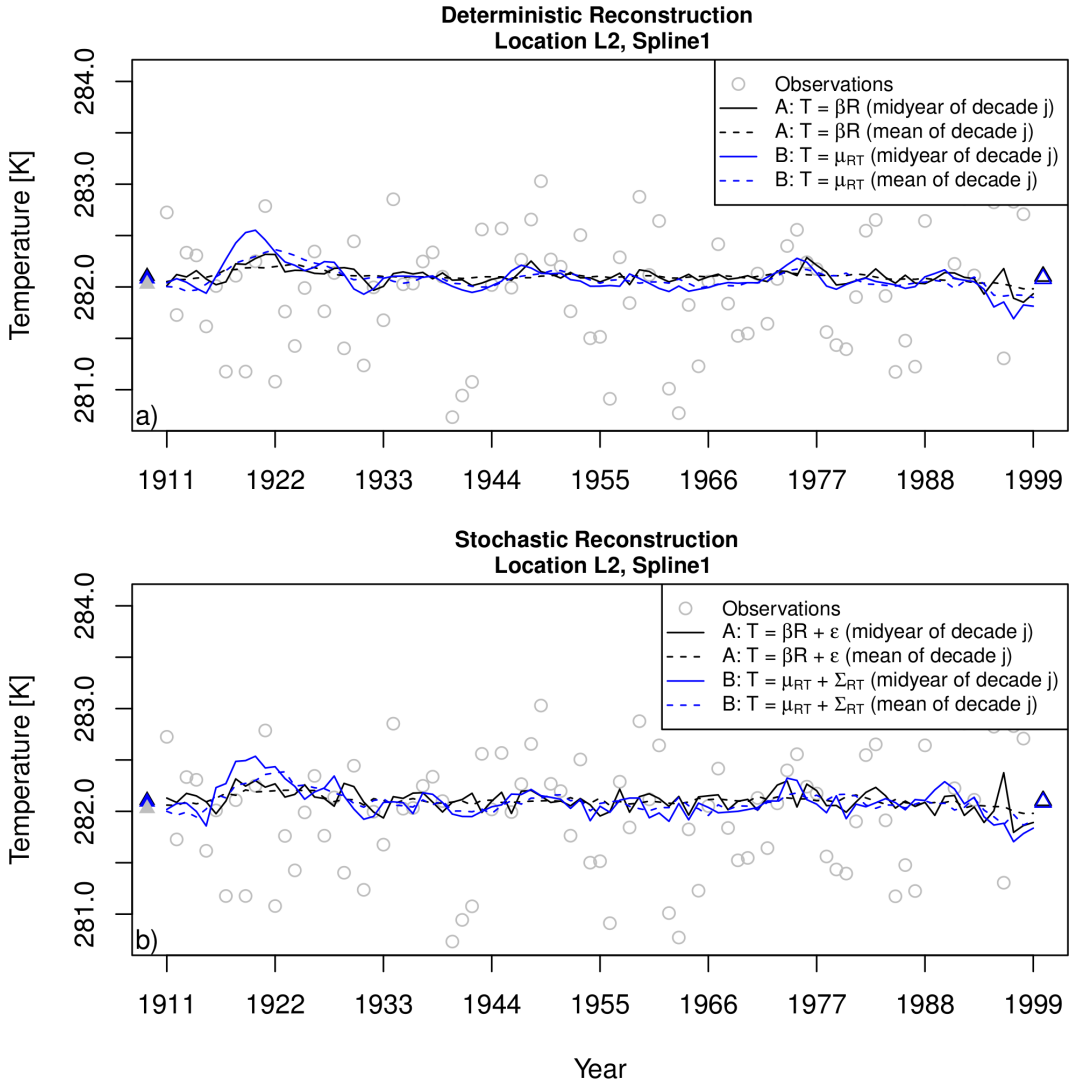

Figure 4.1: Reconstructions of annual average temperature in two meters height exemplarily for location L2 and standardization method spline with (5 knots) for (a) deterministic and (b) stochastic models. Triangles on the left axis indicate means of decade–midyear reconstructions, decade–average on the right.

As for the temperature time series, the RMSE fluctuates more, if the midyear is chosen for reconstruction.

The mean absolute error (MAE) shows similar results as the RMSE and is not shown here. However, the MAE values are much smaller than the RMSE (mean of ca. 0.01 for the midyear reconstructions and ca. 0.06 for decade average reconstructions). So, the MAE is very small indicating a good skill of predicting the climate reference for both models.

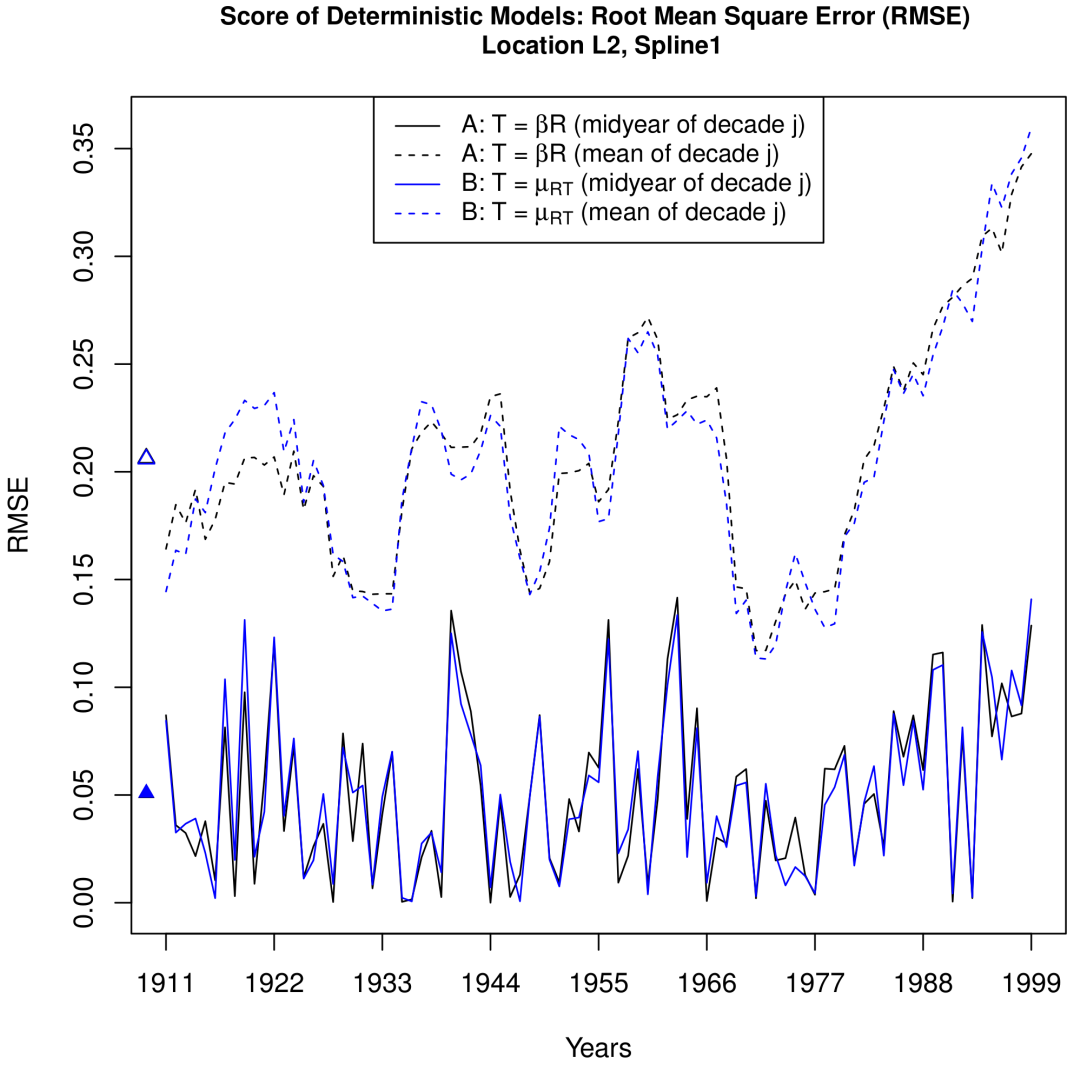

Figure 4.2: RMSE scores of deterministic models exemplarily for location L2 and standardization method spline (5 knots). The smaller the score, the better the prediction. Triangles on the left axis indicate means.

Example CRPS Scores

For the stochastic models, the CRPS score is used for verification. Figure 4.3 shows the CRPS for both models as well as for the reference. The reference represents a simple model for predicting past temperatures: Tthe mean and variance of the verification dataset of time t tilde. Here, we consider the case that the reconstructed decade is averaged to reconstruct the temperature of year i. It becomes clear that the means of the three series (model A, B, and the reference) are approximately the same, indicating a similar ability for modeling the past climate. Furthermore, the CRPS for the conditional model is larger in the beginning and at the end of the time

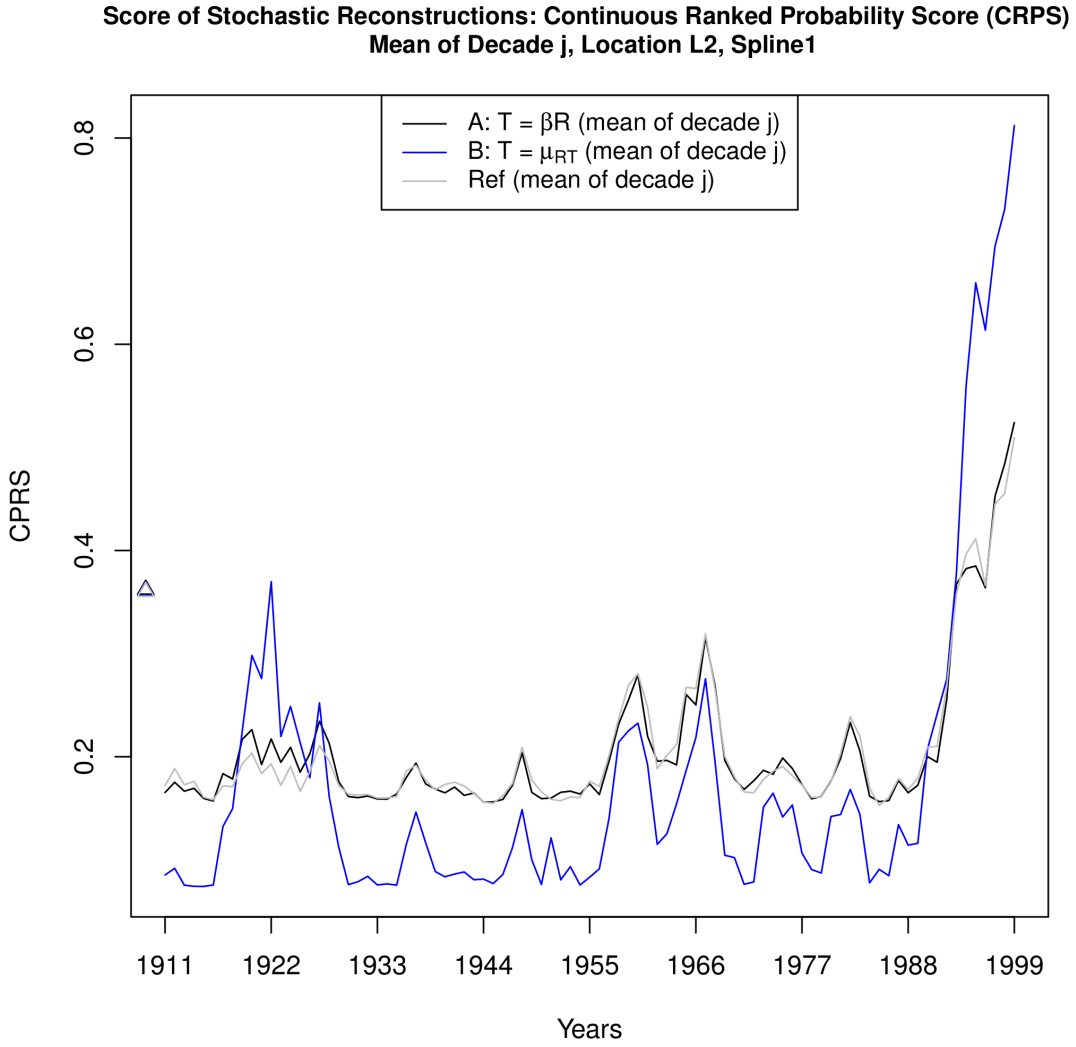

Figure 4.3: CRPS scores of the stochastic models exemplarily for location L2 and standardization method spline (5 knots). The smaller the score, the better the prediction. Triangles on the left axis indicate means.

series compared to that of model B and the reference. In contrast, it is lower for the time in between. The CRPS of model A and the reference do not differ that much from each other throughout the timeline. The decade midyear CRPS scores are not shown here because they depict a similar pattern but with steeper fluctuations.

4.1 Model Comparison

Now, data from all locations and different standardization methods are considererd via skill scores, which aim to compare different model predictions.

Skill Scores of Deterministic Models

Figure 4.4 shows the RMSE and MAE skill scores for the deterministic models. Since we are interested, whether the conditional transfer model B has a better ability of modeling past temperatures than the regression model A, the skill scores are calculated so that positive values indicate that model B is better than A and vice versa. Both deterministic skill scores show that model B predicts past temperatures

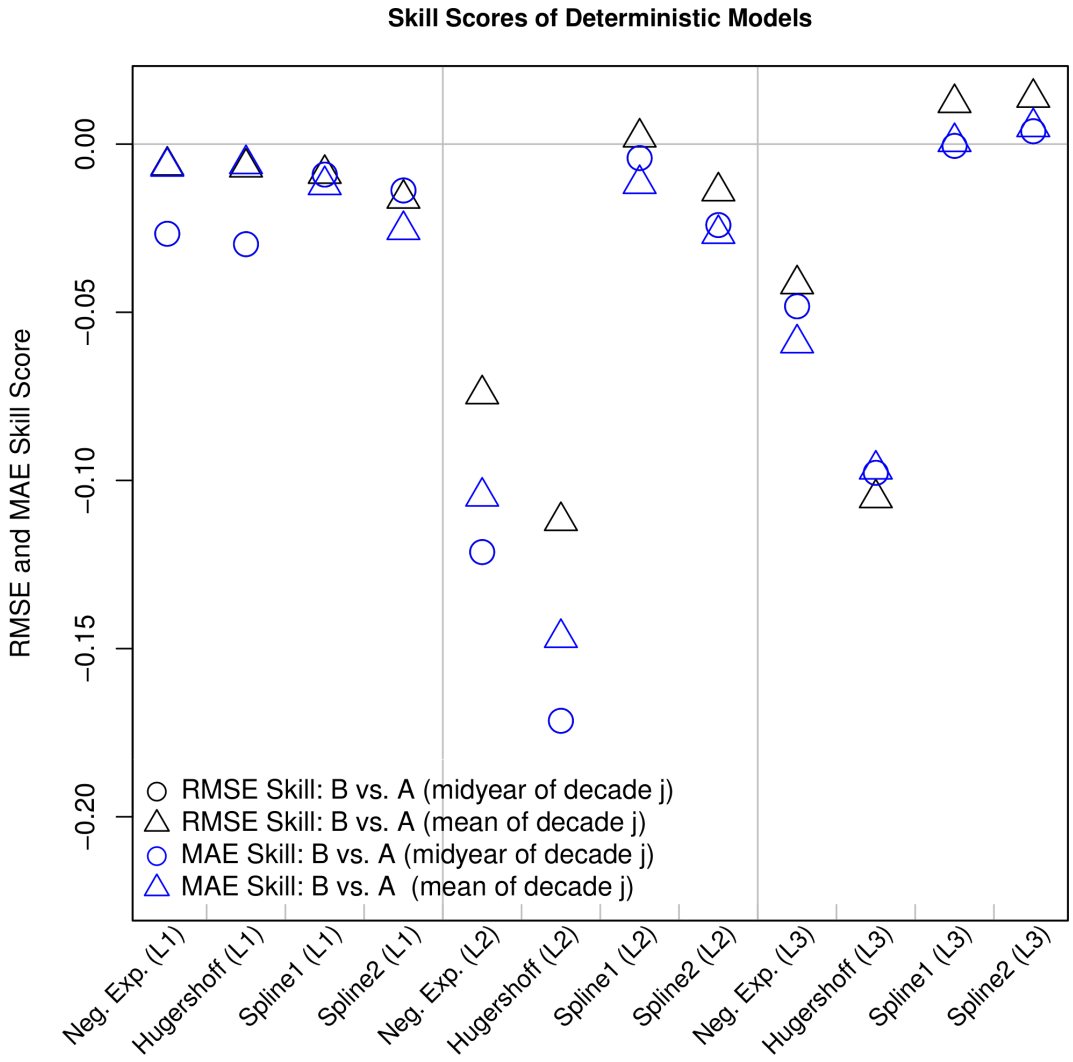

Figure 4.4: RMSE and MAE skill scores of the deterministic conditional transfer model B. Positive values indicate a better skill than the deterministic regression model A. Values around zero represent approximately the same skill. A negative skill stands for a worse prediction ability of model B compared to model A.

not as good as model A. For location L1, both models show approximately the same skill. The worst reconstructions are obtained with tree ring width data from

location L2, with neg. exp. and Hugershoff standardized data with negative skill scores of ca. -0.1 and -0.15. Their skill increases if spline functions are chosen for standardization. The same holds for location L3. However, the skill scores for location L1 indicate a similar skill of model B to predict past temperatures compared to model B. Additionally, the small negative skill scores (minimum of ca. -0.15) indicate that the differences between the two models are not that big.

Stochastic Models versus Climate Reference

Figure 4.5 shows the skill scores of the stochastic models compared to the reference, which is represented by the mean and variance of the observations of the verification time $\tilde{\boldsymbol{t}}_j$. For all locations as well as all standardization methods, model A tends to be as good as the climate average. Except for slightly positive skill scores for location L3, all others are around zero. Model B shows another picture. Here, the skill of reconstructions based on proxy data of location L1 is better (ca. 0.2) than the reference, if the decade average is chosen as reconstructed temperature. This reconstruction method seems to be better for the other locations as well (ca. 0.1 for L2 and 0.15 for L3). In contrast, reconstructions from L2 are worse compared to the reference, especially for negative exponential and Hugershoff standardized tree ring width data. Here, the same reconstruction method, which yields better results compared to the reference, shows a negative skill score of approximately -0.3.

Skill Scores of Stochastic Models

Finally, the skill scores of the two stochastic models are shown in figure 4.6. The conditional model shows positive CRPS skill scores of about 0.1 – 0.2 for reconstructions based on data from location L1 (all standardization methods), L2 and L3 (spline standardized). Again, the skill is negative for reconstructions from location L2 and especially for proxy data which are standardized with an exponential function. The worst result is obtained from location L3 with Hugershoff standardization method. The skill of the stochastic conditional model is worse than -0.3. In general, model B shows better results if the mean of the reconstructed decade is chosen as the temperature of year i. For the other case, both models do not show big differences which is a simliar behavior as in the deterministic case (figure 4.4).

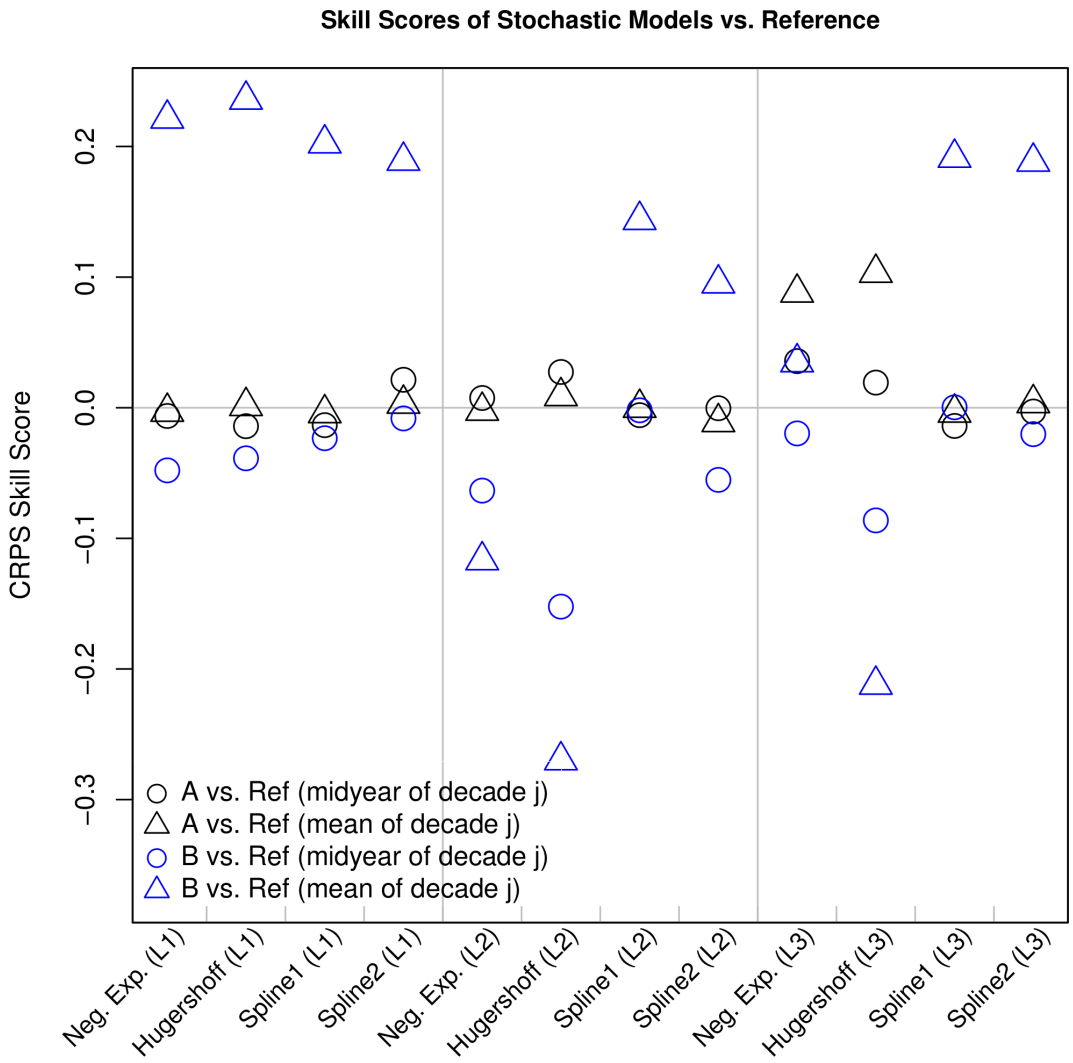

Figure 4.5: CRPS skill scores of the stochastic models compared to the reference. Positive values indicate a better skill than the reference. Values around zero represent approximately the same skill. A negative skill stands for a worse prediction ability of the stochastic models compared to the reference.

4.2 Time Truncation

Because cross validation is applied decade–wise in order to obtain independent training and verification datasets, $d-1=10$ years are removed at the beginning and the end of the initial timeline and the skill scores are calculated again.

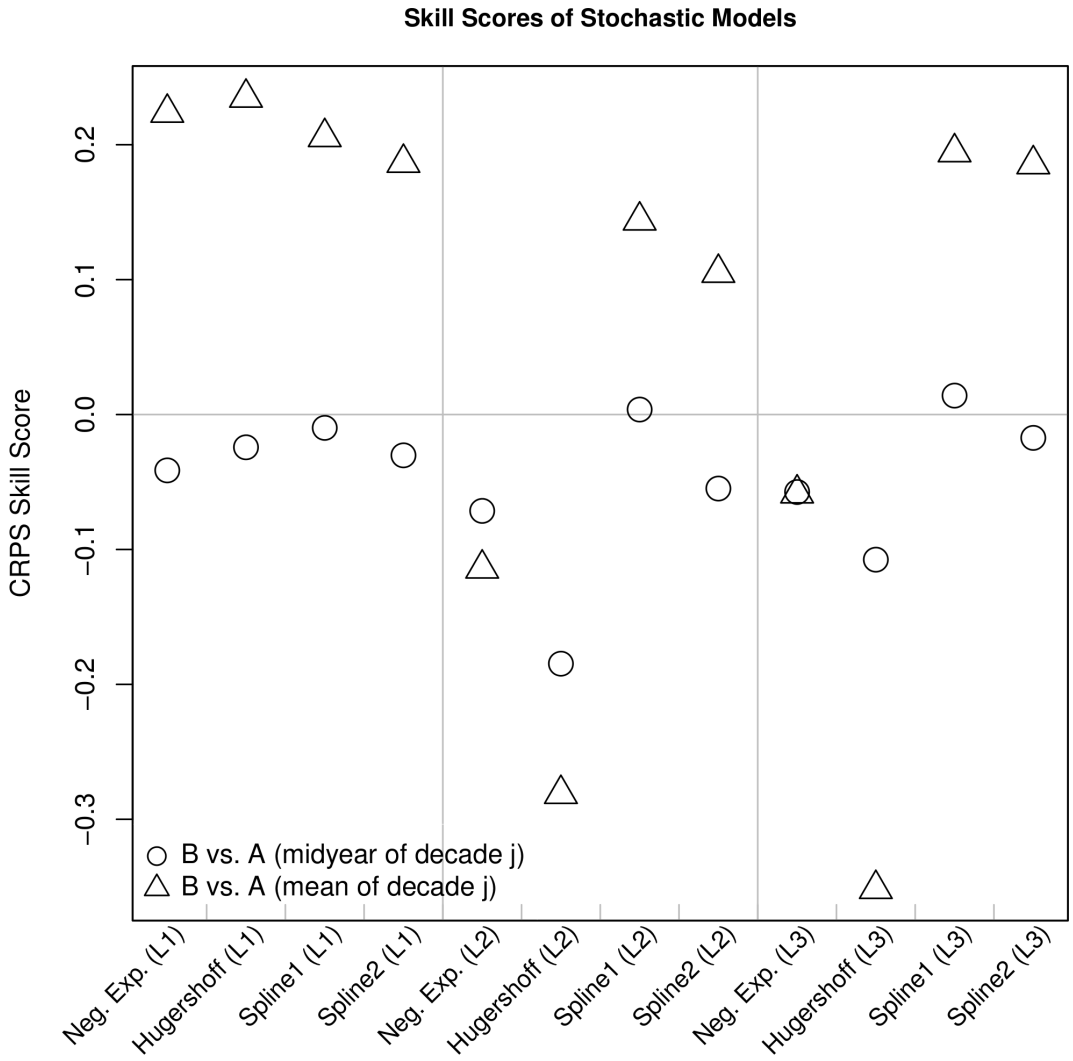

Figure 4.6: CRPS skill scores of the stochastic conditional transfer model B. Positive values indicate a better skill than the stochastic regression model A. Values around zero represent approximately the same skill. A negative skill stands for a worse prediction ability of model B compared to model A.

Skill Scores of Truncated Stochastic Models versus Climate Reference

Figure 4.7 shows the skill scores of the stochastic models compared to the reference but in this case with 10 additional years truncated at the beginning and end. Compared to the CRPS skill scores for the complete reconstructed time $\boldsymbol{t}$, the skill of the stochastic conditional model B increased for all locations around ca. $50-75\,\%$ when the mean of the reconstructed decade is chosen. Moreover, negative skills improve and the worst skill score for the stochastic conditional model B for the complete

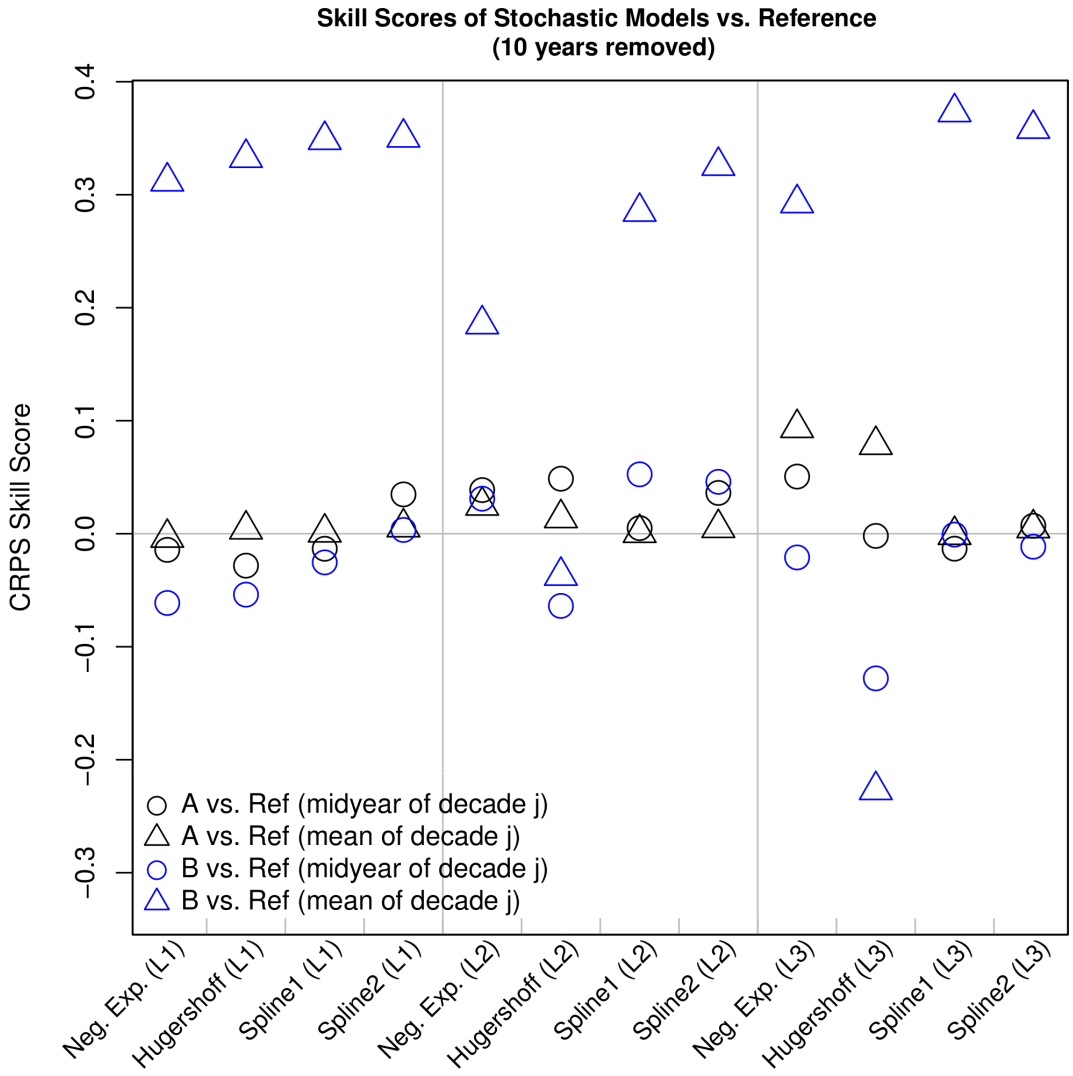

Figure 4.7: CRPS skill scores of the stochastic models against the reference with 10 years removed at the beginning and end of the available time $\boldsymbol{t}$. Positive values indicate a better skill than the reference. Values around zero represent approximately the same skill. A negative skill stands for a worse prediction ability of the stochastic models compared to the reference.

time $\boldsymbol{t}$ (ca. -0.3, L2, Hugershoff, see figure 4.5) is almost zero now. Remarkably, all other skill scores are more or less not affected by time truncation.

Skill Scores of Truncated Stochastic Models

A similar pattern is shown when the skill of the stochastic models compared to each other is examined. Figure 4.8 shows the truncated CRPS skill of the stochastic conditional model B against the stochastic regression model A. Again, the skills

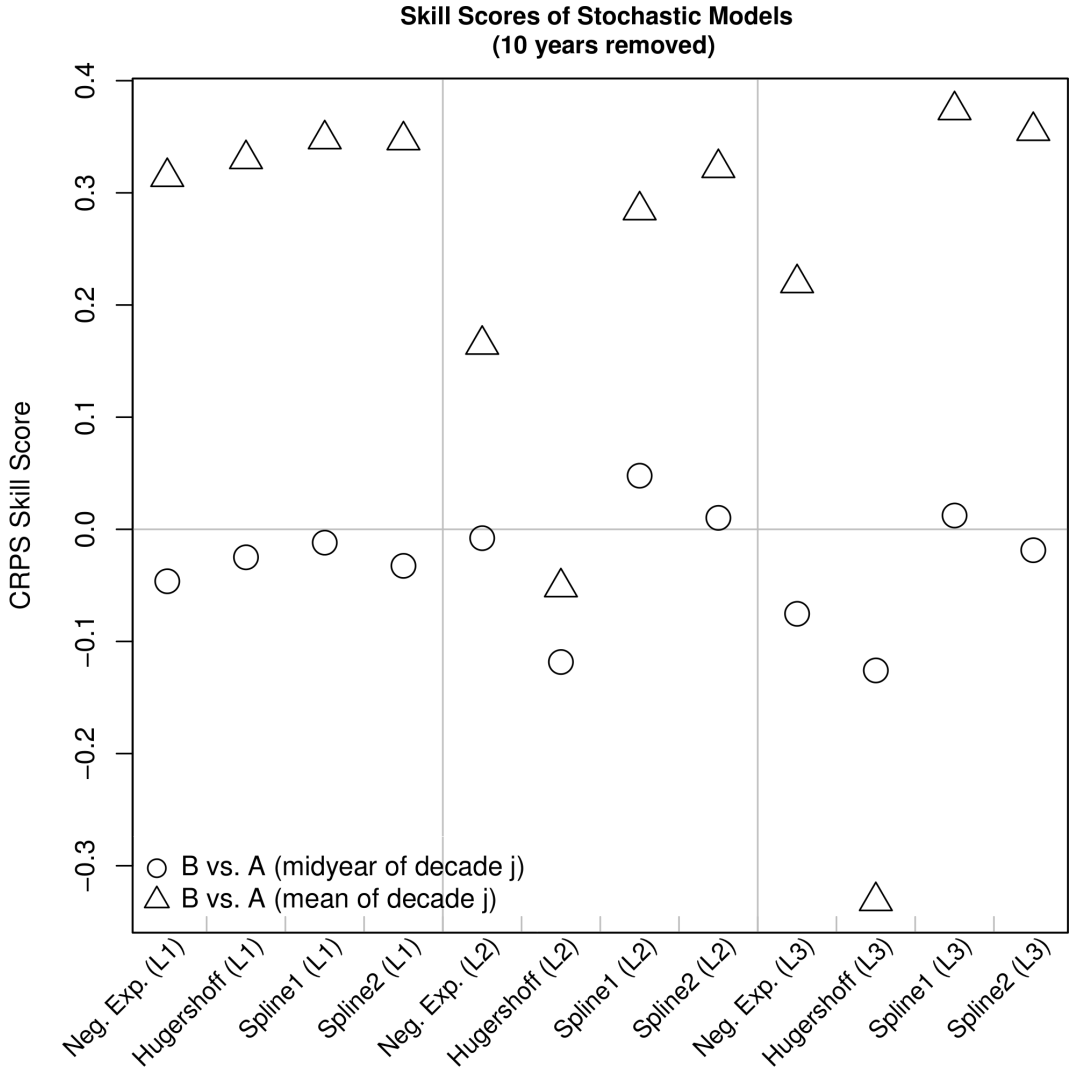

Figure 4.8: CRPS skill scores of the conditional transfer model against the regression transfer model with 10 years removed at the beginning and end of the available time $\boldsymbol{t}$. Positive values indicate a better skill than the stochastic regression model A. Values around zero represent approximately the same skill. A negative skill stands for a worse prediction ability of model B compared to model A.

improved for the conditional model around $50 - 75\,\%$ when the mean is chosen of each reconstructed decade. Now, model B has a maximum skill of almost 0.4 compared to model A, which could lead to a prediction by the stochastic model that is approximately $40\,\%$ better than that of the stochastic prediction by the linear regression model (verbally communicated with Prof. Dr. A. Hense, 2014). Again, the worst prediction skill of the stochastic conditional model comes from location L3 and Hugershoff standardized proxy data. The skill hardly improved compared to the not truncated version (still ca. -0.3). The skill scores of the reconstructions based

on the midyear from each decade seem not to be influenced by truncation which is similar to the model skills compared to the reference in the case of truncation (see. figure 4.5).

5 Discussion

The results can be summarized as follows. First, both models the linear regression (A) as well as the conditional transfer model (B) are able to predict the longtime climate. In the deterministic case, when no noise in the reconstructions is considered, both models depart from the reference in approximately the same size (see exemplarily RMSE score of location L2 in figure 4.2 and RMSE and MAE skill scores of all available data in figure 4.4). If the mean of the calibration decade $\boldsymbol{t}_j$ is chosen for reconstruction, model B seems to improve but shows worse predictions compared to A in any case (see figure 4.4). Second, in considering a stochastic noise, model B seems to improve its prediction skill compared to the reference. Positive CRPS skill scores of ca. 0.2 indicate a better prediction by model B compared to the longterm climate average of the observation data. Furthermore, adding a stochastic component to model A seems to not have no influence, since its CRPS skill score against the reference remains at values around zero, indicating no better prediction skill (see figure 4.5). Third, in comparing both models against each other, model B seems to gain an added value by considering a stochastic noise. Positive skill scores of approximately 0.2 are obtained, when the mean of the reconstructed calibration period $\boldsymbol{t}_j$ is used. In constrast, slightly worse results are shown when the midyear of $\boldsymbol{t}_j$ is chosen (see figure 4.6). Fourth, the conditional transfer model B leads to approximately 40 % better predictions when the reconstructed time is truncated so that 10 years are removed at the beginning and the end of $\boldsymbol{t}$ compared to the reference (see figure 4.7) as well as model A (see figure 4.8). This is interesting especially for reconstructions based on proxy data of location L1. The correlation of the tree ring width and the temperature is almost zero here.
However, the results indicate that we have to distinguish between different standardization techniques. In general, the two exponential functions (negative exponential or Hugershoff curve) do not differ from each other significantly. The same applies for the spline functions with a different number of knots representing variabilities on different timescales. However, there are differences between these two groups of standardization methods. Especially proxy data from location L2 and L3 that are standardized via an exponential function tend to yield worse results then the reconstructions based on splines, in particular those made by model B. In the following, these differences will be examined.

We assume that all tree ring width realizations were drawn from the normal distribution. Figure 5.1 shows a Q–Q plot of the tree ring width site mean values from location L1 (left column) and L2 (right column). Furthermore, the two different standardization methods Hugershoff curve (upper row) and spline function are compared (lower row). It can be seen that the quantiles of the proxy data of

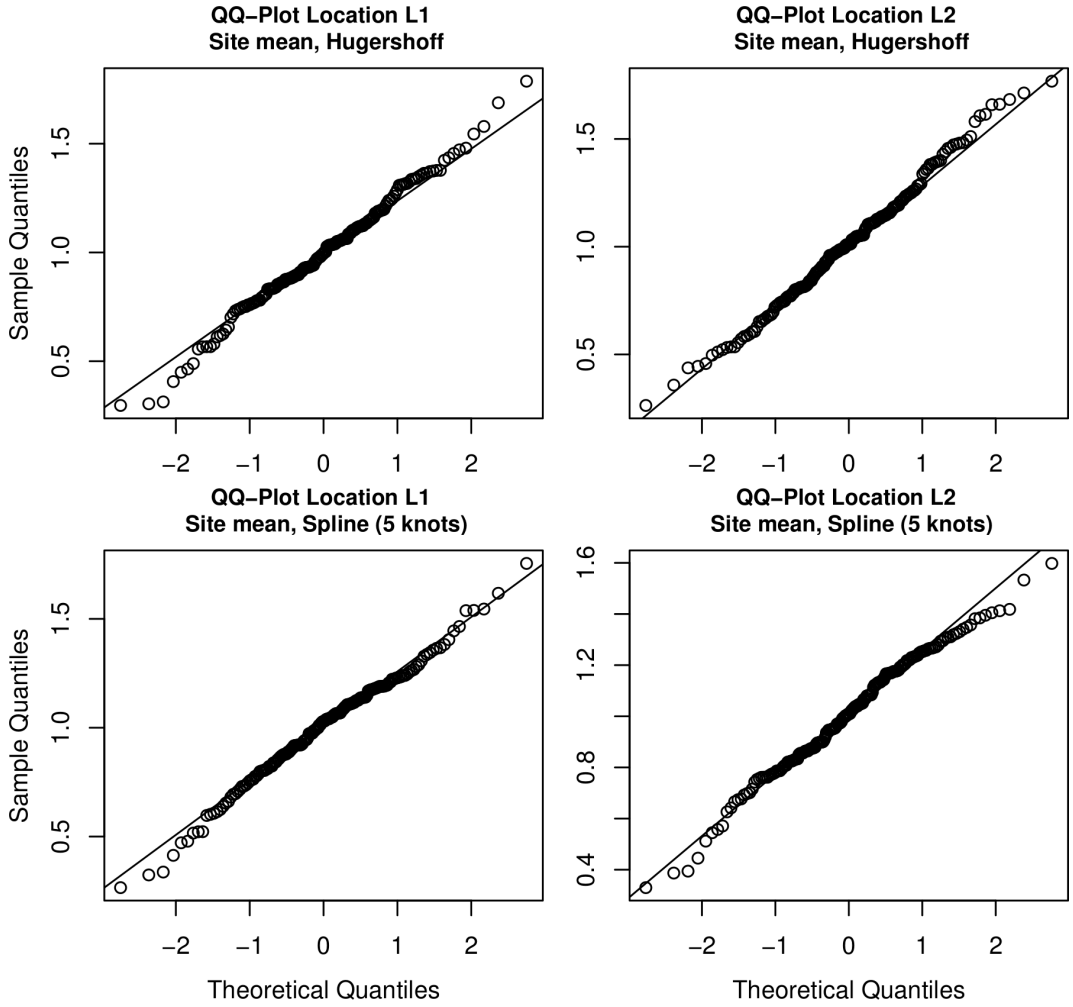

Figure 5.1: Q–Q plot of tree ring width site mean data of location L1 (left column) and L2 (right column) as well as for standardization method Hugershoff curve (upper row) and spline (5 knots, lower row)

both locations (y–axes) follow the distribution of the normal distribution (x–axes), indicated by the black line. Additionally, no significant differences between the two standardization methods are visible. The sample cdfs of the proxy data yield similar results and are not shown. Hence, the assumption of a normal distribution as the underlying distribution for tree ring width data seems valid and does not explain the departures of the model skills for location L2.

It contrast, different standardization techniques result in different autocorrelations of the proxy data which may influence the predictive skill of a transfer model. Figure 5.2 shows the autocorrelation functions for tree ring width site mean data for location L1 (left) and L2 (right) for different standardization methods. The autocorrelation is obtained as follows. First, calculation of the mean tree ring width

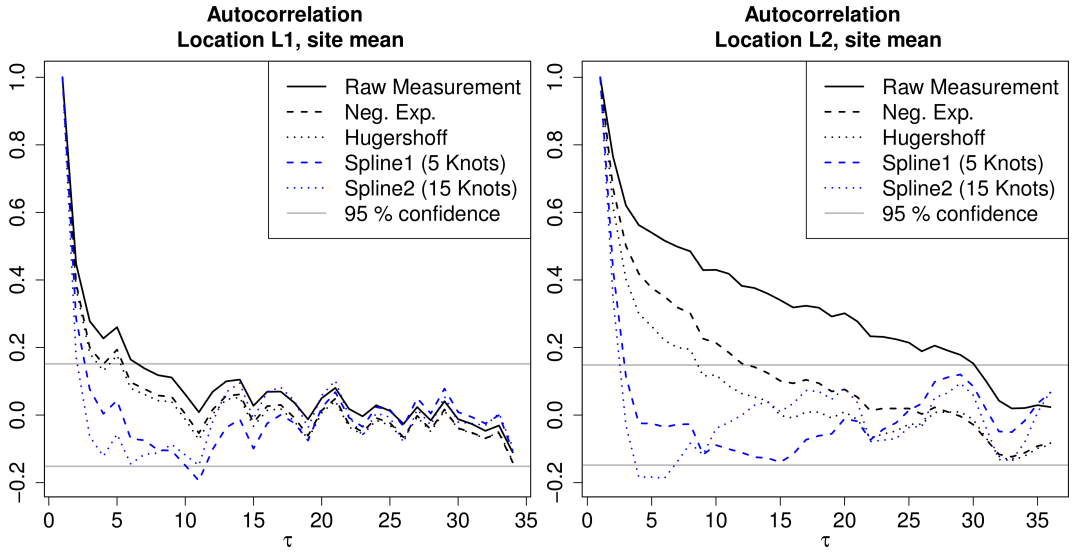

Figure 5.2: Autocorrelation function of tree ring width site mean data of location L1 (left) and L2 (right). First, tree mean is calculated via average over the number of cores per tree. Second, every tree mean time series is standardized. Third, autocorrelation function of every tree mean chronology is calculated. Finally, tree mean autocorrelation functions are averaged over number of trees per location.

series via averaging over all cores per tree (2 in most of the cases). Second, every tree mean time series is standardized. Third, autocorrelation function of every tree mean chronology is calculated. Finally, tree mean autocorrelation functions are averaged over all trees per location. According to Fritts [1976], the maximum value of the lag τ is chosen so that it equals 20 % of the length of the time series.
It becomes clear that for both locations the standardization process yields chronologies with decreased autocorrelations compared to the raw measurement series. However, the measurements of location L2 show comparatively higher autocorrelations compared to those of location L1. This trend remains after standardization of the measurements with the negative exponential function as well as the Hugershoff curve. So, comparatively high autocorrelation values are obtained and lie outside of the 95 % confidence interval up to a lag of about $10 - 13$ years (right plot in figure 5.2), whereas the maximum lag outside the confidence interval from location L1 occurs after $5 - 6$ years (left plot in figure 5.2). These high autocorrelations may lead to a decreased skill of both transfer models to predict past temperatures. In contrast, the autocorrelation functions of the spline standardized tree ring width series show a steeper decrease after a comparatively short time and yield values inside the 95 % confidence interval within five lag years for both locations. This indicates a more successful standardization which may lead to better reconstruction skills.
Obviously, the chosen method of standardization of the proxy data has an influence

on the prediction skills of our paleoclimate transfer models. Although chronologies obtained by spline standardization said to remove climate related growth trends on decadal wavelengths, they seem to improve reconstruction abilities if the autocorrelations in the raw proxy data are high.
These findings may affect the estimated covariances between the environmental and the proxy data, derived via the gLasso method (see sec. 3.4.3.2). Figure 5.3 shows exemplarily for one calibration decade $\boldsymbol{t}_j$ the submatrix $\boldsymbol{\Sigma}_{RT}$ of the estimated block matrix $\boldsymbol{\Sigma}$. Again, location L1 (left column) and L2 (right column) with Hugershoff

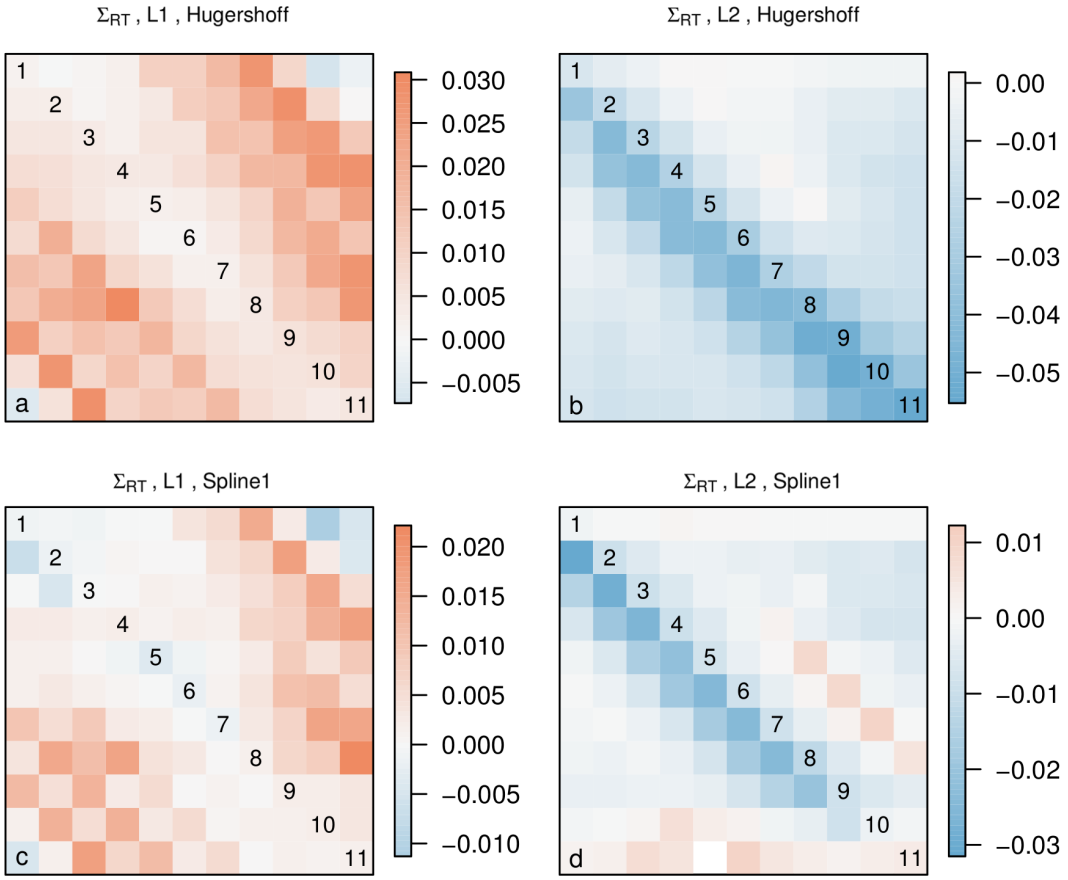

Figure 5.3: Estimated covariance submatrices $\boldsymbol{\Sigma}_{RT}$ for location L1 (left column) and L2 (right column) after standardization via Hugershoff curve (upper row) and spline function (lower row) of on exemplarily calibration decade $\boldsymbol{t}_j$.

(upper row) and spline (5 knots, lower row) standardized tree ring width values are shown. As shown in figure 3.4, these matrices represent the estimated covariabilities in a quantitative way. The diagonal elements stand for the covariances of annual temperatures and tree ring width with zero lag, that is that both variables are from the same year t_i. Every ith lower off-diagonal element shows the covariabilities of temperatures, i year(s) prior to the tree ring width data of year t_i. As expected, the

the first lower off–diagonal of (b) in figure 5.3 indicates the largest covariabilities between temperatures and tree ring width, if the temperature values one year prior to the formation of the tree ring are considered (highest values on the first lower off–diagonal). These values decrease when additional lags, i.e. years, are considered (in this case, tree ring width and temperature values are negative correlated). On the upper off–diagonal elements of (b), the opposite case is shown. The first upper off–diagonal elements show the covariabilities of a tree ring of year i and the temperature of year $i + 1$ (one year in the future). Intuitionally, these values tend to zero which is the case in (b). Obviously, this is not the case for the same time and standardization but tree ring width data from location L1 (a). Here, the covariances tend to zero on the diagonal and increase while going backward and forward in time. For this location, environmental and proxy data are positive correlated. Although this behavior of positive time–lag covariances occurs throughout all years of reconstruction, both models show good prediction skills with proxy data of that location. Furthermore, a difference between standardization methods is visible. Spline standardization reduces the covariabilities more than Hugershoff at both locations (c and d). For location L2, the spline standardization yields even covariabilities of changed sign compared to Hugershoff, which is not the case for location L1.

6 Conclusions

It has been shown that a probabilistic formulation of the paleoclimate reconstruction problem may lead to better reconstructions of the past climate if stochastic noise is considered. However, the obtained reconstructions are not resolved annualy anymore but represent reconstructions of decades.
This is an important finding since analyses obtained using Bayesian methods can be easily extended to more complex problems (Gelman et al. [2004]). Therefore, the conditional transfer model could be extended to an ensemble of models, which would provide more accurate predictions as well as a better assessment of their uncertainty. In doing so, a wider spectrum of the temporal evolution of the covariances could be estimated. For example, via *ensemble kernel dressing*, distributions with more then one maximum could be obtained (Schölzel and Hense [2011]).

Furthermore, by altering prior distributions in an iterative Bayesian framework, parameters may be found which fit better and better to the data (Ellison [1996]). Figure 6.1 (b) illustrates, that a conditional probability may increase the sharpness of a model prediction in the way that its uncertainty is decreased compared to the given prior probability. Here, sharpness refers to the concentration of the predictive

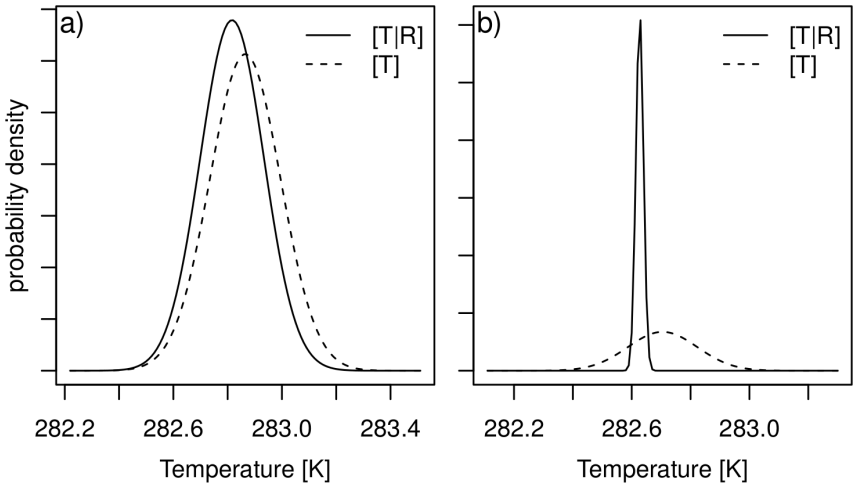

Figure 6.1: Illustration for reducing uncertainty in predictive models exemplarily for location L2. Sampled density functions of normal distributed variables with the derived conditional parameters marginalized from one decade to a one dimensional quantity by applying decade–mean. Spline (a) and Hugershoff (b) standardized tree ring width data.

pdf (Gneiting and Raftery [2007]). In contrast, the same data does not show this effect under another standardization technique (a).

Although the conditional transfer model seems to be not seamless in the way that only decades are reconstructed, it may be modified so that the length of the train-

ing decade is reduced or enlarged. Additionally, proxy data from more sites and longer records shall be used. The use of different species and proxy could yield different results. As motivated in the beginning, the derived reconstructions should be compared with results derived by global or regional circulation models. The uncertainties of the reconstructions could be assessed. As Prof. Dr. H. Storch said: Wir müssen nicht nur unser Wissen kommunizieren, sondern auch die Unsicherheit unseres Wissens[1].

[1] http://www.faz.net/aktuell/wirtschaft/wirtschaftspolitik/klimaforscher-hans-von-storch-hinterm-deich-wuerde-ich-kein-haus-bauen-12584713-p2.html, 22.12.2014

Bibliography

Briffa KR, Jones PD, Bartholin TS, Eckstein D, Schweingruber FH, Karlen W, Zetterberg P, Eronen M. 1992. Fennoscandian summers from AD 500: temperature changes on short and long timescales. *Climate Dynamics* **7**(3): 111–119.

Briffa KR, Jones PD, Schweingruber FH, Karlén W, Shiyatov SG 1996. Tree-Ring variables as proxy-climate indicators: Problems with low-frequency signals. in *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*, Springer Berlin Heidelberg.

Briffa KR, Jones PD, Schweingruber FH, Osborn TJ. 1998a. Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. *Nature* **393**(6684): 450–455.

Briffa KR, Melvin TM, Osborn TJ, Hantemirov RM, Kirdyanov AV, Mazepa VS, Shiyatov SG, Esper J. 2013. Reassessing the evidence for tree-growth and inferred temperature change during the Common Era in Yamalia, northwest Siberia. *Quaternary Science Reviews* **72**: 83–107.

Briffa KR, Osborn TJ, Schweingruber FH, Jones PD, Shiyatov SG, Vaganov EA. 2002. Tree-ring width and density data around the Northern Hemisphere: Part 2, spatio-temporal variability and associated climate patterns. *The Holocene* **12**(6): 759–789.

Briffa KR, Schweingruber FH, Jones PD, Osborn TJ, Shiyatov SG, Vaganov EA. 1998b. Reduced sensitivity of recent tree-growth to temperature at high northern latitudes. *Nature* **391**(6668): 678–682.

Buser CM, Künsch HR, Lüthi D, Wild M, Schär C. 2009. Bayesian multi-model projection of climate: bias assumptions and interannual variability. *Climate Dynamics* **33**(6): 849–868.

Christiansen B Ljungqvist FC. 2012. The extra-tropical Northern Hemisphere temperature in the last two millennia: reconstructions of low-frequency variability. *Climate of the Past* **8**(2): 765–786.

Collins M. 2007. Ensembles and probabilities: a new era in the prediction of climate change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**(1857): 1957–1970.

Cook ER, Briffa KR, Jones PD. 1994. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *International Journal of Climatology* **14**(4): 379–402.

Cook ER, Briffa KR, Meko DM, Graybill DA, Funkhouser G. 1995. The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. *The Holocene* **5**(2): 229–237.

Cubasch U, Wuebbels D, Chen D, Facchini MC, Frame D, Mahowald N, Winther JG 2013. IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex B, Midgley BM, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

DeLuisi JJ, Mateer CL, Theisen D, Bhartia PK, Longenecker D, Chu B. 1994. Northern middle-latitude ozone profile features and trends observed by SBUV und Umkehr, 1979-1990. *Journal of Geophysical Research* **99**(D9): 18901–18908.

Ellison AM. 1996. An Introduction to Bayesian Inference for Ecological Research and Environmental Decision-Making. *Ecological Applications* **6**(4): 1036.

Esper J, Cook ER, Krusic PR, Peters K, Schweingruber FH. 2003. Tests of the RCS Method for Preserving Low-Frequency Variability in Long Tree-Ring Chronologies. *Tree-Ring Research* **59**(2): 81–98.

Evans MN, Reichert BK, Kaplan A, Anchukaitis KJ, Vaganov EA, Hughes MK, Cane MA. 2006. A forward modeling approach to paleoclimatic interpretation of tree-ring data. *Journal of Geophysical Research* **111**(G3).

Fang K, Gou X, Peters K, Li J, Zhang F. 2010. Removing biological trends from tree-ring series: testing modified Hugershoff curves. *Tree-Ring Research* **66**(1): 51–59.

Fleming RA Volney WJA. 1995. Effects of climate change on insect defoliator population processes in Canada's boreal forest: some plausible scenarios. *Water, Air, and Soil Pollution* **82**(1-2): 445–454.

Frazer RA, Duncan WJ, Collar AR. 2007. *Elementary Matrices*, New York: Cambridge University Press.

Friederichs P Hense A. 2005. *Methoden der multivariaten Statistik*.

Friedman J, Hastie T, Tibshirani R. 2008. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9**(3): 432–441.

Fritts HC. 1976. *Tree Rings and Climate*, New York: Academic Press.

Fritts HC, Blasing TJ, Hayden BP, Kutzbach JE. 1971. Multivariate Techniques for Specifying Tree-Growth and Climate Relationships and for Reconstructing Anomalies in Paleoclimate. *J. Appl. Meteor.* **10**: 845–864.

Gelman A, Carlin JB, Stern HS, Rubin DB. 2004. *Bayesian Data Analysis*, Texts in Statistical Science Series, Florida, US: Chapman & Hall/CRC, 2nd ed.

Gneiting T Raftery AE. 2007. Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association* **102**(477): 359–378.

Guiot J, Wu HB, Garreta V, Hatté C, Magny M. 2009. A few prospective ideas on climate reconstruction: from a statistical single proxy approach towards a multiproxy and dynamical approach. *Climate of the Past Discussions* **5**(4).

Haslett J, Whiley M, Bhattacharya S, Salter-Townshend M, Wilson SP, Allen JRM, Huntley B, Mitchell FJG. 2006. Bayesian palaeoclimate reconstruction. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **169**(3): 395–438.

Hersbach H. 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting* **15**(5): 559–570.

Horn RA Zhang F 2005. Basic Properties of the Schur Complement. in *The Schur Complement and its Applications*, ed. Zhang F, New York: Springer, pp. 17–46.

Hughes MK Gaumlich LJ 1996. Multimillenial Dendroclimatic Studies from the Western United States. in *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*, eds. Jones PD, Bradley RS, Jouzel J, Berlin, Heidelberg, New York: Springer, vol. 41 of *NATO ASI SERIES*, pp. 109–124.

Hättenschwiler S, Schweingruber FH, Körner C. 1996. Tree ring responses to elevated CO2 and increased N deposition in Picea abies. *Plant, Cell & Environment* **19**(12): 1369–1378.

Jacoby G, D'Arrigo R, Luckman B 1996. Millenial and Near-Millenial Scale Dendroclimatic Studies in Northern North America. in *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*, eds. Jones PD, Bradley RS, Jouzel J, Berlin, Heidelberg, New York: Springer, vol. 41 of *NATO ASI SERIES*, pp. 67–84.

Jones PD Harris I. 2013. CRU TS3.10: Climatic Research Unit (CRU) Time-Series (TS) Version 3.10 of High Resolution Gridded Data of Month-by-month Variation in Climate (Jan. 1901 - Dec. 2009). *University of East Anglia Climatic Research Unit NCAS British Atmospheric Data Centre* .

Jones PD Mann ME. 2004. Climate over past millennia. *Reviews of Geophysics* **42**(2): 1–42.

Korhola A, Vasko K, Toivonen HT, Olander H. 2002. Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modelling. *Quaternary Science Reviews* **21**(16): 1841–1860.

Kumke T, Schölzel C, Hense A 2004. Transfer functions for paleoclimate reconstructions - theory and methods. in *The KIHZ project: Towards a synthesis of Holocene proxy data and climate models*, Springer, pp. 229–244.

Lauterbacher J, Schmutz C, Gyalistra D, Xoplaki E. 1999. Reconstruction of monthly NAO and EU indices back to AD 1675. *Geophysical Research Letters* **26**(17): 2745–2748.

Litt T, Ohlwein C, Neumann FH, Hense A, Stein M. 2012. Holocene climate variability in the Levant from the Dead Sea pollen record. *Quaternary Science Reviews* **49**: 95–105.

Luckman BH 1996. Reconciling the Glacial and Dendrochronological Records for the Last Millenium in the Canadian Rockies. in *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*, eds. Jones PD, Bradley RS, Jouzel J, Berlin, Heidelberg, New York: Springer, vol. 41 of *NATO ASI SERIES*, pp. 85–108.

Mann ME, Bradley RS, Hughes MK. 1998. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392**(6678): 779–787.

—. 1999. Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophysical Research Letters* **26**(6): 759–762.

Mann ME Rutherford S. 2002. Climate reconstruction using 'Pseudoproxies'. *Geophysical Research Letters* **29**(10): 139–1–139–4.

Masson-Delmotte V, Schulz M, Abe-Ouchi A, Beer J, Ganopolski A, González Rouco JF, Jansen E, Lambeck K, Luterbacher J, Naish T, Osborn T, Otto-Bliesner B, Quinn T, Ramesh R, Rojas M, Shao X, Timmermann A 2013. Information from Paleoclimate Archives. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Min SK, Simonis D, Hense A. 2007. Probabilistic climate change predictions applying Bayesian model averaging. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **365**(1857): 2103–2116.

Misson L. 2004. MAIDEN: a model for analyzing ecosystem processes in dendroecology. *Canadian Journal of Forest Research* **34**(4): 874–887.

Moberg A, Sonechkin DM, Holmgren K, Datsenko NM, Karlén W. 2004. Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature* **433**(7026): 613–617.

Morrison DF. 1976. *Multivariate Statistical Methods*, 2, McGraw-Hill Book Company.

Ohlwein C Wahl E. 2012. Review of probabilistic pollen-climate transfer methods. *Quaternary Science Reviews* **31**: 17–29.

Perkins DL Swetman TW. 1996. A dendroecological assessment of whitebark pine in the Sawtooth - Salmon River region, Idaho. *Candian Journal of Forest Research* **26**: 2123–2133.

Pollack HN. 2004. Borehole climate reconstructions: Spatial structure and hemispheric averages. *Journal of Geophysical Research* **109**(D11).

Rimbu N, Lohmann G, Felis T, Pätzold J. 2001. Arctic Oscillation signature in a Red Sea coral. *Geophysical Research Letters* **28**(15): 2959–2962.

Rind D, Shindell D, Perlwitz J, Lerner J, Lonergan P, Lean J, McLinden C. 2004. The relative importance of solar and anthropogenic forcing of climate change between the Maunder Minimum and the present. *Journal of Climate* **17**(5): 906–929.

Robertson I, Switsur V, Carter A, Lucy D, Baxter L, Pollard A, Aykroyd R, Barker A, Carter A, Switsur V, Waterhouse J. 1999. A kernel-based Bayesian approach to climatic reconstruction. *The Holocene* **9**(4): 495–500.

Röpnack A, Hense A, Gebhardt C, Majewski D. 2013. Bayesian Model Verification of NWP Ensemble Forecasts. *Monthly Weather Review* **141**(1): 375–387.

Schweingruber FH. 1996. *Tree Rings and Environment. Dendroecology*, Bern: Paul Haupt.

—. 2007. *Wood Structure and Environment*, Springer Series in Wood Sciences, Berlin, Heidelberg: Springer.

Schölzel C Hense A. 2011. Probabilistic assessment of regional climate change in Southwest Germany by ensemble dressing. *Climate Dynamics* **36**(9-10): 2003–2014.

Smerdon JE. 2012. Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *Wiley Interdisciplinary Reviews: Climate Change* **3**(1): 63–77.

Smerdon JE, Kaplan A, Zorita E, González-Rouco JF, Evans MN. 2011. Spatial performance of four climate field reconstruction methods targeting the Common Era: Four Method Comparison of CFRs. *Geophysical Research Letters* **38**(11).

Speer JH. 2010. *Fundamentals of Tree Ring Research*, The University of Arizona Press.

Stahle DW Cleavelend MK 1996. Large-Scale Climatic Influences on Baldcypress Tree Growth Across the Southeastern United States. in *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*, eds. Jones PD, Bradley RS, Jouzel J, Berlin, Heidelberg, New York: Springer, vol. 41 of *NATO ASI SERIES*, pp. 67–84.

Treydte K, Esper J, Gärtner H. 2004. Stabile Isotope in der Dendroklimatologie. *Schweizerische Zeitschrift fur Forstwesen* **155**(6): 222–232.

Vasko K, Toivonen HT, Korhola A. 2000. A Bayesian multinomial Gaussian response model for organism-based environmental reconstruction. *Journal of Paleolimnology* **24**(3): 243–250.

von Storch H, Zorita E, Jones JM, Dimitriev Y, González-Rouco F, Tett SFB. 2004. Reconstructing Past Climate from Noisy Data. *Science* **306**(5696): 679–682.

von Storch H Zwiers FW. 1999. *Statistical analysis in climate research*, Cambridge; New York: Cambridge University Press.

Wahl ER Ammann CM. 2007. Robustness of the Mann, Bradley, Hughes reconstruction of Northern Hemisphere surface temperatures: Examination of criticisms based on the nature and processing of proxy climate evidence. *Climatic Change* **85**(1-2): 33–69.

Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*, no. 91 in International Geophysics Series, Academic Press, 2nd ed.

Zerenner T. 2011. *Netzwerkanalyse an Klimadaten*, (Diplomarbeit).

Zorita E, González-Rouco F, Legutke S. 2003. Testing the Mann et al. (1998) Approach to paleoclimate reconstructions in the context of a 1000-Yr control simulation with the ECHO-G coupled climate model. *Journal of Climate* **16**(9): 1378–1390.

Zorita E, von Storch H, Gonzalez-Rouco FJ, Cubasch U, Luterbacher J, Legutke S, Fischer-Bruns I, Schlese U. 2004. Climate evolution in the last five centuries simulated by an atmosphere-ocean model: global temperatures, the North Atlantic Oscillation and the Late Maunder Minimum. *Meteorologische Zeitschrift* **13**(4): 271–289.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Thema

Bayesian Formulation of Uncertainty in Paleoclimate Reconstructions -

A Tree Ring Width Case Study

selbstständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle verwendeten Zitate sind kenntlich gemacht.

Bonn, den 22. Dezember 2014

_____

Christopher Danek