

# Extracting a common high frequency signal from Northern Quebec black spruce tree-rings with a Bayesian hierarchical model

J.-J. Boreux<sup>1</sup>, P. Naveau<sup>2</sup>, O. Guin<sup>2</sup>, L. Perreault<sup>3</sup>, and J. Bernier<sup>1</sup>

<sup>1</sup>The University of Liège, Arlon, Belgium

<sup>2</sup>Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS, France

<sup>3</sup>Institut de Recherche d'Hydro-Quebec, Montréal, Canada

Received: 4 January 2009 – Published in Clim. Past Discuss.: 4 March 2009

Revised: 13 August 2009 – Accepted: 15 August 2009 – Published: 14 October 2009

**Abstract.** One basic premise of dendroclimatology is that tree rings can be viewed as climate proxies, i.e. rings are assumed to contain some hidden information about past climate. From a statistical perspective, this extraction problem can be understood as the search of a hidden variable which represents the common signal within a collection of tree-ring width series. Classical average-based techniques used in dendrochronology have been applied to estimate the mean behavior of this latent variable. Still, depending on tree species, regional factors and statistical methods, a precise quantification of uncertainties associated to the hidden variable distribution is difficult to assess. To model the error propagation throughout the extraction procedure, we propose and study a Bayesian hierarchical model that focuses on extracting an inter-annual high frequency signal. Our method is applied to black spruce (*Picea mariana*) tree-rings recorded in Northern Quebec and compared to a classical average-based techniques used by dendrochronologists (Cook and Kairiukstis, 1992).

past temperatures and precipitation for a given region and/or period for which direct observations are unavailable. One of the most widely known proxy consists in tree-ring widths that possess good skill in representing climate information at the interannual to decadal time scale. An overview on this topic can be found in Cook and Kairiukstis (1992). The fundamental assumption in dendroclimatology is that a climatic signal can be hidden into tree-ring growths. Since the pioneering work of Douglass (1936), dendrochronologists have developed various methods to extract such common signals for different species. A required step in dendrochronology, called standardization, is classically needed to transform ring-width series, that are non-stationary due to tree aging processes, into relative tree-ring indices with unit mean and a constant variance. This can be accomplished by dividing each measured ring width by its expected value, i.e. the growth trend is modelled as a regression function of tree ages. Then a common signal is derived by averaging the ensemble of such tree-ring indices across series for each year. Several methods exist to calculate indices averages (e.g., Melvin et al., 2007). Esper et al. (2002) noticed that the low-frequency climate component can be highly sensitive to the standardization method. Recently, Nicault et al. (2009) proposed a neural network approach to remove the age effect and to estimate regional growth curve via explanatory variables such as tree age and their productivity. They developed a standardization procedure to preserve long-term fluctuations.

In contrast with these past methods, our goal in this paper is neither to reconstruct a series of temperatures or precipitation, nor to propose novel regression schemes based on well-chosen explanatory variables as in Nicault et al. (2009). We prefer to focus on the problem of extracting a common inter-annual high frequency signal from a given tree species and region, without regressing on possible predictants. One

## 1 Introduction

### 1.1 Dendrochronology

In our changing climate, the search for accurate information about the past remains essential to understand and link past, present and future climate variations. Direct measurements are missing beyond the length of instrumental records and proxies are necessary to reconstruct chronologies of



Correspondence to: J.-J. Boreux  
(jj.boreux@ulg.ac.be)

reason for such a choice is based on the intrinsic difficulties in linking tree-ring growths to specific explanatory variables and in interpreting these relationships. Depending on the tree species under study, it is not always clear to dendrochronologists, even today, what are the precise contributions of precipitation, temperatures, soil and hydrological characteristics, competition and other factors, to tree-ring growths. This is particularly true for black spruces in Northern Quebec. Long and reliable instrumental records of precipitation and temperature are not available for this region. By bypassing this selection, our strategy is to let the raw data “speak” for themselves. Of course, our extracted common signal could be interpreted with respect to local measurements of temperatures, precipitation and other hydroclimatological variables, whenever such information would be available. Hence, independently of the estimation step, explanatory variables could be employed in a validation scheme. To some extent, our extraction strategy could be thought within a “blind experiments” framework. The latter is classically used in medical studies to remove the experimenter bias. In dendrochronology, the “experimenter” could be viewed as the dendroclimatologist who can select his/her favorite explanatory variables (precipitation, temperatures, soil information, etc.). In our statistical modeling, neither climatological nor environmental data (besides tree rings) are used in the analysis. Consequently, our hidden signal should not be influenced by the choice of the “experimenter”. Ideally, the extracted common signal should be then linked to climate variables by independent researchers who did not participate in our data analysis. Such a reasoning about design experiments is very common in many research fields (medical studies, nuclear physics, etc.) but it is rarely a topic of interest in climatology. Here we believe that our blind extraction represents an advantage because it can reduce the subjective link between the extracted common tree signal and the climate.

## 1.2 Bayesian Hierarchical Modeling

Assessing uncertainties in any statistical dendrochronological procedure has to be carefully addressed. To tackle this important statistical issue, we opt to work within a Bayesian Hierarchical Modeling (BHM) framework. The main idea of BHMs is to statistically model a complex process and its relationships to observations in several simple components throughout a hierarchy of layers. BHMs handle efficiently the uncertainty assessment of each layer by clearly identifying prior and posterior distributions of underlining processes. Schematically, the prior corresponds to a probability distribution representing knowledge or belief about an unknown quantity *a priori*, that is, before any data have been observed. Then, in the light of relevant data, the prior probability is updated via Bayes’ theorem and becomes the posterior. For an introduction to such models, see e.g. Gelman et al. (2003). In environmental sciences, BHM has become more and more popular during the last two decades. For example, Berliner

et al. (2000) studied long-lead predictions of Pacific Sea Surface Temperatures via Bayesian Dynamic Modeling. Cooley et al. (2005) implemented a BHM to infer glacial retreats in Bolivia using lichen growths as a proxy. Cooley et al. (2007) estimated extreme precipitation return levels by combining BHM and extreme value theory. Concerning dendrochronology, Hooten and Wikle (2007) recently investigated with a BHM shifts in the spatio-temporal growth dynamics of short-leaf pine.

The uncertainty in BHM is spread over different layers, usually three. The base level, called the *data layer*, characterizes observations, e.g. tree ring areas in our case. The second level in the hierarchy, called the *process layer*, models the latent process that drives the growth of such rings, tree-to-tree and regional variations. In this second layer, one can start incorporating temporal processes, e.g. the tree memories. The third level, called the *parameter layer*, consists of the information concerning prior parameters distributions that control the latent process.

What is the interest of BHMs for dendrochronologists? The choice of the Bayesian paradigm allows the use of unobserved variables in a hierarchical structure, while easily modeling uncertainties at each different level of this structure. In particular, expert information can be integrated via probability densities (the priors). In other words, past knowledge, even diffuse or imperfect, from scientists can be taken advantage of. More precisely, each parameter of a Bayesian hierarchical model can be viewed as a random variable and hence, a dialogue with dendrochronologists can be engaged to set the prior distribution of this random variable. If the expert has no prior knowledge then the distribution is set to be very wide (a diffuse prior), otherwise the uncertainty of the parameter can be reduced by using knowledge from past studies. In a following step, the incoming data (tree-ring areas here) are used to update all the parameters of our model. The Bayes’ theorem provides the mathematical formula to perform this updating, i.e. to derive the posterior distributions. In summary, one can see the above Bayesian strategy as an assembly of elementary parts. Its modular character makes it possible to replace prior uncertainty knowledge (set by experts) by posterior distributional information, throughout the incoming data. In this sense, it is an evolutionary construction.

The paper is organized as follows. Section 2 describes the data and the regional characteristics of the site under study. The details of our latent model are presented in Sect. 3. A short discussion about our application is proposed in Sect. 4. Perspectives are given in the conclusion.

## 2 Data and region of interest

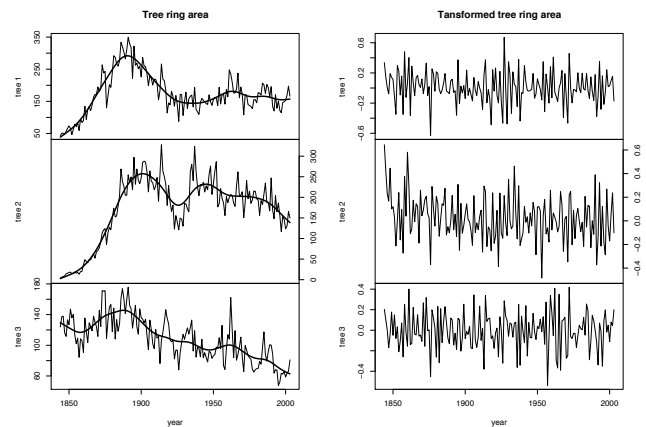
To extract a common tree signal, the dendrochronologist has to make a series of important decisions about the tree species, the region of interest and the sampling procedure



**Fig. 1.** The upper panel corresponds to Northern Quebec. The lower panel is a zoom near the Caniapiscau region and the red star called HM-1 represents the site from which fifteen trees have been sampled.

(e.g., George et al., 2008). Concerning the region choice, Hydro-Quebec, one of the founding agencies involved in this project, has had a strong interest in Northern Quebec because of its hydro-electrical capacities. With this constraint in mind, a mesic site, i.e. with a moderate supply of moisture, close to lake Hurault ( $54^{\circ}15'N$ ,  $70^{\circ}47'W$ ) was chosen, see the red star called HM-1 in the right panel of Fig. 1. This site has the advantages to belong to a climatic homogenous region and of being far away from most human activities. The black spruce (*Picea mariana*) was selected because it is a widespread species in Northern Quebec. Fifteen trees covering a period of 158 years were sampled. These trees were carefully chosen by an expert who removed singular individuals (sick trees, dominated trees, etc.). Each tree provided a ring width series from which annual growth ring areas were estimated. This transformation from ring width to ring area diminishes the geometrical effect impact, basically older trees have thinner rings. The last ring of all sampled trees, albeit missing rings, should correspond to the calendar year. Hence the youngest tree determines the common period length of all trees.

To illustrate the type of dendrochronological times series under study, Fig. 2 shows the temporal behavior of three ring area series, randomly chosen from fifteen trees. The right panels represent those three ring area series. From these three right panels, it is clear that each tree has a different trend and it seems difficult to find a common hidden signal in the low frequency domain. In addition, the variability around the cubic-spline trend in the right panels seems to be stronger after 1880 for trees 1 and 2. This example illustrates the high complexity of separating tree ring areas into their individual growth component and their common hidden component in the low frequency part of these signals. Different techniques (e.g. working with residuals after fitting a reference growth curve) exist to deal with this important issue. In this paper we do not address directly this issue. Instead we apply a simple



**Fig. 2.** Temporal behavior of three ring area time series (randomly chosen from a set of fifteen trees) over the period 1846–2003. The left panels correspond to the measured tree ring areas with a fitted cubic spline trend. The right panels indicate the log difference of the same ring areas, see Eq. (1).

non-parametric transformation to remove trends and to work with stationary time series. This implies that we only focus on inter-annual high frequencies in tree rings. The simple non-parametric transformation is defined as

$$Y_{ts} = \log X_{ts} - \log X_{t-1,s}, \quad (1)$$

with  $t=2, \dots, T$  and  $s=1, \dots, S$  and where  $X_{ts}$  represent the measured annual ring area produced during year  $t$  by tree  $s$  and  $T$  is the length of the temporal sequence and  $S$  the number of trees. Transformation (1) is extensively used in finance (Gencay et al., 2002). Besides its simplicity of implementation, this log-difference has the advantage of removing any smooth (i.e. polynomial) trend, see the right panels of Fig. 2. In addition the change of variability aforementioned in trees 1 and 2 is less pronounced in the right panels of Fig. 2. The drawbacks of using (1) are that, if present, the low frequency part of a possible common signal has been removed and that the time unit  $t$  in  $Y_{ts}$  does not correspond to a year anymore but to a one-year increment. The latter has to be kept in mind when interpreting our results. The former implies that our model described below will only focus on the high frequency part of a possible common signal.

Concerning the interpretation of  $Y_{ts}$  defined as a log-difference between two consecutive ring area values, the following simple facts need to be recalled. Whenever the relative ratio of two inter-annual consecutive ring areas from the same tree is close to one, then  $Y_{ts}$  is close to zero. If this relative ratio is very large (i.e. the ring area from year  $t$  is much larger than the one formed during year  $t-1$ ), then  $Y_{ts}$  has to strongly positive. Conversely, a negative  $Y_{ts}$  represents a large decrease in ring areas between two consecutive years. As exemplified by Fig. 3, working with  $Y_{ts}$  instead of the raw ring areas  $X_{ts}$  allows us to remove long-term trends, to

focus on the inter-annual relative variability and to work with time series that can be assumed to be stationary and Gaussian. One drawback is that we have lost the absolute value of  $X_{ts}$ , i.e. working with the couple  $(X_{ts}, X_{t-1,s})$  is equivalent to analyzing the couple  $(aX_{ts}, aX_{t-1,s})$  for any  $a>0$ , independently of the value of  $a$ . Keeping in mind this drawback and those advantages, the correlation meaning in  $Y_{ts}$  and  $Z_t$  can be viewed as the short term memory in the relative log-transform rate between two consecutive ring areas.

Before closing this section we would like to emphasize that our detrending choice represented by (1) is not unique and others techniques could be used to provide stationary signals. For example, we could have worked with the residuals obtained from the cubic spline fit shown in the left panels of Fig. 2.

### 3 An additive latent model

The random variable  $Y_{ts}$  defined by Eq. (1) is assumed to follow an additive model with a latent variable  $Z_t$

$$Y_{ts} = \mu_s + \lambda_s Z_t + \epsilon_{st}, \quad (2)$$

with  $t=2, \dots, T$  and  $s=1, \dots, S$ , and where  $\mu_s$  corresponds to the mean level of tree  $s$ ,  $Z_t$  represents the hidden regional signal common to all trees and  $\epsilon_{st}$  describes local fluctuations of tree  $s$  during year  $t$ . Tree-to-tree variations captured by  $\epsilon_{st}$  can be due to reserves accumulated by tree  $s$  and other factors that are not directly linked to environmental causes, the latter ones should be represented by  $Z_t$ . For each calendar year  $t$ , the product  $\lambda_s Z_t$  measures how the hidden factor  $Z_t$  contributes to the growth of tree  $s$ . We assume that  $Z_t$  and  $\epsilon_{st}$  are independent processes. With respect to the BHM described in Sect. 1.2, the random variables  $Y_{ts}$  corresponds to the data layer and  $Z_t$  belongs to the process layer.

Before describing the probabilistic structure within  $Z_t$  and  $\epsilon_{st}$ , it is advantageous to rewrite model (2) with obvious vectorial notations

$$\mathbf{Y}_s = \mu_s \mathbf{1} + \lambda_s \mathbf{Z} + \boldsymbol{\epsilon}_s, \quad (3)$$

where  $\mathbf{1}$  is the unit vector of length  $T-1$ . Each tree  $s$  may have a temporal memory that should depend on the hydrological stress or other conditions that are particular to this tree location. Although these tree-to-tree effects can be complex, to keep the inference simple and the risk of over-parametrization low, we opt for a simple zero-mean Gaussian auto-regressive process of order one for  $\boldsymbol{\epsilon}_s$ , i.e.  $\boldsymbol{\epsilon}_s = \phi_s \boldsymbol{\epsilon}_{-s} + \mathbf{V}_s$ . The notation  $\boldsymbol{\epsilon}_{-s}$  corresponds to  $\boldsymbol{\epsilon}_s$  shifted by one year, i.e.  $\boldsymbol{\epsilon}_{-s} = (\epsilon_{s1}, \dots, \epsilon_{s(T-1)})'$ ,  $\phi_s$  represents the auto-regressive coefficient of tree  $s$ , and the random vector  $\mathbf{V}_s$  of length  $T-1$  follows a zero-mean multivariate Gaussian distribution with precision  $\eta_s \times \mathbf{I}$  where  $\mathbf{I}$  is the identity matrix of size  $T-1$ . In other words, all components of vector  $\mathbf{V}_s$  correspond to a standardized normal independent random noise.

To allow the common regional factor  $Z_t$  to have a short year-to-year memory, we assume that the latent  $Z_t$  can be modeled as a zero-mean Gaussian auto-regressive process of order one, i.e.  $\mathbf{Z} = \rho \mathbf{Z}_{-} + \mathbf{U}$  where  $\mathbf{Z}_{-} = (Z_1, \dots, Z_{(T-2)})'$  and  $\mathbf{U}$  represents a zero-mean multivariate normal vector of length  $T-1$  with precision  $\tau \times \mathbf{I}$ .

Our full model counts  $2+4S$  parameters, namely  $(\rho, \tau)$  and  $\boldsymbol{\theta}_s = (\lambda_s, \mu_s, \phi_s, \eta_s)$  with  $s=1, 2, \dots, S$ . We assume that the priors distributions  $[\rho, \tau]$ ,  $[\boldsymbol{\theta}_1], \dots$ , and  $[\boldsymbol{\theta}_S]$  are mutually independent. By writing the joint distribution as a product of conditional distributions with a marginal distribution, the prior for  $(\rho, \tau)$  can take the following form  $[\rho, \tau] = [\rho|\tau][\tau]$ . In a classical way, we assume that the precision parameter  $\tau$  follows a gamma distribution with two hyperparameters that must be fixed to reflect prior beliefs. In our application, a diffuse prior is chosen by setting the two gamma parameters to zero.

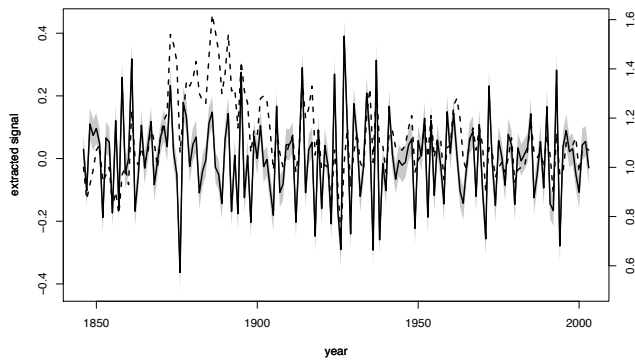
The choice of the auto-regressive coefficient prior  $[\rho|\tau]$  is more delicate. Classically, it is assumed that auto-regressive processes are a priori stationary. This implies that auto-regressive coefficients have to belong to the interval  $[-1, 1]$ . As Bayesian statisticians, we defend the idea that the underlying characteristics of the hidden process  $Z_t$  should not be imposed but arise from the data via the Bayes' rule or via prior knowledge. For this reason, we assume that  $[\rho|\tau]$  follows a zero-mean Gaussian distribution with a precision proportional to  $\tau$ . This multiplicative factor must be fixed between zero and one, mainly to degrade the precision a little. In our application, we work with a diffuse prior by equaling the multiplicative factor to zero.

Concerning the prior of the random vector  $\boldsymbol{\theta}_s = (\lambda_s, \mu_s, \phi_s, \eta_s)$ , we assume conditional independence, i.e.  $[\boldsymbol{\theta}_s] = [\lambda_s|\eta_s][\mu_s|\eta_s][\phi_s|\eta_s][\eta_s]$  where the variable  $\eta_s$  follows a gamma distribution with two hyperparameters (set to zero in our application). The distributions  $[\lambda_s|\eta_s]$ ,  $[\mu_s|\eta_s]$  and  $[\phi_s|\eta_s]$  are assumed to be diffuse Gaussian priors in this paper. As for the auto-regressive coefficient of  $Z_t$ , this means that the auto-regressive coefficient of  $\epsilon_{st}$  are not a priori assumed to be in the interval  $[-1, 1]$ .

To compute the posteriors of the latent vector  $Z_t$  and of the  $2+4S$  parameters, we implement the Gibbs sampler described in the Appendix. The Bayesian inference was carried out with the open source R statistical software (R Development Core Team, 2009). Our programs are available upon request.

### 4 Results and discussion

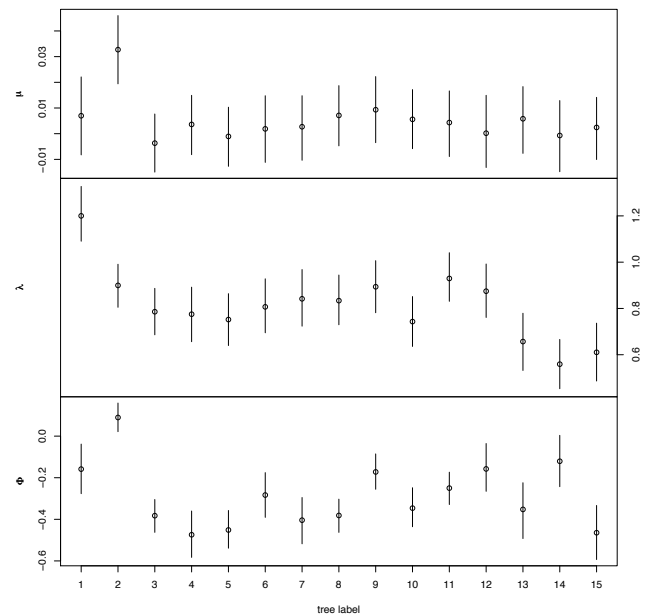
The solid line in Fig. 3 shows the estimated posterior median value of the common factor  $Z_t$  over the period 1846–2003. The shaded area corresponds to the 90% credible regions (CR). Note that the value of  $Z_t$  and  $\lambda_s$  are estimated up to a constant because it is always possible in (2) to multiply  $Z_t$  by a constant and divide the  $\lambda_s$  by the same constant



**Fig. 3.** The solid line corresponds to the estimated posterior median value of the common signal  $Z_t$  from (2) over the period 1846–2003. The shaded area corresponds to the 90% credible regions. The dashed line represents the so-called tree-growth index which is an arithmetic mean of ratios over all trees. Each ratio is derived by dividing ring thickness over a temporally smoothed tree signal for each tree (e.g., Cook and Kairiukstis, 1992).

without being able to identify this multiplicative factor. In Fig. 3, we compare our BHM results with a classical technique employed by dendrochronologists. The output of this procedure is represented by the dashed line, a so-called tree-growth index which is an arithmetic mean of ratios over all trees. Each ratio is derived by dividing ring thickness over a temporally smoothed tree signal for each tree (e.g., Cook and Kairiukstis, 1992). Up to a constant (this explains the two different scales for the y-axis), the classical tree-growth index behaves similarly to  $Z_t$  by staying in the CR over a long time period. From about 1875 to 1900, there is a discrepancy between  $Z_t$  and the classical tree-growth index, the latter producing higher values during this period. Although fairly localized in time, this difference indicates that this classical technique by not providing confidence intervals shows its limitations. Still, this comparison between the two extracted signals makes us believe that our BHM approach is capable of providing meaningful outputs for dendrochronologists because they do not contradict past results and offer another statistical approach to this community of scientists.

Concerning the memory within  $Z_t$ , the posterior distribution of the autoregressive coefficient  $\rho$  indicates a negative correlation because its 25%, 50% and 75% posterior quantiles are equal to  $-0.40$ ,  $-0.36$  and  $-0.32$ , respectively. In addition to CRs, our methods allow the practitioner to derive a finer analysis of her/his tree ring data. For example, an analysis tree-by-tree can be undertaken. For each of the fifteen trees, Fig. 4 displays the posterior mean and 90% CRs of the parameters  $\mu_s$ ,  $\lambda_s$  and  $\phi_s$ , respectively. The mean posterior value of  $\mu_s$  mostly oscillates around zero for all trees. Overall, each tree but tree 2 appears to have a mild negative inter-annual memory, all autoregressive coefficients (but tree 2) shown in the bottom panel of Fig. 4 have a  $\phi_s$  posterior



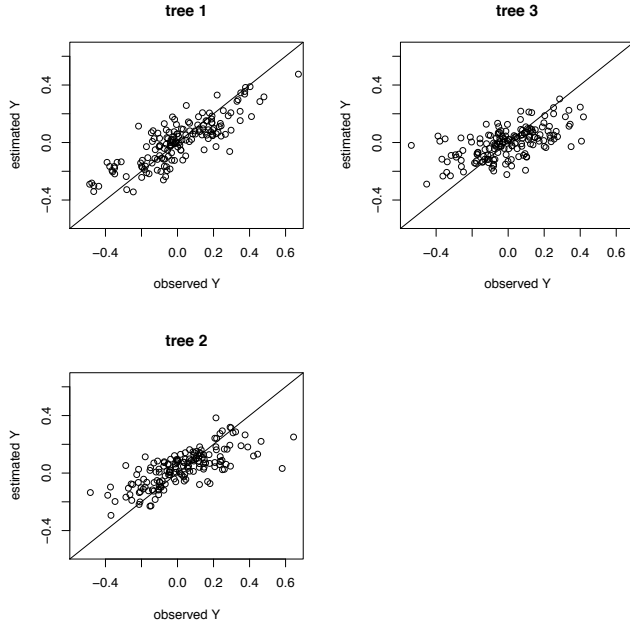
**Fig. 4.** For each of the fifteen trees, the posterior mean of  $\mu_s$ ,  $\lambda_s$  and  $\rho_s$  are represented by circles in the top, middle and bottom panels, respectively. Vertical bars correspond to the 90% credible regions.

median around  $-0.4$ . The central panel clearly points out tree 1 which seem to contribute the most to  $Z_t$ .

To check the quality of our estimation, Fig. 5 displays for trees 1, 2 and 3 (shown in Fig. 2), the observed  $Y_{t,s}$  versus the naive estimate  $\hat{Y}_{t,s}$  obtained by plugging our median posterior parameter values in (2) without noise. As expected, the relationships appear to be linear. The same result holds for the other trees.

## 5 Conclusions

To summarize our findings, we have implemented a hierarchical Bayesian model to estimate a common hidden signal in high frequency component of trees. This latent signal should be viewed as a representation of the regional pressure affecting black spruce trees over our studied area in Northern Quebec. The hierarchical structure provides another way to model the temporal structure associated to tree memories at the regional and tree-to-tree levels. This model attempts to quantify the contribution of a high frequency common hidden signal to each tree growth. This could help selecting trees with regard to a possible climatological interpretation in a reconstruction context. Compared with past approaches, our hidden signal was strongly correlated to the estimate obtained with the most traditional procedure. This confirms a past method derived by dendrochronologists, while bringing the benefits of a BHM approach. As a further step in this analysis, it would be of interest to integrate low frequency



**Fig. 5.** For each of the three randomly chosen trees described in Fig. 2, the observed  $Y_{ts}$  versus its estimator from model (2) is plotted.

in Eq. (2). One possibility is to bypass transformation (1) by making the term  $\mu_s$  in (2) varying in time. For example,  $\mu_{ts}$  could be modeled by Bayesian splines. Besides the complexity of such an approach, the main difficulty is our limited sample size (fifteen trees). Another aspect is the handling of missing values and consequently avoid the limitation brought by the age of the youngest tree. In addition, ongoing field trips should provide a much larger sample of tree rings and allows us to extend our BHM procedure in future research. In this context, our present work should rather be viewed as an addition of a simple statistical procedure to the mathematical toolbox of dendroclimatologists rather than a comprehensive study of black spruce trees in Northern Quebec.

The combination of the additive model described by (2) and the Bayesian paradigm allows the practitioner to easily generate the full posterior distribution of the hidden signal, and consequently it is possible to simulate realizations of the relative log-transform ratio between two consecutive ring areas. Such simulations could help simulating the way tree growth in response to climatic forces that drives the common inter-annual variations. Another interesting perspective of the Bayesian approach resides in the possibility to compute predictive posterior densities for future years. Since we can derive the posterior density of the extracted signal, the predictive posterior  $[z_{t+1}|\mathbf{y}_t, \theta]$  for an unobserved year  $t+1$  can be obtained by computing the hidden state posterior density at time  $t+1$ .

## Appendix A

### Gibbs sampling procedure

- *Step 0:* Initialize the vector  $\mathbf{Z}|\rho, \tau, z_0$  of length  $T$  from multivariate normal distribution with mean  $z_0\mathbf{B}_0$  and variance  $\tau^{-1}\mathbf{B}\mathbf{B}^T$  where

$$\mathbf{B}_0^t \equiv [\rho, \rho^2, \dots, \rho^T] \text{ and } \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \rho & 1 & 0 & \dots & 0 \\ \rho^2 & \rho & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^T & \rho^{T-1} & \dots & \rho & 1 \end{bmatrix}$$

- *Step 1:* Draw the precision  $\tau|\mathbf{z}, z_0, \rho$  from a gamma distribution with parameters  $a + \frac{T+1}{2}$  and  $(b + \frac{1}{2} \sum_{t=1}^T (z_t - \rho z_{t-1})^2)^{-1}$  where  $a$  and  $b$  are prior parameters (e.g.  $a=b=0$ )

- *Step 2:* Draw the correlation coefficient  $\rho|\mathbf{z}, z_0, \tau$  from a normal distribution with mean  $\frac{k_\rho m_\rho + \sum_{t=1}^T z_{t-1} z_t}{k_\rho + \sum_{t=1}^T z_{t-1}^2}$  and pre-

cision  $\left[ \tau \left( k_\rho + \sum_{t=1}^T z_{t-1}^2 \right) \right]^{-1/2}$  where  $k_\rho$  and  $m_\rho$  are prior parameters (e.g.  $k_\rho=m_\rho=0$ )

- *Step 3:* For  $s=1, 2, \dots, S$ .

- *Step 3.1:* Let  $\psi_s$  represent  $\mu_s, \lambda_s$  or  $\varphi$ . Draw  $\psi_s|\mathbf{y}_s, \mathbf{z}, z_0, y_{0s}$  from a normal distribution with mean  $\frac{k_\psi m_\psi + \mathbf{f}_s^T \mathbf{g}_s}{k_\psi + \mathbf{g}_s^T \mathbf{g}_s}$  and correlation  $[k_\psi + \mathbf{g}_s^T \mathbf{g}_s]^{-1/2}$  where  $k_\psi$  et  $m_\psi$  are prior parameters which are invariant from tree to tree (e.g.  $k_\psi=m_\psi=0$ ). The vectors  $\mathbf{f}_s$  and  $\mathbf{g}_s$  depend on handling parameters.

- *Step 3.2:* Draw precision  $\eta_s|\mu_s, \lambda_s, \varphi_s, \mathbf{y}_s, \mathbf{z}, z_0, y_{0s}$  from a gamma distribution with parameters  $c + \frac{T+3}{2}$  and  $[d + 0.5\mathbf{v}^T \mathbf{v}]^{-1}$  where  $c$  and  $d$  are prior parameters (e.g.  $c=d=0$ )

- *Step 4.:* Draw vector  $\mathbf{U}|\tau, \eta_s, \mathbf{L}_s, \mathbf{R}_s$  from a multivariate normal distribution with mean  $\omega$  and covariance  $\Omega^{-1}$  and set  $\mathbf{z}=z_0\mathbf{B}_0+\mathbf{B}\mathbf{u}$ . The mean  $\omega$  and matrix  $\Omega^{-1}$  relate vector  $\mathbf{L}$  and matrix  $\mathbf{R}$  which depend on previous parameters.

- *Step 5:* Return to step 1.

Note that when the Gamma hyper parameters are theoretically equal to zero, this means that they are set to a very small value in real computations, e.g.  $a=b=.001$  in our case.

**Acknowledgements.** This study is founded by the Hydro-Quebec, the CRSNG and OURANOS. The authors would like to thank Joel Guiot and Delphine Grancher for interesting discussions about the statistical aspects of dendrochronology, Yves Bégin (INRS-ETE) and his colleagues for their data and their scientific expertise and James Merleau for his helpful suggestions. The ANR Escarcel, AssimilEx and AQCWA projects and the MAIF foundation are also acknowledged by Philippe Naveau and Ophélie Guin. Finally, this paper is dedicated to the memory of Dominique Joly.

Edited by: H. Goosse



Publication of this paper was granted by  
EDD (Environnement, Développement Durable)  
and INSU (Institut des Sciences de l'Univers) at CNRS.

## References

- Berliner, L., Wike, C., and Cressie, N.: Long-lead prediction of Pacific SSTs via Bayesian Dynamic Modeling, *J. Climate*, 13, 3953–3968, 2000.
- Cooley, D., Naveau, P., Jomelli, V., Rababtel, A., and Grancher, D.: A bayesian hierarchical extreme value model for lichenometry, *Environmetrics*, 16, 1–20, 2005.
- Cooley, D., Nychka, D., and Naveau, P.: Bayesian spatial modeling of extreme precipitation return levels, *J. Am. Stat. Assoc.*, 102(479), 824–840, 2007.
- Cook, E. R. and Kairiukstis, A.: *Methods of dendrochronology*, Kluwer Academic Publishers, 1992.
- Douglass, A. E.: *Climatic cycles and tree-growth*, Carnegie Institution of Washington publication 289, vol. 3, 1936.
- Esper, J., Cook, E. R., and Schweingruber, F. H.: Low-Frequency Signals in Long Tree-Ring Chronologies for Reconstructing Past Temperature Variability, *Science*, 295, 2250–2254, 2002.
- Gencay, R., Selcuk, F., and Whitcher, B.: *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Academic Press, San Diego, 2002.
- Gelman, A., Carlin, J., Stern, H., and Rubin, D.: *Bayesian Data Analysis*, 2nd edn., Chapman & Hall, 2003.
- George, S., Meko, D. M., and Evans, M. E.: Regional tree growth and inferred summer climate in the Winnipeg River basin, Canada, since AD 1783, *Quaternary Res.*, 70, 158–172, 2008.
- Gelman, A.: Inference and monitoring convergence, in: *Markov Chain Monte Carlo in Practice*, edited by: Gilks, W. R., Richardson, S., and Spiegelhalter, D. J., Chapman and Hall, 1996.
- Hooten, M. B. and Wike, C. K.: Shifts in the spatio-temporal growth dynamics of shortleaf pine, *Environ. Ecol. Stat.*, 14(3), 207–227, 2007.
- Nicault, A., Guiot, J., Edouard, J.-L., and Brewer, S.: Preserving long-term fluctuations in standardisation of tree-ring series by Adaptive Regional Growth Curve (ARGC), *Dendrochronologia*, in press, 2009.
- Melvin, T. M., Briffa, K. R., Nicolussi, K., and Grabner, M.: Time-varying-response smoothing, *Dendrochronologia*, 25, 65–69, 2007.
- R Development Core Team: *R: A Language and Environment for Statistical Computing*, ISBN 3-900051-07-0, <http://www.R-project.org>, R Foundation for Statistical Computing, 2009.