Wavelet-Based Multiresolution Analysis of Wivenhoe Dam Water Temperatures

D. B. Percival,[1] S. M. Lennox,[2] Y.-G. Wang,[2,3] and R. E. Darnell[2]

S. M. Lennox, Y.-G. Wang and R. E. Darnell, CSIRO Mathematics, Informatics and Statistics, Long Pocket Laboratories, 120 Meiers Road, Indooroopilly, QLD 4068, Australia.

D. B. Percival, Applied Physics Laboratory, Box 355640, University of Washington, Seattle, WA 98195–5640, USA. (dbp@apl.washington.edu)

[1]Applied Physics Laboratory, University of Washington, Seattle, Washington, USA.

[2]CSIRO Mathematics, Informatics and Statistics, Long Pocket Laboratories, Indooroopilly, Queensland, Australia.

[3]current affiliation: Centre for Applications in Natural Resource Mathematics (CARM), School of Mathematics and Physics, University of Queensland, St Lucia, Queensland, Australia.

Abstract. Water temperature measurements from Wivenhoe Dam offer a unique opportunity for studying fluctuations of temperatures in a subtropical dam as a function of time and depth. Cursory examination of the data indicate a complicated structure across both time and depth. We propose simplifying the task of describing these data by breaking the time series at each depth into physically meaningful components that individually capture daily, subannual and annual (DSA) variations. Precise definitions for each component are formulated in terms of a wavelet-based multiresolution analysis. The DSA components are approximately pairwise uncorrelated within a given depth and between different depths. They also satisfy an additive property in that their sum is exactly equal to the original time series. Each component is based upon a set of coefficients that decomposes the sample variance of each time series exactly across time and that can be used to study both time-varying variances of water temperature at each depth and time-varying correlations between temperatures at different depths. Each DSA component is amenable for studying a certain aspect of the relationship between the series at different depths. The daily component in general is weakly correlated between depths, including those that are adjacent to one another. The subannual component quantifies seasonal effects and in particular isolates phenomena associated with the thermocline, thus simplifying its study across time. The annual component can be used for a trend analysis. The descriptive analysis provided by the DSA decomposition is a useful precursor to a more formal statistical analysis.

1. Introduction

The Queensland Bulk Water Supply Authority (henceforth ‘Seqwater’) is the bulk water supplier for South-East Queensland (SEQ), Australia. The overall mission of Seqwater is to manage catchments, water storages and treatment services to ensure the quantity and quality of water supplies to SEQ (see `www.seqwater.com.au` for details). To help fulfill this mission, Seqwater recently upgraded their ongoing monitoring program by the permanent installation of YSI 6955 vertical profiling systems within the Lake Wivenhoe dam to monitor a number of water quality indicators at different depths every 2 hours (see `www.ysi.com` for details about the profiler). In addition to temperature (the focus of this paper), these water quality indicators include pH, turbidity, dissolved oxygen, specific conductivity, blue green algae and chlorophyll-a. This upgrade was in response to a need for a greater frequency in sampling because of concerns that algae blooms, conductivity spikes, anoxic events and lake turnovers might be inadequately captured and/or represented under the old monitoring program. The ability to collect data more frequently and automatically both expands the scope of the old monitoring program (in which the time between samples might be up to 3 weeks) and reduces costs involved with the need for a greater frequency in sampling.

For this paper we conduct a detailed study of temperature because it is an important driver for other water quality indicators. We examine a 600 day segment of temperatures collected by the profiling system at depths of 1, 5, 10, 15 and 20 meters at two hour intervals starting on 1 October 2007. These data offer a unique chance to study the depth/temporal evolution of dam temperatures in a subtropical climate. Figure 1 shows plots of the temperature series at these five depths. A cursory examination of this figure indicates a complicated structure both across time and down different depths. We propose simplifying the task of describing these data by breaking each series up into components that individually capture the daily, subannual (seasonal) and annual (DSA) variations in the series. As discussed in Section 2, we formulate precise definitions for each of these components in terms of a wavelet-based multiresolution analysis (MRA). The DSA components are such that (1) they are approximately pairwise uncorrelated; (2) they satisfy an additive property in that their sum is exactly equal to the original time series; and (3) they are based upon coefficients that can be used to decompose the sample variance of each time series exactly across time and that are amenable for studying the relationships between the series at different depths. Our analysis is mainly descriptive, but provides insight into what components would be needed for a complete statistical model of water temperatures as a function of time and depth. Our results are also of potential interest for comparison with physical models of how dam water temperatures evolve over depth and across time.

The remainder of this paper is organized as follows. Section 2 gives an overview of standard wavelet analysis and the adaptations we have made. Section 3 describes the preparations we have made to the data prior to our analysis. Section 4 presents our analysis, followed by a discussion in Section 5. We summarize our main results and technical contributions in Section 6. Appendices A to E contain some technical details.

2. Wavelet-Based Analysis and Its Adaptation for Dam Water Temperatures

The analysis of dam water temperatures we present in this paper is an adaptation of standard wavelet analysis. Prior to describing our adaptations in Section 2.2, we review the key ideas behind wavelet analysis for time series in the following subsection, with technical details deferred to Appendix A.

2.1. Overview of wavelet analysis of time series

Let $\mathbf{X}$ denote a column vector whose elements $X_t, t = 0, 1, \ldots, N-1$, represent a time series of N regularly sampled observations; i.e., the time associated with X_t is $t_0 + t\,\Delta$, where t_0 is the time at which X_0 was observed, and Δ is the sampling time between adjacent observations ($\Delta = 2$ hours for the water temperature time series). The wavelet analysis of a time series is based upon a linear transformation of $\mathbf{X}$, expressed as

$$\widetilde{\mathbf{W}} = \widetilde{\mathcal{W}}\mathbf{X}. \tag{1}$$

Here $\widetilde{\mathcal{W}}$ is a matrix that takes the time series and produces a vector of so-called maximal overlap discrete wavelet transform (MODWT) coefficients $\widetilde{\mathbf{W}}$ (*Percival and Guttorp*, 1994). This type of wavelet transform is essentially the same as ones going under the names 'undecimated DWT' (*Shensa*, 1992), 'shift invariant DWT' (*Beylkin*, 1992; *Lang et al.*, 1995), 'wavelet frames' (*Unser*, 1995), 'translation invariant DWT' (*Coifman and Donoho*, 1995; *Liang and Parks*, 1996; *Del Marco and Weiss*, 1997), 'stationary DWT' (*Nason and Silverman*, 1995), 'time invariant DWT' (*Pesquet et al.*, 1996) and 'non-decimated DWT' (*Bruce and Gao*, 1996).

There are two types of MODWT coefficients in $\widetilde{\mathbf{W}}$, namely, wavelet coefficients and scaling coefficients. While each element X_t in $\mathbf{X}$ is associated with a time index t, each wavelet coefficient $\widetilde{W}_{j,t}$ in $\widetilde{\mathbf{W}}$ has two indices, namely, a so-called scale index (or level) j, where $j = 1, 2, \ldots, J_0$, and a time index t (as explained below, we select $J_0 = 9$ as the maximum level to be entertained for the water temperature time series, but this choice is application dependent). The time index for $\widetilde{W}_{j,t}$ can be related to the time index for X_t and says that, in forming this coefficient, we are only making use of values in $\mathbf{X}$ centered at a particular time. The scale index j indicates how many values from $\mathbf{X}$ are in effect being used to form $\widetilde{W}_{j,t}$. If j is small (large), then $\widetilde{W}_{j,t}$ depends mainly upon a small (large) number of values from $\mathbf{X}$. A complementary interpretation of the level j is as an index for an interval of frequencies f defined by $1/(2^{j+1}\,\Delta) < f \leq 1/(2^j\,\Delta)$. With this interpretation, we say that $\widetilde{W}_{j,t}$ is summarizing the frequency content in a subset of values from $\mathbf{X}$ over the interval of frequencies $\mathcal{I}_j = (1/(2^{j+1}\,\Delta), 1/(2^j\,\Delta)]$.

The scaling coefficients are the other type of coefficients in $\widetilde{\mathbf{W}}$. Like the wavelet coefficients, each scaling coefficient has a level index and a time index t, but the former assumes only the single value J_0. We denote the scaling coefficients by $\widetilde{V}_{J_0,t}$. The index t in $\widetilde{V}_{J_0,t}$ has the same interpretation as in $\widetilde{W}_{j,t}$, but the associated interval of frequencies is now $\mathcal{I}_0 = [0, 1/(2^{J_0+1}\,\Delta)]$. Note that the union of $\mathcal{I}_j$, $j = 0, 1, \ldots, J_0$, is $[0, 1/(2\,\Delta)]$, which comprises all of the physically meaningful frequencies in a Fourier decomposition of $\mathbf{X}$. Collectively, we can think of the wavelet and scaling coefficients as forming localized Fourier analyses of $\mathbf{X}$, where the first index on a coefficient indicates the interval of frequencies with which the coefficient is associated, while the second index t indicates what part of the time series is being looked at. The scaling coefficients capture the localized low-frequency variations in $\mathbf{X}$, whereas the wavelet coefficients do the same over the frequency intervals $\mathcal{I}_j$, $j = 1, 2, \ldots, J_0$.

Let us now place all the wavelet coefficients in $\widetilde{\mathbf{W}}$ that are associated with level j into the vector $\widetilde{\mathbf{W}}_j$, and all the scaling coefficients into the vector $\widetilde{\mathbf{V}}_{J_0}$. Each of these vectors has the same number of elements N as the original time series $\mathbf{X}$ (hence, in Equation (1), the matrix $\widetilde{\mathcal{W}}$ is of dimension $(J_0+1)N \times N$, and the vector $\widetilde{\mathbf{W}}$ has $(J_0+1)N$ elements). Let $\|\mathbf{X}\|^2 \equiv \sum_t X_t^2$ denote the square of the Euclidean norm of the vector $\mathbf{X}$. One important property of the wavelet transform of $\mathbf{X}$ is that the MODWT coefficients preserve the sum of squares of the original data; i.e., $\|\widetilde{\mathbf{W}}\|^2 = \|\mathbf{X}\|^2$. Since the union of the elements of $\widetilde{\mathbf{W}}_1$, $\widetilde{\mathbf{W}}_2$, ..., $\widetilde{\mathbf{W}}_{J_0}$ and $\widetilde{\mathbf{V}}_{J_0}$ comprises all the elements of $\widetilde{\mathbf{W}}$, we also have

$$\|\mathbf{X}\|^2 = \sum_{j=1}^{J_0} \|\widetilde{\mathbf{W}}_j\|^2 + \|\widetilde{\mathbf{V}}_{J_0}\|^2. \tag{2}$$

The interpretation of $\|\widetilde{\mathbf{W}}_j\|^2$ is that it is the part of $\|\mathbf{X}\|^2$ attributable to localized Fourier coefficients associated with the frequency interval $\mathcal{I}_j$; on the other hand, $\|\widetilde{\mathbf{V}}_{J_0}\|^2$ is associated with the low-frequency interval $\mathcal{I}_0$. Letting $\overline{X} = \sum_t X_t/N$ represent the sample mean of $\mathbf{X}$, we can express its sample variance as

$$\begin{aligned}\hat{\sigma}_X^2 \equiv \frac{1}{N}\sum_{t=0}^{N-1}\left(X_t - \overline{X}\right)^2 = \frac{1}{N}\|\mathbf{X}\|^2 - \overline{X}^2 &= \sum_{j=1}^{J_0}\frac{1}{N}\|\widetilde{\mathbf{W}}_j\|^2 + \left(\frac{1}{N}\|\widetilde{\mathbf{V}}_{J_0}\|^2 - \overline{X}^2\right) \\ &\equiv \sum_{j=1}^{J_0}\hat{\sigma}_j^2 + \hat{\sigma}_0^2, \end{aligned} \tag{3}$$

where $\hat{\sigma}_j^2$ and $\hat{\sigma}_0^2$ can be interpreted as sample variances associated with $\widetilde{\mathbf{W}}_j$ and $\widetilde{\mathbf{V}}_{J_0}$ (the nature of the wavelet transform is such that the sample mean of $\widetilde{\mathbf{V}}_{J_0}$ is $\overline{X}$ also, whereas the coefficients in $\widetilde{\mathbf{W}}_j$ can be considered as coming from a population whose theoretical mean value is zero). We thus can break up the sample variance of $\mathbf{X}$ into J_0+1 parts, J_0 of which (the $\hat{\sigma}_j^2$'s) are attributable to fluctuations in the intervals of frequencies $\mathcal{I}_j$, and the last ($\hat{\sigma}_0^2$), to fluctuations in $\mathbf{X}$ over the low-frequency interval $\mathcal{I}_0$. We refer to the decomposition of $\hat{\sigma}_X^2$ afforded by Equation 3 as a wavelet-based analysis of variance (ANOVA).

In addition to a wavelet-based ANOVA, we can use the MODWT coefficients to obtain a wavelet-based additive decomposition known as a multiresolution analysis (MRA). Formally the MRA follows from the fact that we can readily recover $\mathbf{X}$ from its MODWT coefficients $\widetilde{\mathbf{W}}$ via the synthesis equation

$$\mathbf{X} = \widetilde{\mathcal{W}}^T \widetilde{\mathbf{W}}. \tag{4}$$

By an appropriate partitioning of both $\widetilde{\mathcal{W}}$ and $\widetilde{\mathbf{W}}$, we can rewrite the synthesis equation as

$$\mathbf{X} = \sum_{j=1}^{J_0} \widetilde{\mathcal{D}}_j + \widetilde{\mathcal{S}}_{J_0}, \tag{5}$$

where $\widetilde{\mathcal{D}}_j$ and $\widetilde{\mathcal{S}}_{J_0}$ are N-dimensional vectors known as, respectively, the jth level 'detail' and the J_0th level 'smooth.' The vector $\widetilde{\mathcal{D}}_j$ depends just upon $\widetilde{\mathbf{W}}_j$ and those rows in $\widetilde{\mathcal{W}}$ used to create $\widetilde{\mathbf{W}}_j$ from $\mathbf{X}$, so we can interpret $\widetilde{\mathcal{D}}_j$ as the portion of the additive decomposition due to fluctuations in the interval of frequencies $\mathcal{I}_j$; an analogous argument says that we can interpret $\widetilde{\mathcal{S}}_{J_0}$ as the part of the MRA due to low-frequency fluctuations. The components of an MRA are intended to capture distinct aspects of a time series and, if proper care is taken, can be regarded as approximately pairwise uncorrelated.

2.2. Wavelet analysis adapted for use with dam water temperatures

Two important physical drivers of dam water temperature time series can ultimately be traced to the daily rotation of the earth and to the revolution of the earth about the sun. We seek an additive decomposition of the series with components that isolate diurnal and annual variations. Such a decomposition should facilitate analysis of water temperatures because we can then study their physically motivated components individually. Since $\Delta = 2$ hours here, the frequency intervals $\mathcal{I}_3$, $\mathcal{I}_2$ and $\mathcal{I}_1$ correspond to $[0.75, 1.5]$, $[1.5, 3]$ and $[3, 6]$ cycles per day. Any purely periodic daily variation in a time series that is sampled every two hours can be expressed exactly with a Fourier decomposition involving a constant and sinusoids with (at most) six frequency components, namely, the fundamental frequency $f_1 = 1$ cycle/day and its five harmonics $f_k = kf_1$, $k = 2, 3, \ldots, 6$ cycles/day. This fact suggests that, in a wavelet-based MRA, daily fluctuations are captured primarily in details $\widetilde{\mathcal{D}}_1$, $\widetilde{\mathcal{D}}_2$ and $\widetilde{\mathcal{D}}_3$. On the other hand, the smooth $\widetilde{\mathcal{S}}_9$ in a level $J_0 = 9$ MRA captures fluctuations that are lower in frequency than $1/(2^{10}\,\Delta) \doteq 4.3$ cycles/year. Empirically, as shown in Fig. 2, $\widetilde{\mathcal{S}}_9$ is preferable to either $\widetilde{\mathcal{S}}_8$ or $\widetilde{\mathcal{S}}_{10}$ as a representation of interannual fluctuations: the former is arguably undersmoothed (containing fluctuations better ascribed to intra-annual variations), while the latter is somewhat oversmoothed (hence distorting the interannual fluctuations). With the choice of $\widetilde{\mathcal{S}}_9$, we can lump together the remaining details $\widetilde{\mathcal{D}}_4$, $\widetilde{\mathcal{D}}_5$, $\ldots, \widetilde{\mathcal{D}}_9$ in a level $J_0 = 9$ MRA into a component that captures frequency fluctuations lower than those associated with daily variations, but higher than those with annual variations, leading to the following the modified MRA:

$$\mathbf{X} = \mathcal{D} + \mathcal{S} + \mathcal{A}, \tag{6}$$

where

$$\mathcal{D} = \widetilde{\mathcal{D}}_1 + \widetilde{\mathcal{D}}_2 + \widetilde{\mathcal{D}}_3, \;\; \mathcal{S} = \widetilde{\mathcal{D}}_4 + \widetilde{\mathcal{D}}_5 + \cdots + \widetilde{\mathcal{D}}_9 \;\text{ and }\; \mathcal{A} = \widetilde{\mathcal{S}}_9.$$

We refer to $\mathcal{D}$, $\mathcal{S}$ and $\mathcal{A}$ as the daily, subannual (or seasonal) and annual components and to the modified MRA as the DSA decomposition. We denote the tth elements of $\mathcal{D}$, $\mathcal{S}$ and $\mathcal{A}$ by $\mathcal{D}_t$, $\mathcal{S}_t$ and $\mathcal{A}_t$.

We can formulate an ANOVA corresponding to the DSA decomposition in two ways. An obvious approach is to just combine together the squared wavelet coefficients from each of the levels involved in forming the daily and subannual components; however, the statistical properties of such a combination are difficult to ascertain because we need to know the relative influence of squared coefficients from the different $\widetilde{\mathbf{W}}_j$'s. A second approach, which leads to a more tractable ANOVA and is described in detail in Appendix B, is to define a new transform, say, $\mathbf{U} = \mathcal{U}\mathbf{X}$, with a corresponding synthesis equation $\mathbf{X} = \mathcal{U}^T\mathbf{U}$. Here $\mathcal{U}$ has dimension $3N \times N$, and $\mathbf{U}$ contains three types of transform coefficients, which we place in the N-dimensional vectors $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$. These coefficients lead to the sum of squares decomposition

$$\|\mathbf{X}\|^2 = \|\mathbf{D}\|^2 + \|\mathbf{S}\|^2 + \|\mathbf{A}\|^2, \tag{7}$$

where

$$\|\mathbf{D}\|^2 = \sum_{j=1}^{3} \|\widetilde{\mathbf{W}}_j\|^2, \;\; \|\mathbf{S}\|^2 = \sum_{j=4}^{9} \|\widetilde{\mathbf{W}}_j\|^2 \text{ and } \|\mathbf{A}\|^2 = \|\widetilde{\mathbf{V}}_{J_0}\|^2.$$

In the same way that the sum of squares decomposition of Equation (2) led to the ANOVA of Equation (3), the above gives us an ANOVA based upon the $\mathcal{U}$ transform:

$$\hat{\sigma}_X^2 = \frac{1}{N}\|\mathbf{D}\|^2 + \frac{1}{N}\|\mathbf{S}\|^2 + \left(\frac{1}{N}\|\widetilde{\mathbf{A}}\|^2 - \overline{X}^2\right) \equiv \hat{\sigma}_D^2 + \hat{\sigma}_S^2 + \hat{\sigma}_A^2, \tag{8}$$

where

$$\hat{\sigma}_D^2 = \sum_{j=1}^{3} \hat{\sigma}_j^2, \;\; \hat{\sigma}_S^2 = \sum_{j=4}^{9} \hat{\sigma}_j^2 \text{ and } \hat{\sigma}_A^2 = \hat{\sigma}_0^2$$

A manipulation of the synthesis equation leads to exactly the same additive decomposition as given by Equation (6). In essence, we have 'collapsed' the $3N$ wavelet coefficients in $\widetilde{\mathbf{W}}_1$, $\widetilde{\mathbf{W}}_2$ and $\widetilde{\mathbf{W}}_3$ into the N coefficients $\mathbf{D}$, and, using just $\mathbf{D}$, we can determine $\mathcal{D} = \widetilde{\mathcal{D}}_1 + \widetilde{\mathcal{D}}_2 + \widetilde{\mathcal{D}}_3$; likewise, the $6N$ wavelet coefficients in $\widetilde{\mathbf{W}}_4, \widetilde{\mathbf{W}}_5, \ldots, \widetilde{\mathbf{W}}_9$ collapse into the N-dimensional vector $\mathbf{S}$, and we only need $\mathbf{S}$ in order to form $\mathcal{S}$. Henceforth we refer to $\mathcal{U}$ as the DSA transform. We refer to $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$ collectively as the DSA transform coefficients (or just DSA coefficients) and individually as the daily, subannual and annual coefficients.

3. Data Preparation

The monitoring system at Wivenhoe Dam is designed to measure water temperature and other variables at depths of $1, 2, \ldots, 20$ m every two hours (to simplify tables and figures presented later on, we concentrate on the representative depths of 1, 5, 10, 15 and 20 m). For the most part, this protocol was successfully adhered to, but, as can be seen from Fig. 1, there are a number of gaps in the data, and there is also some jitter in the collection times (e.g., a measurement is collected a minute later than anticipated). Jittering is unlikely to impact our analysis significantly, but gaps in the data are more problematic. There is wavelet methodology for handling gappy time series but currently only for univariate time series (see, e.g., *Hall and Turlach*, 1997, *Sardy et al.*, 1999, *Mondal and Percival*, 2010, and *Porto et al.*, 2010). Since we are interested in the relationships between time series collected at different depths, we have elected to fill in the gaps using a scheme documented in Appendix C. The gap-filled series are then amenable to analysis via the techniques discussed in the previous two sections.

The nature of the water temperature data also dictates that we pay close attention to how the MODWT and the DSA transforms handle boundary conditions. The procedure we used is described in Appendix D and is designed to minimize distortions that can arise in the analysis at the starts and ends of the various time series.

4. Data Analysis

Here we present our analysis of the Wivenhoe Dam water temperature time series based upon the wavelet and DSA transforms described in Section 2. Figure 3 shows the DSA decomposition for water temperature time series X_t at depths 1, 5, 10, 15 and 20 m (these decomposition are based on interpolated series; the uninterpolated series are shown in Fig. 1). In terms of explaining the variability in X_t, the relative importance of the three DSA components is qualitatively easy to see from this figure, where the distance between adjacent vertical tick marks represents 2 degrees Celsius in all fifteen plots. For each depth, the annual variation is clearly the dominant component, with the subannual variation being second in importance. Overall the daily component seems to contribute the least to the overall variability of X_t, although there are some limited stretches of time over which the daily component apparently has greater variability than the subannual component.

To quantify the relative importance of the three components globally (i.e., when considered across the entire 600 day stretch of data) and to explore the relationship between the MODWT and DSA coefficients, let us first consider the wavelet-based ANOVA given by Equation (3). Figure 4 shows the sample wavelet variances $\hat{\sigma}_j^2$ versus levels $j = 1, 2, \ldots, 9$, for the five depths, along with $\hat{\sigma}_0^2$ (the variance associated with the scaling coefficients $\widetilde{V}_9$). The wavelet variances for depths of 15 and 20 m are quite similar in their overall patterns, and those for 1 and 5 m are also, except for some divergence at levels $j = 6, 7$ and 8. The 10 m depth has a pattern that represents a transition between the patterns at shallower and deeper depths. While the absolute levels are different, the gross patterns of variability in the wavelet variances are by and large the same at all depths: an increase from $j = 1$ to $j = 3$ (with the single exception of 15 m), followed by a drop between $j = 3$ and 4, and a tendency to increase after that. As noted previously, the fundamental frequency of a periodic time series with a period of a day is trapped by the nominal frequency interval $\mathcal{I}_3$ associated with level $j = 3$, while its associated harmonics are contained in $\mathcal{I}_1$ and $\mathcal{I}_2$. The fact that, with the exception of 15 m, the wavelet variance at level $j = 3$ is larger than those at levels 1 and 2 indicates that the fluctuations with a frequency content close to the fundamental frequency are stronger than those associated with the harmonics. The sum of the wavelet variances indexed by $j = 1, 2$ and 3 accounts for the portion of the total variance of X_t attributable to daily variations; similarly, the sum of those indexed by $j = 4, \ldots, 9$ accounts for the variance ascribable to subannual variations. The rest of the variance of X_t is accounted for by the variance of the annual coefficients, which is the same as that of the level $J_0 = 9$ scaling coefficients and is shown for each depth in the upper right-hand corner of Fig. 4. The top part of Table 1 gives the DSA analysis of variance for the water temperature data at the five depths (the bottom part has the corresponding standard deviations). At each depth, the variance attributable to annual coefficients is one or two orders of magnitude greater than that for subannual coefficients, which in turn is greater by at least a factor of two than the variances attributable to daily coefficients. The variance of the annual coefficients decreases monotonically with depth, while the variances of the subannual and daily coefficients decrease also, with minor exceptions to this general pattern.

As shown in Figure 3, the patterns of the annual components for the five depths are qualitatively similar, but there are some interesting differences (aside from the overall decrease in temperature with increasing depth). The vertical dotted lines in these plots indicate the locations of the peak values in 2008 and 2009 and the minimum in 2008. The dates of the peaks and minima in 2008 increase with depth, with the peak and minimum at 20 m occurring about a month later than the ones at 1 m. The dates of the peaks in 2009 also increase with depth, but now the 20 m peak occurs about three months later than 1 m peak. If we subtract the height of the 2008 peak from the 2009 peak, the differences decrease with increasing depth and switch from being positive to a negative value at 20 m. The annual components at different depths thus do not consistently track one another in their fine details across the 600 days of data, and there are noticeable variations from one year to the next in the annual pattern, even though we have observed less than two full years of data.

Figure 3 suggests that the variability associated with subannual and daily components is not constant across time. We can explore the time-varying variability by studying the subannual and daily coefficients from the DSA transform. The square of any individual coefficient is a time-localized contribution to the overall variance of the time series. We can track changes in variance across time by locally smoothing the squared coefficients. Figure 5 shows plots of the squared coefficients after applying a Gaussian-shaped smoother with an effective bandwidth of about a month (solid curves), along with lower and upper limits of pointwise 95% confidence intervals (CIs); see Appendix E for details. When averaged over all 600 days, the variance of the coefficients typically decreases with depth for both components (the 600-day average variances are indicated by the horizontal lines). The CIs for the variance fluctuations in the daily coefficients rule out the hypothesis that the variance is constant across time; the same holds for the subannual component, but less dramatically so. There are statistically significant fluctuations in the variance of the daily coefficients at a depth of 1 m, but, at greater depths, the relative fluctuations are greater (e.g., about three orders of magnitude difference between the largest and smallest variances at 10 m). Thus, while variance of the 1 m daily coefficients is relatively stable across time, the same cannot be said for the lower depths. The opposite pattern seems to be the case for the subannual fluctuations: the three lower depths seem to have more homogenous variances than the two shallower ones.

Let us now turn to a study of the global cross-correlations between the DSA coefficients at the 5 depths. We have a total of 15 sets of coefficients in all, so there are $\binom{15}{2} = 105$ cross-correlations to consider. Of these, 75 are between two different types of coefficients, either at the same depth or different depths. These between-type cross-correlations are generally small: 6 are between 0.1 and 0.15, while the remaining 69 are less than 0.1 and greater than -0.03. The fact that these cross-correlations are so small lends credence to the claim that the DSA transform is separating the X_t series into coefficients whose different types (i.e., daily, subannual or annual) are approximately uncorrelated. The remaining 30 cross-correlations involve pairs of within-type coefficients at different depths and are shown in Table 2, along with the cross-correlations between the X_t series themselves. The daily cross-correlations tend to be quite small, with the largest (0.22) being between depths of 5 and 10 m. In particular, there seems to be little correlation between the surface (1 m) and other depths. If we lag one of the daily coefficients by $\pm 2, \pm 4, \ldots, \pm 22$ hours and look at the cross-correlations between it and the other coefficients, there is virtually no difference between these and the unlagged cross-correlations. This rules out the hypothesis that there might be a simple lead/lag relationship between the daily coefficients at different depths. On the other hand, as is to be expected from an examination of the first column of Fig. 3, there are strong cross-correlations between annual coefficients, with the correlation decreasing as the distance between the depths increases. There is little difference between these cross-correlations and the corresponding ones for the X_t series themselves (see the two tables in the bottom row of Table 2). The cross-correlations between subannual coefficients are all positive and are larger (smaller) than between the corresponding daily (annual) components. Thus, while the original time series X_t are highly correlated, the DSA transform allows us to quantify the fact that this correlation is largely due to the annual pattern and to examine how the series are related on a daily and subannual basis once the annual pattern has been taken away. The weaker cross-correlations between sub-annual components might be explained by atmospheric events. The severity of an event could determine how many depths are affected. If a rainfall event is strong enough, the surface and middle layers might mix, resulting in similar changes at all depths; however, a weaker event might only affect the surface conditions, and not the deeper depths.

Figures 6 and 7 explore the consistency across time in the cross correlations between different depths in the daily and subannual coefficients. Here we compute sample cross-correlation coefficients on a month-by-month basis. There are some interesting changes in the subannual correlations at the deeper depths (Fig. 6). For example, there is a stretch of high positive correlations between the 15 and 20 m coefficients from February to September in 2008, followed by a gradual decline after that. This stretch of high correlation seems to coincide with a stretch of decreased variability at both depths as evidenced in Figs. 3 and 5. The cross-correlations in the daily coefficients in Fig. 7 tend to be smaller and to be less time dependent than those for the subannual coefficients. In particular, the cross-correlations between the 1 m coefficients and those at different depths are small overall, indicating little direct daily co-temporaneous relationship between temperatures near the surface and those at deeper levels. (We also looked at cross-correlations on a week-by-week basis for the daily coefficients, but focusing on shorter intervals did not yield correlations that were markedly stronger.)

The periods of high correlation seem to be well aligned with known periods of stratification in the lake. Stratification often occurs during the summer period when the surface water is heated, creating a warm and well mixed surface layer (epilimnion). The deeper water remains cold, well mixed and much denser, thus creating a thermocline (a range of depths that show a rapid change in temperature) between the surface and bottom layers. The surface water cools leading up to winter, creating a much denser and cooler surface layer that will exchange with the bottom, resulting in an overturn of the lake. This mixing process is evident in Figure 6 during autumn/winter, where the monthly correlations between depths are positive and strong. This mixing process is also associated with the periods of decreased variability. The correlations between 1 m and 20 m depth appear to be showing a period of overturn between and April and September 2008. The sudden decrease in correlation after September 2008 might identify the beginning of the stratification of the lake. With surface and bottom temperatures separated by a thermocline, we would expect there to be lower correlations between the surface and bottom temperatures.

5. Discussion

The biggest contributors to the overall variability of each temperature series X_t are the annual coefficients, which determine the annual component in the DSA decomposition. Even though the available data span just 600 days, it is evident from our analysis that the annual component at a particular depth can vary considerably from year to year. More data are needed to develop an overall depth/time model for the annual component, but a study of sparsely sampled historical data could potentially help identify explanatory covariates that might drive the distortions from a purely periodic pattern (e.g., indicators of weather patterns, average surface temperatures and total depth of water in the dam).

The second biggest contributors to the variance of X_t are the subannual coefficients, while the daily coefficients are the weakest contributors. The subannual coefficients show some indication of increased variability at 1 m depth when compared to the temperatures at 20 m, yet the reverse is true for the variability of the daily coefficients. The smaller variability in daily coefficients occurs at 1 m, with the largest variability at a depth of 20 m. The surface temperatures are affected very much by atmospheric conditions such as wind and air temperature. Changes in atmospheric conditions will result in changes in the surface temperatures, thus creating a less stable system on the subannual scale. The bottom depths, however, are not as strongly related to the atmospheric conditions, and this is particularly the case when lake stratification has occurred. A substantial change in surface temperatures would be required – or a minor change for an extended period of time – to have a significant impact on the bottom temperatures, resulting in much more stable conditions at the deeper depths.

Finally we note that global statistics do not necessary reflect localized patterns in the time series. To see this, let us consider the daily coefficients. These coefficients correspond to what is contributing to the level $j = 1$, 2 and 3 wavelet variances. Figure 4 shows these variances track each other quite closely at depths of 1 and 5 m and, to a lesser extent, at depths of 15 and 20 m. The upper left-hand parts of Figs. 5 and 7 indicate that this global similarity for 1 and 5 m does not translate into similarity in localized variability or significant correlation between daily coefficients. By contrast, the global similarity for 15 and 20 m exhibited at the three wavelet variances is matched by a similar pattern in localized variability in the lower left-hand plots of Fig. 5 and by significant positive correlations in the lower right-hand plot of Fig. 7. The fact that similarity between global summary statistics might or might not correspond to similarity between localized measures stresses the need for a localized analysis such as is afforded by the DSA transform.

6. Summary

As can be see from a cursory examination of Fig. 1, Wivenhoe Dam water temperatures vary in a complex manner across both depth and time. We can simplify the task of describing these data through our proposed DSA decomposition, which is a variation on wavelet-based MRA. The motivation for this variation is to combine components from the usual MRA into components that capture daily, subannual and annual fluctuations. The partitioning afforded by the DSA transform leads to a simple way of quantifying the key sources of variability in the data, yielding a component-based description of how water temperatures vary across time and how they are related at different depths. This approach is largely descriptive, but addresses some of the questions that could be answered more formally through a statistical modeling approach. Our exploratory analysis suggests what components would be needed in a formal depth/time model to address questions of interest to scientists (e.g., how exactly the thermocline manifests itself across depth/time in terms of correlations). An item for future work is to study the other water quality indicators collected by the profiling system (particularly chlorophyll-a, turbidity, dissolved oxygen and specific conductivity) and their relationship to temperature.

In addition to our analysis of water temperatures, our paper makes four technical contributions. We propose a frequency-domian method for constructing a filter that is collectively combines the wavelet coefficients across different levels into a single set of coefficients that can be used to track inhomogeneity of variance across time (Appendix B). We devise a scheme for filling in gaps in the water temperature data based upon the DSA decomposition (Appendix C), and we propose a method for handling boundary conditions that is appropriate for our data (Appendix D). Finally, we adapt the statistical theory for the standard 'boxcar windowed' wavelet variance estimator to work with a 'Gaussian windowed' variance estimator based upon the daily and subannual coefficients from the DSA transform (Appendix E).

Appendix A: Wavelet-Based Analysis of Time Series

Here we provide some technical details about standard wavelet analysis of time series to complement our discussion in Section 2.1 (see *Percival and Walden*, 2000, for further details using notation consistent with usage below).

The starting point in a wavelet-based analysis of $\{X_t\}$ is a Daubechies wavelet filter $\{\tilde{h}_{1,l}, l = 0, 1, \ldots, L_1 - 1\}$, where, for convenience, we define $\tilde{h}_{1,l} = 0$ for $l < 0$ or $l \geq L_1$. By definition, this filter must satisfy three properties:

$$\sum_l \tilde{h}_{1,l} = 0, \;\; \sum_l \tilde{h}_{1,l}^2 = 1/2 \;\text{ and }\; \sum_l \tilde{h}_{1,l}\tilde{h}_{1,l+2n} = 0, \quad n = \pm 1, \pm 2, \ldots. \tag{A1}$$

We denote the transfer function (i.e., discrete Fourier transform (DFT)) for $\{\tilde{h}_{1,l}\}$ by

$$\widetilde{H}_1(f) \equiv \sum_l \tilde{h}_{1,l} e^{-i2\pi f l},$$

and its associated squared gain function by $\widetilde{\mathcal{H}}_1(f) \equiv |\widetilde{H}_1(f)|^2$. Both functions are periodic with a period of unity, and, since $\widetilde{H}_1(-f) = \widetilde{H}_1^*(f)$ and $\widetilde{\mathcal{H}}_1(-f) = \widetilde{\mathcal{H}}_1(f)$, we need only be concerned about $f \in [0, 1/2]$ (here z^* denotes the complex conjugate of z). The wavelet filter in turn is used to define a scaling filter $\tilde{g}_{1,l} \equiv (-1)^{l+1}\tilde{h}_{1,L_1-l-1}$. We denote its corresponding transfer and squared gain functions by $\widetilde{G}_1(f)$ and $\widetilde{\mathcal{G}}_1(f)$. (In dealing with a time series with a sampling interval of $\Delta \neq 1$, we must map the interval $[0, 1/2]$ of standardized frequencies over to the interval $[0, 1/(2\,\Delta)]$ of physically meaningful frequencies. It is convenient to let $f \in [0, 1/2]$ denote a standardized frequency in what follows, but then to let $f \in [0, 1/(2\,\Delta)]$ denote a physical frequency when dealing with an actual time series.)

The simplest wavelet filter is the Haar wavelet filter, which has width $L_1 = 2$ and filter coefficients $\tilde{h}_{1,0} = 1/2$ and $\tilde{h}_{1,1} = -1/2$. The Haar scaling filter is $\tilde{g}_{1,0} = \tilde{g}_{1,1} = 1/2$, The squared gain functions for the Haar wavelet and scaling filters are given by $\widetilde{\mathcal{H}}_1(f) = \sin^2(\pi f)$ and $\widetilde{\mathcal{G}}_1(f) = \cos^2(\pi f)$. These functions are shown in Figure 8(a) versus $f \in [0, 1/2]$. The wavelet filter is a high-pass filter with a nominal pass-band defined by $f \in (1/4, 1/2]$, whereas $\{\tilde{g}_{1,l}\}$ is a low-pass filter with pass-band dictated by $f \in [0, 1/4]$. Note that, for all f,

$$\widetilde{\mathcal{H}}_1(f) + \widetilde{\mathcal{G}}_1(f) = 1. \tag{A2}$$

Figure 8(b) shows the squared gain functions for the Daubechies 'least asymmetric' (LA) wavelet and scaling filters of width $L_1 = 8$, which are the ones used in the analysis presented in this paper. These filters are better approximations to ideal high- and low-pass filters than the Haar filters, where the ideal filters would have

$$\mathcal{H}_1(f) = \begin{cases} 0, & f \in [0, 1/4], \\ 1, & f \in (1/4, 1/2], \end{cases} \quad \text{and} \quad \mathcal{G}_1(f) = \begin{cases} 1, & f \in [0, 1/4], \\ 0, & f \in (1/4, 1/2]. \end{cases}$$

The figure also suggests that Equation (A2) still holds for the LA(8) filters, which in fact is true for all wavelet and related scaling filters.

Using just the basic wavelet and scaling filters $\{\tilde{h}_{1,l}\}$ and $\{\tilde{g}_{1,l}\}$, we can create so-called 'higher-level' wavelet and scaling filters. We denote these by $\{\tilde{h}_{j,l}, l = 0, 1, \ldots, L_j - 1\}$ and

$\{\tilde{g}_{j,l}, l = 0, 1, \ldots, L_j - 1\}$, where $j = 2, 3, \ldots$ is the level index, and $L_j = (2^j - 1)(L_1 - 1) + 1$ (the basic filters are thus associated with level $j = 1$). We denote the squared gain functions for these filters by $\widetilde{\mathcal{H}}_j(f)$ and $\widetilde{\mathcal{G}}_j(f)$. The filter $\{\tilde{h}_{j,l}\}$ is approximately a band-pass filter with a pass-band given by $f \in (1/2^{j+1}, 1/2^j]$, while $\{\tilde{g}_{j,l}\}$ is approximately a low-pass filter with pass-band $f \in [0, 1/2^{j+1}]$. An extension to Equation (A2) is

$$\sum_{j=1}^{J_0} \widetilde{\mathcal{H}}_j(f) + \widetilde{\mathcal{G}}_{J_0}(f) = 1 \tag{A3}$$

for all f and any $J_0 \geq 1$. The plausibility of this equation for the LA(8) wavelet is illustrated in the top portion of Fig. 9.

Upon filtering $\{X_t\}$ with $\{\tilde{h}_{j,l}\}$, $j = 1, \ldots, J_0$, and $\{\tilde{g}_{J_0,l}\}$, we obtain the MODWT wavelet and scaling coefficients:

$$\widetilde{W}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} X_{t-l \bmod N} \text{ and } \widetilde{V}_{J_0,t} \equiv \sum_{l=0}^{L_j-1} \tilde{g}_{J_0,l} X_{t-l \bmod N}, \quad t = 0, 1, \ldots, N-1,$$

which form the elements of the vectors $\widetilde{\mathbf{W}}_j$ and $\widetilde{\mathbf{V}}_{J_0}$; here '$t - l \bmod N$' should be interpreted as '$(t - l) \bmod N$' (for integer u, we define $u \bmod N$ to be u if $0 \leq u \leq N - 1$; if not, its definition is $u + nN$, where n is the unique integer such that $0 \leq u + nN \leq N - 1$). While creating $\widetilde{W}_{j,t}$ formally involves L_j values from the time series, many of the $\tilde{h}_{j,l}$ coefficients are quite close to zero. The effective width of $\{\tilde{h}_{j,l}\}$ is 2^j, which is better indication than L_j of how much of the time series is influencing $\widetilde{W}_{j,t}$ (likewise, the effective width of $\{\tilde{g}_{J_0,l}\}$ is 2^{J_0}). The first $L_j - 1$ coefficients of $\widetilde{\mathbf{W}}_j$ involve a linear combination of values from both the beginning and end of the time series, as do the first $L_{J_0} - 1$ coefficients of $\widetilde{\mathbf{V}}_{J_0}$ These so-called 'boundary' coefficients can be difficult to interpret and hence merit further consideration (see Appendix D). The relationship between the vectors $\widetilde{\mathbf{W}}_j$ and $\mathbf{X}$ can be expressed as $\widetilde{\mathbf{W}}_j = \widetilde{\mathcal{W}}_j \mathbf{X}$, where $\widetilde{\mathcal{W}}_j$ is an $N \times N$ matrix whose elements are dictated by the filter $\{\tilde{h}_{j,l}\}$; likewise, we can write $\widetilde{\mathbf{V}}_{J_0} = \widetilde{\mathcal{V}}_{J_0}\mathbf{X}$, where the matrix $\widetilde{\mathcal{V}}_{J_0}$ depends just on $\{\tilde{g}_{J_0,l}\}$. Stacking $\widetilde{\mathcal{W}}_1$, $\widetilde{\mathcal{W}}_2$, …, $\widetilde{\mathcal{W}}_{J_0}$ and $\widetilde{\mathcal{V}}_{J_0}$ together yields the $(J_0+1)N \times N$ matrix $\widetilde{\mathcal{W}}$ in Equation (1) expressing the MODWT.

Two key descriptors for a time series that the MODWT provides are the ANOVA of Equation (2) and the MRA of Equation (5). The ANOVA follows from an application of Parseval's theorem and Equation (A3). As noted in the discussion surrounding Equations (4) and (5), appropriate partitioning of $\widetilde{\mathcal{W}}$ and $\widetilde{\mathbf{W}}$ yields the details and smooth comprising the MRA, namely,

$$\widetilde{\mathcal{D}}_j = \widetilde{\mathcal{W}}_j^T \widetilde{\mathbf{W}}_j \ \text{ and } \ \widetilde{\mathcal{S}}_{J_0} = \widetilde{\mathcal{V}}_{J_0}^T \widetilde{\mathbf{V}}_{J_0}.$$

Based upon the above, we can write the tth elements $\widetilde{\mathcal{D}}_{j,t}$ and $\widetilde{\mathcal{S}}_{J_0,t}$ of $\widetilde{\mathcal{D}}_j$ and $\widetilde{\mathcal{S}}_{J_0}$ explicitly as

$$\widetilde{\mathcal{D}}_{j,t} \equiv \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} \widetilde{W}_{j,t+l \bmod N} \ \text{ and } \ \widetilde{\mathcal{S}}_{J_0,t} \equiv \sum_{l=0}^{L_{J_0}-1} \tilde{g}_{J_0,l} \widetilde{V}_{J_0,t+l \bmod N}, \quad t = 0, 1, \ldots, N-1.$$

The components of an MRA are intended to capture distinct aspects of a time series and ideally should be approximately pairwise uncorrelated (the approximation improves as the width L_1 is increased, which is one reason for preferring the LA(8) wavelet over the Haar wavelet).

Appendix B: Construction of DSA Transform

We can cast the DSA transform as a special case of the following theorem (the proof of which is in *Percival*, 2010).

Theorem 1: *Let $\{X_t, t = 0, 1, \ldots, N-1\}$ be a real-valued time series, and let $\{a_{m,l}\}$, $m = 1, \ldots, M$, be a set of M filters with corresponding squared gain functions $\mathcal{A}_m(f)$ such that $\sum_{m=1}^{M} \mathcal{A}_m(\frac{k}{N}) = 1$, $k = 0, 1, \ldots, N-1$. Define $Y_{m,t} \equiv \sum_l a_{m,l} X_{t-l \bmod N-1}$ and $Z_{m,t} \equiv$*

$\sum_l a_{m,l} Y_{m,t+l \bmod N-1}$, $t = 0, 1, \ldots, N-1$. *Then we have the following decompositions:*

$$\sum_{m=1}^{M} \sum_{t=0}^{N-1} Y_{m,t}^2 = \sum_{t=0}^{N-1} X_t^2 \ \textit{and} \ \sum_{m=1}^{M} Z_{m,t} = X_t.$$

We note in passing that the component $\{Z_{m,t}\}$ of the additive decomposition depends only on the squared gain function $\mathcal{A}_m(f)$ and not on the phase function for the filter $\{a_{m,l}\}$.

To construct the DSA transform, consider the following three squared gain functions:

$$\mathcal{A}_1(f) \equiv \sum_{j=1}^{3} \widetilde{\mathcal{H}}_j(f), \ \ \mathcal{A}_2(f) \equiv \sum_{j=4}^{9} \widetilde{\mathcal{H}}_j(f) \ \text{ and } \ \mathcal{A}_3(f) \equiv \widetilde{\mathcal{G}}_9(f).$$

It follows from Equation (A3) with $J_0 = 9$ that $\mathcal{A}_1(f) + \mathcal{A}_2(f) + \mathcal{A}_3(f) = 1$ for all f, as required by Theorem 1. The bottom part of Figure 9 shows $\mathcal{A}_1(f)$, $\mathcal{A}_2(f)$ and $\mathcal{A}_3(f)$ based upon the $\widetilde{\mathcal{H}}_j(f)$'s and $\widetilde{\mathcal{G}}_9(f)$ arising from the LA(8) filters. The corresponding filtering operations are implemented in the frequency domain by simply multiplying the DFT of $\{X_t\}$ (denote this as $\{\mathcal{X}_k\}$) by $\{\mathcal{A}_m^{1/2}(\frac{k}{N}), k = 0, 1, \ldots, N-1\}$ and then taking the inverse DFT of the resulting sequence $\{\mathcal{A}_m^{1/2}(\frac{k}{N})\mathcal{X}_k\}$. The squared gain function for each implicitly defined filter $\{a_{m,l}\}$, $m = 1, 2$ and 3, is given by $\{\mathcal{A}_m(\frac{k}{N})\}$, hence satisfying the conditions required by Theorem 1 and thus providing the desired sum of squares decomposition stated by Equation (7). The outputs from the filtering operations that are obtained from the inverse DFTs form the elements of the N-dimensional vectors $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$. An additional advantage of this frequency-domain approach is that the filters have a zero phase function, which makes it easy to align the elements of $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$ with those of $\mathbf{X}$; for details, see Section 4.8 of *Percival and Walden*, 2000.

Appendix C: Gap-Filling via Stochastic Interpolation

The dam water temperature measurements have time-varying features acting on the daily and subannual components of the DSA decomposition. Any gap-filling scheme must pay careful attention to what is going on around each gap in both components. In addition, filling in the gaps using realizations from locally adapted stochastic models allows us to evaluate the effect of the gap-filling scheme by generating many different realizations.

Accordingly, we start by linearly interpolating the gappy time series to produce a gap-free series, which we subject to the DSA decomposition. Noting the start/stop locations of a particular gap in the original time series, we then go to the same locations in the $\mathcal{D}$ and $\mathcal{S}$ components. In the case of $\mathcal{D}$, we locate K values before – and K values after – the start/stop locations in $\mathcal{D}$ that correspond to actual measured values in the original time series (we set $K = 36$ so that data from at least three days before and after the gap are utilized). Using least squares, we then fit a harmonic model to these $2K$ values using sine and cosine terms with a fundamental frequency of 1 cycle per day and with $L = 3$ of its harmonics. The values currently in the gap in $\mathcal{D}$ are replaced by an extrapolation from the fitted harmonic model, with the addition of a sample from a Gaussian white noise process whose variance is dictated by the sum of squares of the residuals from the least squares fit.

In the case of $\mathcal{S}$, spectral analysis of its various subseries suggests that the correlation structure is relatively constant across time, but that this component is subject to fluctuations in its variance. Accordingly, we fill in a gap by sampling from a multivariate Gaussian distribution with a mean vector and covariance matrix dictated by (1) conditioning on the two values observed just before and after the gap, (2) an estimate of the autocorrelation sequence for $\mathcal{S}$ and (3) a localized variance estimate based on $K = 36$ actual values before – and K values after – the start/stop locations of the gap (for details, see Appendix B of *Percival et al., 2008*).

Letting $\widetilde{\mathcal{D}}$ and $\widetilde{\mathcal{S}}$ represent the altered versions of $\mathcal{D}$ and $\mathcal{S}$, the gap-filled series is taken to be $\widetilde{\mathcal{D}}+\widetilde{\mathcal{S}}+\mathcal{A}$, where $\mathcal{A}$ is from the DSA decomposition of the linearly interpolated series (note that the components in the DSA decomposition of the gap-filled time series will not in general be equal to $\widetilde{\mathcal{D}}$, $\widetilde{\mathcal{S}}$ and $\mathcal{A}$). Figure 11 shows three stochastic interpolations of the dam water temperature series at 10 m. While this figure indicates that the gappy filling procedure is visually reasonable, the scheme is inherently univariate and cannot mimic cross-correlations in the time series at different depths. This defect is mitigated somewhat by the facts that, with a few notable exceptions, most of the gap lengths are small and that any assessment that an observed cross-correlation based on gap-filled data is significantly different from zero will tend to be conservative.

Appendix D: Boundary Conditions for Wavelet Transforms

As is true for the discrete Fourier transform, the MODWT and the DSA transform treat a time series $\{X_t, t = 0, 1, \ldots, N-1\}$ such that X_t for $t < 0$ or $t \geq N$ is implicitly defined to be $X_{t \bmod N}$; i.e., the unobserved values $X_{-1}, X_{-2}, \ldots$ that are needed to compute certain transform coefficients are taken to be equal to $X_{N-1}, X_{N-2}, \ldots$. If there is a significant mismatch between the beginning and end of a time series, certain transform coefficients (termed 'boundary' coefficients) can be adversely affected, leading to undesirable artifacts in the wavelet-based MRA or DSA decomposition near $t = 0$ and $t = N-1$. To reduce these artifacts, we need better surrogates for $X_{-1}, X_{-2}, \ldots$.

One approach that sometimes yields better surrogates is to form a new time series of length $2N$ by taking $\{X_t\}$ and tacking on its time-reversed version, yielding

$$\{X'_t, t = 0, 1, \ldots, 2N-1\} \equiv \{X_0, X_1, \ldots, X_{N-2}, X_{N-1}, X_{N-1}, X_{N-2}, \ldots, X_1, X_0\}.$$

The boundary coefficients for $\{X'_t\}$ should be less prone to introducing artifacts in an MRA or DSA decomposition because the beginning and end of $\{X'_t\}$ might match up better than those for $\{X_t\}$. For the time series of interest here, the reflection trick does not work well because of rapid increases or decreases at the beginning and/or end of some series, leading to an undesirable cusp in $\{X'_t\}$. We can handle a increase or decrease that is approximately linear at the end of $\{X_t\}$ by tacking on a reversed and flipped upside-down version of the original series; i.e., we construct

$$\{Y_t, t = 0, 1, \ldots, 2N-1\} \equiv \{X_0, X_1, \ldots, X_{N-2}, X_{N-1}, c - X_{N-1}, c - X_{N-2}, \ldots, c - X_0\},$$

where c is a constant. To set c, assume $X_t \approx \alpha + \beta t$ for t close to $N-1$. Since $Y_t = X_t$ for $t \leq N-1$ and $Y_t = c - X_{2N-1-t}$ for $t \geq N$, setting $c = 2\alpha + \beta(2N-1)$ ensures continuity of the approximation across the two regions. In particular, if α and β are determined solely based upon X_{N-2} and X_{N-1}, then $c = 3X_{N-1} - X_{N-2}$. We can handle the fact that the beginning and end of $\{Y_t\}$ need not match up by tacking on its time-reversed version to create a series $\{Y'_t\}$ of length $4N$ for use with the MODWT or DSA transform.

To handle approximate linear increases or decreases at both ends of $\{X_t\}$, we construct the following time series $\{Z_t\}$ of length $3N$:

$$\{a - X_{N-1}, \ldots, a - X_1, a - X_0, X_0, X_1, \ldots, X_{N-2}, X_{N-1}, b - X_{N-1}, b - X_{N-2}, \ldots, b - X_0\},$$

where a and b are constants that can be set as before. Assuming $X_t \approx \alpha_0 + \beta_0 t$ for t close to 0 and $X_t \approx \alpha_1 + \beta_1 t$ for t close to $N-1$, the appropriate settings are $a = 2\alpha_0 - \beta_0$

and $b = 2\alpha_1 + \beta_1(2N - 1)$. Determination of α_0 and β_0 using just X_0 and X_1 yields $a = 3X_0 - X_1$; likewise, $b = 3X_{N-1} - X_{N-2}$ when based upon just X_{N-2} and X_{N-1}. Again, for use with the MODWT or DSA transform, we can tack on a time-reserved version of $\{Z_t\}$ to create a series $\{Z'_t\}$ of length $6N$.

Although formally the MRAs and DSA decompositions for $\{X'_t\}$, $\{Y'_t\}$ and $\{Z'_t\}$ consist of components of length, respectively, $2N$, $4N$ and $6N$, we need only extract those portions that correspond to the original series $\{X_t\}$; i.e., the portions corresponding to $\{X'_t, t = 0, 1, \ldots, N-1\}$, $\{Y'_t, t = 0, 1, \ldots, N-1\}$ and $\{Z'_t, t = N, N+1, \ldots, 2N-1\}$. Figure 10 compares the $\mathcal{A}$ component of the DSA decomposition (i.e., the smooth $\widetilde{\mathcal{S}}_9$ of the corresponding MRA) based upon $\{X_t\}$, $\{X'_t\}$, $\{Y'_t\}$ and $\{Z'_t\}$ (with $a = 3X_0 - X_1$ and $b = c = 3X_{N-1} - X_{N-2}$). Arguably the component based upon $\{Z'_t\}$ gives the best representation of the large-scale behavior of the time series at its beginning and end.

The different definitions for the boundary coefficients have an impact on the wavelet-based analysis of variance. For the reflection-based approach, the sample means $\overline{X}$ for $\{X_t\}$ and $\{X'_t\}$ are identical by construction, as are their sample variances $\hat{\sigma}^2_X$, so the empirical wavelet variance for $\{X'_t\}$ can serve as an analysis of the sample variance of the original series also. The sample variances of $\{Y'_t\}$ and $\{Z'_t\}$, say $\hat{\sigma}^2_Y$ and $\hat{\sigma}^2_Z$, are related to $\hat{\sigma}^2_X$ via

$$\begin{aligned}
\hat{\sigma}^2_Y &= \hat{\sigma}^2_X + \frac{c^2}{4} - c\overline{X} + \overline{X}^2 \\
\hat{\sigma}^2_Z &= \hat{\sigma}^2_X + \frac{2(a^2 - ab + b^2) - 4(a+b)\overline{X} + 8\overline{X}^2}{9}.
\end{aligned}$$

We can use these equations to translate the wavelet-based decomposition of $\hat{\sigma}^2_Y$ or $\hat{\sigma}^2_Z$ into a decomposition of $\hat{\sigma}^2_X$ if we are willing to make the ad hoc assumption that the correction terms should be applied solely to the ANOVA component due to $\hat{\sigma}^2_0$ in Equation (3) or to $\hat{\sigma}_A^2$ in Equation (8). The justification for this assumption is that the obvious difference between $\{X_t\}$ and either $\{Y_t'\}$ or $\{Z_t'\}$ is in the artificial creation of large-scale variations in the latter, and such variations are captured by the scaling coefficients in the MODWT or the annual coefficients in the DSA transform.

Appendix E: Variance Estimators Based on Daily and Subannual Coefficients

Let C_t stand for the tth element of either the daily coefficients $\mathbf{D}$ or subannual coefficients $\mathbf{S}$ from the DSA transform of a water temperature time series $\mathbf{X}$. Consider a weighted sum of squares of M consecutive coefficients, which, for convenience (and without loss of generality), we take to be indexed by $t = 0, 1, \ldots, M-1$:

$$\hat{\sigma}_C^2 \equiv \sum_{t=0}^{M-1} g_t C_t^2$$

where the g_t's are a set of nonnegative weights such that $\sum_t g_t = 1$; here we set $M = 801$ and set g_t approximately equal to $f(t-400)$, where f is the probability density function for a Gaussian random variable (RV) with mean zero and variance σ^2, with the choice $\sigma = 180/\sqrt{\pi}$ giving a bandwidth measure $\Delta / \sum_t g_t^2$ of 30 days (recall that $\Delta = 2$ hours; the g_t weights are very close to $f(t-400)$ – but not exactly so – because they are actually generated via convolutions carried out in the frequency domain). Under the assumption that the observed coefficients are a realization of a portion $C_0, C_1, \ldots, C_{M-1}$ of a stationary process with mean zero and variance σ_C^2, we can regard $\hat{\sigma}_C^2$ as an estimator of the variance σ_C^2 (the assumption that the process has zero mean is reasonable because of differencing operations embedded in the wavelet filters used to construct the DSA transform – for details, see Chapter 8, *Percival and Walden*, 2000). Following a standard approach, we assume that $\hat{\sigma}_C^2$ has approximately the same distribution as the RV $\sigma_C^2 \chi_\eta^2/\eta$, where χ_η^2 is a chi-square RV with η degrees of freedom. We can estimate η via

$$\hat{\eta} = \frac{\hat{\sigma}_C^4}{\sum_{\tau=-(m-1)}^{m-1} \hat{s}_\tau^2 \sum_{l=0}^{M-|\tau|-1} g_{l+|\tau|} g_l},$$

where $\hat{s}_\tau$ is the biased estimator of the autocovariance sequence for $C_0, C_1, \ldots, C_{M-1}$ after multiplication by a Parzen lag window:

$$\hat{s}_\tau = \frac{w_{m,\tau}}{M} \sum_{t=0}^{M-|\tau|-1} C_{t+|\tau|} C_t \text{ and } w_{m,\tau} = \begin{cases} 1 - 6\left(\tau/m\right)^2 + 6\left(|\tau|/m\right)^3, & |\tau| \leq m/2; \\ 2\left(1 - |\tau|/m\right)^3, & m/2 < |\tau| < m; \\ 0, & |\tau| \geq m; \end{cases}$$

here we set $m = 30$. The approximate 95% confidence intervals for the various σ_C^2 shown in Figure 5 are given by

$$\left[\frac{\hat{\eta}\hat{\sigma}_C^2}{Q_{\hat{\eta}}(0.975)}, \frac{\hat{\eta}\hat{\sigma}_C^2}{Q_{\hat{\eta}}(0.025)}\right],$$

where $Q_\eta(p)$ is the $p \times 100\%$ percentage point from the χ_η^2 distribution.

Acknowledgments. The first author acknowledges support from CSIRO for a visiting scientist position that allowed him to carry out the work described in this paper. The authors also thank Sean Gibson of Seqwater for his efforts in data preparation.

References

Beylkin, G., On the representation of operators in bases of compactly supported wavelets, *SIAM Journal on Numerical Analysis*, *29*, 1716–1740, 1992.

Bruce, A. G., and H.–Y. Gao, *Applied Wavelet Analysis with S-PLUS*, Springer, New York, 1996.

Coifman, R. R., and D. L. Donoho, Translation-invariant de-noising, in *Wavelets and Statistics* (Lecture Notes in Statistics, Volume 103), edited by A. Antoniadis and G. Oppenheim, Springer–Verlag, New York, 125–150, 1995.

Davison, A. C., and D. V. Hinkley, *Bootstrap Methods and their Applications*, Cambridge University Press, Cambridge, England, 1997.

Del Marco, S., and J. Weiss, Improved transient signal detection using a wavepacket-based detector with an extended translation-invariant wavelet transform, *IEEE Transactions on Signal Processing*, *45*, 841–850, 1997.

Hall, P., and B. A. Turlach, Interpolation methods for nonlinear wavelet regression with irregularly spaced design, *Annals of Statistics*, *25*, 1912–1925, 1997.

Lang, M., H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, Nonlinear processing of a shift invariant DWT for noise reduction, in *Wavelet Applications II* (Proceedings of the SPIE 2491), edited by H. H. Szu, SPIE Press, Bellingham, Washington, 640–651, 1995.

Liang, J., and T. W. Parks, A translation-invariant wavelet representation algorithm with applications, *IEEE Transactions on Signal Processing*, *44*, 225–232, 1996.

Mondal, D., and D. B. Percival, Wavelet variance analysis for gappy time series, *Annals of the Institute of Statistical Mathematics*, in press, 2010.

Nason, G. P., and B. W. Silverman, The stationary wavelet transform and some statistical applications, in *Wavelets and Statistics* (Lecture Notes in Statistics, Volume 103), edited by A. Antoniadis and G. Oppenheim, Springer–Verlag, New York, 281–299, 1995.

Percival, D. B., Discrete wavelet transforms based on zero-phase Daubechies filters, manuscript in preparation, 2010.

Percival, D. B., and P. Guttorp, Long-Memory processes, the Allan variance and wavelets, in *Wavelets in Geophysics*, edited by E. Foufoula–Georgiou and P. Kumar, Academic Press, San Diego, 325–344, 1994.

Percival, D. B., D. A. Rothrock, A. S. Thorndike, and T. Gneiting, The variance of mean sea-ice thickness: Effect of long-range dependence, *J. Geophys. Res.* 113, C01004, doi:10.1029/2007JC004391, 2008.

Percival, D. B., and A. T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge, England, 2000.

Pesquet, J.–C., H. Krim, and H. Carfantan, Time-invariant orthonormal wavelet representations, *IEEE Transactions on Signal Processing*, *44*, 1964–1970, 1996.

Porto, R. F., P. A. Morettin, D. B. Percival, and E. C. Q. Aubin, Wavelet shrinkage for regression models with random design and correlated errors, under review, 2010.

Sardy, S., D.B. Percival, A.G. Bruce, H.-Y. Gao and W. Stuetzle, Wavelet shrinkage for unequally spaced data, *Statistics and Computing*, *9*, 65–75, 1999.

Shensa, M. J., The discrete wavelet transform: wedding the à trous and Mallat algorithms, *IEEE Transactions on Signal Processing*, *40*, 2464–2482, 1992.

Unser, M., Texture classification and segmentation using wavelet frames, *IEEE Transactions on Image Processing*, *4*, 1549–1560, 1995.

	1 m	5 m	10 m	15 m	20 m
$\hat{\sigma}^2_D$	0.07	0.07	0.03	0.01	0.01
$\hat{\sigma}^2_S$	0.58	0.38	0.08	0.05	0.06
$\hat{\sigma}^2_A$	11.74	10.56	9.88	9.14	7.98
$\hat{\sigma}^2_X$	12.39	11.00	9.99	9.20	8.06
$\hat{\sigma}_D$	0.26	0.26	0.17	0.10	0.11
$\hat{\sigma}_S$	0.76	0.61	0.28	0.23	0.25
$\hat{\sigma}_A$	3.43	3.25	3.14	3.02	2.83
$\hat{\sigma}_X$	3.52	3.32	3.16	3.03	2.84

Table 1. Decomposition of sample variance of water temperature series into variances of daily, subannual and annual coefficients (upper part of table), along with corresponding standard deviations in degrees C (lower part) (see Equation (8)).

D_t

	1 m	5 m	10 m	15 m
5 m	−0.09			
10 m	0.03	0.22		
15 m	0.05	0.01	0.06	
20 m	0.05	−0.18	−0.07	0.12

S_t

	1 m	5 m	10 m	15 m
5 m	0.61			
10 m	0.20	0.48		
15 m	0.21	0.28	0.56	
20 m	0.05	0.04	0.19	0.43

A_t

	1 m	5 m	10 m	15 m
5 m	0.99			
10 m	0.92	0.95		
15 m	0.79	0.84	0.96	
20 m	0.68	0.74	0.89	0.97

X_t

	1 m	5 m	10 m	15 m
5 m	0.97			
10 m	0.89	0.94		
15 m	0.77	0.83	0.96	
20 m	0.66	0.73	0.88	0.97

Table 2. Global cross-correlations between daily, subannual and annual coefficients at different depths, along with cross-correlations between original series.

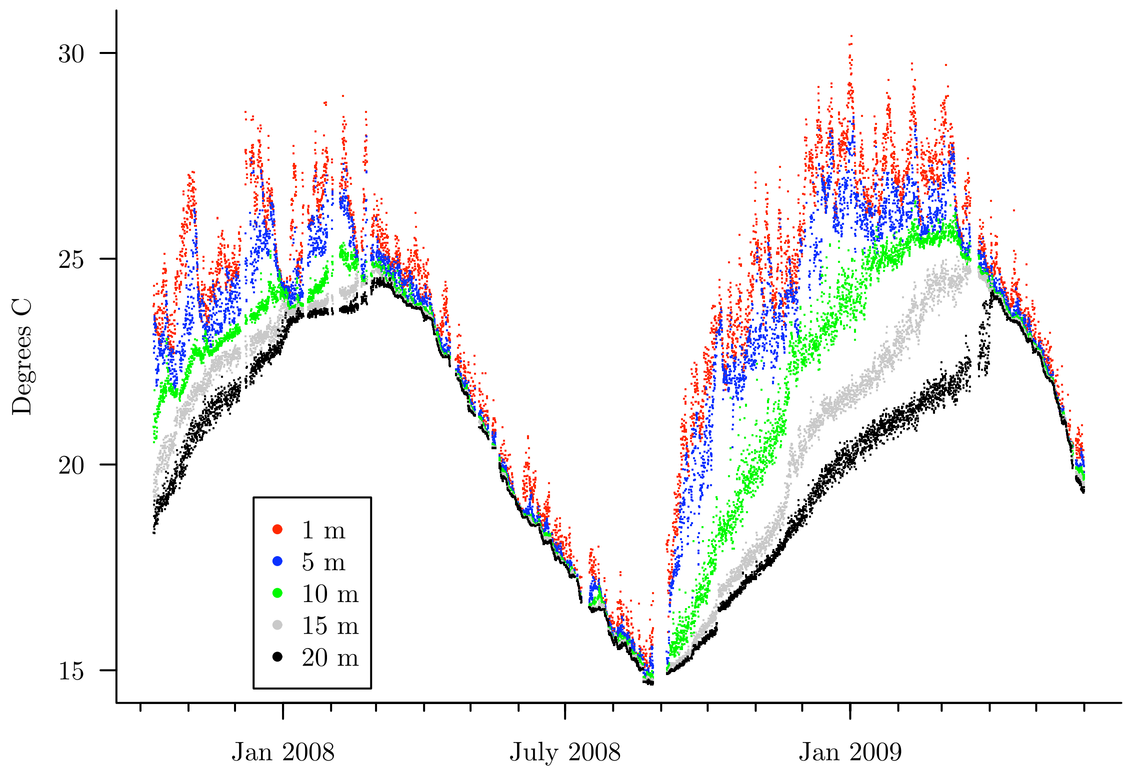

Figure 1. Water temperature time series from Wivenhoe Dam as recorded at five depths.

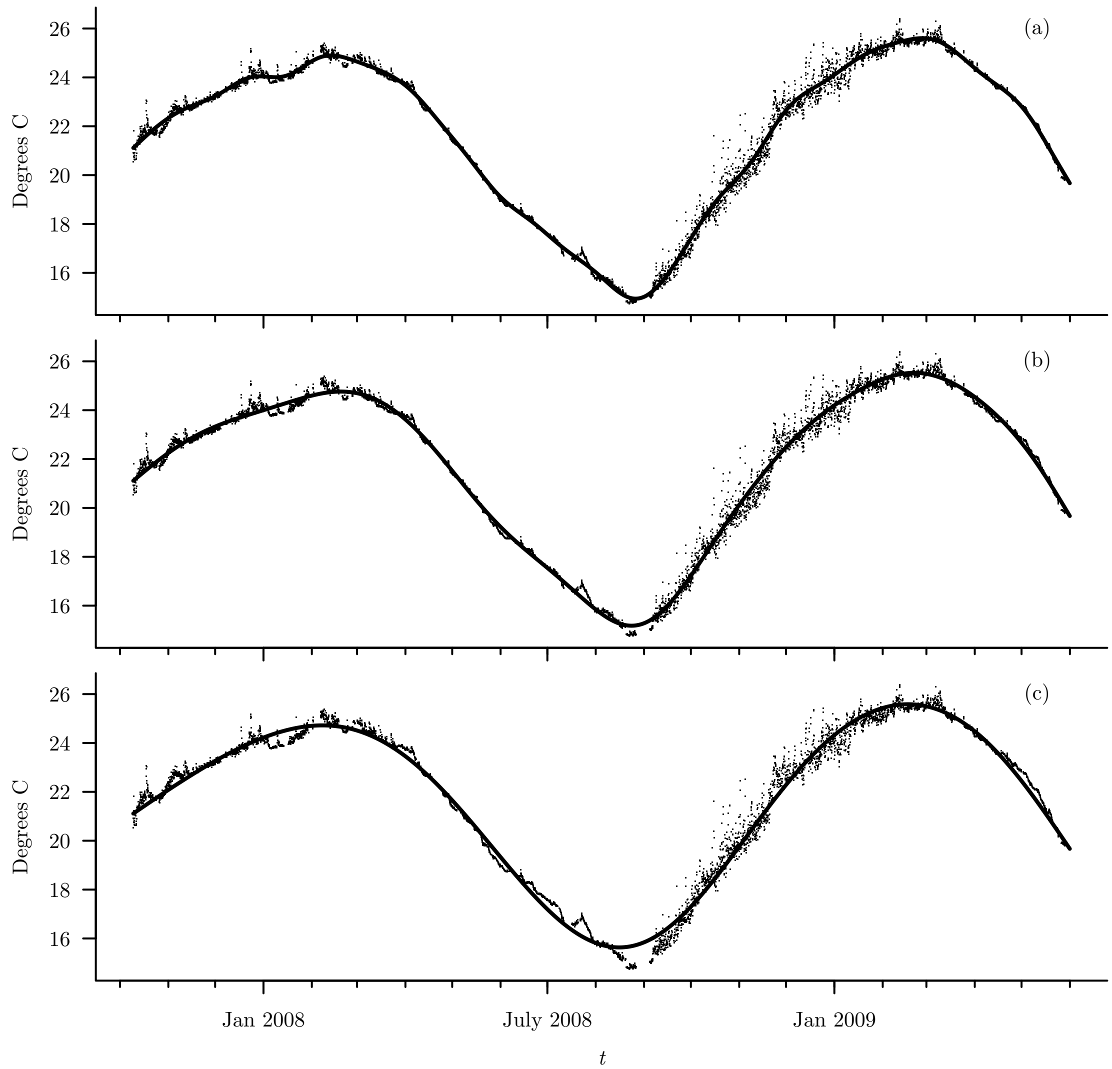

Figure 2. Comparison of smooths $\widetilde{\mathcal{S}}_{J_0}$ (solid curves) based upon the LA(8) wavelet for the 10 m water temperature series (dots) with (a) $J_0 = 8$, (b) $J_0 = 9$ and (c) $J_0 = 10$. Arguably some parts of the $J_0 = 8$ smooth are better regarded as a subannual variation (e.g., the month-long dip following the start of 2008), while the $J_0 = 10$ smooth appears to be oversmoothing the data over some long stretches (e.g. March to July of 2008).

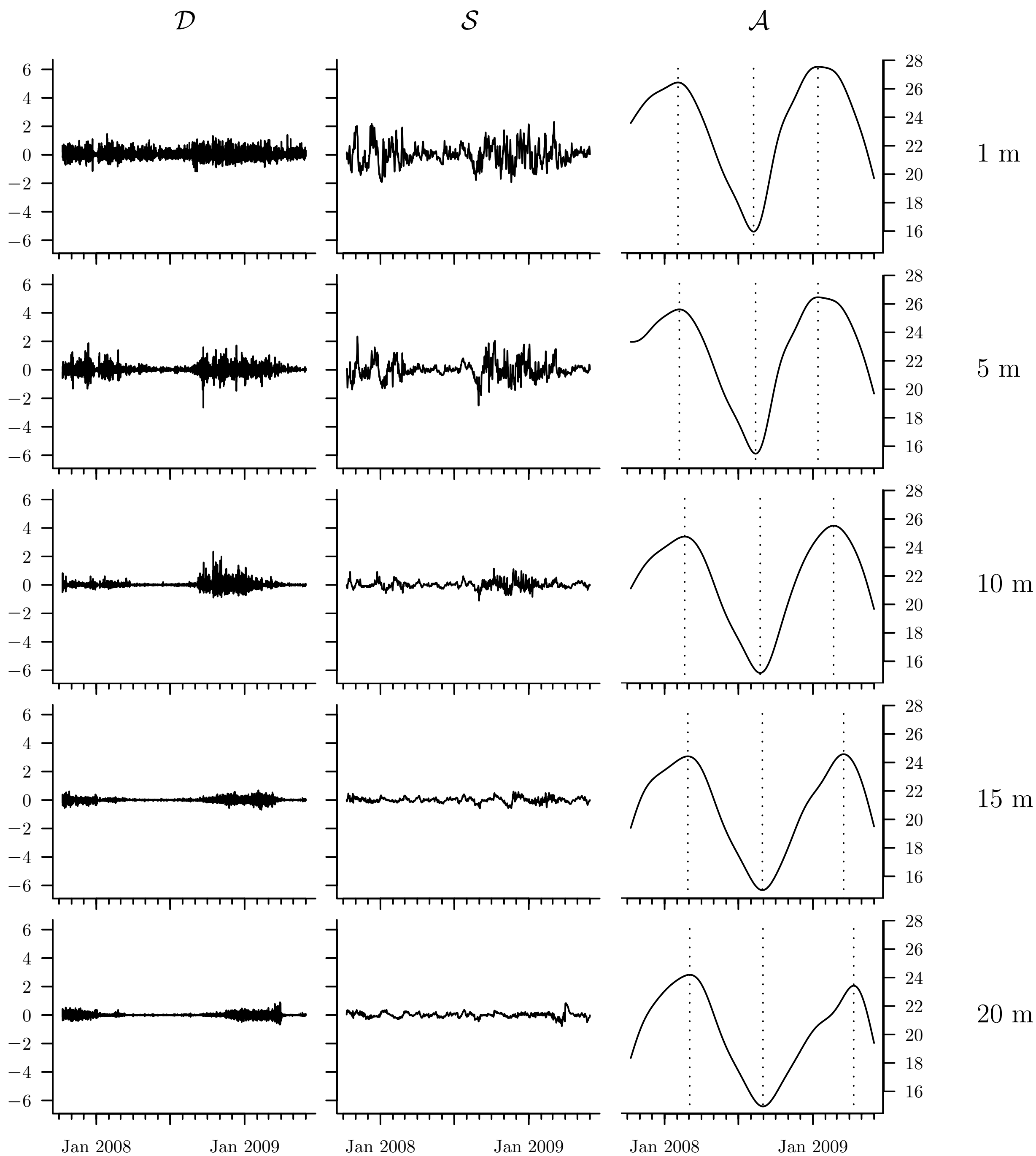

Figure 3. DSA decomposition based upon the LA(8) wavelet for 1, 5, 10, 15 and 20 m depths (top to bottom rows). The daily, subannual and annual components are shown in the left, middle and right columns. The distance between vertical tick marks represents a temperature change of 2 degrees Celsius in all fifteen plots.

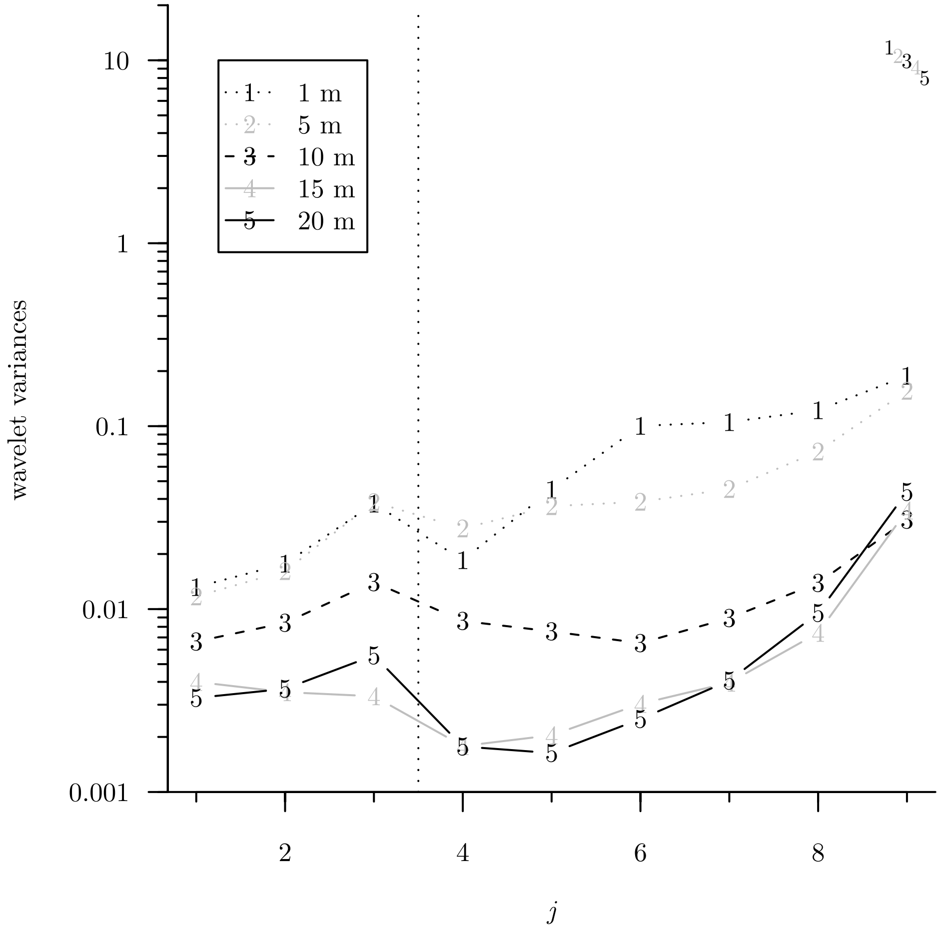

Figure 4. Wavelet variances $\hat{\sigma}_j^2$ based upon the LA(8) wavelet for five depths and nine levels j, along with variances of scaling coefficients $\hat{\sigma}_0^2$ for each depth. The nine wavelet variances for each depth are connected by lines, while the variance of the corresponding scaling coefficients is shown as a single character in the upper right-hand corner of the plot. Wavelet variances indexed by $j = 1, 2$ and 3 make up the daily component in the DSA decomposition (plotted to left of vertical dotted line); the remaining six wavelet variances make up the subannual component. The variance of the scaling coefficients is associated with the annual component. The sum of the nine wavelet variances along with the variance of the scaling coefficients for a particular depth is exactly equal to the variance of the time series for that depth (see Equation (3)).

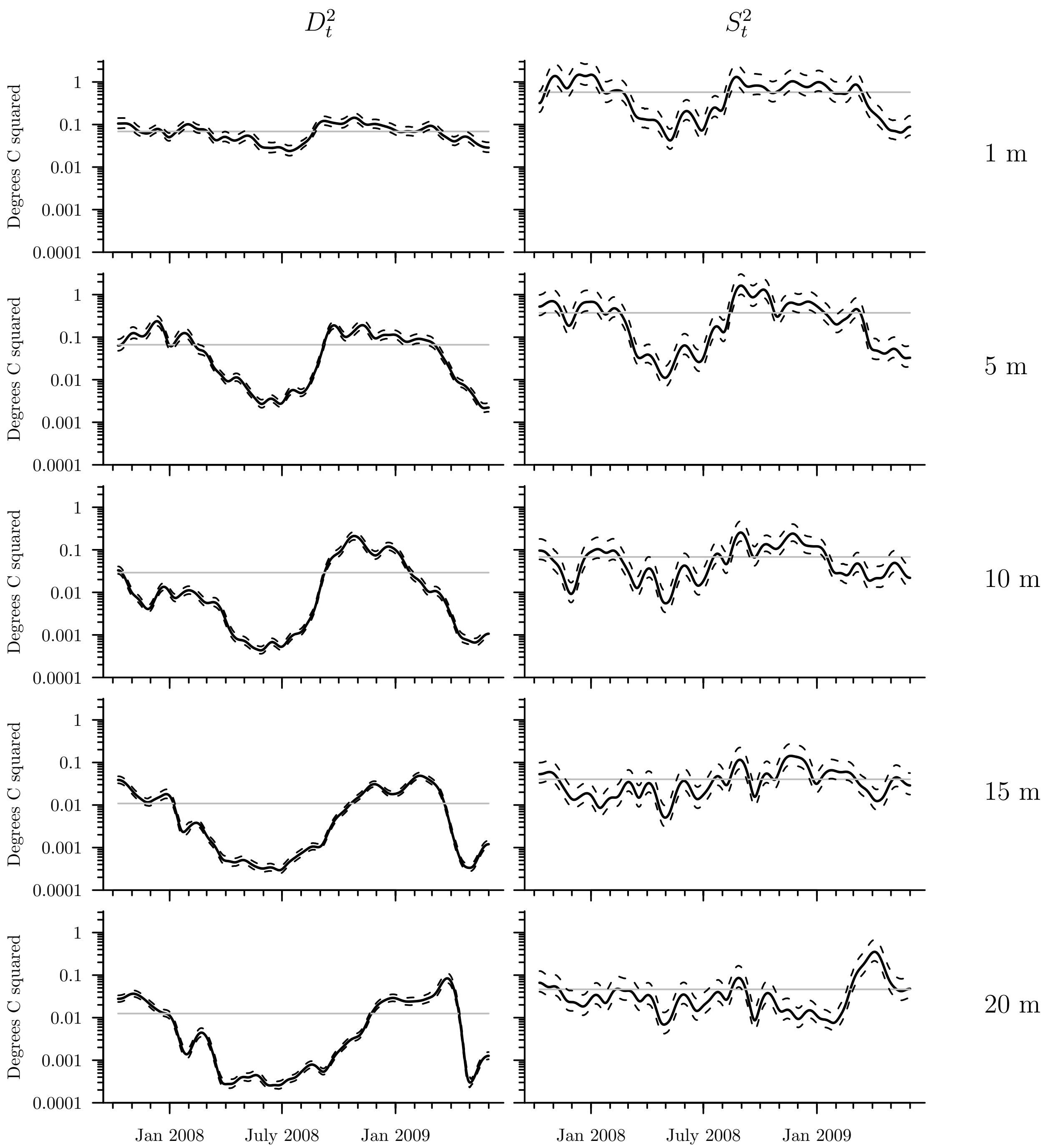

Figure 5. Variance of daily (left-hand column) and subannual (right) coefficients smoothed over 30 days for 5 depths (from top to bottom, 1, 5, 10, 15 and 20 m). The upper and lower dashed lines depict 95% confidence intervals.

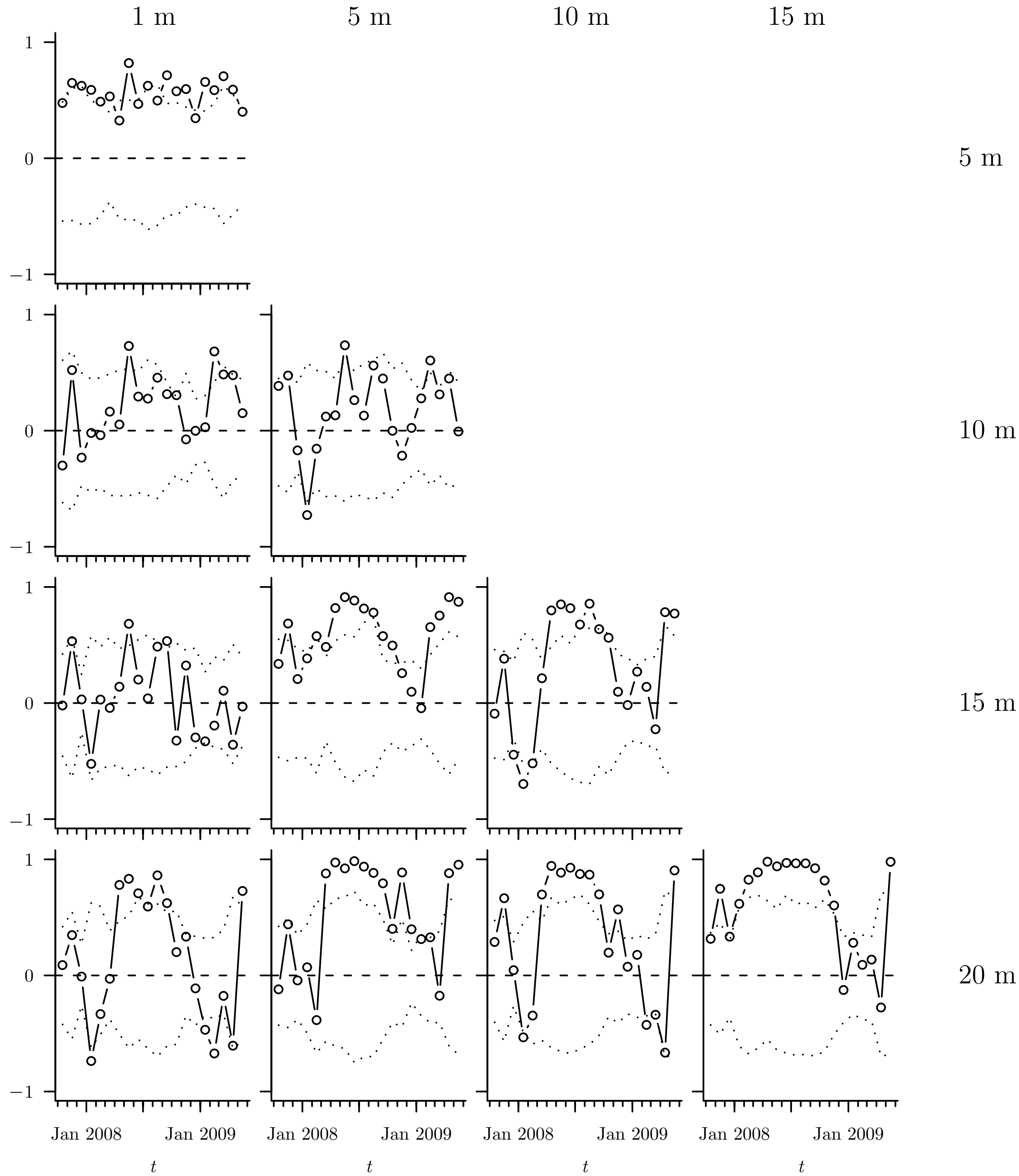

Figure 6. Month-by-month correlations between 5 depths for subannual coefficients (circles). The upper and lower dotted lines depict 95% confidence intervals computed via an autoregressive bootstrapping procedure (*Davison and Hinkley*, 1997) operating under the null hypothesis that the true correlations are zero (i.e., anything correlation falling above (below) the upper (lower) dotted line can be regarded as significantly different from zero at the 95% confidence level).

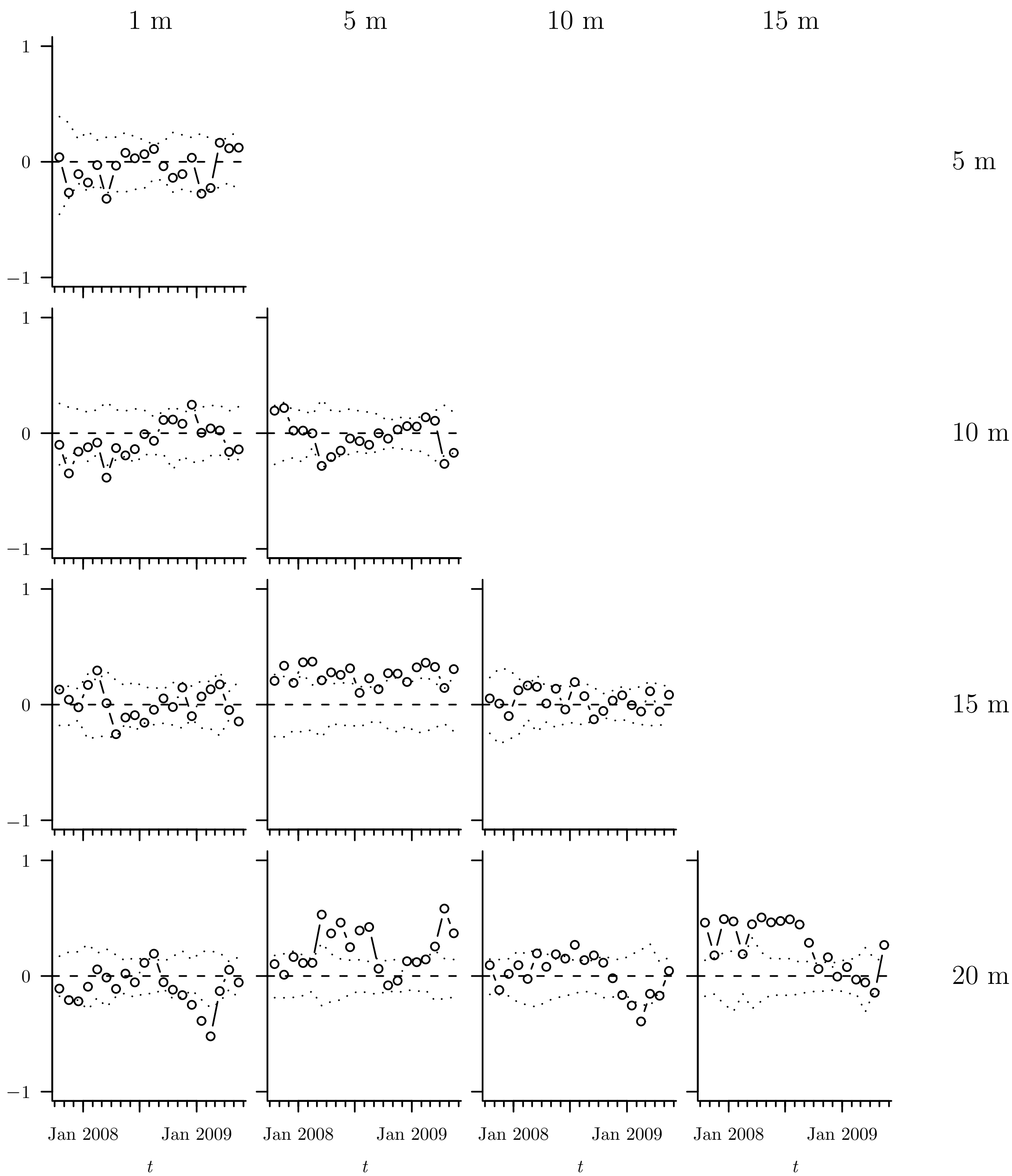

Figure 7. Month-by-month correlations between 5 depths for daily coefficients. The upper and lower dotted lines depict 95% confidence intervals formed in the same manner as in Fig. 6.

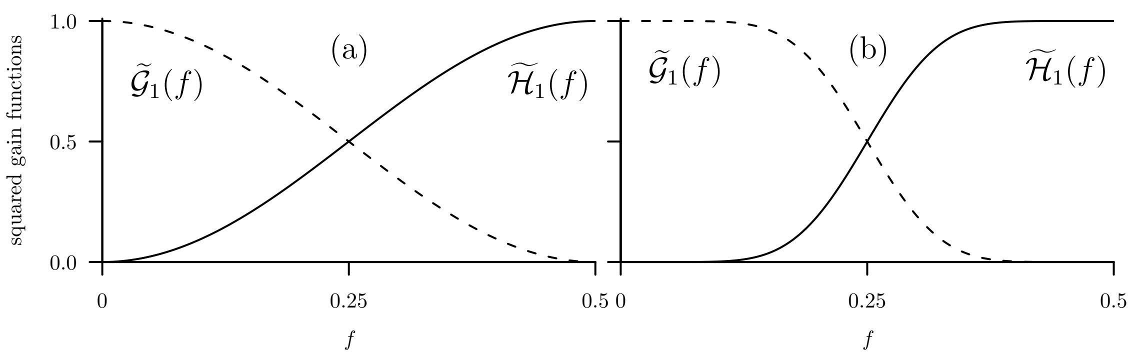

Figure 8. Squared gain functions $\widetilde{\mathcal{H}}_1(f)$ and $\widetilde{\mathcal{G}}_1(f)$ for Haar wavelet and scaling filters (plot (a), solid and dotted curves, respectively). Plot (b) shows the corresponding functions for the LA(8) filters.

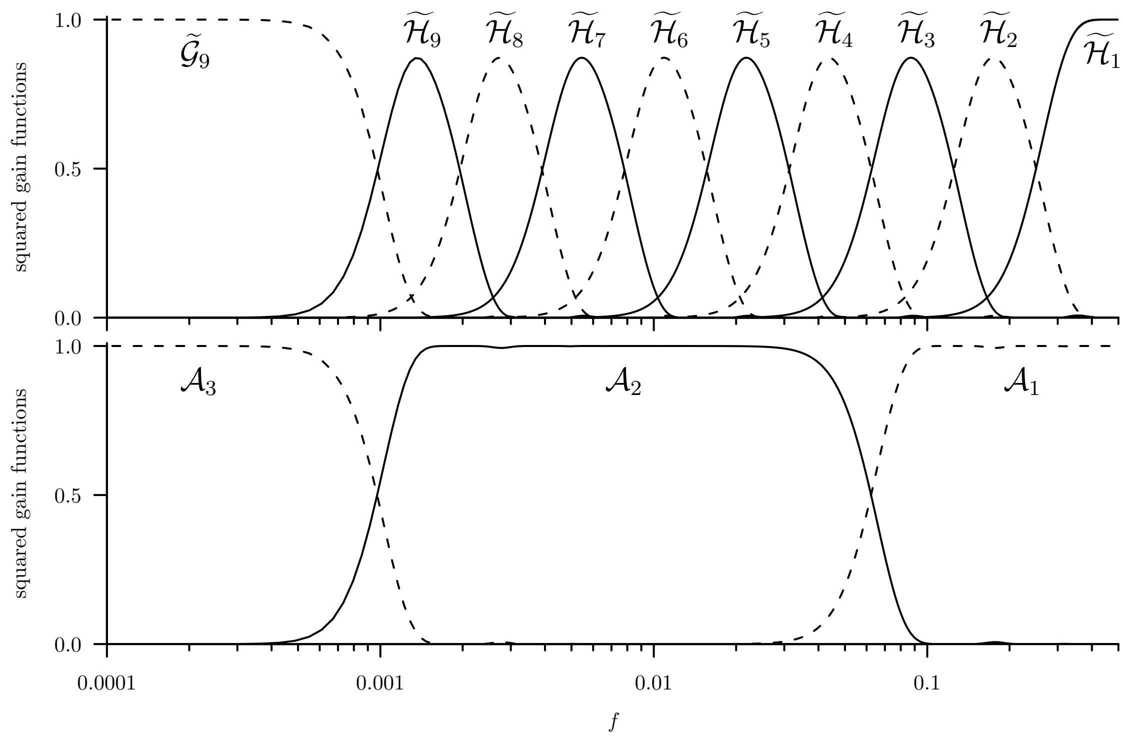

Figure 9. Squared gain functions $\widetilde{\mathcal{H}}_j(f)$, $j = 1, \ldots, 9$, and $\widetilde{\mathcal{G}}_9(f)$ based upon the LA(8) wavelet and filters (top plot, from right to left, alternating solid and dashed curves), and squared gain functions $\mathcal{A}_1(f)$, $\mathcal{A}_2(f)$ and $\mathcal{A}_3(f)$ associated with the DSA decomposition (bottom plot, from right to left).

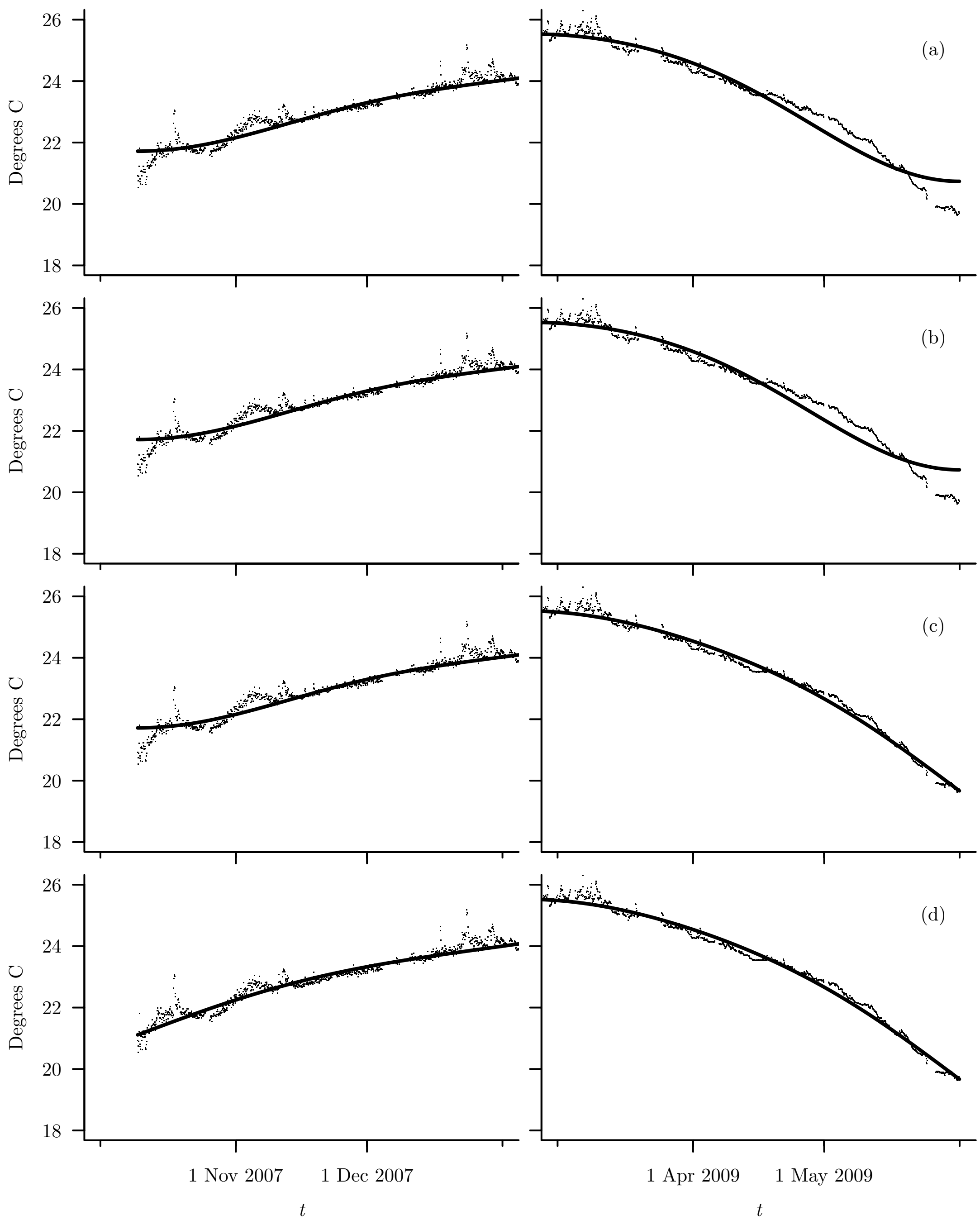

Figure 10. Comparison of beginning (left-hand column) and end (right-hand column) of annual component $\mathcal{A}$ (solid curves) for 10 m time series (dots) created using (a) original series only, (b) series extended by reflection, (c) series extended at end by flipping and reflection and and (d) series extended at beginning and end by flipping and reflection.

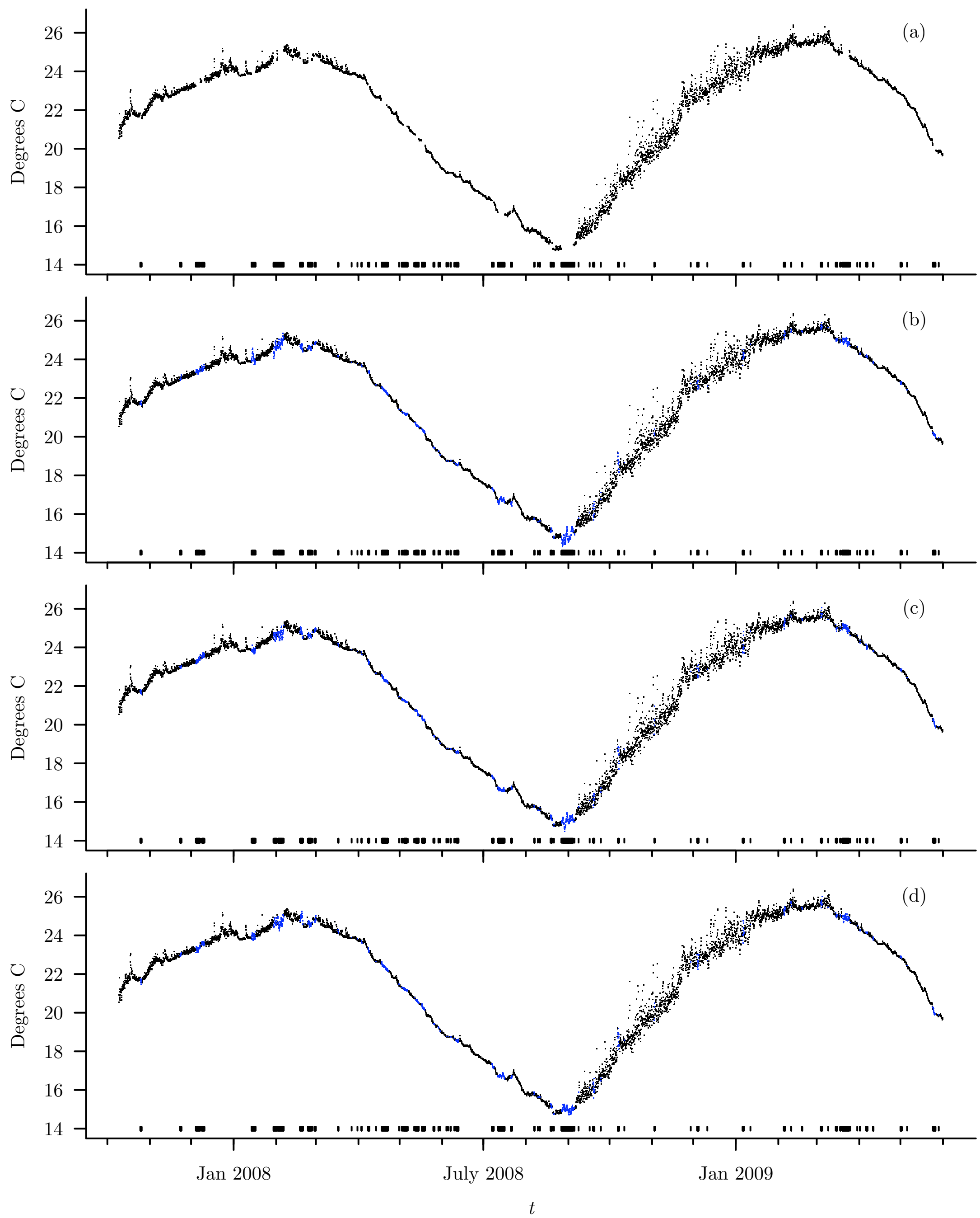

Figure 11. Three stochastic interpolations (bottom three plots) of the 10 m time series (top plot, without interpolation). The row of vertical hatches at the bottom of each plot indicates the locations of the gaps.