

## Chapter 4

# Uncertainty, Emergence, and Statistics in Dendrochronology

Edward R. Cook and Neil Pederson

**Abstract** Some fundamental concepts of dendrochronological analysis are reviewed in the context of statistically modeling the climatically related environmental signals in cross-dated tree-ring series. Significant uncertainty exists due to our incomplete mechanistic understanding of radial growth of most tree species in the natural world, one where environmental effects are unobserved, uncontrolled, and steadily changing over time. This biological uncertainty cascades into the realm of statistical uncertainty in ways that are difficult to quantify even though the latter may be well constrained by theory. Therefore, great care must be taken to apply the many well-developed and tested statistical methods of dendrochronology in ways that reduce the probability of making false inferences. This is especially true in the case of biological emergence. This is a special case of uncertainty that arises from the way in which trees as complex organisms can have properties expressed in their ring widths that are impossible to predict from a basic understanding of lower-level physiological processes. Statistical modeling must be conducted in ways that allow for the discovery of such phenomena and, at the same time, protect from the incorrect acceptance of spurious emergent properties. To reduce the probability of the latter, we argue that model verification be an important part of any dendrochronological inquiry based on statistics. Correlation and response function analysis is used to illustrate some of the concepts discussed here. The value of empirical signal strength statistics as predictors of climatic signal strength in tree rings is also investigated.

**Keywords** Dendrochronology · Uncertainty · Emergence · Statistics · Response functions · Verification

---

E.R. Cook (✉)

Tree-Ring Laboratory, Lamont-Doherty Earth Observatory, 61 Route 9W, Palisades, NY 10964, USA

e-mail: drdendro@ldeo.columbia.edu

## 4.1 Introduction

The science of dendrochronology resides in a physical and biological world dominated by *uncertainty* and *emergence*. The relative importance of each will vary considerably from study to study, but neither uncertainty nor emergence can be assumed to be absent from most, if not all, tree-ring studies, especially those that seek to identify and extract environmental signals from tree rings. The closest thing we have to absolute certainty in dendrochronology is the assignment of calendar year dates to annual tree rings by an experienced tree-ring scientist using some accepted method of crossdating (e.g., Huber 1943; Douglass 1946; Ghent 1952; Stokes and Smiley 1968; Baillie and Pilcher 1973; Heikkinen 1984; Wigley et al. 1987; Schweingruber et al. 1990; Yamaguchi 1991; Yamaguchi and Allen 1992; Fowler 1998). Without this foundation, dendrochronology ceases to exist as a legitimate science.

Yet even here, dendrochronologists have not been immune from accusations of less dating certainty than is claimed (i.e., *zero* dating uncertainty), hence the importance of the paper by LaMarche and Harlan (1973) that documented the accuracy of the tree-ring dating of semiarid site bristlecone pine in California for calibration of the radiocarbon ( $^{14}\text{C}$ ) timescale. A similar study by Pilcher et al. (1984) further illustrated the power of crossdating for absolute dating of tree rings in their documentation of the dating accuracy of the long European oak tree-ring chronology. In turn, high-precision  $^{14}\text{C}$  measurements of independently crossdated bristlecone pine and European oak wood over the same time period have yielded very similar long-term  $^{14}\text{C}$  variations (Linick et al. 1985). This and the confirmation of the radiocarbon Suess (1965) ‘wiggles’ in dendrochronologically dated wood (de Jong et al. 1979) have independently validated the robustness of crossdating as a geochronological dating tool. These studies have collectively established crossdating as the most precise and accurate dating method in geochronology.

Since that time, quality control programs like COFECHA (Holmes 1983) have become indispensable tools for objectively testing the quality of crossdating and the correct assignment of calendar year dates to tree rings. These computer-assisted tools have convinced all but the most intransigent critics of dendrochronology that crossdating works as claimed when it is applied by properly trained individuals. This being the case, it is useful to briefly define and discuss now both uncertainty and emergence in generic terms before proceeding to the heart of this chapter. Simply put, *uncertainty* means that we do not know as much as we would like about the information contained in the tree rings we are studying, and *emergence* means that those same tree rings are likely to contain totally unexpected and inherently unpredictable information about environmental effects on tree growth. The effects of these principles on dendrochronological research will be examined in the context of statistically identifying and extracting climatic signals from tree rings, but they truly span the entire science of dendrochronology and its myriad applications.

## 4.2 Uncertainty

*Uncertainty* arises from the fact that tree-ring analysis is usually applied as a retrospective study of radial tree growth formed in an unobserved, uncontrolled, and steadily changing environment over some years in the past. Consequently, we rarely know with much certainty the kinds and details of the environmental signals contained in tree rings prior to analysis. And even after our best analyses, we can only make our best ‘guess’ of what those signals are. This acknowledgment of uncertainty ought not be viewed negatively, because this is the way science works. The great Nobel Prize-winning physicist Richard Feynman viewed it as a necessary part of science: ‘Nothing is certain or proved beyond all doubt. And as you develop more information in the sciences, it is not that you are finding out the truth, but that you are finding out that this or that is more or less likely.’ (Feynman 1999, p. 248). So, here we have a clear statement that scientific ‘truth’ is really probabilistic and, thus, always associated with some degree of uncertainty, which in Feynman’s view should be embraced as central to scientific inquiry.

This fact does not mean that dendrochronologists cannot make useful a priori inferences about the likely environmental signals in tree-ring series. Thus, we might be able to accurately infer the most likely dominant signal(s) expressed in the growth variations of an annual ring width series based on the tree species being studied (e.g., Douglas-fir [*Pseudotsuga menziesii*] or white spruce [*Picea glauca*]), the growth metric being studied (e.g., ring width or maximum latewood density), and the growth environment in which it is growing (e.g., lower or upper forest border limits). For example, ring widths of semiarid site conifers growing at a lower forest border limit site are likely to reflect variations in available soil moisture supply and evapotranspiration demand (Fritts 1971), and maximum latewood densities of high-elevation conifers are likely to reflect variations in growing season temperatures (Schweingruber et al. 1987). However, we must be honest in saying that such inferences are nothing more than educated guesses that must be modeled and verified (Fritts 1976; Snee 1977) in some justifiable way, with the understanding that we could still be wrong. In addition, we should always be ready for surprises; i.e., unexpected discoveries that may point the way to rich new research opportunities.

Thus far, this discussion has dealt essentially with biological uncertainty because we are dealing with the problem of identifying environmental signals in a biological time series, again with an emphasis here on climatic influences on tree growth. However, a completely different form of uncertainty also exists in tree-ring data that is associated with the development of annual tree-ring chronologies most frequently used for further study. This uncertainty is statistical rather than biological, although some of the statistical uncertainty in a chronology may be generated by the tree growth properties of a particular tree species (e.g., variable ring boundaries and poor circuit uniformity). Typically, such chronologies are mean-value functions of crossdated tree-ring series from many individual trees on a site. Given that the tree-ring data used in the mean-value function all crossdate to an acceptable degree,

perhaps as determined by COFECHA (Holmes 1983), this is an explicit indication that some perhaps very complicated common environmental signal exists in those tree rings. In turn, we can expect the averaging process to concentrate the common signal in the mean-value function by averaging out the noise. The act of averaging is a powerful way of getting rid of unwanted noise in tree rings, but it is never perfect; i.e., the noise is never completely eliminated. Therefore, some measure of empirical signal strength is useful because it tells us how well we have estimated the underlying common signal and eliminated the unwanted noise.

The pioneering dendroclimatologist Edmund Schulman was aware of signal strength issues and suggested that the ratio of the mean sensitivity of the average chronology (MSc) to the average mean sensitivity of the individual series (MSs) in the chronology be used as a signal strength diagnostic (Schulman 1956, pp. 20–24). This ratio he defined as coefficient  $R$ . With no noise present in the series,  $MSc = MSs$ , so  $R = 1.0$ . With any noise present between series,  $MSc < MSs$  because the averaging process reduces variance, so  $R < 1.0$ . Thus, the lower the average chronology mean sensitivity is relative to the average of the individual series, the lower the empirical signal strength. In Schulman's examples for semi-arid site conifers,  $R > 0.80$  was common. Schulman even described how to use  $R$  for evaluating the signal strength in subsets of tree-ring series from a site to assess the long-term stability of the chronology signal. This idea presaged the subsample signal strength (SSS) statistic derived by Wigley et al. (1984).

Mean sensitivity has continued to be used as a diagnostic signal strength statistic, along with tree-ring chronology standard deviation and first-order autocorrelation (e.g., Fritts and Shatz 1975; DeWitt and Ames 1978). These statistics do not provide explicit estimates of statistical uncertainty in the chronology mean-value function, however. So, with the growing power and availability of computers in the 1960s, Harold C. Fritts introduced the use of analysis of variance (ANOVA) to quantitatively describe the sources of tree-ring chronology uncertainty (Fritts 1963), and through his %Y term, its relative signal strength. An excellent example of the interpretive use of ANOVA in tree-ring research can be found in Fritts (1969). Later, the average correlation between series (RBAR) was shown to be effectively equivalent to %Y as a measure of percent variance in common between series (Fritts 1976, p. 294). These results were in turn extended to include explicit estimates of signal-to-noise ratio in tree-ring chronologies (Cropper 1982a). Shortly thereafter, Wigley et al. (1984) explicitly derived the theory underlying the use of RBAR as an estimate of percent common variance between series, demonstrated its mathematical equivalence to %Y, and extended those results to the derivation of the expressed population signal (EPS), which provides an estimate of how closely a mean chronology based on a finite number of trees expresses its hypothetically perfect chronology based on an infinite number of trees. In addition, Wigley et al. (1984) derived the subsample signal strength statistic, which quantifies the changing uncertainty in a tree-ring chronology due to changing sample size. For more details and additional extensions of these extremely valuable and widely used measures of empirical signal strength, see Briffa and Jones (1990). A complementary method of expressing tree-ring chronology uncertainty is the development of annual confidence intervals

for tree-ring chronologies. This is easily done by using parametric methods or the data-adaptive bootstrap (Cook 1990).

Quantifying generic statistical uncertainty in annual tree-ring chronologies is standard practice now in dendrochronology, principally through the use of the RBAR, EPS, and SSS statistics. The methods for doing so are theoretically sound and well tested, but unfortunately they still tell us exactly nothing about the true strength of the environmental signal(s) in any given tree-ring chronology. It is up to the dendrochronologist to determine what those signals are, but considerable biological uncertainty still exists in knowing what they truly are and how they are expressed in the ring widths. Process-based forward models of cambial growth based on first principles of tree physiology have successfully modeled the effects of climate on radial growth of certain tree species (e.g., Fritts et al. 1991; Fritts and Shashkin 1995; Fritts et al. 1999; Shashkin and Vaganov 1993; Anchukaitis et al. 2006; Evans et al. 2006; Vaganov et al. 2006). The results to date are very promising, but the challenge remains to make these models more adaptable to the likely presence of biological uncertainty and emergence in many tree-ring studies.

### 4.3 Emergence

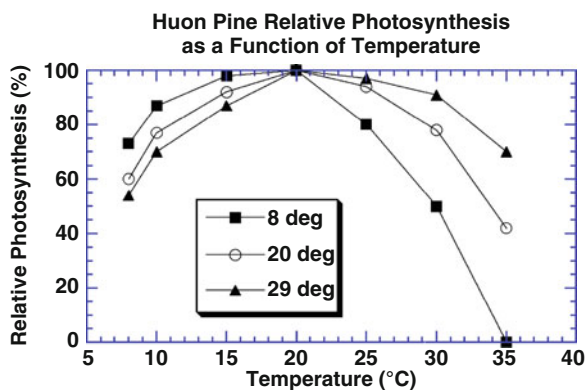
*Emergence* is the greatest source of biological uncertainty in dendrochronology because it represents a property or signal in a tree-ring series that cannot be predicted, even in principle, from our best understanding of the fundamental physiological and environmental processes that control radial growth in trees. Emergent properties arise in tree rings as a function of the inherent complexity of trees as living things and the ways in which they interact with and are constrained by their operational environment (Fritts 1976, pp. 48–50).

Examples of emergence in the fields of dendroclimatology and dendroecology are the almost universal positive correlation of March temperatures with subsequent radial growth of eastern hemlock (*Tsuga canadensis*), which appears to be independent of site characteristics and location within that species' range (Cook and Cole 1991), the almost universal positive correlation between December–January temperatures and subsequent radial growth of high-elevation red spruce (*Picea rubens*) in the northern Appalachian Mountains (Cook and Johnson 1989), the importance of winter temperatures on annual radial growth of six northern range margin tree species (Pederson et al. 2004), and the phylogenetic separation of *Quercus* species by subgenus (*Erthrobalanus* vs. *Leucobalanus*) in the West Gulf Coast forests of Texas and Louisiana, even when these oak subgenera grow together on the same site (Cook et al. 2001). None of these well-replicated statistical properties would have been predicted based on previous studies, and it is unlikely that they would ever be deduced from our best understanding of the underlying physiological processes that directly result in annual ring formation. The interactions between tree genetics and the tree's operational environment render such emergent properties inherently difficult, if not impossible, to predict. It also does not matter that some of these emergent climate signals, like the March temperature response in eastern hemlock, are of no

immediate interest for climate reconstruction. The presence of such signals is telling us something much more fundamental about how a given tree species has adapted to its climatic environment over time. That in itself is worth knowing.

Other forms of emergence in dendrochronology are arguably the effectively linear association commonly found between tree rings and local climate and the existence of crossdating itself. It is well known that the physiological processes and rates of reactions that lead to ring formation are often strongly nonlinear. A classic example is the nonlinear relationship between net photosynthesis and daily temperature (Fig. 4.1), which is caused by the differential effects of increasing temperature on rates of primary photosynthesis and dark respiration (Kozłowski et al. 1991, p. 189). The former becomes more quickly limited by increasing temperature than the latter, which results in the roughly quadratic form of the net photosynthesis curves as a function of temperature. A similar nonlinear relationship is found between net photosynthesis and leaf water potential (Kozłowski et al. 1991, p. 39). Remarkably, the aggregate expression of the many interacting nonlinear, rate-limited processes within a tree is often an effectively linear response to local climate (e.g., precipitation and temperature) in its ring widths.

This being said, certain kinds of nonlinearity may be found in tree-ring/climate relationships when a particular growth-limiting factor such as available soil moisture becomes saturated and tree growth is no longer responsive to it. Consequently, nonlinear artificial neural networks (ANNs; Guiot et al. 1995) may perform better than linear models in modeling and reconstructing precipitation and drought signals in tree-ring chronologies (e.g., D'Odorico et al. 2000; Ni et al. 2002, although there apparently can be a problem of overfitting using ANNs in relatively low signal-to-noise cases (e.g., Woodhouse 1999). The fact that the linear models developed in these examples still produced acceptable precipitation reconstructions suggests that the overall nonlinearity of tree-ring/climate relationships is not large. Vaganov et al. (1999) and Anchukaitis et al. (2006) also demonstrated how a mechanistic nonlinear model can in certain cases outperform best-fit linear statistical models when a change in local climate alters the way in which the trees respond. Such changes in climate response can also be modeled by using dynamic linear regression modeling



**Fig. 4.1** Relative photosynthesis as a function of temperature for Huon pine in Tasmania. Three curves are based on different initial temperature acclimatizations. Adapted from Read and Busby (1990)





species. However, its occurrence over a broad range of taxa growing in extremely diverse habitats worldwide indicates that crossdating is a property of tree growth that frequently *emerges* from the radial growth increments of trees being subjected to a highly variable set of microenvironmental conditions that contain within it a secondary set of common external growth-limiting factors.

The principle of emergence in biological systems might be regarded as a fundamental flaw in the current state of biological theory, one that could be eliminated if biologists knew how to develop exquisite mathematical models with high predictive skill, like those produced by physicists using classical reductionist methods to explain much of the physical world. However, the physical world is also filled with emergent properties that can never be predicted from underlying theory, even in principle. This fact was pointed out by the great evolutionary biologist Ernst Mayr, in a May 1997 interview published in the magazine *Natural History* (Angier 1997):

It's now so clear that every time you have a more complex system, new qualities appear that you could not have predicted from the components. That's the principle of emergence. I once gave a lecture in Copenhagen, and I said something I now realize to be wrong. I said emergence is characteristic only of biology. That was in 1953, when emergence was very suspect, nobody believed it. The famous physicist Niels Bohr got up to object, and I thought he'd say emergence was metaphysical and supernatural and all sorts of things. Instead he said, 'We have emergence all over the inanimate world,' and he gave the famous example of water. If you know all the characteristics of hydrogen and all the characteristics of oxygen, you still couldn't predict that the product would be liquid.

Mayr went on to say, 'So that's the end of reductionism' (Angier 1997), a highly provocative statement that is admittedly not readily accepted by many physicists and biologists who are still committed to reductionism in science.

Among others, Hatcher and Tofts (2004) have countered Ernst Mayr's provocative statement in their recent paper *Reductionism Isn't Functional*. In it, the authors propose a solution to the emergence-versus-reductionism problem based on a mathematical theorem in concurrency theory, a branch of theoretical computer science. This theorem 'proves that all possible systems can be reasoned about in terms of their subcomponents' based on methods of process algebra. Doing the modeling this way replaces underlying functions with objects, 'i.e., every object in the system is represented by an object in the model' (Hatcher and Tofts 2004). This approach accepts the validity of constitutive reductionism in biology, which states that 'systems are composed of systems or entities at a lower level and conform to the laws governing the latter' (Hatcher and Tofts 2004). If this statement is true, then object-based methods of process algebra naturally lead to the conclusion that 'all systems can be explained in terms of the sum (composition) of their parts' (Hatcher and Tofts 2004). This conclusion thus formally rejects the fundamental premise of emergence, i.e., the whole is greater than the sum of its parts, which allows for completely unexpected properties to emerge from the collective. Because constitutive reductionism is based on 'objects' and not 'functions,' it differs from classical explanatory reductionism used in theoretical physics to deduce higher-order properties from more fundamental processes. In this sense, constitutive reductionism is a weaker form of reductionism, with little or no implied deductive ability.



For object-based models to work in the way described by Hatcher and Tofts (2004), ‘every object in the system’ must be known. This we maintain is the fundamental flaw of constitutive reductionism if one were to apply it to tree-ring research. We do not have a complete understanding of the evolution of a given tree species that has resulted in its genotypic definition as a species, its phenotypic expression as an organism, and its expressible range of phylogenetic responses to its operational environment. Consequently, many properties of tree-ring series can only be discovered by experiment. So in our view, the proposed method of Hatcher and Tofts (2004) does not solve the emergence-versus-reductionism problem in tree-ring analysis in favor of constitutive reductionism, because there will never be enough known ‘objects in the system’ to model tree growth at the genus and species levels needed for dendrochronology. Nor is it likely that their approach will even work in the physical world, where theory has historically been based on mathematically rigorous forms of explanatory reductionism.

Nobel Prize-winning physicist Robert Laughlin and coauthor David Pines (Laughlin and Pines 2000) argue convincingly in their paper *The Theory of Everything* that reductionist methods ‘have succeeded in reducing all of ordinary physical behavior to a simple, correct Theory of Everything only to discover that it has revealed *exactly nothing* about many things of great importance’ (emphasis added), and they give several examples in physics where this is undoubtedly the case. Laughlin and Pines (2000) go on to state that ‘emergent physical phenomena regulated by higher organizing principles have a property, namely their *insensitivity to microscopics*, that is directly relevant to the broad question of what is knowable in the deepest sense of the term’ (emphasis added). This means that even if we know the exact laws of physics at microscopic scales, we cannot use them to deduce the macroscopic properties of something as common as a crystalline solid, because the macroscopic world is ‘regulated by higher organizing principles’ that are distinct from those at the microscopic level. This harkens back to Ernst Mayr’s earlier reference to Neils Bohr’s comment about water. Consequently, ‘living with emergence means, among other things, *focusing on what experiment tells us* about candidate scenarios for the way a given system might behave before attempting to explore the consequences of any specific model’ (Laughlin and Pines 2000; emphasis added). Even earlier, Nobel Prize-winning physicist Philip Anderson (1972) made similar arguments in his paper *More is Different*, which stimulated much of the thinking that went into the paper by Laughlin and Pines (2000). Those interested in this very important topic are strongly encouraged to read the papers by Anderson (1972), Laughlin and Pines (2000), and the recently published book *A Different Universe: Remaking Physics from the Bottom Down* (Laughlin 2005).

Emergence is an inherent macroscopic property of complex systems in both the physical and biological worlds that cannot be ignored or rejected by reductionist arguments. Its presence in dendrochronological studies should therefore be expected and, indeed, welcomed because it means that we can expect surprises in our tree-ring studies that make the science of dendrochronology fun and thought provoking. Thus, we should approach our analyses with a good deal of exploration in mind and follow statistician John Tukey’s exhortation that we be *data detectives* (Tukey 1977).

## 4.4 Statistics

Our rather lengthy, somewhat philosophical, discussions of uncertainty and emergence have set the stage now for our discussion of statistical analysis in dendrochronology. Dendrochronology is predominantly an empirical science that is deeply rooted in statistical modeling and its attendant probabilistic nature. Therefore, biological uncertainty and emergence are generally still expressed through statistical descriptions of the data. This approach places an added burden on the results and interpretations of dendrochronological analyses because, as we all know, statistical analysis can lead to the incorrect acceptance of apparently strong, yet utterly false, statistical associations. Hence the famous quote from Mark Twain's autobiography (Twain 1924, p. 246): 'There are three kinds of lies: lies, damned lies, and statistics.' Clearly, Mark Twain did not have much trust in statistics!

In statistical jargon, the generic way in which we test empirical associations is through *hypothesis testing*. We start out by assuming that no experimental treatment effect or association exists between the statistical samples or variables being compared. This is our *null hypothesis*. The *alternate hypothesis* is simply the converse of the null hypothesis; i.e., that the experimental treatment effect or association is detectable with some level of confidence. In standard statistical notation, the null and alternate hypotheses are expressed as  $H_0: \rho = 0$  and  $H_a: \rho \neq 0$ , respectively, where  $\rho$  is the statistic being tested, such as the Pearson correlation coefficient  $r$ . We then test the statistical significance of the treatment effect or association via an appropriate statistical test. When the null hypothesis is falsely rejected, this is known as a Type-1 error in statistics. In other words, we have concluded that the tested effect or association is true, when in fact it is not true. This is a serious mistake that can lead us down the wrong scientific path. To counterbalance Type-1 error, there is Type-2 error, which is the false rejection of the alternate hypothesis when it is true. This mistake is also serious, but in a different way, because now we may have missed an important scientific discovery.

The balancing act between Type-1 and Type-2 errors is generally based on the chosen  $\alpha$ -level probability used for rejecting the null hypothesis. Commonly,  $\alpha = 0.05$  is selected (a 1-in-20 chance of being wrong), which is equivalent to the 95% significance level ( $1 - \alpha$ ), but there is no theoretical reason for picking any particular  $\alpha$ -level. It all depends on how willing the analyst being wrong is in not accepting the null hypothesis as true. The '1-in-20 chance of being wrong' example also illustrates the way in which statistical hypothesis testing is generally weighted towards accepting the null hypothesis. What also matters is the choice of a one-tailed or two-tailed hypothesis test. If the sign of the outcome has no a priori expectation (e.g., the correlation  $r$  may be either positive or negative to be statistically significant), a two-tailed hypothesis test is used. Conversely, if the sign of the outcome can only be positive or negative to be meaningfully significant, a one-tailed hypothesis test is used. In this case, the alternate hypothesis notation changes from  $H_a: \rho \neq 0$  to either  $H_a: \rho < 0$  or  $H_a: \rho > 0$  to account for the sign of the outcome being important. In our later examples of modeling the climate signals in tree rings, we do not assume that we know the signs of the statistical associations a priori and, thus, we

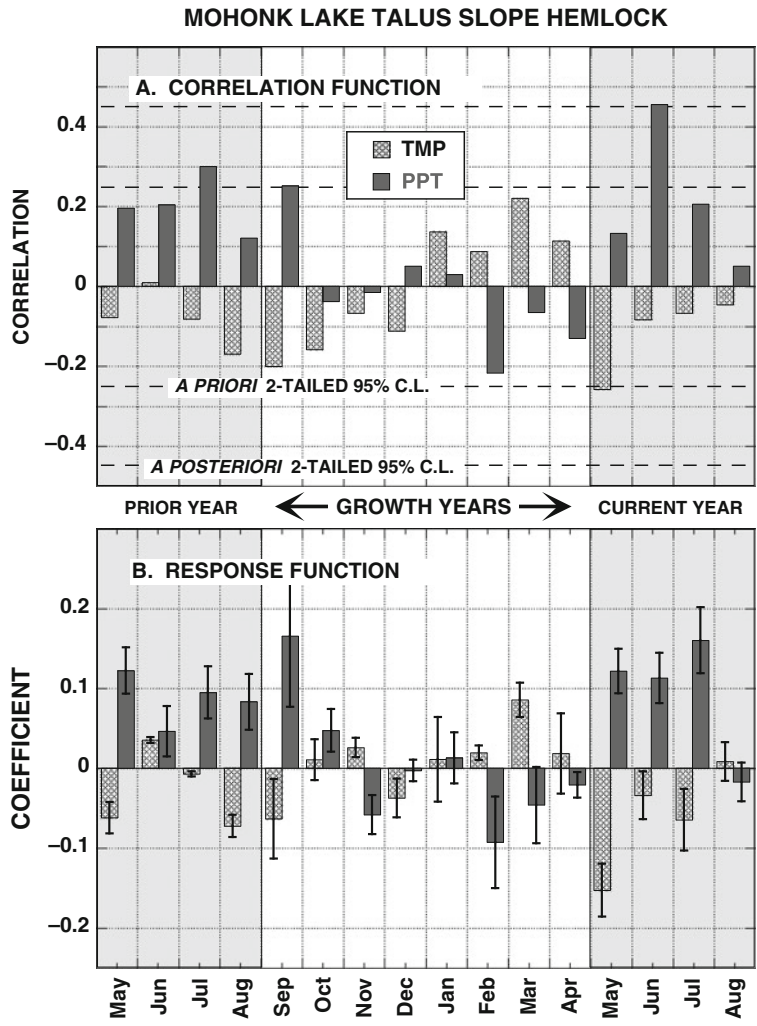
will use a two-tailed hypothesis test. For examples of using a one-tailed hypothesis test in modeling the effects of temperature on tree growth in dendrochronology, see Pederson et al. (2004).

The interpretation of any given  $\alpha$ -level probability is complicated by the fact that we often mine tree-ring data for statistically significant associations with climate through multiple tests of the same statistic. This is because we do not know what to expect more often than we would like. For example, if we wish to calculate the monthly correlations between tree rings and climate over, say, a 16-month dendroclimatic year (Fritts 1976; Blasing et al. 1984) beginning in the May of the previous year and ending in August of the current year of growth, and we do this for both temperature and precipitation, we are making 32 tests simultaneously and then looking for any statistically significant correlations to tell us how to interpret the tree-ring series as a function of climate. Unfortunately, this *multiplicity* of tests weakens the a priori  $\alpha$ -level probability for any given correlation out of the 32 candidates based on a standard  $t$ -test for  $r$ , and one can therefore no longer legitimately claim that any given correlation is truly significant at the a priori  $1 - \alpha$  level. Most reported associations between tree rings and climate in the literature ignore this multiplicity problem in describing statistically significant months in climate correlation functions.

A simple correction for the effects of multiplicity on significance levels can be made as  $P = 1 - (1 - p)^m$ , where  $p$  is the a priori probability (same as our  $\alpha$ -level described above),  $m$  is the number of tests being made, and  $P$  is the resulting a posteriori probability (Mitchell et al. 1966; Yamaguchi 1994). In our hypothetical correlation function example, if  $p = 0.05$  and  $m = 32$ , then  $P = 0.806$ , which does not instill much confidence in our statistical interpretations if the correlations barely exceed the a priori 95% significance level to begin with. To achieve an a posteriori probability  $P = 0.05$  requires an approximate a priori probability  $p \sim 0.001$  in our hypothetical case. This is equivalent to an a priori 95% significance level for  $r$  of  $\sim 0.25$  and an a posteriori 95% significance level of  $\sim 0.46$  for all 32 correlations considered jointly. This correction for multiplicity would render many correlations between tree rings and climate reported in the literature (including ours!) not statistically significant. As will be shown by example below, we should not err on the side of excessive statistical rigor at this stage, because those same climate correlations may yet have significant dendroclimatic meaning even if most of them do not exceed the a priori 95% significance level, let alone the stringent correction for multiplicity. So what do we mean here? An example of correlation and response function analysis applied to an annual tree-ring chronology will serve as an illustration.

## 4.5 Correlation and Response Function Analysis

Consider the case where few, if any, simple correlations are statistically significant ( $p < 0.05$ ) after correction for multiplicity. Figure 4.3 shows such an example for an eastern hemlock (*Tsuga canadensis*) tree-ring chronology from a xeric, quartzite conglomerate, talus slope site located near Mohonk Lake, Ulster County, New York (41.76°N, 74.15°W; elev. 379 m). Correlation and response functions (Fritts 1976;



**Fig. 4.3** Correlation function (*top panel*) and response function (*bottom panel*) analyses of an eastern hemlock tree-ring chronology located near Mohonk Lake, New York

Blasing et al. 1984) have been computed for this hemlock chronology by using monthly climate data from the Mohonk Lake cooperative weather station located about one kilometer from the tree-ring site at nearly the same elevation. The tree-ring and climate records cover the periods 1850–1996 and 1896–2004, respectively. So the joint period covered by the two datasets is 1896–1996. The correlation and response functions shown here were calculated and tested as follows:

1. The correlation and response functions span a ‘dendroclimatic year’ (Fritts 1976) that begins in May of the previous growing season and ends in August of the current growing season. This 16-month year is arranged to include two complete

radial growth seasons (roughly May–August) and the intervening ‘dormant’ season. As a consequence, the first year of climate data was used in creating the prior year portion of the 1897 dendroclimatic year, which resulted in the first analysis year being 1897.

2. The tree-ring data used were first prewhitened with a best-fit first-order autoregressive (AR) model (Box and Jenkins 1976) to remove the effect of persistence on degrees of freedom and the determination of statistical significance. The monthly climate data had very little, if any, persistence in the monthly variables. Consequently, the climate data were not prewhitened and the first analysis year remained 1897.
3. The correlation and response functions were computed by using data only over the 1931–1996 interval. This 66-year interval provides 64 degrees of freedom for the simple correlations, which yields a two-tailed a priori 95% confidence limit of  $\pm 0.25$  and an a posteriori 95% confidence limit of  $\pm 0.45$  (each shown as dashed lines in Fig. 4.3a).
4. The response function was based on retaining as candidate predictors the eigenvectors of climate with eigenvalues  $> 1.0$  (the EV1 rule; Guttman 1954; Kaiser 1960) and the best-fit regression model was determined by using the minimum Akaike information criterion (AIC; Akaike 1974; Hurvich and Tsay 1989). The actual goodness-of-fit is expressed in terms of the classical coefficient of multiple determination or  $R^2$  statistic used in regression analysis as a measure of explained variance.
5. The pre-1931 data were reserved for statistical validation tests of the response function estimates. The validation tests used were the square of the Pearson correlation coefficient (RSQ), the reduction of error (RE), and coefficient of efficiency (CE) (Cook et al. 1999).

The correlation function (Fig. 4.3a) reveals that only 4 out of the 32 monthly correlations have exceeded the two-tailed a priori 95% confidence limits (prior-July, prior-September, and current June precipitation, and current May temperature), while only one month (current June precipitation) exceeds the a posteriori limit. Does this result mean that only current June precipitation truly matters to these trees? Based on the most rigorous statistical considerations described above, the answer would appear to be ‘Yes.’ However, there may be much more physiological meaning in the structure of the correlation function than purely statistical considerations would suggest, because the response function (Fig. 4.3b), which was explicitly estimated from the matrix of these monthly correlations, has many more ‘significant’ coefficients, 24 to be exact. This result in itself is tricky to interpret because of the way the response function and its confidence limits are calculated as a regression-weighted, linear combination of fitted climate eigenvectors (Fritts 1976; Guiot et al. 1982; Fekedulegn et al. 2002). Indeed, even the estimation of the response function confidence limits has been somewhat controversial. There is good reason to believe that the original method of Fritts (1976) produced confidence limits that were a bit too narrow. This problem has been rectified in two different ways: (1) by using bootstrap resampling to generate empirical confidence intervals (Guiot

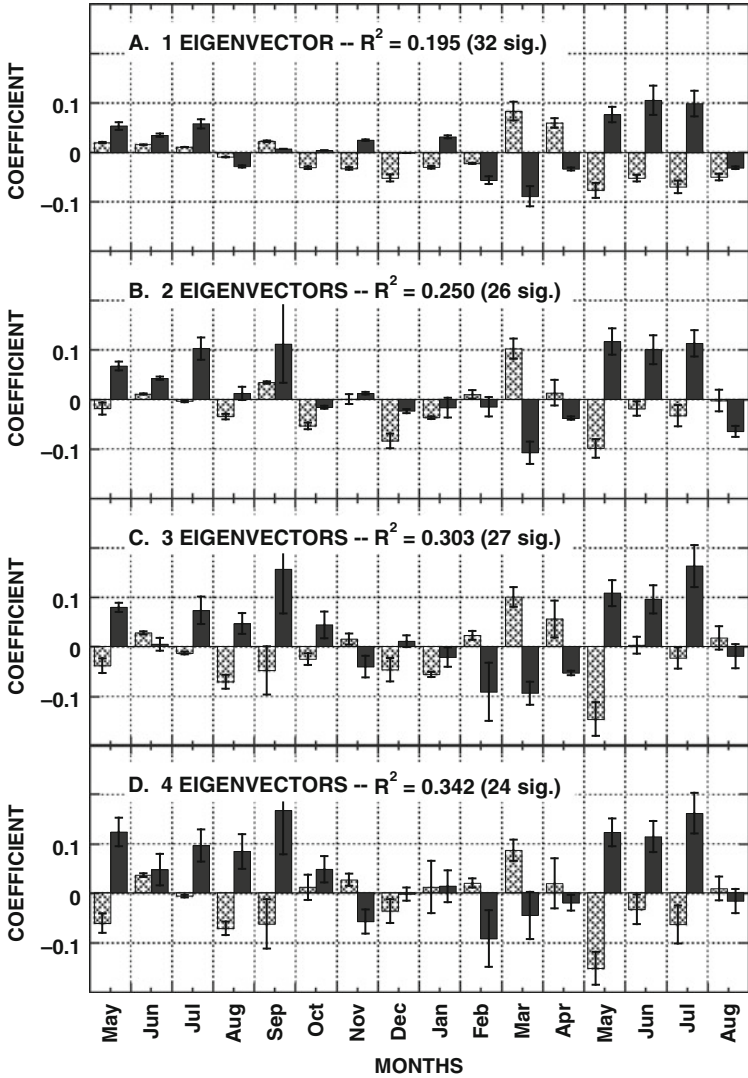
1991; Biondi and Waikul 2004) and (2) by using a  $t$ -test derived from principal components regression theory (Fekedulegn et al. 2002). We use the  $t$ -test described by Fekedulegn et al. (2002) here.

The number of climate eigenvectors included in the response function (four in our example) also affects the number of resulting ‘significant’ variables. This effect is illustrated in the four steps used in calculating the final response function. The Step-1 response function (Fig. 4.4a) is based on only one eigenvector that explains 19.5% of the variance. Yet, all 32 monthly variables are ‘significant’ based on the calculated 95% confidence limits. Step-2 cumulatively explains 25% of the variance, but now has only 26 ‘significant’ variables. Step-3 cumulatively explains 30.3% of the variance and has 27 ‘significant’ variables. Final Step-4 cumulatively explains 34.2% of the variance and now has only 24 significant variables. So, from Step-1 to Step-4, the explained variance increases by 75%, but the number of significant variables decreases by 25%. A similar result can be found in Fritts (1976, Fig. 7.13, p. 367). This tendency for an inverse dependence between the number of eigenvectors in a response function model and the resulting number of ‘significant’ monthly variables means that one must be careful about interpreting response functions based on few climate eigenvectors because the number of ‘significant’ months may be inflated; this will be especially the case when only one climate eigenvector is used.

Using simulation methods, Cropper (1982b) illustrated how the number of ‘significant’ months can be inflated in response functions. This inflation occurs in part because the monthly patterns of the climate eigenvectors are determined by the intercorrelations between the monthly variables themselves. Some of these climate intercorrelations will be based on true physical associations (e.g., temperature may be inversely correlated with precipitation for a given month or season). However, some of the intercorrelations will also be unique to the analysis period or occur by chance alone, and thus have little or no true physical meaning. Yet, they will show up in the monthly patterns of the climate eigenvectors and may be carried into the response function when they are part of the climate eigenvectors that best explain tree growth. We must always keep in mind that the climate eigenvectors are mathematically defined orthogonal modes that are not constrained to have any physical meaning whatsoever (although they will often have some in practice) and certainly do not have *any* inherent biological meaning. So the number of statistically significant months in a response function is not a good diagnostic for determining the successful application of the method. Response function monthly confidence intervals also do not take into account the multiplicity problem described earlier. Gray et al. (1981) tackled this problem by using the binomial distribution to determine the minimum number of significant ( $p < 0.05$ ) months needed for a response function to have overall significance ( $P < 0.05$ ). For response functions based on 32 monthly coefficients (Fig. 4.3b), one needs a minimum of four to five significant coefficients ( $p < 0.05$ ) to claim that the response function is significant (overall response function  $P < 0.05$ ), which in our example is the case. Our correlation function would barely pass this test, however.

This finding brings up another issue that can greatly affect the estimation of response functions. Fekedulegn et al. (2002) note that response function analysis





**Fig. 4.4** The stepwise development of the hemlock response function shown in Fig. 4.3b. The progressive changes in fractional variance explained ( $R^2$ ) and the number of significant coefficients are indicated

is highly sensitive to (1) the number of eigenvectors retained as candidate predictors in regression analysis and (2) the criterion for entering eigenvectors into the regression model. Fritts (1976), Guiot et al. (1982), and Fekedulegn et al. (2002) argue for retaining a large number of candidate eigenvectors, ones that may explain up to 90–95% of the total variance in the climate data correlation matrix. An objective



way for doing so is to use the cumulative eigenvalues product (PVP in French) criterion (Guiot et al. 1982), which is based on the cumulative product of the eigenvalues, one way of calculating the determinant of a matrix. The actual cutoff occurs where PVP drops below 1.0, the value of the determinant of a random correlation matrix of equal rank. In contrast, the EV1 cutoff occurs at the point where the remaining eigenvectors cannot explain as much variance as those extracted from an equal-rank random correlation matrix. Both PVP and EV1 are based on appealing asymptotic arguments that relate to expected values of random correlation matrices. Yet, they yield quite different results in practice.

In our example, the EV1 cutoff retained 13 of 32 eigenvectors as candidate predictors, which cumulatively explained 73.4% of the total climate variance. If the PVP cutoff had been used, it would have retained 25 of the 32 eigenvectors (equal to 96.5% of the total variance) as candidate predictors. So, is one cutoff better than the other? Support for EV1 comes from the fact that Monte Carlo estimates of eigenvalue confidence limits based on ‘Rule N’ (Preisendorfer et al. 1981; Preisendorfer 1988) always select a cutoff that is consistent with the asymptotic argument underlying EV1. On the other hand, there may be some useful climate information in the deleted eigenvectors below the EV1 cutoff, which is the principal argument for using PVP (or some other criterion) to retain more candidate eigenvectors (cf. Jolliffe 1973). This argument is appealing because the total information in the original climate correlation matrix is contained in the complete set of eigenvectors. See Fritts (1976, p. 357) for this mathematical equivalence. However, the increasing orthogonality constraints imposed on the higher-order eigenvectors are likely to distort the climatic meaning of those modes increasingly away from physical reality, which may make them more sensitive to chance correlations within the original intercorrelation matrix of climate variables. This and our ‘Rule N’ argument above are the reasons why we prefer the EV1 cutoff.

The choice of EV1 or PVP can also be argued in terms of Type-1 and Type-2 errors. Choosing EV1 is a more conservative choice because it reduces the number of candidate predictors and the likely inflation of  $R^2$  (Rencher and Pun 1980). However, the premium paid for protecting against inflated  $R^2$  using EV1 is the possible loss of additional useful climate information in the higher-order climate eigenvectors that PVP would retain. Thus, EV1 reduces the chance of Type-1 error and increases the chance of a Type-2 error in response functions by eliminating more eigenvectors from the candidate predictor pool. In contrast, PVP increases the chance of Type-1 error and decreases the chance of Type-2 error by allowing more potentially spurious candidate eigenvectors to be included in the response function.

Given the selected cutoff used to retain candidate eigenvectors, the criterion for entering the eigenvectors into the regression-based response function model ultimately determines the final form of the response function. Fritts (1976), Guiot et al. (1982), and Fekedulegn et al. (2002) argue for somewhat lenient entry criteria. Fritts (1976) used an  $F$ -level = 1.0 cutoff for entering eigenvectors, which in our example has a probability  $p = 0.30$ . Guiot et al. (1982) suggested  $p = 0.50$  for relatively short datasets ( $n = 30$  years) when the number of variables is comparable to the number of observations and 0.10–0.20 for longer datasets like that used here ( $n =$

66). Fekedulegn et al. (2002) suggested using  $p = 0.15$  for entering eigenvectors. Here, we argue for the use of the minimum Akaike information criterion (Akaike 1974), with a correction for small sample bias (Hurvich and Tsay 1989). This is a totally objective way of determining the order of the model, one that is based on sound and well-tested information theoretic principles. For this reason, we prefer the minimum AIC cutoff. Interestingly, the original uncorrected AIC (Akaike 1974) also allows variables to enter into a model if  $p < 0.15$  (Jones 1985), the same as the variable entry criterion of Fekedulegn et al. (2002). So the use of the AIC should not change the response function results nearly as much as the chosen eigenvector cutoff.

Table 4.1 provides more detail on the four steps used in calculating our hemlock response function. Eigenvectors #3, #9, #10, and #6 were added to the model in that order, with a cumulative explained variance of 34.2% and a minimum AIC of  $-16.63$ . Perhaps the most interesting result here is what is not included in the model; i.e., climate eigenvectors #1 and #2, which together explain 19.2% of the total variance in the correlation matrix of temperature and precipitation at Mohonk Lake. Even though those eigenvectors are the two most important modes of covariance among the monthly temperature and precipitation variables, they correlate extremely poorly with hemlock growth ( $r = -0.01$  and  $0.06$ , respectively). This result illustrates another aspect of response function analysis. The most important modes of monthly climate variability defined by the eigenvectors need not relate to the needs of tree growth. Thus, it is always dangerous to impose a priori expectations on how trees respond to climate. We should allow the climate response of a tree-ring series to objectively *emerge* from our analysis.

The bottom two rows of Table 4.1 provide additional steps in the response function model that would have occurred if the PVP criterion were used as the cutoff for candidate eigenvectors instead of EV1. As was mentioned earlier, PVP retains 25 of the 32 eigenvectors (96.5% of the total variance) as candidate predictors. In this case, eigenvectors #15 and #18 would have also been added to the model, with a minimum AIC now at  $-19.37$ . This result suggests that PVP is better than EV1 because the final model with six eigenvectors has a smaller AIC. Is this true? In this

**Table 4.1**    The four regression steps used in calculating the hemlock response function

STEP	EIG	CORR	T-STAT	PROB	PART $R^2$	$R^2$	AIC
1	3 (7.2)	0.441	3.931	0.0003	0.195	0.195	$-10.09$
2	9 (4.6)	0.235	1.932	0.0549	0.055	0.250	$-12.57$
3	10 (4.1)	$-0.232$	$-1.907$	0.0580	0.054	0.303	$-15.20$
4	6 (5.2)	0.197	1.605	0.1093	0.039	0.342	$-16.63$
5	15 (2.8)	0.197	1.605	0.1094	0.039	0.381	$-18.21$
6	18 (2.1)	0.183	1.490	0.1371	0.034	0.413	$-19.37$

STEP = response function step; EIG = the eigenvector number entered and its (percent variance); CORR = simple correlation of the eigenvector with tree rings; T-STAT = Student's  $t$  statistic for Corr; PROB = Probability of  $t$ -stat; PART  $R^2$  = partial  $R^2$  or fractional variance contributed by each step;  $R^2$  = cumulative fractional variance of the model; AIC = Akaike information criterion.

case at least, the answer appears to be ‘No,’ but how do we know? Recall that we withheld the 1897–1930 data from response function estimation for model validation purposes. We can use these data now to objectively test each response function model for skill prior to the 1931–1996 calibration period. For comparison, we have also done this using only the significant climate months in the correlation function as predictors: the single month that exceeds the a posteriori 95% significance level (current June precipitation) and the four months that exceed the a priori 95% significance level (prior-July, prior-September, and current June precipitation, and current May temperature). The results of these runs are shown in Table 4.2.

Note that the calibration period  $R^2$  increases as the number of predictors in the model increases. This is expected even by chance alone (Morrison 1976). In particular, there is a big jump in  $R^2$  from one predictor (current June precipitation) to four predictors (prior-July, prior-September, and current June precipitation, and current May temperature) and a corresponding substantial reduction in the AIC. This jump in  $R^2$  is also strongly maintained in the form of increased skill in the verification period of the four-predictor model; i.e., the RSQ, RE, and CE statistics are all substantially higher for the four-predictor model compared to that based on only one predictor. So the four monthly variables selected by a priori testing of the monthly correlations are collectively more important to eastern hemlock growth than the one variable selected by the stringent a posteriori test. In this case anyway, the multiplicity problem discussed earlier is not a problem at all!

The results of the full response function tests based on using all 32 monthly climate variables in the eigenanalysis, and using either the EV1 or PVP cutoffs for retaining candidate eigenvectors, are even more interesting. As we saw previously, the response function based on the candidate predictor pool selected by PVP resulted in a higher  $R^2$  and smaller AIC than those selected by EV1. Yet, the verification statistics are somewhat better for the full response function based on the EV1 pool (especially for RE and CE). This result illustrates that while the AIC is useful for

**Table 4.2** Comparisons of climate models used to estimate the Mohonk Lake hemlock chronology

# PREDICTORS	Calibration period statistics			Verification period statistics		
	NEIG	$R^2$	AIC	RSQ	RE	CE
1	1/1	0.209	−11.27	0.094	0.039	0.039
4	2/2	0.343	−21.34	0.226	0.173	0.173
32	EV1 13/4	0.342	−16.63	0.269	0.251	0.251
32	PVP 25/6	0.413	−19.37	0.248	0.111	0.110

The calibration period is 1931–1996 and the verification period is 1897–1930.

# PREDICTORS = number of monthly climate variables used in each principal components regression model; NEIG = the number of candidate climate eigenvectors/the number of climate eigenvectors entered into the model;  $R^2$  = cumulative fractional variance of the model; AIC = Akaike information criterion; RSQ = square Pearson correlation; RE = verification reduction of error; CE = verification coefficient of efficiency. Higher RSQ, RE, and CE mean better verification of the fitted model.

selecting the best subset model within a given pool of candidate predictors, it does not mean that the minimum AIC model among all competing models from differing candidate predictor pools will also select the model that verifies best. How one chooses the candidate predictor pool for response functions matters, as Fekedulegn et al. (2002) point out. This example also supports our contention that higher-order eigenvectors may be less reliable predictors of tree growth because of increasing orthogonality constraints on those extracted modes. But we have also found that response functions of other tree species sampled near Mohonk Lake sometimes verified better using PVP (see below), so our case for using EV1 is certainly not closed. More provocatively, the overall monthly *structure* (i.e., month-to-month evolution) of the correlation function appears to be physiologically meaningful to these trees, even when only 4 out of 32 variables pass the a priori 95% significance level. How might this be?

We argue that the answer lies in the difference between how the monthly correlations are treated statistically here versus how monthly climate actually affects tree growth. The monthly correlations are tested in the correlation function as if adjacent months of climate are completely independent of each other. Yet, the positive and negative correlations between tree rings and precipitation/temperature during the prior and current growing season months (Fig. 4.3a) almost certainly reflect the sensitivity of our eastern hemlock chronology to overall changes in growing season moisture supply and evapotranspiration demand that may span adjacent months. Response functions attempt to address this possible interaction between months both within and between climate variables by exploiting the eigenstructure of the climate correlation matrix, apparently with some success here. This effect is best revealed by comparing the current growing season May–June–July correlations in Fig. 4.3a with the coefficients of the response function for the same months based on the single most important climate eigenvector in Fig. 4.4a. The response function coefficients are more uniformly positive for precipitation and negative for temperature over those growing season months. This result implies that the simple correlations are not measuring the strength of the tree growth response to climate as fully as the response function. The significant verification statistics of the final response function model support this conclusion (Table 4.2; EV1 results). Thus, the appearance of more statistically significant response function coefficients during the growing season months (Fig. 4.3b) does not appear to be a statistical artifact of the method in this example.

## 4.6 Response Functions and Empirical Signal Strength

As was noted earlier, our successful demonstration of the EV1 cutoff does not necessarily mean that it will always perform better than the PVP cutoff. To demonstrate this, we have calculated response functions for a total of seven tree species, all located near Mohonk Lake and its cooperative weather station. The tree species are eastern hemlock (*Tsuga canadensis*; TSCA), pitch pine (*Pinus rigida*; PIRI), chestnut oak (*Quercus prinus*; QUPR), black oak (*Quercus velutina*; QUVE), pignut

hickory (*Carya glabra*; CAGL), tulip poplar (*Liriodendron tulipifera*; LITU), and black birch (*Betula lenta*; BELE). This selection covers a diverse range of conifer and deciduous hardwood tree genera and subgenera. The tree-ring chronologies are based on crossdated ring width series from 6 to 14 trees/12–20 cores per species (Table 4.3). In addition, 6–10 paired cores per tree were included to allow for the calculation of within-tree correlations. The total sample size per chronology is admittedly somewhat modest (especially for PIRI), but the selected samples allowed for a consistent assessment of chronology signal strength over the same 1931–1996 period used for climate calibration. This enabled us to directly compare the empirical signal strength statistics with the response function results. For comparative purposes, all ring width series were also standardized with fixed 50-year smoothing splines (Cook and Peters 1981). In addition, the response functions and signal strength statistics were calculated from the tree-ring chronologies after removal of autocorrelation based on best-fit, low-order autoregressive models (the residual chronology from program ARSTAN; Cook 1985). Doing so eliminated different levels of autocorrelation in the original chronologies that might obscure the interspecies comparisons presented here.

Table 4.3 provides a suite of 10 descriptive statistics for each tree-ring chronology. MS is mean sensitivity; SD is standard deviation; R1 is first-order autocorrelation; ESR is Edmund Schulman's *R*; RTOT is the average correlation between all series, including within-tree replicate cores; RWT is the average correlation of the within-tree replicate cores; RBT is the average correlation of only the between-tree cores; REFF is a weighted average correlation based on RWT and RBT; EPS is the expressed population signal; and SNR is the signal-to-noise ratio. Refer to our earlier discussion of these statistics and also Briffa and Jones (1990). Explicit mathematical definitions of these statistics are also included in an Appendix of this chapter for those who are unfamiliar with them.

MS, SD, R1, and the correlation-based signal strength statistics for our tree species all fall in the typical range occupied by tree-ring data from eastern North America (ENA) (cf. Fritts and Shatz 1975; DeWitt and Ames 1978). For example, our average MS is 0.166, which compares well with 0.175 for ENA tree-ring chronologies in DeWitt and Ames (1978, their Table II, p. 10). In contrast, average MS for western North America chronologies is 0.365 in DeWitt and Ames (1978) and 0.390 in Fritts and Shatz (1975). Estimates of common signal strength for ENA chronologies (mean %Y = 0.289) in DeWitt and Ames (1978) are much lower than those presented here (mean RBT = 0.460 and REFF = 0.527), but this is an unfair comparison. Both %Y and the RBAR-based signal strength statistics (RBT, RWT, REFF) are highly sensitive to how the tree-ring data have been processed. Estimating signal strength from prewhitened tree-ring data, as was done here, will in general increase those statistics. The method of detrending will also have an effect. In contrast, MS is relatively insensitive to how the tree-ring data are processed; e.g., MS of standardized (detrended) tree-ring indices is the same as that of the original ring width measurements. The MS statistics in Table 4.3 also show the danger in using this statistic as a direct measure of climate sensitivity. All of our tree species have been subjected to the same macroclimatic effects and most are

**Table 4.3** Chronology descriptive statistics and measures of empirical signal strength

SPECIES	T/C/P	MS	SD	R1	ESR	RTOT	RWT	RBT	REFF	EPS	SNR
TSCA	10/16/6	0.161	0.165	0.334	0.671	0.439	0.623	0.429	0.484	0.926	12.509
PIRI	6/12/6	0.245	0.248	0.316	0.605	0.522	0.718	0.502	0.569	0.923	12.016
QUPR	9/18/9	0.107	0.152	0.520	0.633	0.544	0.687	0.535	0.634	0.955	21.441
QUVE	14/20/6	0.106	0.131	0.511	0.716	0.439	0.723	0.429	0.456	0.940	15.626
CAGL	11/20/9	0.152	0.156	0.275	0.664	0.433	0.727	0.418	0.470	0.938	15.250
LITU	10/20/10	0.176	0.223	0.569	0.759	0.503	0.695	0.492	0.581	0.953	20.248
BELE	11/20/9	0.214	0.252	0.603	0.606	0.422	0.589	0.414	0.497	0.936	14.613

See the text and the Appendix in the chapter for details.

Column headings: SPECIES = tree species used: eastern hemlock (TSCA), pitch pine (PIRI), chestnut oak (QUPR), black oak (QUVE), pignut hickory (CAGL), tulip poplar (LITU), black birch (BELE); T/C/P = number of trees/number of cores/number of trees with paired cores; MS = mean sensitivity; SD = standard deviation; R1 = first-order autocorrelation; ESR = Edmund Schulman's *R*; RTOT = average correlation between all series including within-tree replicate cores; RWT = average correlation of the within-tree replicate cores; RBT = average correlation of only the between-tree cores; REFF = weighted average correlation based on RWT and RBT; EPS = expressed population signal; SNR = signal-to-noise ratio.

growing on well-drained to xeric sites. The greatest exception is the tulip poplar site growing along a stream bottom. Yet, the range of MS is quite large, with PIRI and BELE having more than twice the MS of QUPR and QUVE. This result suggests that the expressible range of MS in trees growing in a common macroclimatic environment can be highly species dependent, which complicates the use of MS as a general measure of climate sensitivity.

Some signal strength statistics have been used in the past as qualitative predictors of climate sensitivity of tree-ring chronologies (e.g., MS and RBT). Since they are based on the same prewhitened tree-ring data as the response functions, the signal strength statistics can be tested as predictors of the response function modeling results (cf. Cropper 1982a). This was done for response functions based on the EV1 and PVP eigenvector cutoffs, with model selection determined by the minimum AIC. Those results are provided in Table 4.4. With respect to calibration  $R^2$ , all of the response functions calibrated a significant amount of tree-ring variance, with PVP always outperforming EV1 because the former resulted in the entry of more model predictors. This difference ranges from one to five additional predictors and 0.037 to 0.164 in additional fractional variance explained. The verification statistics (RSQ, RE, CE) tell a more mixed story. As before, the hemlock (TSCA) response function verifies strongly for all three statistics, with EV1 verifying somewhat better than PVP. The next best result is for chestnut oak (QUPR), which has a statistically significant verification RSQ ( $p < 0.01$ ) for the PVP model. Tulip poplar (LITU) and black birch (BELE) also have verification RSQs that are weakly significant ( $p < 0.10$ ), again for the PVP model, and pignut hickory (CAGL) performs slightly better with EV1. None of the other species/models verify in any useful way for either EV1 or PVP, although the RE and CE tend to be less negative for the EV1 models. Taken together, these results marginally support PVP over EV1 as an eigenvector cutoff criterion, but the difference is not large.

Given the way that certain tree-ring statistics have often been used as predictors of climate sensitivity (e.g., MS; Fritts and Shatz 1975), we have tested that capacity using the EV1 and PVP response function modeling results. The statistics used as predictors of response function  $R^2$  are MS, REFF, SNR, and ESR in Table 4.3. The other signal strength statistics are either biased by the high within-tree core correlations (RTOT) or are absorbed in the estimate of REFF (RWT, RBT). EPS also includes REFF in its estimate and asymptotes quickly towards 1.0, making it not very sensitive for our tests. These comparisons assume that there are no ‘species effects’ in our results; i.e., the joint distributions of our predictors of climate sensitivity with the response function results are independent of the species being tested. This is unlikely to be the case here because of our diverse taxa, but without within-species replication we have no way of directly testing for ‘species effects.’ Also, we have only seven cases to test (5 degrees of freedom), so no claims of statistical significance will be made. However, the results are interesting enough to warrant additional study using many more within- and between-species tests.

Figure 4.5 shows the four x-y scatterplots with fitted bivariate regression curves and simple correlations. REFF correlates with  $R^2$  (Fig. 4.5a) at levels that suggest a predictive relationship between them:  $r$  equals 0.644 for EV1 and 0.548 for PVP.

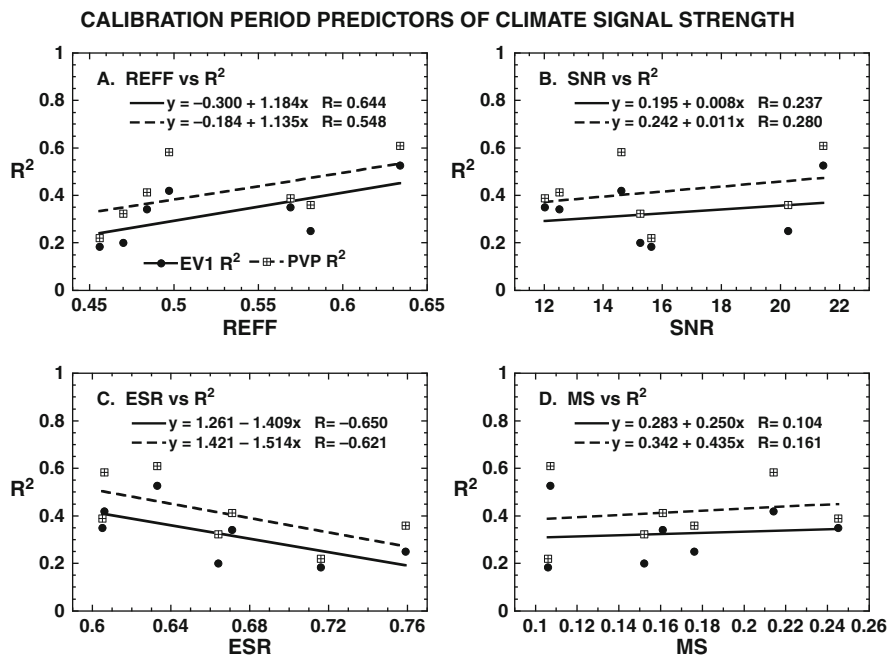


Table 4.4 Response function modeling statistics for seven tree species growing near Mohonk Lake

SPECIES	CUTOFF	NEIG	EIGENVECTORS	R <sup>2</sup>	AIC	RSQ	RE	CE
TSCA	EV1	4	<b>3,9,10,6</b>	0.342	-16.63	0.269	0.251	0.251
	PVP	6	<b>3,9,10,6,15,18</b>	0.413	-19.37	0.248	0.111	0.110
PIRI	EV1	4	<b>3,11,8,10</b>	0.350	-17.44	0.001	-0.349	-0.349
	PVP	5	<b>3,11,19,8,10</b>	0.389	-19.05	0.002	-0.403	-0.403
QUPR	EV1	8	<b>6,3,10,11,13,9,7,8</b>	0.527	-28.16	0.062	-0.152	-0.154
	PVP	11	<b>6,3,10,11,13,9,15,7,8,20,19</b>	0.610	-32.22	0.159	-0.079	-0.081
QUVE	EV1	3	<b>3,6,10</b>	0.184	-4.75	0.003	-0.114	-0.114
	PVP	4	<b>3,6,10,14</b>	0.221	-5.51	0.008	-0.114	-0.114
CAGL	EV1	2	<b>3,6</b>	0.200	-8.36	0.040	-0.042	-0.043
	PVP	5	<b>3,6,21,14,25</b>	0.323	-12.35	0.020	-0.345	-0.356
LITU	EV1	3	<b>3,6,10</b>	0.250	-10.37	0.038	-0.073	-0.082
	PVP	5	<b>3,25,15,6,10</b>	0.360	-16.06	0.064	-0.355	-0.367
BELE	EV1	4	<b>3,6,9,10</b>	0.420	-24.93	0.005	-0.249	-0.262
	PVP	9	<b>3,6,21,9,10,17,22,5,13</b>	0.584	-33.95	0.051	-0.493	-0.508

The tree species are eastern hemlock (*Tsuga canadensis*; TSCA), pitch pine (*Pinus rigida*; PIRI), chestnut oak (*Quercus prinus*; QUPR), black oak (*Quercus velutina*; QUVE), pignut hickory (*Carya glabra*; CAGL), tulip poplar (*Liriodendron tulipifera*; LITU), and black birch (*Betula lenta*; BELE). The three most important eigenvectors—#3, #6, and #10—are highlighted as bold numbers.

SPECIES = tree species abbreviations; CUTOFF = eigenvalue cutoff used to retain candidate eigenvectors; NEIG = number of eigenvectors entered into each response function model. EIGENVECTORS = list of entered eigenvectors in their order of entry; R<sup>2</sup> = cumulative fractional variance of the model; AIC = Akaike information criterion; RSQ = square Pearson correlation; RE = verification reduction of error; CE = verification coefficient of efficiency. Higher RSQ, RE, and CE mean better verification of the fitted model.



**Fig. 4.5** Relationships between four empirical measures of signal strength (REFF, SNR, ESR, MS) and explained variance ( $R^2$ ) by response functions based on the EV1 and PVP eigenvector cutoff criteria. The signal strength and  $R^2$  statistics are described in the text and their values come from Tables 4.3 and 4.4

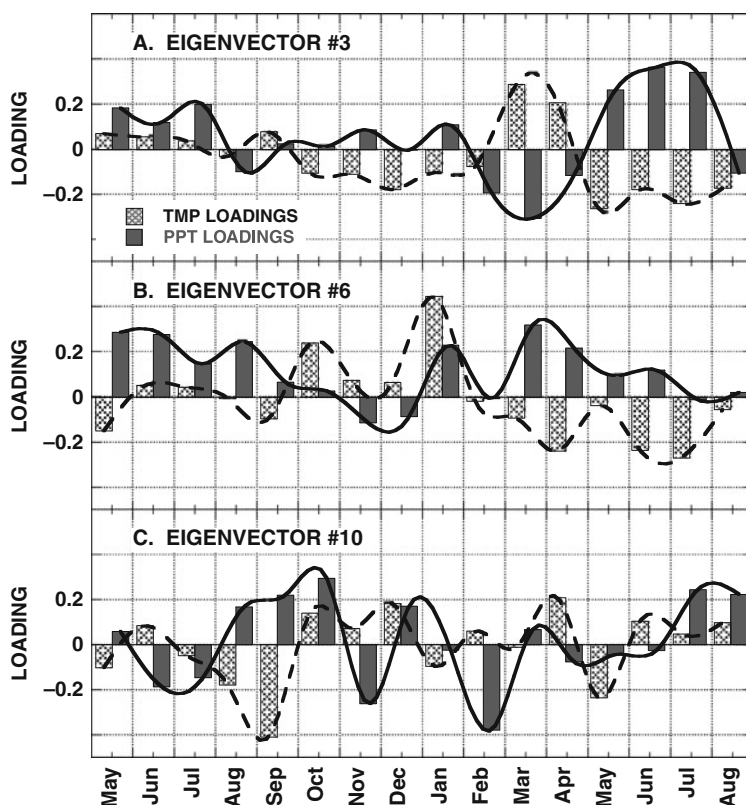
That relationship weakens considerably when SNR is used to predict  $R^2$  (Fig. 4.5b), probably because of the strongly nonlinear behavior of SNR. Interestingly, ESR also correlates highly with  $R^2$  (Fig. 4.5c), but in an inverse sense:  $r$  equals  $-0.650$  for EV1 and  $-0.621$  for PVP. As far as we know, ESR has never been used before as a diagnostic of chronology signal strength since Schulman (1956) first described it, but intuitively we would have expected positive correlations like those between REFF and  $R^2$ . If this inverse relationship between ESR and  $R^2$  were to hold up, it would suggest that loss of MS in the mean-value function is good. This is counterintuitive if we accept the premise that high common MS among trees in the chronology is a true measure of signal strength. Thus, the ESR result could be reflecting a ‘species effect’ here, because MS varies considerably between the tree species used: 0.106–0.245 (Table 4.3). Interestingly, the weakest relationship of all is between chronology MS and  $R^2$  (Fig. 4.5d):  $r$  equals 0.104 for EV1 and 0.161 for PVP. Again, we could have a ‘species effect’ here that weakens this relationship because of differences in how conifers and hardwoods produce secondary growth in the form of annual tree rings. In any case, the results presented here provide limited support for REFF as a predictor of climate sensitivity. This result supports the findings of Cropper (1982a), but many more test cases must be studied to determine just how useful it really is for that purpose.

## 4.7 Additional Response Function Interpretations

Table 4.4 lists the actual eigenvectors used in each response function in order of entry into the models. Given the disappointing verification results, these tree species have still clearly ‘voted’ for the most important cross-taxa climate eigenvectors. For six of the seven species, eigenvector #3 enters first; for the remaining species it enters second (refer to Table 4.4). The second most important is climate eigenvector #6, which enters either first or second in five of seven species models, all deciduous hardwoods (refer to Table 4.4). The third most important is climate eigenvector #10, which often enters third or fourth into the models (refer to Table 4.4). After that, the selected eigenvectors vary much more between species and models.

We stated earlier that the more important climate eigenvectors probably have some physical meaning due to the intercorrelations between monthly temperature and precipitation, but they are not constrained to have any biological meaning at all. While this is true, the results of our response function analyses also indicate that eigenvectors #3, #6, and #10, at least, are likely to have significant biological meaning to the trees, given their associations with ring width. For this reason, we will examine these climate eigenvectors for some biological meaning.

Climate eigenvector #3 (Fig. 4.6a) reveals an oppositional pattern between monthly temperature and precipitation that is strongest during the May–July current growing season months. This pattern is probably physically based in the sense that higher rainfall in the warm-season months should result in lower temperatures due to increased cloudiness and vice versa. It also makes biological sense as a predictor of tree growth, because our trees are growing on well-drained sites where they should have a natural sensitivity to moisture supply and evapotranspiration demand during the growing season. Thus, above-average precipitation and below-average temperature should jointly contribute to above-average radial growth, and this model applies more or less equally well to all seven tree species. Interestingly, there is a reversal in the precipitation/temperature association during the preceding March–April months that presages the start of the radial growth season. This observation indicates that warm/dry spring conditions tend to precede cool/wet late-spring/summer conditions at Mohonk Lake and vice versa. The physical meaning of this spring–summer pattern is unclear and its biological significance is even less certain. The March–April pattern could be totally unrelated to actual tree growth and simply carried into the response function because of its association with the more biologically meaningful May–July pattern in eigenvector #3. Conversely, warm/dry March–April conditions could also help in initiating early physiological activity in the trees prior to the radial growth season; i.e., it may have some phenological significance. The fact that the March–April pattern greatly diminishes in the final hemlock response function (cf. Fig. 4.4a,d) supports the former argument, but the highly significant positive March temperature response is still probably real for this species (Cook and Cole 1991). In contrast, the May–July pattern remains largely intact through all four steps of the response function calculation. Prior to March, the eigenvector loadings are uniformly lower, which means that they are more likely to be there by chance alone.



**Fig. 4.6** Climate eigenvectors Nos. 3, 6, and 10. These are shown to describe the modes of climate that are most important in estimating the response functions of the seven tree species from near Mohonk Lake. A cubic spline has been applied to each eigenvector, separately for the precipitation (*solid bars*) and temperature (*hatched bars*) loadings, in order to highlight the more general structure of each eigenvector. On-line version shows this figure in color

Climate eigenvector #6 (Fig. 4.6b) is most strongly associated with the radial growth of the deciduous hardwoods. It has a significantly different pattern of monthly climate loadings that emphasizes a tendency for both prior summer and current spring precipitation to be jointly above or below average. It is unclear how much of this structure is due to orthogonality constraints alone, but it appears to be biologically meaningful nonetheless to the tree species being tested here. The added importance of prior growing season rainfall to radial growth the following year occurs even when autoregressive persistence has been removed from the tree-ring chronologies (see Fig. 4.3b for hemlock). Thus, this carryover effect of climate on ring width, from one year to the next, is not related to autocorrelation per se. Rather, it appears to be related to a discrete physiological event in a given year, like foliar budset, whose effect on radial growth can be delayed one or more years. Nonstructural carbon, which is an important component of bud construction and early leaf development, may also play a role here, as it has a residence time up to

3–5 years in oaks (Trumbore et al. 2002). The relative importance of this climate eigenvector in the response functions may also reflect some phylogenetic differences in the climate response of hardwoods and conifers. For example, the implied greater dependence on above-average precipitation during the March–April months may be related to critical early springwood vessel enlargement, particularly in the *Quercus* species. The other interesting feature of eigenvector #6 is the prominent positive loading in January. This phenomenon may be associated with the winter temperature sensitivity of certain tree species described by Pederson et al. (2004).

Climate eigenvector #10 (Fig. 4.6c) is more complicated (perhaps due to orthogonality constraints) and, therefore, more difficult to interpret from a tree physiological perspective. However, its presence in most of the response functions indicates that it does have some true biological meaning. It mainly emphasizes monthly climate variability during the late prior growing season, mainly during the prior August–October months. As a predictor of tree growth, it also appears to be more important for the hardwoods.

This evaluation of the three most important climate eigenvectors has revealed a well-understood dependence between tree growth and moisture supply/evapotranspiration demand during the current growing season of trees growing on well-drained sites, especially eigenvector #3. This common signal among all seven tree species is why they all crossdate significantly; from COFECHA (Holmes 1983), the mean correlation between all series is 0.55 (range: 0.30–0.69). Therefore, a forward model of cambial growth parameterized to model this basic water relations' effect on tree growth should produce useful first-order estimates of these tree-ring series (see Vaganov et al. Chapter 3, this volume). Details at the genus and species level might be missed by the forward model, however, unless additional climate response information suggested by eigenvector #6 (and perhaps eigenvector #10) is part of the model. Interestingly, the most poorly correlated series among the seven is pitch pine (PIRI), whose response function does not include eigenvector #6. The lack of this signal, common to all other species, may help explain why this pitch pine chronology crossdates relatively weakly with the others ( $r = 0.30$ ).

## 4.8 Some Implications for Climate Reconstruction

The response function model results indicate that the seven tree species tested have statistically significant climate information in them, but with varying degrees of strength and fidelity. Based on the examinations of the most important climate eigenvectors entered into the response function models, it appears that the dominant common signal among all species is a May–July growing season response to above-average precipitation and below-average temperature. Climate eigenvectors #3 and #6 also indicate that this response occurs in both the current and prior growing seasons. Consequently, it should be possible to exploit both years of climate information in the tree rings for reconstructing past growing season climate. We have investigated this potential by reconstructing May–July total precipitation at Mohonk Lake using the mutual climate information in all seven tree-ring series. We

also did this for each species separately to see how much might be gained by using a multi-species assemblage.

Using the same principal components regression program that calculated the response functions, we reversed the order of dependence and used the seven tree-ring chronologies as predictors of May–July total precipitation. The tree-ring chronologies were lagged on themselves by one year ( $t$  and  $t+1$  for each series) to allow for the carryover of the climate signal from one year to the next, thus creating a matrix of 14 candidate predictors. Lagging the tree-ring data resulted in the loss of one year for calibration (now 1931–1995), and the use of the now unlagged climate data resulted in a gain of one year for verification (now 1896–1930). The 14 candidate predictors were next screened for simple correlation with May–July precipitation over the calibration period. Those that did not correlate significantly ( $p < 0.05$ ; one-tailed test) were removed from this candidate predictor pool. Of the 14  $t$ ,  $t+1$  candidates, only 3 were rejected: PIRI  $t+1$ , QUPR  $t+1$ , and QUVE  $t+1$ . Those that passed the screening had correlations with May–July precipitation ranging from 0.272 to 0.468. The retained 11 candidate predictors were subjected to principal components analysis and the EV1 cutoff was used to retain the candidate eigenvectors. The EV1 cutoff retained 3 of 11 eigenvectors, which accounted for 65.5% of the total variance. These were used as candidate predictors in regression analysis using the minimum AIC for selecting the best-fit model. See Table 4.5 for details. The result was the entry of the first two tree-ring eigenvectors, which is consistent with what we found earlier in our evaluation of the most important climate eigenvectors associated with tree growth. The reconstruction model

**Table 4.5** Comparisons of May–July total precipitation reconstructions at Mohonk Lake

PREDICTORS	Calibration period statistics			Verification period statistics		
	NEIG	$R^2$	AIC	RSQ	RE	CE
ALL (11)	3/2	0.436	−30.83	0.126*	−0.047	−0.112
TSCA (2, $t$ , $t+1$ )	2/1	0.266	−15.91	0.098*	0.040*	0.020*
PIRI (1, $t$ )	1/1	0.213	−11.36	0.072*	−0.033	−0.097
QUPR (1, $t$ )	1/1	0.195	−9.88	0.257*	0.256*	0.244*
QUVE (1, $t$ )	1/1	0.128	−4.72	0.051	0.014*	−0.002
CAGL (2, $t$ , $t+1$ )	2/1	0.284	−17.48	0.019	−0.137	−0.155
LITU (2, $t$ , $t+1$ )	2/1	0.237	−13.42	0.069*	0.008*	−0.008
BELE (2, $t$ , $t+1$ )	2/1	0.286	−17.69	0.011	−0.262	−0.282

The calibration period is 1931–1995 and the verification period is 1896–1930.

PREDICTORS = the tree-ring variables used in each principal components regression model that passed the correlation screening with climate (all but the first row are single-species predictors); NEIG = the number of candidate tree-ring eigenvectors/the number of tree-ring eigenvectors entered into the model;  $R^2$  = cumulative fractional variance of the model; AIC = Akaike information criterion; RSQ = square Pearson correlation; RE = verification reduction of error; CE = verification coefficient of efficiency. Higher RSQ, RE, and CE mean better verification of the fitted model.

\*Means significant  $p < 0.05$ , RE  $> 0$ , or CE  $> 0$ .

produced a highly significant calibration  $R^2$  of 0.436 and significant ( $p < 0.05$ ) verification RSQ, but the RE and CE were weakly negative.

The individual species results are an interesting contrast to the joint reconstruction model (Table 4.5). None of the single-species models calibrated as much variance as the joint model. Single-species  $R^2$  ranged from 0.128 for QUVE to 0.286 for BELE. Thus, a significant amount of additional calibrated variance has been captured by using the mutual climate information contained in the seven tree species chronologies. In terms of verification RSQ, the joint model outperforms all single-species models except that for QUPR. The RE and CE statistics show more variability between the models, but most are  $\pm 0.05$ , a range that is indeterminate with respect to assessing differences in the significance of the models (Gordon and LeDuc 1981). The QUPR model is the most anomalous single-species model. It has substantially higher verification statistics even though its calibration  $R^2$  is the second lowest one. Since the 1950s, the forests around Mohonk Lake have been subjected to episodic gypsy moth (*Lymantria dispar*) infestations that have resulted in the occasional defoliation of certain tree species, and chestnut oak is among the most favored host species. Years of notable gypsy moth population buildup and defoliation in the vicinity of Mohonk Lake were 1957, 1965–1966, 1971, 1981, and 1987–1988 (Smiley and Huth 1982; Huth 2005). So it is possible that the weak QUPR calibration is due to gypsy moth defoliation effects on the radial growth of chestnut oak.

Overall, the results presented in Table 4.5 support the use of multiple trees species for reconstructing past climate. This finding is not terribly surprising, because, as we have seen from our response function modeling results (Table 4.4) and the results from other studies (e.g., Graumlich 1993; Cook et al. 2001), tree species at the genus and subgenus levels can have different phylogenetic responses in their ring widths to nearly the same macroclimatic influences on growth. These statistically expressed differences in response to a common climate forcing are likely to enter into any climate reconstruction if it is based on only one tree species. The introduction of certain non-climatic biases into such reconstructions is therefore likely. This ought to be avoided whenever practical through the use of multiple tree species in climate reconstructions.

## 4.9 Concluding Remarks

Tree-ring analysis is one of the most powerful tools available for the study of environmental change and the identification of fundamental relationships between tree growth and climate. At every stage of analysis there is both statistical and biological uncertainty. In statistical analysis, we wish to reduce the uncertainty of our inferences as best as our data and analysis methods can reasonably allow, always taking into account the fact that we could be horribly wrong. To reduce the chances of making false statistical inferences, some form of model validation should therefore be conducted whenever possible (Snee 1977). Validation based on the analysis of withheld data as described here and on new tree-ring data from the same or



phylogenetically related species (e.g., Cook and Cole 1991; Graumlich 1993; Cook et al. 2001) are two recommended approaches.

Biological uncertainty is a more difficult problem to deal with because it can be due to emergent properties of tree growth that are essentially unpredictable from lower-level processes. For this reason, process-based mechanistic models may never be sufficiently complete to model the finer details of tree radial growth due to climate over a broad range of tree species and environmental conditions. Even so, simplified mechanistic models can work quite well in modeling the essentials of ring width variation due to climate for a variety of tree species (Anchukaitis et al. 2006; Evans et al. 2006; Vaganov et al. Chapter 3, this volume). Therefore, to deal with the possibility of emergent properties of tree growth that might be missed by a process-based model, we suggest that mechanistic and statistical models be used jointly to model the response of trees to climate and to test each other, because neither approach is likely to provide all the answers.

Finally, tree-ring analysis should also always contain a significant amount of data exploration, because the inherent biological uncertainty of tree growth in uncontrolled natural environments will never be eliminated. This, in our opinion, is good because it means that there will be a lot of interesting discoveries to make in the future! In turn, the results of exploratory data analysis can—and should—lead to numerous confirmatory tests of newly revealed associations that would have been otherwise missed if only a previously planned analysis were conducted. This does not eliminate the need or desire for a good a priori experimental design. Rather, it expands the scope of statistical analyses to allow for the unforeseen. As the father of exploratory data analysis, John Tukey, put it, ‘restricting one’s self to the planned analysis—failing to accompany it with exploration—loses sight of the most interesting results too frequently to be comfortable’ (Tukey 1977, p. 3).

**Acknowledgements** This chapter is a contribution to the meeting ‘Tree Rings and Climate: Sharpening the Focus,’ held at the University of Arizona, Tucson, on April 6–9, 2004. We thank the organizing committee members (Malcolm Hughes, Henry Diaz, and Tom Swetnam) for their kind support and encouragement. This chapter is based on the generous long-term support of the Lamont-Doherty Tree-Ring Laboratory (TRL) by the National Science Foundation and the National Oceanic and Atmospheric Administration Office of Global Programs. The U.S. Department of Energy Global Change Education Program also supported N. Pederson in his PhD dissertation research at the Lamont TRL that contributed to this chapter. We also thank the Mohonk Preserve (Paul Huth and John Thompson) for permission to sample trees used in this study and for access to the Mohonk Lake meteorological data, and to the remarkable naturalist Daniel Smiley of Mohonk who made all of this possible. Lamont-Doherty Earth Observatory Contribution No. 7205.

## Appendix

Basic chronology statistics for tree-ring series of length  $n$ :

Arithmetic Mean:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Standard Deviation:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Lag-1 Autocorrelation:

$$r_1 = \frac{\sum_{i=2}^n (x_i - \bar{x})(x_{i-1} - \bar{x})}{(n - 1)s_x^2}$$

Mean Sensitivity:

$$ms = \frac{1}{n - 1} \sum_{i=2}^n \left| \frac{2(x_i - x_{i-1})}{(x_i + x_{i-1})} \right|$$

Empirical signal strength statistics for  $m$  tree-ring series from  $t$  trees of length  $n$ :

Mean chronology:

$$\bar{X} = \frac{1}{m} \sum_{j=1}^m x_{i,j}$$

for each year,  $i = 1, n$

Chronology MS:

$$MS_c = \frac{1}{n - 1} \sum_{i=2}^n \left| \frac{2(\bar{X}_i - \bar{X}_{i-1})}{(\bar{X}_i + \bar{X}_{i-1})} \right|$$

Average  $ms$  of the  $m$  series:

$$\overline{MS}_s = \frac{1}{m} \sum_{j=1}^m ms_j$$

Edmund Schulman's  $R$ :

$$ESR = \frac{\overline{MS}_s}{MS_c}$$

For  $t$  trees and  $m$  total cores, calculate all possible between-series Pearson cross-correlations as

Pearson correlation:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1) s_x s_y}$$

where  $r_{xy}$  is the correlation between cores  $x$  and  $y$ .

From Briffa and Jones (1990), let  $c_i$  equal the number of within-tree cores for a given tree and let RBAR be a given average correlation statistic. Then,

RBAR for all series:

$$\text{RTOT} = \frac{1}{\text{NTOT}} \sum_{i=1}^t \sum_{l=1}^t \sum_{j=1}^{c_i} r_{ilj}$$

$i \neq l$

where

$$\text{NTOT} = \frac{1}{2} \left[ \sum_{i=1}^t c_i \right] \left\{ \left[ \sum_{i=1}^t c_i \right] - 1 \right\}$$

RBAR within trees:

$$\text{RWT} = \frac{1}{\text{NWT}} \sum_{i=1}^t \left[ \sum_{j=2}^{c_i} r_{ij} \right]$$

where

$$\text{NWT} = \sum_{i=1}^t \frac{1}{2} c_i (c_i - 1)$$

RBAR between trees:

$$\text{RBT} = \frac{1}{\text{NBT}} (\text{RTOT} \cdot \text{NTOT}) - \text{RWT} \cdot \text{NWT}$$

where  $\text{NBT} = \text{NTOT} - \text{NWT}$

The effective RBAR (REFF) is a weighted average of RBT and RWT that includes the added signal strength due to the within-tree correlations.

RBAR effective:

$$\text{REFF} = \frac{\text{RBT}}{\text{RWT} + \frac{1 - \text{RWT}}{\text{CEFF}}}$$

where

$$\frac{1}{\text{CEFF}} = \frac{1}{t} \sum_{i=1}^t \frac{1}{c_i}$$

The signal-to-noise ratio and expressed population signal are estimated from REFF now as

Signal-to-noise ratio:

$$\text{SNR} = \frac{t \cdot \text{REFF}}{1 - \text{REFF}}$$

Expressed population signal:

$$\text{EPS} = \frac{t \cdot \text{REFF}}{t \cdot \text{REFF} + (1 - \text{REFF})}$$

## References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* AC-19(6):716–723
- Anchukaitis KJ, Evans MN, Kaplan A, Vaganov EA, Hughes MK, Grissino-Mayer HD, Cane MA (2006) Forward modeling of regional-scale tree-ring patterns in the southeastern United States and the recent emergence of summer drought stress. *Geophys Res Lett* 33(4):L04705. doi:10.1029/2005GL025050
- Anderson PW (1972) More is different. *Science* 177:393–396
- Angier N (1997) Ernst Mayr at 93. *Nat Hist* 106(4):8–11
- Baillie MGL, Pilcher JR (1973) A simple cross-dating program for tree-ring research. *Tree-Ring Bull* 33:7–14
- Biondi F, Waikul K (2004) DENDROCLIM2002: a C++ program for statistical calibration of climate signals in tree-ring chronologies. *Comput Geosci* 30:303–311
- Blasing TJ, Solomon AM, Duvick DN (1984) Response functions revisited. *Tree-Ring Bull* 44: 1–15
- Box GEP, Jenkins GM (1976) *Time series analysis: forecasting and control*. Holden Day, San Francisco, 553 pp
- Briffa KR, Jones PD (1990) Basic chronology statistics and assessment. In: Cook ER, Kairiukstis LA (eds) *Methods of dendrochronology: applications in the environmental sciences*. International Institute for Applied Systems Analysis, Kluwer Academic Publishers, Boston, pp. 137–152
- Cook ER (1985) *A time series analysis approach to tree ring standardization*. PhD Dissertation, University of Arizona, Tucson, 185 pp
- Cook ER (1990) Bootstrap confidence intervals for red spruce ring-width chronologies and an assessment of age-related bias in recent growth trends. *Can J Forest Res* 20: 326–1331
- Cook ER, Cole J (1991) Predicting the response of forests in eastern North America to future climatic change. *Climatic Change* 19:271–282
- Cook ER, Johnson AH (1989) Climate change and forest decline: a review of the red spruce case. *Water Air Soil Poll* 48:127–140
- Cook, ER, Peters K (1981) The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bull* 41: 43–53

- Cook ER, Meko DM, Stahle DW, Cleaveland MK (1999) Drought reconstructions for the continental United States. *J Climate* 12:1145–1162
- Cook ER, Glitzenstein JS, Krusic PJ, Harcombe PA (2001) Identifying functional groups of trees in west Gulf Coast forests (USA): a tree-ring approach. *Ecol Appl* 11(3): 883–903
- Cropper JP (1982a) Comment on climate reconstructions from tree rings. In: Hughes MK, Kelly PM, Pilcher JR, LaMarche VC Jr (eds) *Climate from tree rings*. Cambridge University Press, Cambridge, pp 65–67
- Cropper JP (1982b) Comment on response functions. In: Hughes MK, Kelly PM, Pilcher JR, LaMarche VC Jr (eds) *Climate from tree rings*. Cambridge University Press, Cambridge, pp 47–50
- Cropper JP, Fritts HC (1982) Density of tree-ring grids in western North America. *Tree-Ring Bull* 42:3–10
- de Jong AFM, Mook WG, Becker B (1979) Confirmation of the Suess wiggles 3200–3800 BC. *Nature* 280:48–49
- DeWitt E, Ames M (1978) Tree-ring chronologies of eastern North America. Chronology series IV, vol 1. Laboratory of Tree-Ring Research, University of Arizona, Tucson, Arizona, 42 pp + tables
- D'Odorico P, Revelli R, Ridolfi L (2000) On the use of neural networks for dendroclimatic reconstructions. *Geophys Res Lett* 27(6):791–794
- Douglass AE (1946) Precision of ring dating in tree-ring chronologies. In: Laboratory of Tree-Ring Research Bulletin 3. University of Arizona Bull 17(3):1–21
- 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. *J Geophys Res-Biogeosci* 111:G03008. doi:10.1029/2006JG000166
- Fekedulegn BD, Colbert JJ, Hicks RR Jr, Schuckers ME (2002) Coping with multicollinearity: an example on application of principal components regression in dendroecology. Research paper NE-721. U.S. Department of Agriculture, Forest Service, Northeastern Research Station, Newtown Square, Pennsylvania, 43 pp
- Feynman RP (1999) The relation of science and religion. In: *The pleasure of finding things out*. Helix Books, Perseus, Cambridge, Massachusetts, pp 245–257
- Fowler A (1998) XMATCH98: an interactive tree-ring crossdating program. Occasional paper 38. University of Auckland, Department of Geography, pp 1–29
- Fritts HC (1963) Computer programs for tree-ring research. *Tree-Ring Bull* 25(3–4):2–7
- Fritts HC (1965) Tree-ring evidence for climatic changes in western North America. *Mon Weather Rev* 93(7):421–443
- Fritts HC (1969) Bristlecone pine in the White Mountains of California. Growth and ring-width characteristics. Papers of the Laboratory of Tree-Ring Research 4. University of Arizona Press, Tucson, Arizona, pp 1–44
- Fritts HC (1971) Dendroclimatology and dendroecology. *Quaternary Res* 1(4):419–449
- Fritts HC (1976) *Tree rings and climate*. Academic, London, 567 pp
- Fritts HC, Shashkin AV (1995) Modeling tree-ring structure as related to temperature, precipitation, and day length. In: Lewis TE (ed) *Tree rings as indicators of ecosystem health*. CRC, Boca Raton, pp 17–57
- Fritts HC, Shatz DJ (1975) Selecting and characterizing tree-ring chronologies for dendroclimatic analysis. *Tree-Ring Bull* 35:31–40
- Fritts HC, Vaganov EA, Sviderskaya IV, Shashkin AV (1991) Climatic variation and tree-ring structure in conifers: empirical and mechanistic models of tree-ring width, number of cells, cell size, cell-wall thickness and wood density. *Climate Res* 1(2):97–116
- Fritts HC, Shashkin A, Downes GM (1999) A simulation model of conifer ring growth and cell structure. In: Wimmer R, Vetter RE (eds) *Tree-ring analysis: biological, methodological, and environmental aspects*. CABI, Oxon, UK, pp 3–32
- Ghent AW (1952) A technique for determining the year of the outside ring of dead trees. *Forest Chron* 28(4):85–93

- Gordon GA, LeDuc SK (1981) Verification statistics for regression models. Seventh conference on probability and statistics in atmospheric sciences, Monterey, California, USA
- Gray BM, Wigley TML, Pilcher JR (1981) Statistical significance and reproducibility of tree-ring response functions. *Tree-Ring Bull* 41:21–35
- Graumlich LJ (1993) Response of tree growth to climatic variation in the mixed conifer and deciduous forests of the upper Great Lakes region. *Can J Forest Res* 23:133–143
- Guiot J (1991) The bootstrapped response function. *Tree-Ring Bull* 51:39–41
- Guiot J, Berger AL, Munaut AV (1982) Response functions. In: Hughes MK, Kelly PM, Pilcher JR, LaMarche VC (eds) *Climate from tree rings*. Cambridge University Press, Cambridge, UK, pp 38–45
- Guiot J, Keller T, Tessier L (1995) Relational databases in dendroclimatology and new non-linear methods to analyse the tree response to climate and pollution. In: Ohta S, Fujii T, Okada N, Hughes MK, Eckstein D (eds) *Tree rings: from the past to the future*. Proceedings of the international workshop on Asian and Pacific dendrochronology. Forestry and Forest Products Research Institute Scientific Meeting Report, vol 1, pp 17–23
- Guttman L (1954) Some necessary conditions for common-factor analysis. *Psychometrika* 19: 149–161
- Hatcher MJ, Tofts C (2004) Reductionism isn't functional. Unpublished paper. Trusted Systems Laboratory, HP Laboratories, Bristol. [www.hpl.hp.com/techreports/2004/HPL-2004-222.pdf](http://www.hpl.hp.com/techreports/2004/HPL-2004-222.pdf)
- Heikkinen HJ (1984) Tree-ring patterns: a key-year technique for cross-dating. *J Forest* 82: 302–305
- Holmes RL (1983) Computer-assisted quality control in tree-ring dating and measurement. *Tree-Ring Bull* 43:69–78
- Huber B (1943) Über die Sicherheit jahrringchronologische Datierung (On the accuracy of dendrochronological dating). *Holz als Roh und Werkstoff*, 6(10/12):263–268
- Hurvich CM, Tsay CL (1989) Regression and time series modeling in small samples. *Biometrika* 76:297–307
- Huth P (2005) Personal communication on an updated history of gypsy moth defoliation in Ulster County, New York. Daniel Smiley Research Center, Mohonk Preserve, Mohonk Lake, New Paltz, New York
- Jolliffe IT (1973) Discarding variables in a principal component analysis II: real data. *Appl Stat* 22:21–31
- Jones RH (1985) Time series analysis—time domain. In: Murphy AH, Katz RW (eds) *Probability, statistics and decision making in the atmospheric sciences*. Westview, Boulder, Colorado, pp 223–259
- Kaiser HF (1960) The application of electronic computers to factor analysis. *Educ Psychol Meas* 20:141–151
- Kozlowski TT, Kramer PJ, Pallardy SG (1991) *The physiological ecology of woody plants*. Academic, London, 657 pp
- LaMarche VC Jr, Harlan TP (1973) Accuracy of tree-ring dating of bristlecone pine for calibration of the radiocarbon time scale. *J Geophys Res* 78(36):8849–8858
- Laughlin RB (2005) *A different universe: remaking physics from the bottom down*. Basic Books, New York
- Laughlin RB, Pines D (2000) The theory of everything. *Proc Natl Acad Sci USA* 97:28–31
- Linick TW, Suess HE, Becker B (1985) La Jolla measurements of radiocarbon on south German oak tree ring chronologies. *Radiocarbon* 27(1):20–30
- Mitchell JM Jr, Dzerdzeevskii B, Flohn H, Hofmeyr WL, Lamb HH, Rao KN, Wallén CC (1966) Climatic change. WMO technical note no 79
- Morrison DF (1976) *Multivariate statistical methods*, 2nd edn. McGraw-Hill, New York
- Ni F, Cavazos T, Hughes MK, Comrie AC, Funkhouser G (2002) Cool-season precipitation in the southwestern USA since AD 1000: comparison of linear and nonlinear techniques for reconstruction. *Int J Climatol* 22:1645–1662

- Pederson N, Cook ER, Jacoby GC, Peteet DM, Griffin KL (2004) The influence of winter temperatures on the annual radial growth of six northern range margin tree species. *Dendrochronologia* 22:7–29
- Pilcher JR, Baillie MGL, Schmidt B, Becker B (1984) A 7272-year tree-ring chronology for western Europe. *Nature* 312(5990):150–152
- Preisendorfer RW (1988) Principal component analysis in meteorology and oceanography. Elsevier Science, Amsterdam, 426 pp
- Preisendorfer RW, Zwiers FW, Barnett TP (1981) Foundations of principal components selection rules. SIO reference series 81-4, Scripps Institution of Oceanography, La Jolla, California, USA, 192 pp
- Read J, Busby JR (1990) Comparative responses to temperature of the major canopy species of Tasmanian cool temperate rainforest and their ecological significance. II. Net photosynthesis and climate analysis. *Aust J Bot* 38:185–205
- Rencher AC, Pun FC (1980) Inflation of  $R^2$  in best subset regression. *Technometrics* 22:49–53
- Schulman E (1956) Dendroclimatic changes in semiarid America. University of Arizona Press, Tucson, 142 pp
- Schweingruber FH, Bräker OU, Schär E (1987) Temperature information from a European dendroclimatological sampling network. *Dendrochronologia* 5:9–33
- Schweingruber FH, Eckstein D, Serre-Bachet F, Bräker OU (1990) Identification, presentation and interpretation of event years and pointer years in dendrochronology. *Dendrochronologia* 8:9–38
- Shashkin AV, Vaganov EA (1993) Simulation model of climatically determined variability of conifers' annual increment (on the example of common pine in the steppe zone). *Russ J Ecol* 24:275–280
- Smiley D, Huth P (1982) Gypsy moth defoliation in vicinity of Mohonk Lake 1956–1981. Unpublished research report. Mohonk Preserve, Inc
- Snee RD (1977) Validation of regression models: methods and examples. *Technometrics* 19: 415–428
- Stokes MA, Smiley TL (1968) An introduction to tree-ring dating. University of Chicago Press, Chicago, Illinois, 73 pp
- Suess HE (1965) Secular variations of the cosmic-ray-produced carbon-14 in the atmosphere and their interpretation. *J Geophys Res* 70:5937–5952
- Trumbore S, Gaudinski JB, Hanson PJ, Southon JR (2002) Quantifying ecosystem-atmosphere carbon exchange with a  $^{14}\text{C}$  label. *Eos* 83:265, 267–268
- Tukey JW (1977) Exploratory data analysis. Addison-Wesley, London, 688 pp
- Twain M (1924) Mark Twain's autobiography, vol 1. Harper & Brothers, New York, 368 pp
- Vaganov EA, Hughes MK, Kirdyanov AV, Schweingruber FH, Silkin PP (1999) Influence of snowfall and melt timing on tree growth in subarctic Eurasia. *Nature* 400:149–151
- Vaganov EA, Hughes MK, Shashkin AV (2006) Growth dynamics of tree rings: images of past and future environments. Springer-Verlag, Berlin, Heidelberg, New York
- Visser H, Molenaar J (1988) Kalman filter analysis in dendroclimatology. *Biometrics* 44(4): 929–940
- Wigley TML, Briffa KR, Jones PD (1984) On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *J Clim Appl Meteorol* 23:201–213
- Wigley TML, Jones PD, Briffa KR (1987) Cross-dating methods in dendrochronology. *J Archaeol Sci* 14:51–64
- Woodhouse CA (1999) Artificial neural networks and dendroclimatic reconstructions: an example from the Front Range, Colorado, USA. *Holocene* 9(5):521–529
- Yamaguchi DK (1991) A simple method for cross-dating increment cores from living trees. *Can J Forest Res* 21:414–416
- Yamaguchi DK (1994) More on estimating the statistical significance of cross-dating positions for 'floating' tree-ring series. *Can J Forest Res* 24(2):427–429
- Yamaguchi DK, Allen GA (1992) A new computer program for estimating the statistical significance of cross-dating positions for 'floating' tree-ring series. *Can J Forest Res* 22(9):1215–1221