

Finding best regression approach for description of climate-growth relationships by floating time spans of varying width

W. Beck

*Federal Research Centre for Forestry and Forest Products, Institute for Forest Ecology and Forest Inventory
- Section Ecology -, Alfred-Möller-Straße 1, 16225 Eberswalde, Germany
E-mail: wbeck@bfh-inst7.fh-eberswalde.de*

Introduction

Climate growth relationships – actuality and research needs

Mathematical description of the climate-growth system (acc. to Fritts 1976) is a central topic of tree ring research. Already Douglass, the founder of dendrochronology understood tree ring width series as proxy series for climatic data. Since the beginning, numerous attempts were undertaken to investigate and to describe climate-growth relationships. Recently heat and drought during the summer of 2003 attracted the public interest to the problems of climate change and forests' long-term vitality. Effects of long lasting heat and drought are seen to be potential threats for the continuity of the existing forests. Additionally and simultaneously to increasing climatic stress we have to ascertain that the enduring deposition of eutrophying substances and the increase of carbon-dioxide concentration may cause hypertrophic growth rates and therefore raise demands for water supply. Forestry in Germany as well as in Europe is faced to the difficulties that the environmental conditions of forests continuity are changing rapidly and that the knowledge to set adaptive silvicultural measures into action is not available at the required extent.

Summarizing all facts and needs of the current situation, it can be concluded that the elaboration of quantified knowledge on the effects of climatic elements on tree growth in an all-embracing breadth is of outstanding importance.

Knowledge of the effects of climatic elements on tree growth – fields of lacking knowledge

To solve the task mentioned above open fields of lacking knowledge must be appointed:

- ranges and limits of adaptability of tree species against the background of more and more extreme climatic courses,
- climatic sensitivity of our main tree species in various regions, their local climates and soil conditions,
- modification of climatic sensitivity of a tree species by local site conditions,
- changing of competition dynamics in mixed forests against the background of climate change,
- composition of forest structures which can react plastically to climatic stress (far-reaching adaptability; mixed species; diversification of risks).

To complete these fields it is necessary to explain the preponderant portion of variance in tree ring series and to possess appropriate techniques to decompose the climatic signal in tree ring index series.

Analysis of the climate-growth system and mathematical description of climate-effect relationships

The following considerations start with the availability of tree ring index series. The methods used therefore, that means the linear aggregate model of tree ring series (Cook 1985), trend calculation and elimination, the detrending and standardisation are assumed to be known and will no further discussed here.

The tree ring index value of a certain year is a result of all the environmental factors acting on the tree stand during the whole year. In any mathematical attempt to analyse climate-effect relationships these year-related values of tree ring index are faced by climatic variables which are available for all days within the year. Owing to the impracticality to apply daily data, methods of data pooling must be applied, otherwise the problem would be unmanageable. The general criterion for such a data-pooling method would be that time spans inside the year (length and position) must be found which correspond to the real climatic demands of the trees. Monthly variables (precipitation sums or averaged temperatures) are accounted to be not fully appropriate to analyse climate-growth-relationships.

Currently wide spread methods to analyse Climate-Effect Relationships:

Single factor analysis:

Values of a climatic variable of a few months assumed to be important for tree growth are summarised (seasonalised) and compared to tree ring widths. Choice of variables and limits of the seasons are defined by the experience of the scientist.

Single year analysis:

A discontinuous method which investigates the climatic course in event years and pointer years.

Year to year comparison:

The increment of the current year is compared with the average of 10 preceding years. The applied climatic variable here is the De Martonne-index.

Multiple regression analysis:

Numerous approaches are referred for instance by Fritts (1976).

Response functions (Fritts 1976):

The substantial methodical progress consists in a control of the multicollinearity between climatic variables by PCA and in the reduction of the number of predictors by use of eigenvectors. The high number of primary variables, mostly low portion of explained variance and sometimes a complicated interpretability of the results are to be mentioned as disadvantages. Schweingruber (1993) dealing with the response-function method stated: Climate-growth relationships cannot be static.

Bootstrapped Response function (Guiot 1991); Evolutionary and Moving Intervals (Biondi & Waikul 2004):

Advantages: enhanced reliability and robustness of parameters; time dependent changes of impact strength of predictors can be shown.

State-space models; Kalman filter technique (Visser & Molenaar 1988, van Deusen & Koretz 1988):

Time series is understood as a state-space. The advantage of this method is that time-variant impact strength of variables on ring width is the determining theoretical assumption. Unfortunately there is no updated software for modern operating systems available.

Common to all mentioned continuous statistical procedures above is that the choice of variables is done by the authors, according to their background of ecological experience. Mostly monthly limits are preset. Therefore the real climatic demands of trees cannot be fully hit. The conclusion from this situation is that a new procedure must be developed which objectifies climatic data pooling and variable selection.

Development of the method of 'floating time spans of varying width'

The task consists in the determination of position and width of intra-annual time spans which explain the variance of tree ring index series at the best possible rate. Preferably few restrictive constraints should be adopted.

Realisation:

Basic climatic data (temperature, precipitation) in daily resolution were supplied by the German Weather Service (DWD) for the period from 1951 till 2003. These data were proofed for errors and homogeneity by the Potsdam Institute for Climate Impact Research (PIK). As most suitable weather station is chosen the one which is closest to the particular sample plot (horizontal and altitudinal distance).

All climatic variables included in the model are relative values related to the averaged cumulative yearly course obtained from the period from 1951 till 2003. The following predictive variables are applied:

- X1: relative precipitation balance;
- X2: $X1^2$; preliminary investigations have shown a potential nonlinearity;
- X3: relative temperature balance;
- X4: $X1 \times X3$; temperature and precipitation within a period are mostly intercorrelated; interaction term $X1 \times X3$ is used to compensate the effects of intercorrelation.

The time range to search for the optimal time span starts with 1 April and ends with 31 October. This range covers more than the vegetation period. Assuming a minimum width of time span of 21 days 194 different positions within the superior time range have to be proofed. The width of the time span will be increased by one day until the maximum width of 183 days (a half year) is reached. Applying this width of time span, 32 different positions have to be proofed. Altogether a very high, but finite number of combinations (18,419 cases) from time span width and position is produced. The optimum time span for calculating climatic variables is identified by '*complete enumeration*' of all these combinations. For each case a multiple regression analysis is calculated and its results are evaluated. As a criterion to evaluate and compare quality of all obtained solutions, the coefficient of determination is used at first. As regression analyses is applied here on time series, the synchrony of

measured series and modelled series must be kept. Therefore the coefficient of Gleichläufigkeit is calculated additionally. Sensitivity of tree ring index series is an important property and reflects in which size climate acts on tree growth. The magnitude of this impact strength must be reproduced by the regression model. A ratio of sensitivity of modelled series divided by the sensitivity of the measured series is calculated and used as a third control parameter.

In the following, the chronology “Ebersberg Forest” (tree species: Norway spruce (*Picea abies*), level II plot, Bavaria) is used as an example to demonstrate the methods and the workflow of the purpose-built computer program. After a first search-run within the range between 1 April and 31 October the time span from 17 August to 3 October is found to fit best the tree ring index series (Fig. 1).

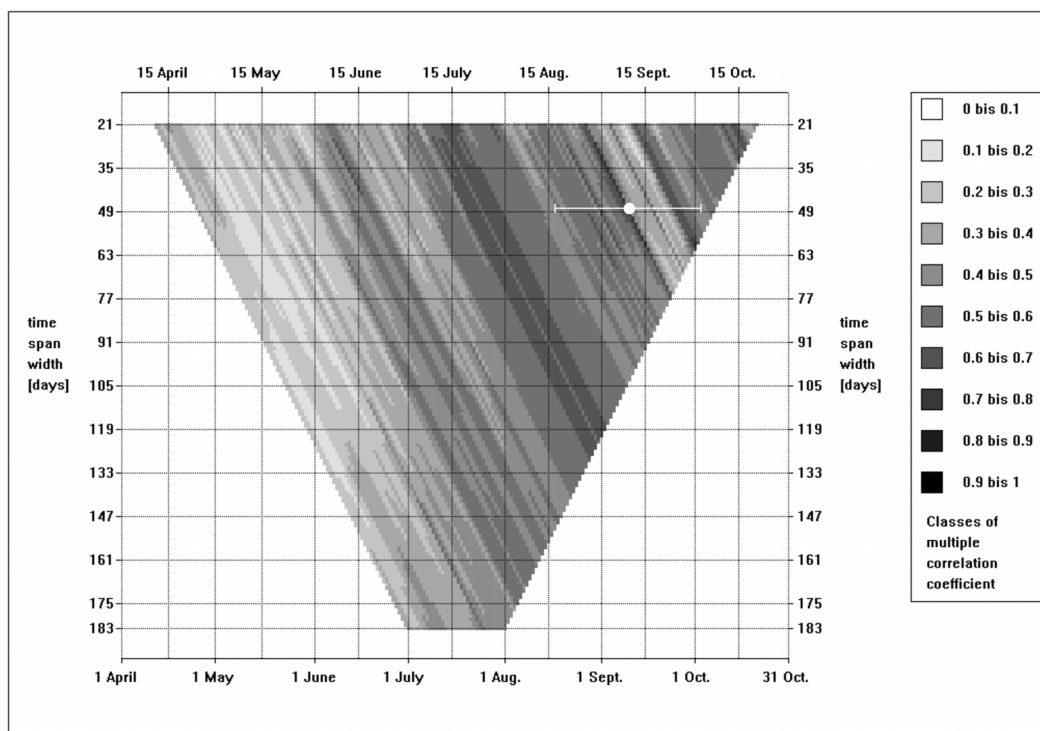


Figure 1: Results of a first searching run within the vegetation period; each rectangle indicates with its position the dated middle of a time span and with its grey level the class of multiple correlation coefficient

53,09% of the variance is explained by the four variables (mentioned above) within this time span and by the tree ring index value of the preceding year. This additional predictor is included only during the very first run and acts as a proxy value instead of the physiological initial state just before the beginning of the running year. Evaluating this initial result it has to be presumed that not only this first found time span is important for tree growth during vegetation period but other time spans may also be important and may probably enlarge the portion of explainable variance. That is why the search-run is repeated under recurrent use of the predictors of the already found time span. This step – the repetition of the search-run under inclusion of the already found time spans with their predictors – is carried out until the remaining time gaps within the range from 1 April till 31 October are smaller than 21 days (minimum time span width). The result of such a repeated search within vegetation period is shown in Figure 2.

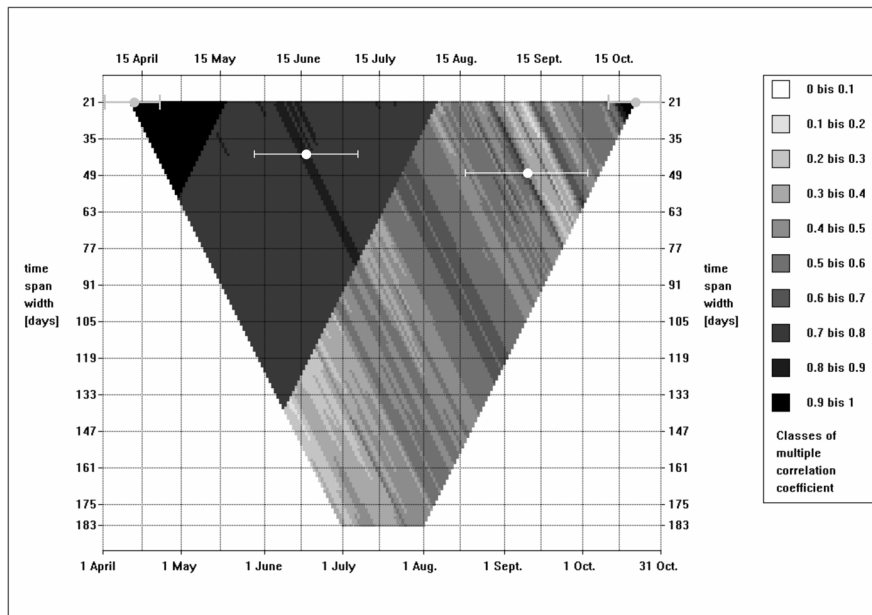


Figure 2: Results of repeated search within the vegetation period; four time spans are found and indicated by fat white points (dated middle) and by white bars (limits of the time spans)

The preponderant portion of variance, often nearly completely, is explained. However, the sufficiency and significance of the regression model at this stage is restricted. Growth rate of a running year is not only the result of factors acting in this vegetation period, but is preconditioned by preceding events, at least by the weather conditions during late winter and spring. A second time range is defined therefore from 1 December till 31 March. Here a last searching run with time spans of widths from 21 days till 84 days is carried out (Fig. 3).

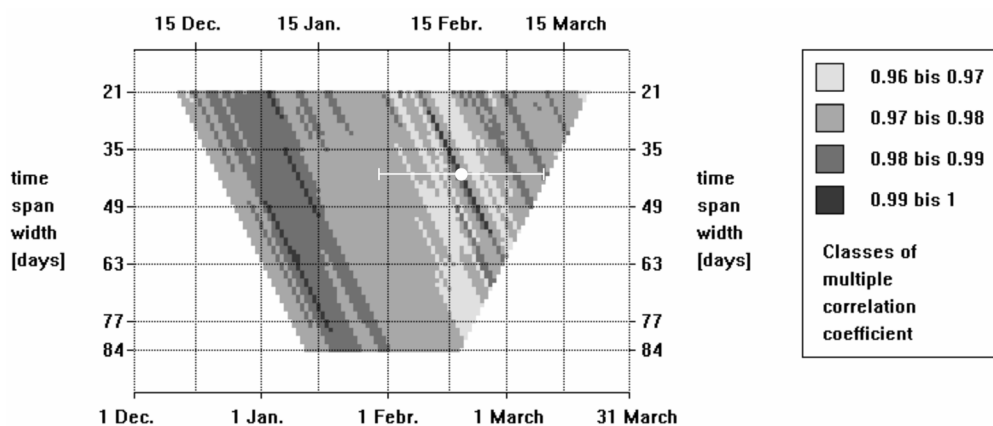


Figure 3: Results of the last search within the second time range, the winter and spring time

After that a full regression model with the maximum number of predictors exists. The next step consists of the application of an interactive backward selection strategy, because it is to assume that not all of the included predictors will score the significance level. Therefore a t-test is applied and all predictors which fail the test will be excluded and the reduced regression model will be recalculated. This stepwise reduction is done until the model includes only significant predictors. All steps of building up the regression model and the following backward selection are recorded (number of predictors, coefficient of determination, sensitivity, Gleichläufigkeit, Akaike information criterion) and summarized to a graph (Fig. 4).

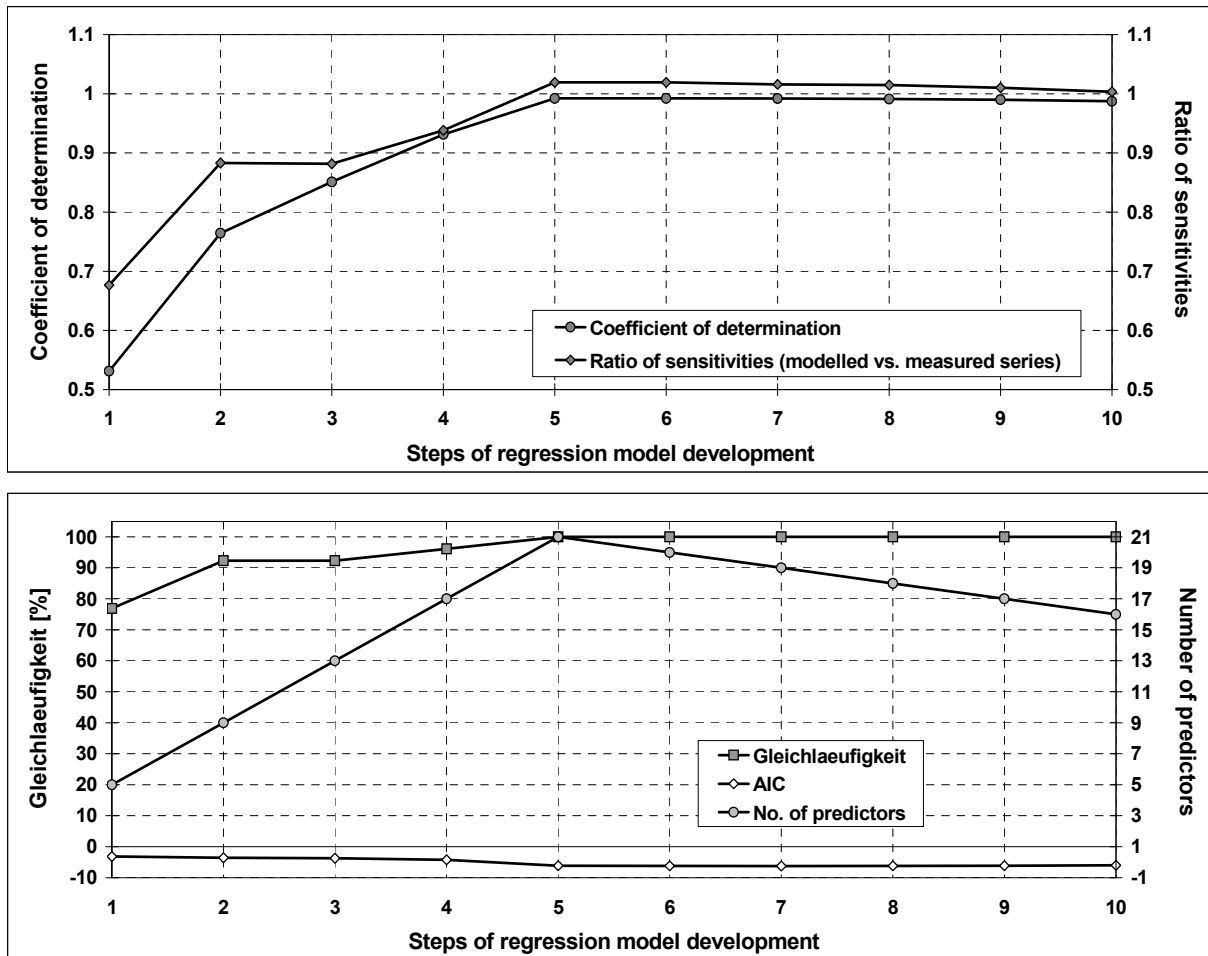


Figure 4: Built up (steps 1 to 5) and backward selection (steps 6 to 10) of the regression model; Number of predictors drops down from 21 to 16. Coefficient of determination, Gleichlaeufigkeit and ratio of sensitivities nearly keep their level during backward selection.

One of the most interesting questions is to ask for the impact strength of the found predictors within their time spans. To answer this question the following statistical parameters for each predicting variable are calculated:

- mean
- mean+standard deviation
- mean-standard deviation
- absolute maximum
- absolute minimum

To investigate the impact strength of one predictor the values of all the other variables in the model are preset to their mean and for the predictor under investigation the corresponding values of mean+standard deviation, mean-standard deviation, absolute maximum and absolute minimum are applied successively. In this way the corresponding tree ring index values are generated from regression model. Because tree ring index values (TRI) of the time series oscillate around the mean of 1,0 they can be transformed into relative growth deviations (RGD): $RGD_t = 100\% \cdot (TRI_t - 1,0)$. This procedure is applied for all included predictors and displays simultaneously a complete analysis of model sensitivity (Fig. 5).

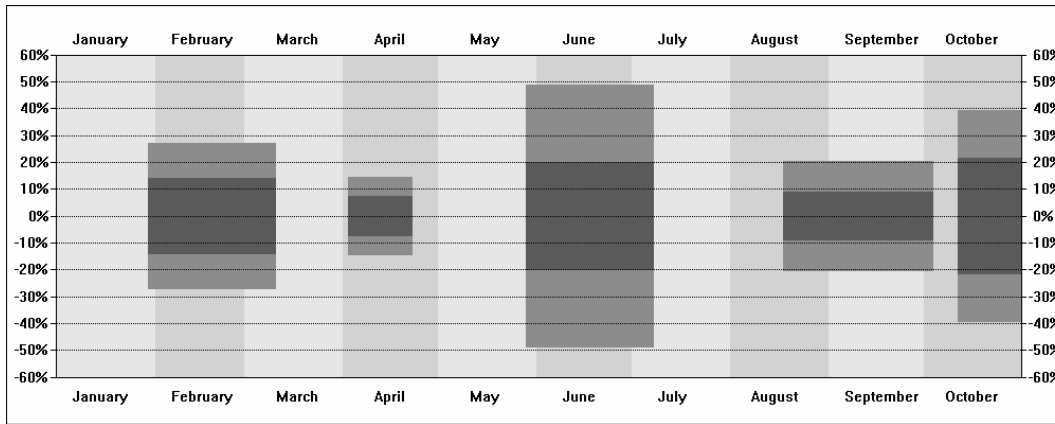


Figure 5a:
mean
relative
precipitation
balance

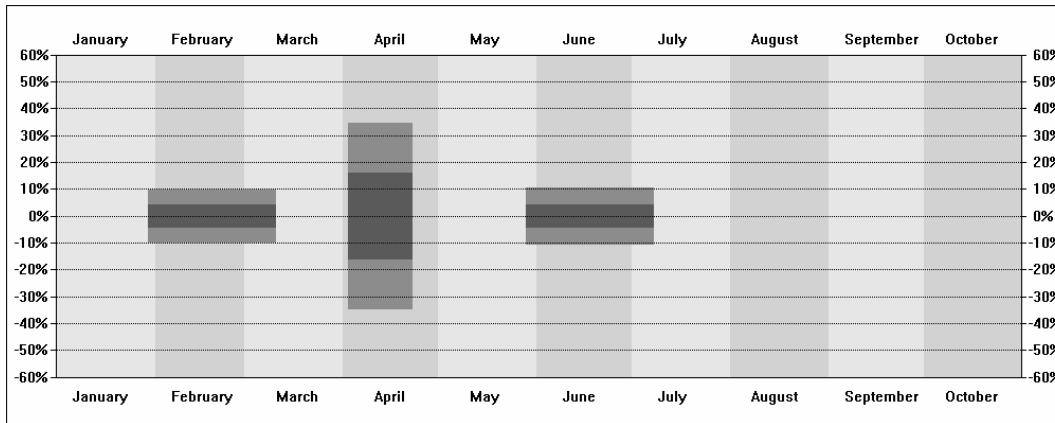


Figure 5b:
mean
relative
precipitation
balance -
quadratic

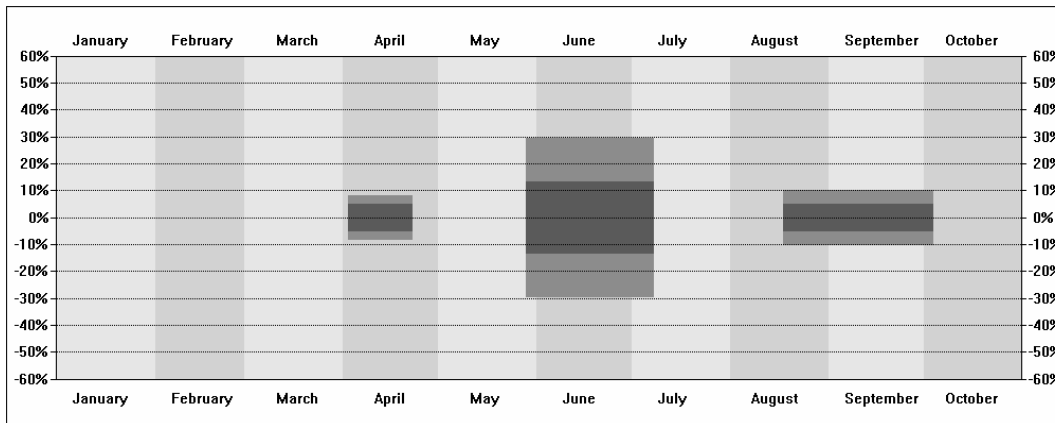


Figure 5c:
mean
relative
temperature
balance

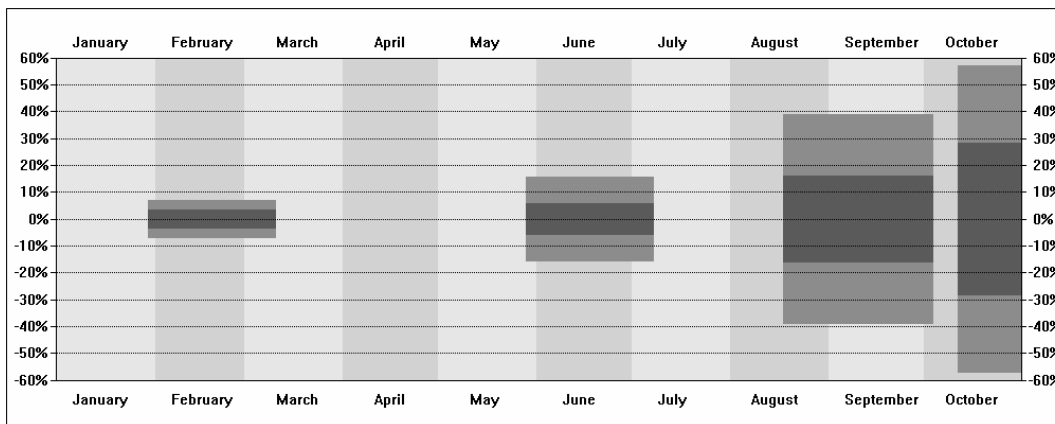


Figure 5d:
interaction
terms,
products
of mean
relative
precipitation
and temp-
erature
balance

Figure 5: Impact strength and position of the found climatic predictors; vertical axis: relative increment deviation [%] caused by the climatic variables; dark-grey bars: increment deviations caused by the predictors changing between mean + standard deviation and mean - standard deviation; mean-grey bars: increment deviations caused by the predictors changing between absolute minimum and maximum values.

Other graphs for further information include a comparison of time series of original (measured) and modelled tree ring index values (Fig. 7) and the residuals related to the predictor variables (Fig. 8).

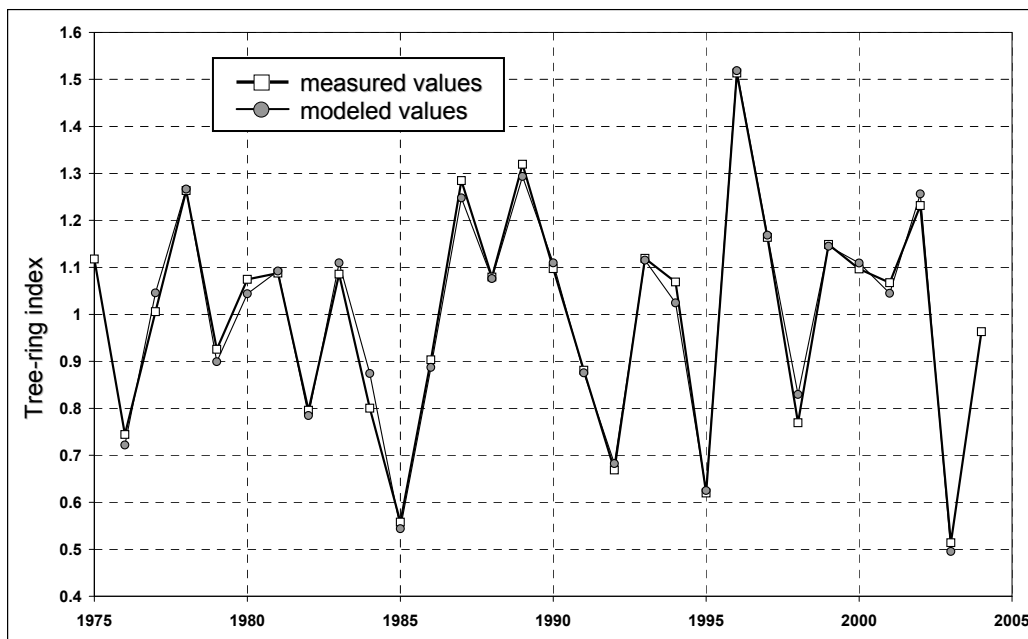


Figure 7: Comparison of the time series of measured and modelled values of tree ring index

What can this method perform? - Discussion

Tree ring index series can be reproduced by the model with high degrees of fidelity, especially if time series length is below 35 years. Variables included into the model are chosen by a self-controlling algorithm. Therefore the choice of variables is objectified. Found intra-annual time spans of climatic variables correspond to the real demands of the trees.

The multiple application of the described method to many different time series and tree species can enlarge our knowledge about the climate-growth system substantially. Especially the information about position and width of time spans wherein climatic elements are important for tree growth can improve our knowledge on the growth process. The results can also adjust some of our perceptions. In this sense it can be understood as a learning tool. Otherwise the regression model works not dynamically. If time series are too long or if environmental condition change rapidly the fidelity of results is lowered.

Application and generalisation of the results obtained from multiple use of the new method

The mode of action of climatic influences on tree growth can be compared between

- tree species growing under comparable environmental conditions such as soil properties, water supply, elevation, deposition, etc.,
- differenced site conditions including deposition of pollutants at the same tree species (numerous combinations),
- different regions with their different local climates at the same tree species,
- pure stands of a tree species and mixed stands with this species and other mixtures of species.

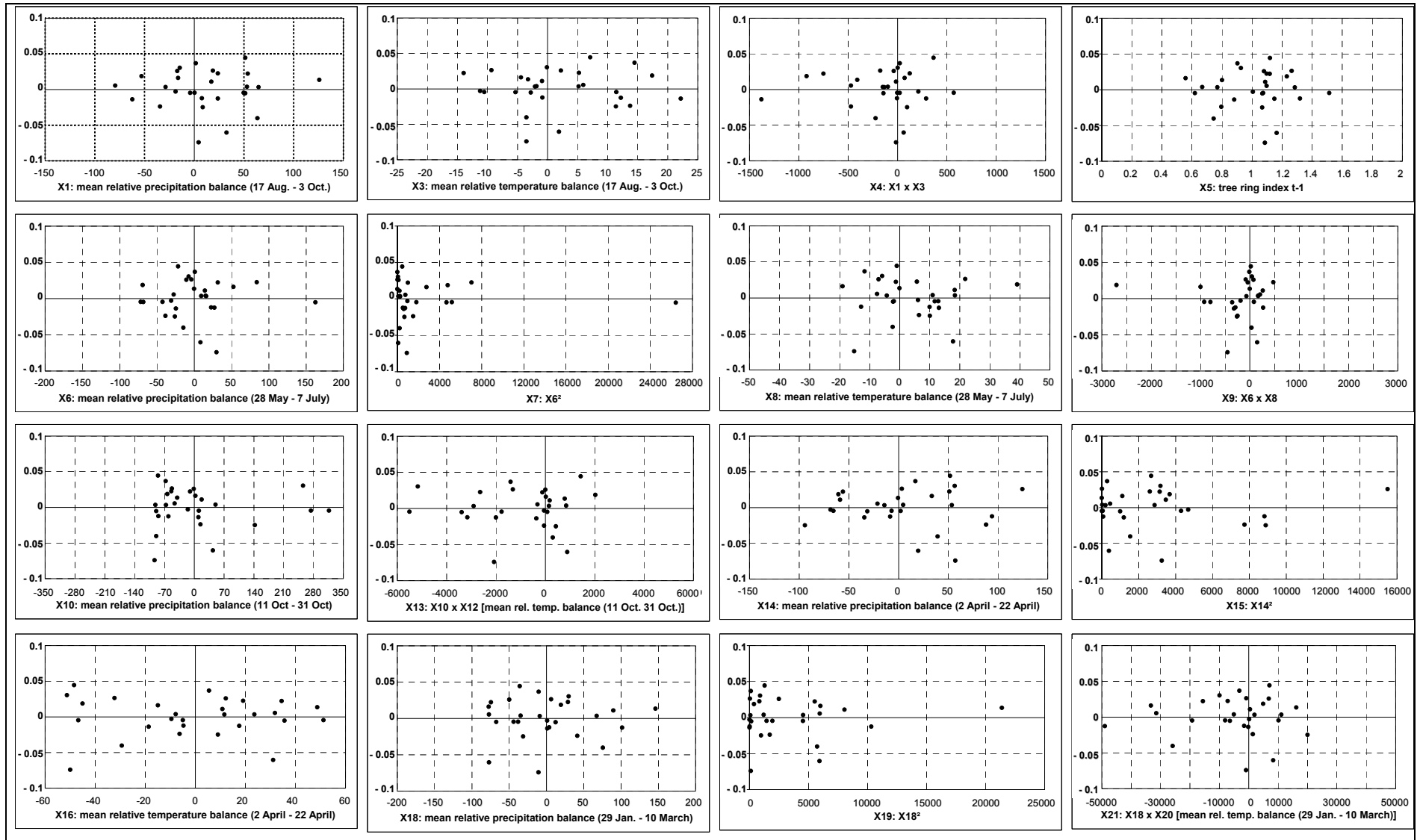


Figure 8: Residuals of the regression model related to all 16 included predictors; The predictors X2, X11, X12, X17 and X20 were excluded during backward selection because they failed the t-test.

If all the modifying effects of these causation complexes on impact strength of climate on tree growth are known, then we are able to forecast how tree and stand growth go on under conditions of progressive warming and different scenarios of precipitation supply.

The algorithm presented here produces satisfying results within the analysed time series range but it doesn't work dynamically. Time (state) dependent changes of the effect of climatic variables – that's the reality of the climate-growth system – cannot be performed. An algorithm which combines the advantages of objectifying of climatic data pooling and variable selection with moving and evolutionary intervals (pseudo-dynamic approach) or the state-space model (dynamic approach) would be appropriate to the true nature of the climate-growth system.

Acknowledgements

I am grateful to Mr. Dietrich, Bayerische Landesanstalt für Wald und Forstwirtschaft for maintaining my intention to take increment cores from the Bavarian Level II plots. At the sample plots Mr Wörle assisted me when coring. Special tribute is paid to Mrs Lasch, Potsdam Institute for Climate Impact Research for provision of climatic data from suitable weather stations uncomplicated and very quickly. Many thanks to Mr Seidling, Federal Research Centre for Forestry and Forest Products, for reviewing and discussion of an earlier draft of this paper.

References

- Biondi, F. and Waikul, K. (2004): DENDROCLIM2002: A C++ program for statistical calibration of climate signals in tree-ring chronologies. *Computers & Geosciences* 30: 303-311
- Cook, E.R. (1985): A Time Series Analysis Approach to Tree-Ring Standardization. Unpublished Ph.D. Dissertation. University of Arizona, Tucson, AZ, USA
- Fritts, H.C. (1976): *Tree Rings and Climate*. Academic Press, London, New York, San Francisco, 567p.
- Guiot, J. (1991): The bootstrapped response function. *Tree-Ring Bulletin* 51: 39-41
- Schweingruber, F.H. (1993): *Jahrringe und Umwelt: Dendroökologie*. Erweitertes Scriptum zur Vorlesung Dendrochronologie am Botanischen Institut der Universität Basel. Birmensdorf: Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft: 474
- Van Deusen, P.C., Koretz, J. (1988): Theory and programs for dynamic modelling of tree rings from climate. USDA Forest Service, General Technical Report SO-70: 18
- Visser, H., Molenaar, J. (1988): Kalman Filter Analysis in Dendroclimatology. *Biometrics*, Vol. 44, No. 4.: 929-940