Wavelet-Based Multiresolution Analysis of Wivenhoe Dam Water Temperatures

Don Percival

Applied Physics Laboratory
Department of Statistics
University of Washington
Seattle, Washington, USA

http://faculty.washington.edu/dbp

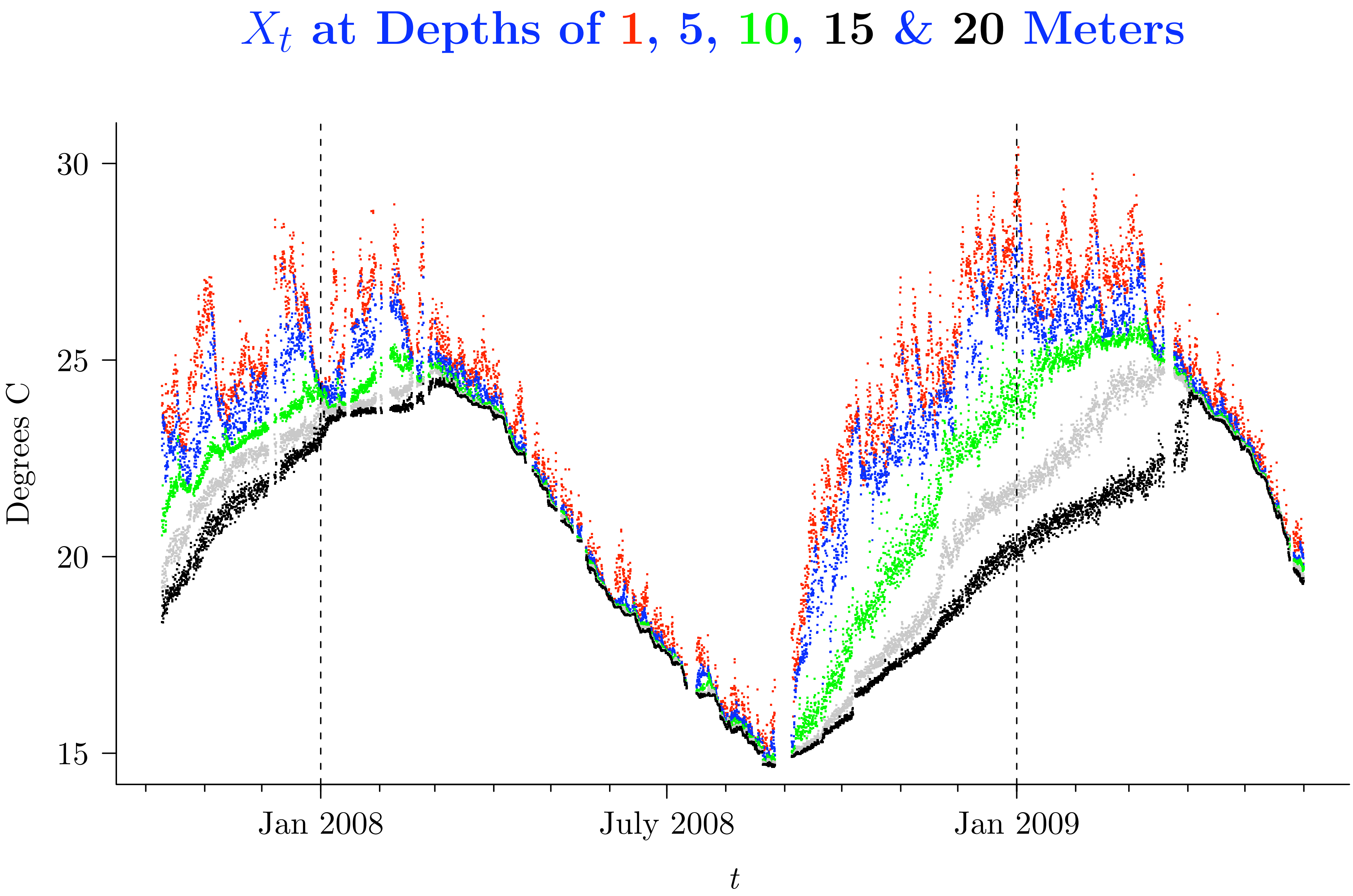
Collaborative effort with Sarah Lennox, You-Gan Wang and Ross Darnell, with support from Seqwater personnel

Background: I

- Queensland Bulk Water Supply Authority (Seqwater) manages catchments, water storages and treatment services to ensure quality and quantity of water supplied to Southeast Queensland
- ongoing monitoring program recently upgraded with permanent installation of vertical profilers at Lake Wivenhoe dam
- each profiler monitors water quality indicators every two hours at different depths at a fixed location (temperature, pH, ...)
- leads to a unique opportunity to study fluctuations in these indicators in a subtropical dam as a function of time and depth
- will concentrate on a 600+ day segment of temperature fluctuations X_t recorded at dam wall (temperature is regarded as important driver for other water quality indicators)

Lake Wivenhoe

Photograph: Andrew Watkinson (Seqwater)


X_t at Depths of 1, 5, 10, 15 & 20 Meters
Degrees C
30
25
20
15
Jan 2008
July 2008
Jan 2009
t

Background: II

- complicated structure both across time and down depth
- Q: how can we best quantify variations in data?
- propose to simplify task by breaking X_t into components capturing daily, subannual & annual (DSA) variations
- can formulate precise definitions for each component in terms of a wavelet-based multiresolution analysis (MRA)
- DSA components are such that they
  - are approximately pairwise uncorrelated
  - sum up to original X_t exactly and
  - based upon coefficients that decompose sample variance of X_t exactly across time

Background: III

- some questions our approach can help address:
  1. how does variance change across time?
  2. are variations from one day to the next more prominent than variations from, say, one month to the next?
  3. how repeatable are variations in annual cycle at each depth?
  4. what are the pairwise relationships between depth series over different spans of time (e.g., day-to-day, month-to-month)?
- resulting analysis mainly descriptive, but provides insight into components needed for a formal statistical model

Overview of Remainder of Talk

- give overview of standard wavelet analysis
- describe adaptations for analysis of dam temperatures
- discuss preparations to data prior to analysis
- present key results of our analysis (complete analysis documented in recently completed manuscript)

Basic Description of Wavelet Analysis: I

- let $\mathbf{X}$ be a column vector with elements $X_t, t = 0, 1, \ldots, N-1$
- $\mathbf{X}$ represents time series of N 'regularly sampled' observations, i.e., time associated with X_t is $t_0 + t\,\Delta$
  - t_0 is the time at which X_0 was observed
  - Δ is the sampling time between adjacent observations (e.g., $\Delta = 2$ hours for water temperature time series)
  - t is time index for element X_t
- wavelet analysis of $\mathbf{X}$ is a linear transformation, expressed as

$$\widetilde{\mathbf{W}} = \widetilde{\mathcal{W}}\mathbf{X},$$

where $\widetilde{\mathcal{W}}$ is a matrix transforming $\mathbf{X}$ into a vector of maximal overlap discrete wavelet transform (MODWT) coefficients $\widetilde{\mathbf{W}}$

Basic Description of Wavelet Analysis: II

- $\widetilde{\mathbf{W}}$ contains two types of MODWT coefficients
  - wavelet coefficients
  - scaling coefficients
- let's focus first on wavelet coefficients, denoted by $\widetilde{W}_{j,t}$
- while each X_t in $\mathbf{X}$ has just a time index t, each wavelet coefficient $\widetilde{W}_{j,t}$ in $\widetilde{\mathbf{W}}$ has two indices:
  - level index j, where $j = 1, 2, \ldots, J_0$
  - time index t, where $t = 0, 1, \ldots, N - 1$
- maximum level J_0 depends upon the particular application (for water temperature series, $J_0 = 9$ as discussed later)

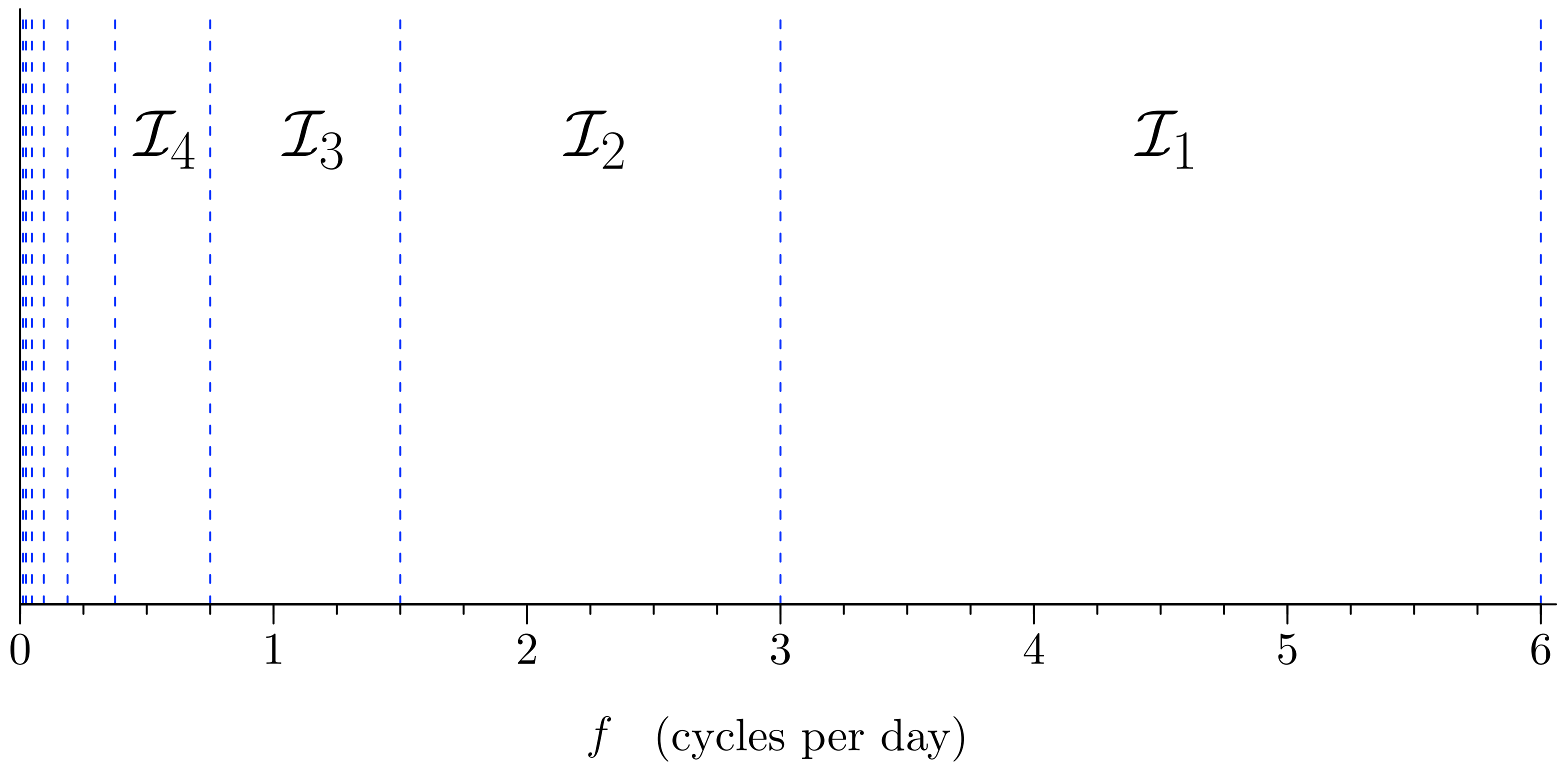
Basic Description of Wavelet Analysis: III

- index t for $\widetilde{W}_{j,t}$ says that formation of this coefficient involves only parts of $\mathbf{X}$ centered about a particular time
- index j tells us how many values in $\mathbf{X}$ are in effect being used to form $\widetilde{W}_{j,t}$
- if j is small (*large*), $\widetilde{W}_{j,t}$ depends mainly on a small (*large*) number of values from $\mathbf{X}$
- another interpretation is that j is an index for the interval of frequencies f given by

$$\mathcal{I}_j = \left(\frac{1}{2^{j+1}\,\Delta}, \frac{1}{2^j\,\Delta}\right]$$

- $\widetilde{W}_{j,t}$ is that part of a localized Fourier decomposition of $\mathbf{X}$ associated with frequencies $f \in \mathcal{I}_j$ (localization dictated by t)

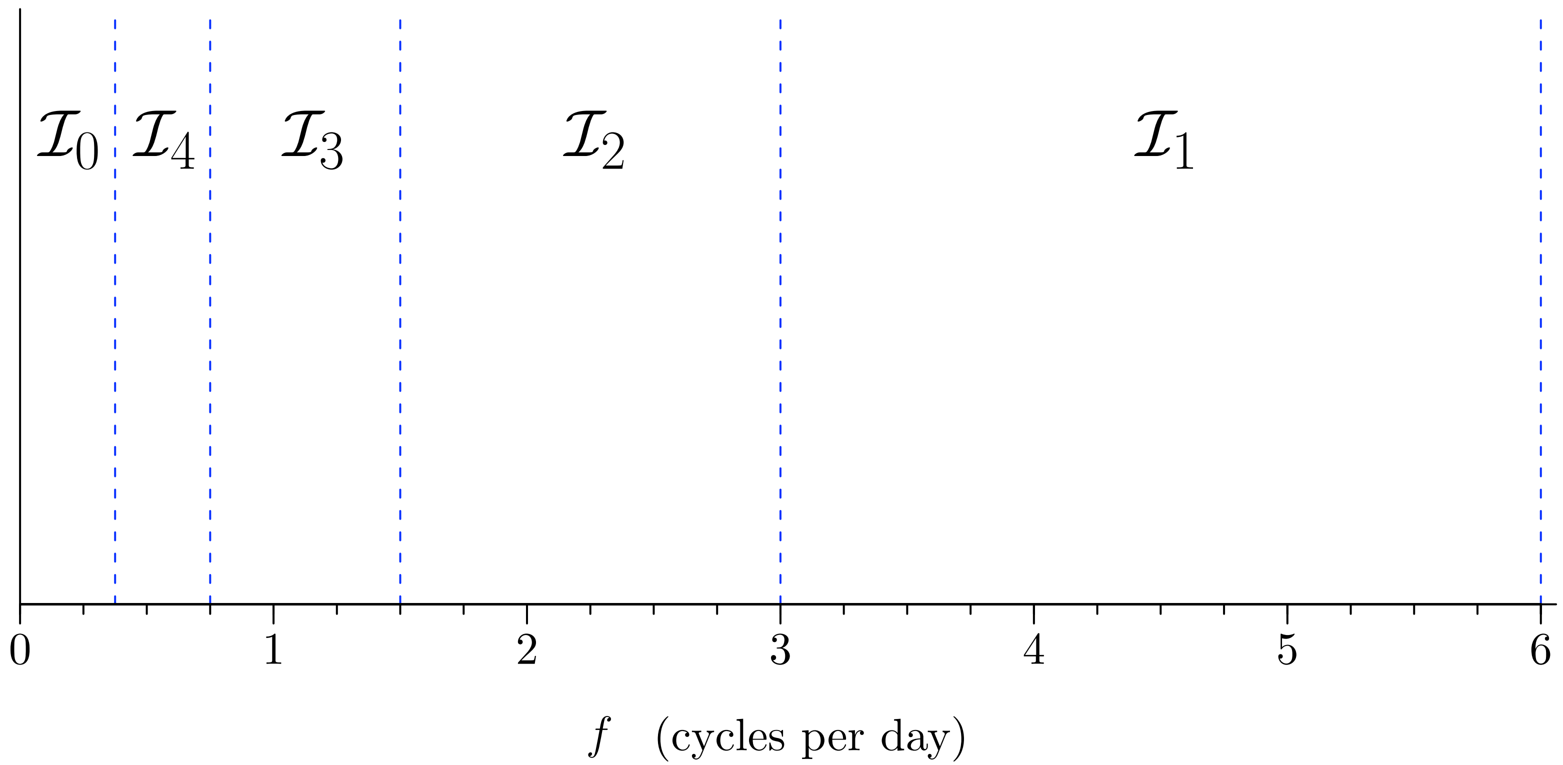
Frequency Intervals $\mathcal{I}_j$ When $\Delta = 2$ hours



Basic Description of Wavelet Analysis: IV

- let's now focus on scaling coefficients $\widetilde{V}_{J_0,t}$ in $\widetilde{\mathbf{W}}$
- each scaling coefficient also has a level index and a time index, but level index can assume only single value J_0
- time index t on $\widetilde{V}_{J_0,t}$ has same interpretation as for $\widetilde{W}_{j,t}$
- interval of frequencies associated with $\widetilde{V}_{J_0,t}$ is
$$\mathcal{I}_0 = \left[0, \frac{1}{2^{J_0+1}\,\Delta}\right]$$
- union of $\mathcal{I}_j$, $j = 0, 1, \ldots, J_0$, is $[0, 1/(2\,\Delta)]$, i.e., all physically meaningful frequencies in Fourier decomposition of $\mathbf{X}$
- scaling coefficients capture localized low-frequency variations in $\mathbf{X}$, whereas wavelet coefficients do the same over frequency intervals $\mathcal{I}_j$, $j = 1, 2, \ldots, J_0$

Frequency Intervals $\mathcal{I}_j$ When $\Delta = 2$ hours & $J_0 = 4$



Wavelet-based Analysis of Variance: I

- place all wavelet coefficients in $\widetilde{\mathbf{W}}$ associated with level j into vector $\widetilde{\mathbf{W}}_j$ & all scaling coefficients into vector $\widetilde{\mathbf{V}}_{J_0}$
- denote the square of Euclidean norm of $\mathbf{X}$ as $\|\mathbf{X}\|^2 \equiv \sum_t X_t^2$
- important property of MODWT of $\mathbf{X}$ is that $\|\widetilde{\mathbf{W}}\|^2 = \|\mathbf{X}\|^2$
- since $\widetilde{\mathbf{W}}$ is the union of $\widetilde{\mathbf{W}}_1$, $\widetilde{\mathbf{W}}_2$, ..., $\widetilde{\mathbf{W}}_{J_0}$ and $\widetilde{\mathbf{V}}_{J_0}$, also have

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

- can interpret $\|\widetilde{\mathbf{W}}_j\|^2$ as part of $\|\mathbf{X}\|^2$ attributable to localized Fourier coefficients associated with frequency interval $\mathcal{I}_j$, and $\|\widetilde{\mathbf{V}}_{J_0}\|^2$ as being associated with low-frequency interval $\mathcal{I}_0$

Wavelet-based Analysis of Variance: II

- let $\overline{X} = \sum_t X_t/N$ denote sample mean of $\mathbf{X}$, and consider its sample variance:

$$\hat{\sigma}_X^2 = \frac{1}{N}\sum_{t=0}^{N-1}\left(X_t - \overline{X}\right)^2 = \frac{1}{N}\|\mathbf{X}\|^2 - \overline{X}^2$$

- can reexpress the above as

$$\hat{\sigma}_X^2 = \sum_{j=1}^{J_0}\frac{1}{N}\|\widetilde{\mathbf{W}}_j\|^2 + \left(\frac{1}{N}\|\widetilde{\mathbf{V}}_{J_0}\|^2 - \overline{X}^2\right) \equiv \sum_{j=1}^{J_0}\hat{\sigma}_j^2 + \hat{\sigma}_0^2,$$

  where $\hat{\sigma}_j^2$ and $\hat{\sigma}_0^2$ are sample variances associated with $\widetilde{\mathbf{W}}_j$ and $\widetilde{\mathbf{V}}_{J_0}$ (sample mean of $\widetilde{\mathbf{V}}_{J_0}$ is $\overline{X}$ also, whereas wavelet coefficients come from populations with zero means)

Wavelet-based Analysis of Variance: III

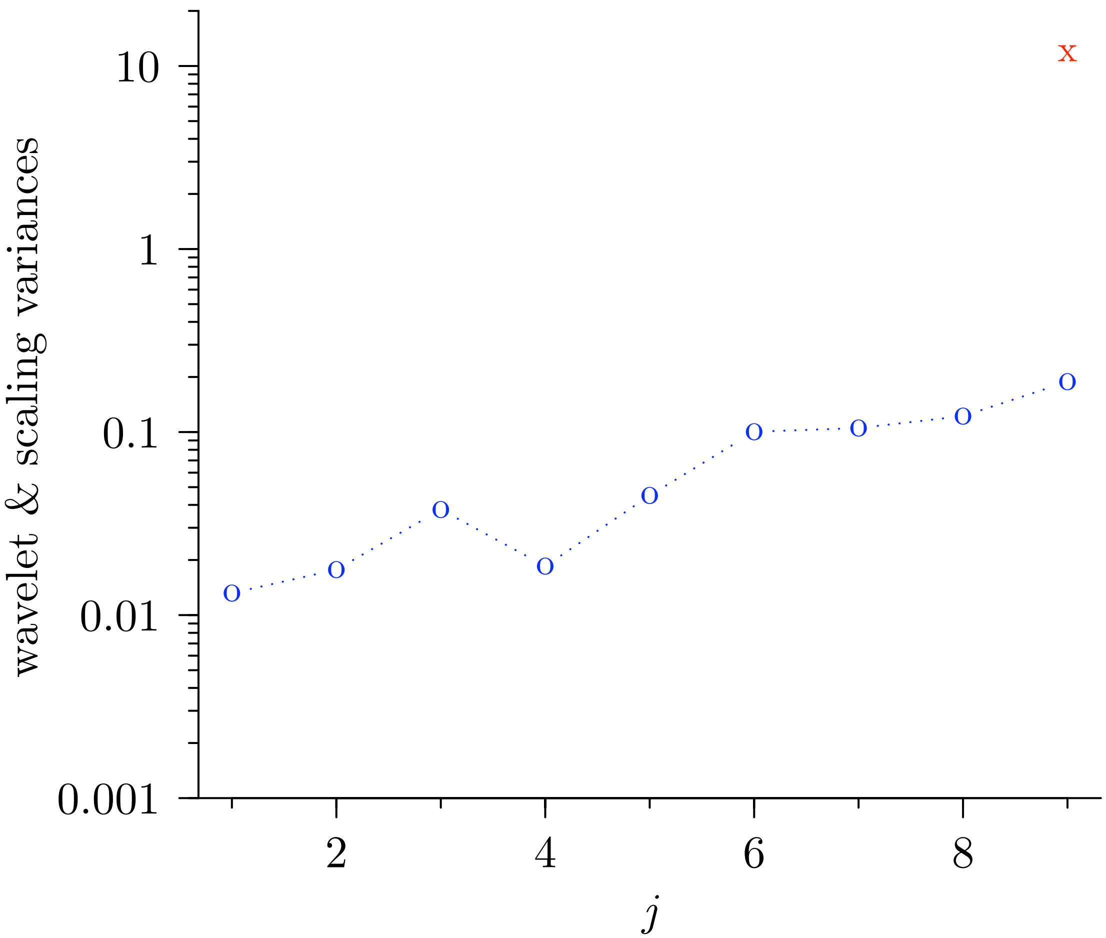
- can break up sample variance of $\mathbf{X}$ into $J_0 + 1$ parts, J_0 of which (the $\hat{\sigma}_j^2$'s) are attributable to fluctuations in intervals of frequencies $\mathcal{I}_j$, and the last ($\hat{\sigma}_0^2$), to fluctuations over low-frequency interval $\mathcal{I}_0$
- refer to decomposition of $\hat{\sigma}_X^2$ afforded by

$$\hat{\sigma}_X^2 = \sum_{j=1}^{J_0} \hat{\sigma}_j^2 + \hat{\sigma}_0^2$$

as a wavelet-based analysis of variance (ANOVA)

Wavelet-based ANOVA for 1 m Water Temperatures

- o's are $\hat{\sigma}_j^2$'s while x is $\hat{\sigma}_0^2$ (will offer interpretation later)



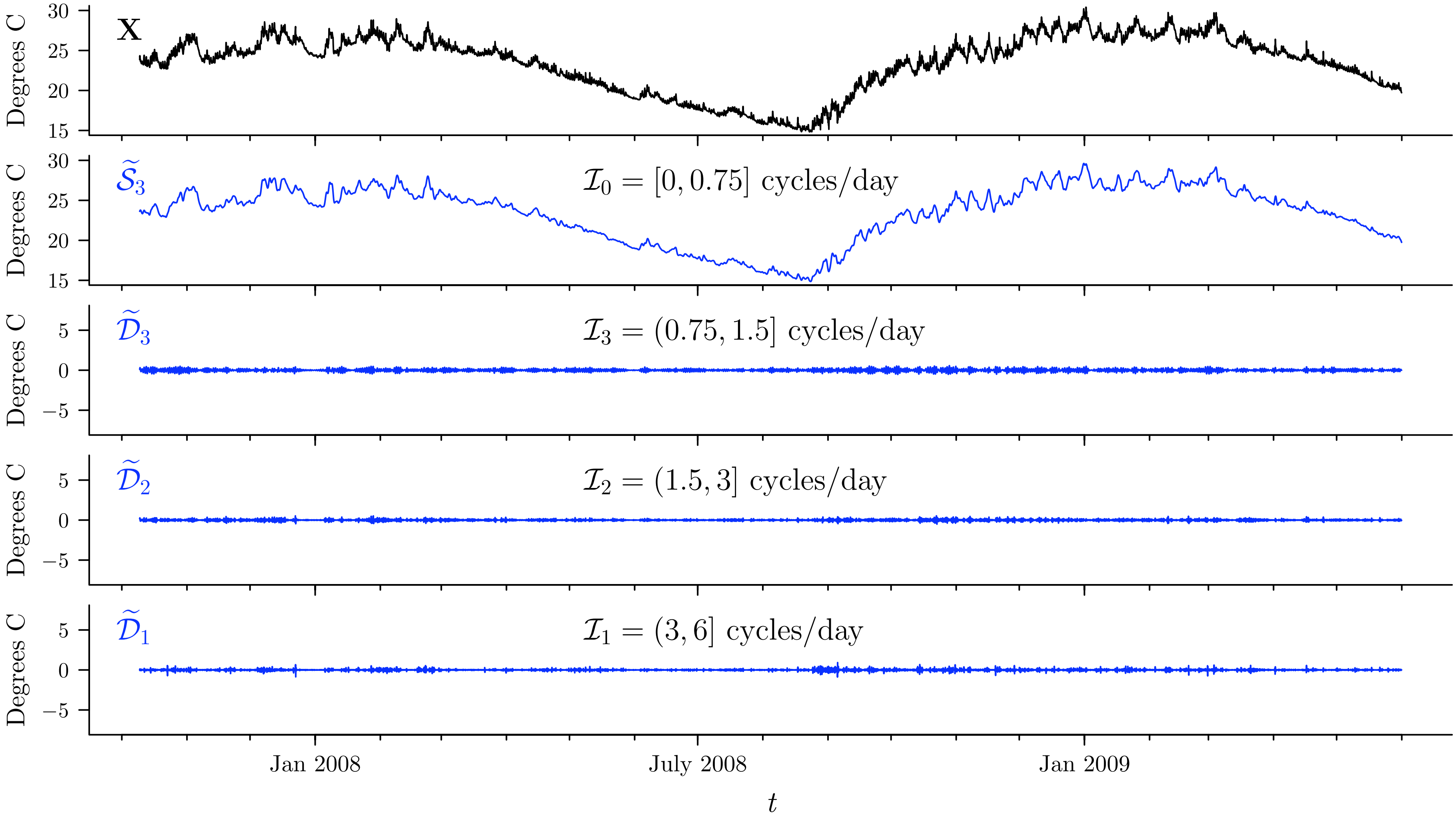
Multiresolution Analysis (MRA)

- can use the MODWT coefficients to obtain a wavelet-based additive decomposition known as a multiresolution analysis
- start with fact that $\mathbf{X}$ can be recovered from its MODWT coefficients $\widetilde{\mathbf{W}} = \widetilde{\mathcal{W}}\mathbf{X}$ via synthesis equation $\mathbf{X} = \widetilde{\mathcal{W}}^T\widetilde{\mathbf{W}}$
- partitioning both $\widetilde{\mathcal{W}}$ and $\widetilde{\mathbf{W}}$ allows rewriting $\mathbf{X} = \widetilde{\mathcal{W}}^T\widetilde{\mathbf{W}}$ as

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

- $\widetilde{\mathcal{D}}_j$ is a 'detail' series depending just on $\widetilde{\mathbf{W}}_j$ and those rows in $\widetilde{\mathcal{W}}$ used to create $\widetilde{\mathbf{W}}_j$ from $\mathbf{X}$
- $\widetilde{\mathcal{D}}_j$ captures part of $\mathbf{X}$ attributable to fluctuations in $\mathcal{I}_j$
- $\widetilde{\mathcal{S}}_{J_0}$ is a 'smooth' series capturing low-frequency fluctuations

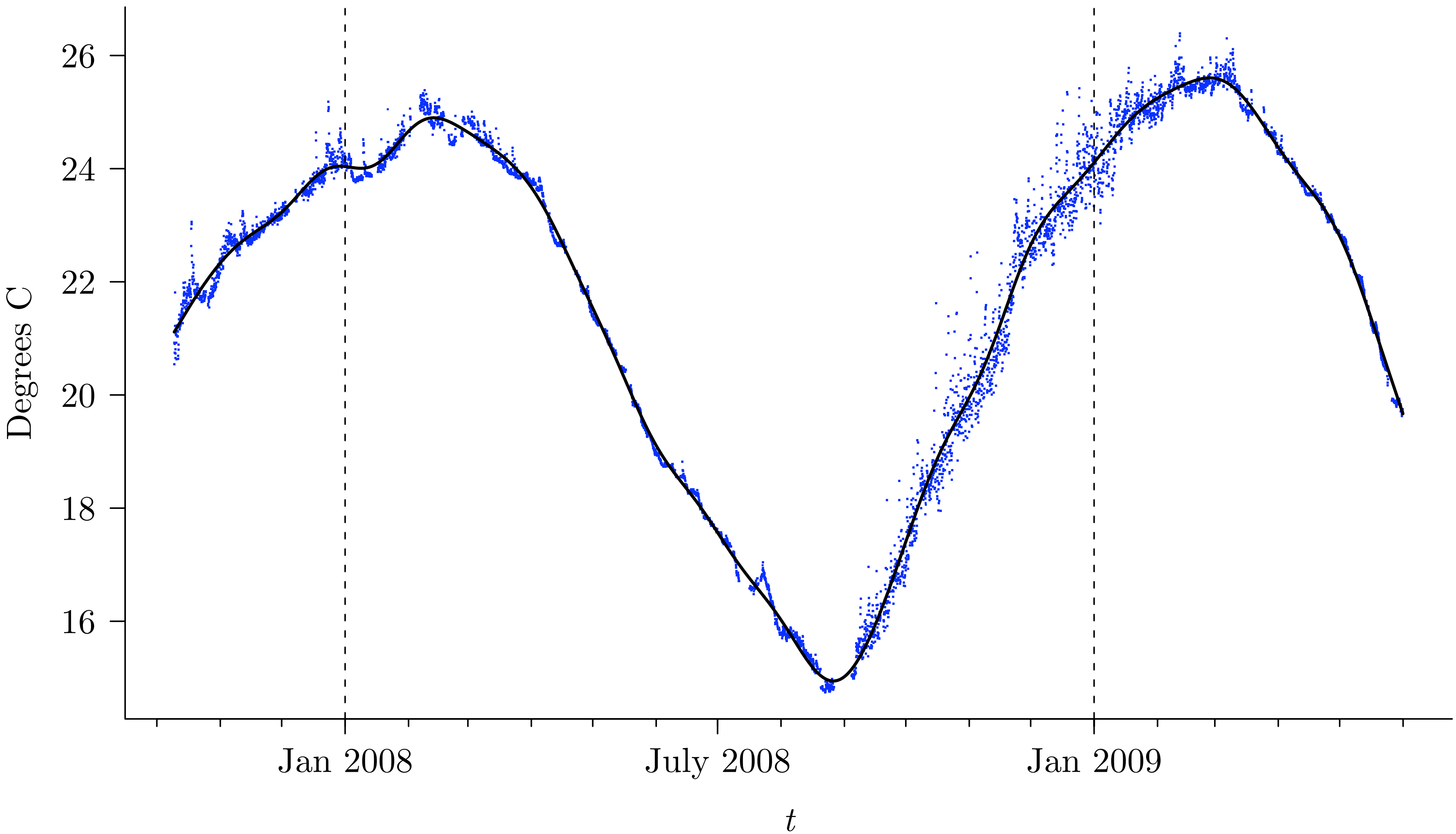
$J_0 = 3$ MRA for 1 m Water Temperatures



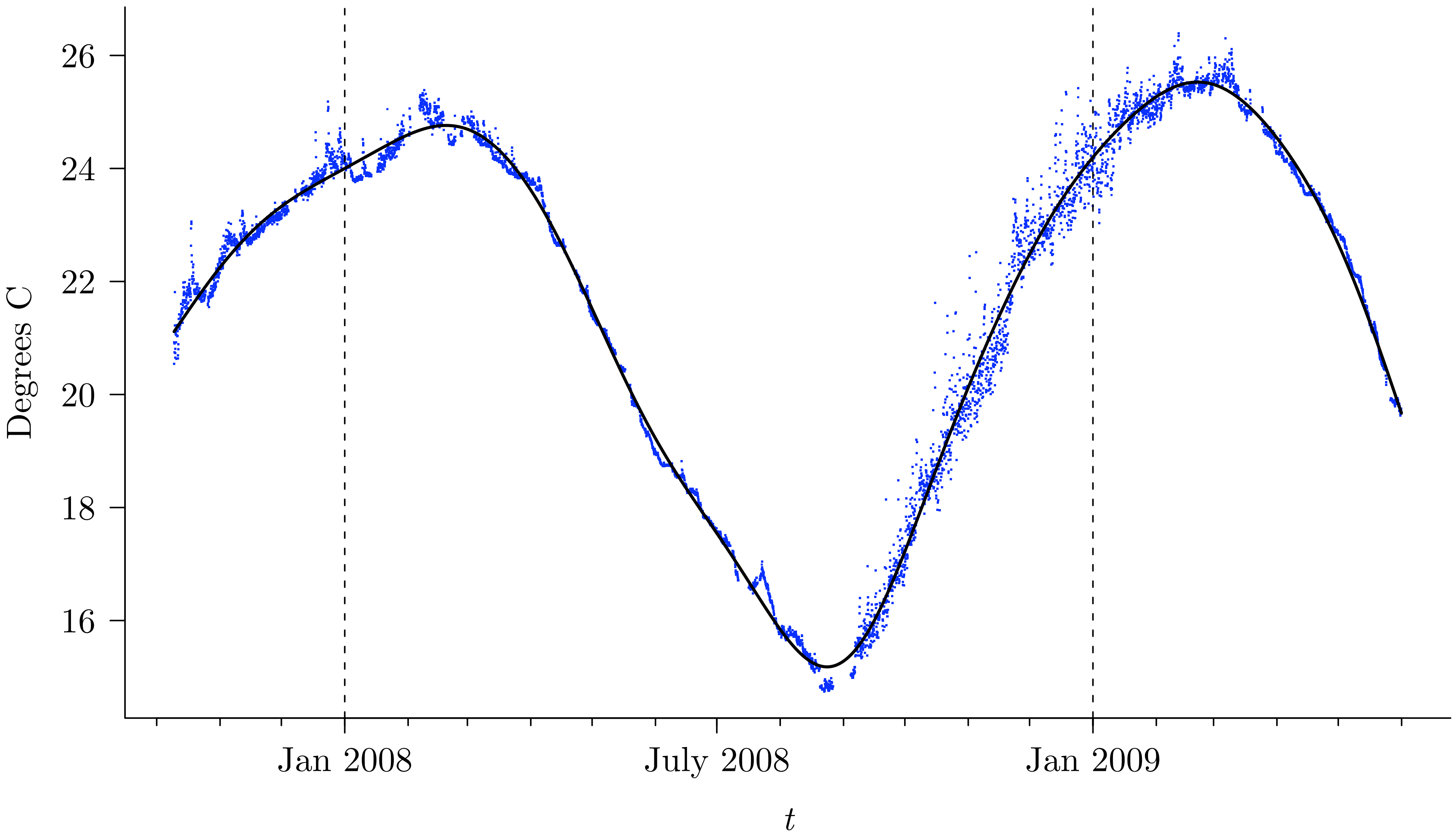
Adaptation of MRA for Water Temperatures: I

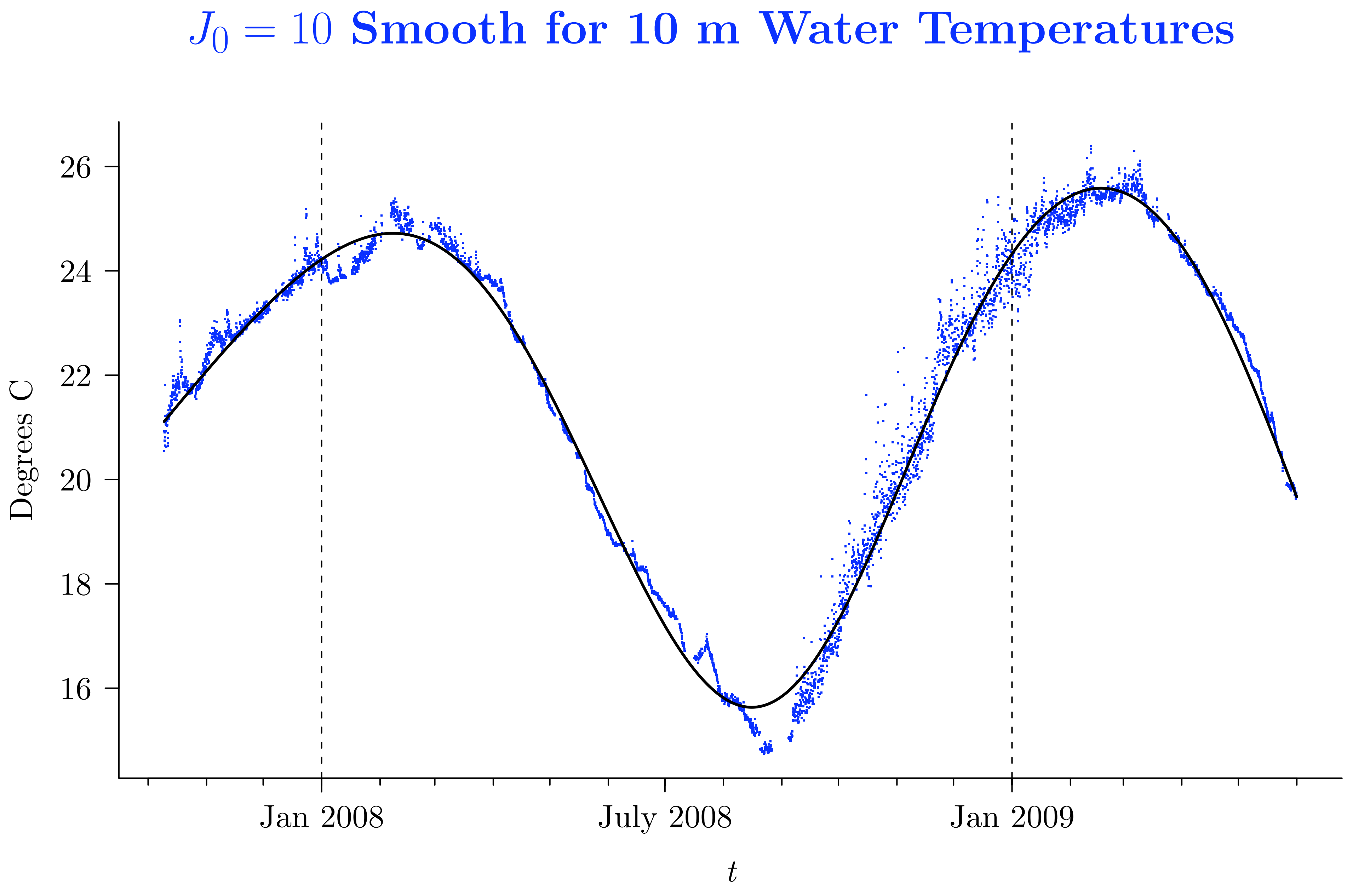
- idea behind MRA is to break $\mathbf{X}$ up into components $\widetilde{\mathcal{D}}_j$ and $\widetilde{\mathcal{S}}_{J_0}$ capturing different aspects of $\mathbf{X}$ (in statistical terms, components should be approximately pairwise uncorrelated)
- J_0 usually chosen so that $\widetilde{\mathcal{S}}_{J_0}$ captures prominent large-scale (low-frequency) fluctuations in $\mathbf{X}$
- setting $J_0 = 9$ means that $\widetilde{\mathcal{S}}_9$ captures fluctuations lower in frequency than 4.3 cycles/year
- empirically $\widetilde{\mathcal{S}}_9$ is preferable to either $\widetilde{\mathcal{S}}_8$ or $\widetilde{\mathcal{S}}_{10}$ in capturing interannual variations
  - $\widetilde{\mathcal{S}}_8$ is arguably undersmoothed, containing fluctuations better ascribed to intra-annual variations
  - $\widetilde{\mathcal{S}}_{10}$ is somewhat oversmoothed, hence distorting the interannual fluctuations

$J_0 = 8$ Smooth for 10 m Water Temperatures



$J_0 = 9$ Smooth for 10 m Water Temperatures

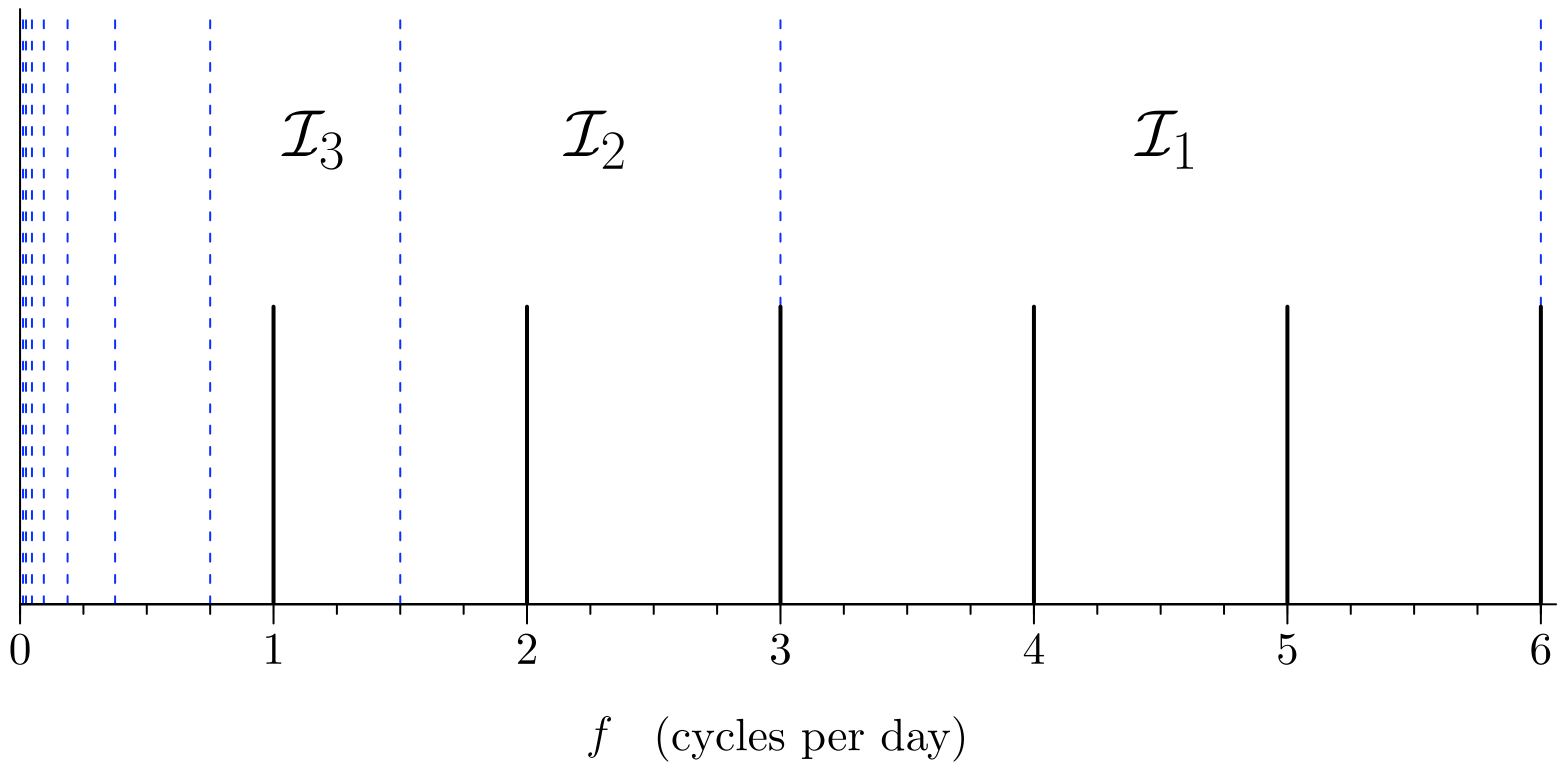



$J_0 = 10$ Smooth for 10 m Water Temperatures
Degrees C
26
24
22
20
18
16
Jan 2008
July 2008
Jan 2009
t

Adaptation of MRA for Water Temperatures: II

- with $J_0 = 9$, would have 9 detail series and smooth $\widetilde{\mathcal{S}}_9$ in usual wavelet-based MRA, but desirable to have decomposition with physically motivated components
- two important physical drivers for dam water temperatures are
  - revolution of earth about sun (influences annual variations)
  - daily rotation of earth (influences diurnal variations)
- seek additive decomposition isolating these variations
- setting $J_0 = 9$ results in $\widetilde{\mathcal{S}}_9$ capturing annual variations
- any purely periodic daily variation in time series with $\Delta =$ 2 hours can be expressed exactly with a Fourier decomposition involving a constant and sines & cosines with (at most) 6 frequencies, namely, fundamental frequency $f_1 = 1$ cycle/day and five harmonics $f_k = kf_1$, $k = 2, 3, 4, 5$ and 6 cycles/day.

Frequency Intervals $\mathcal{I}_j$ When $\Delta = 2$ hours



Adaptation of MRA for Water Temperatures: III

- daily fluctuations are captured primarily in $\widetilde{\mathcal{D}}_1$, $\widetilde{\mathcal{D}}_2$ and $\widetilde{\mathcal{D}}_3$
- accordingly, let's define a daily component as

$$\mathcal{D} = \widetilde{\mathcal{D}}_1 + \widetilde{\mathcal{D}}_2 + \widetilde{\mathcal{D}}_3$$

- $\mathcal{A} = \widetilde{\mathcal{S}}_9$ defines the annual component
- combine remaining details into a 'subannual' component

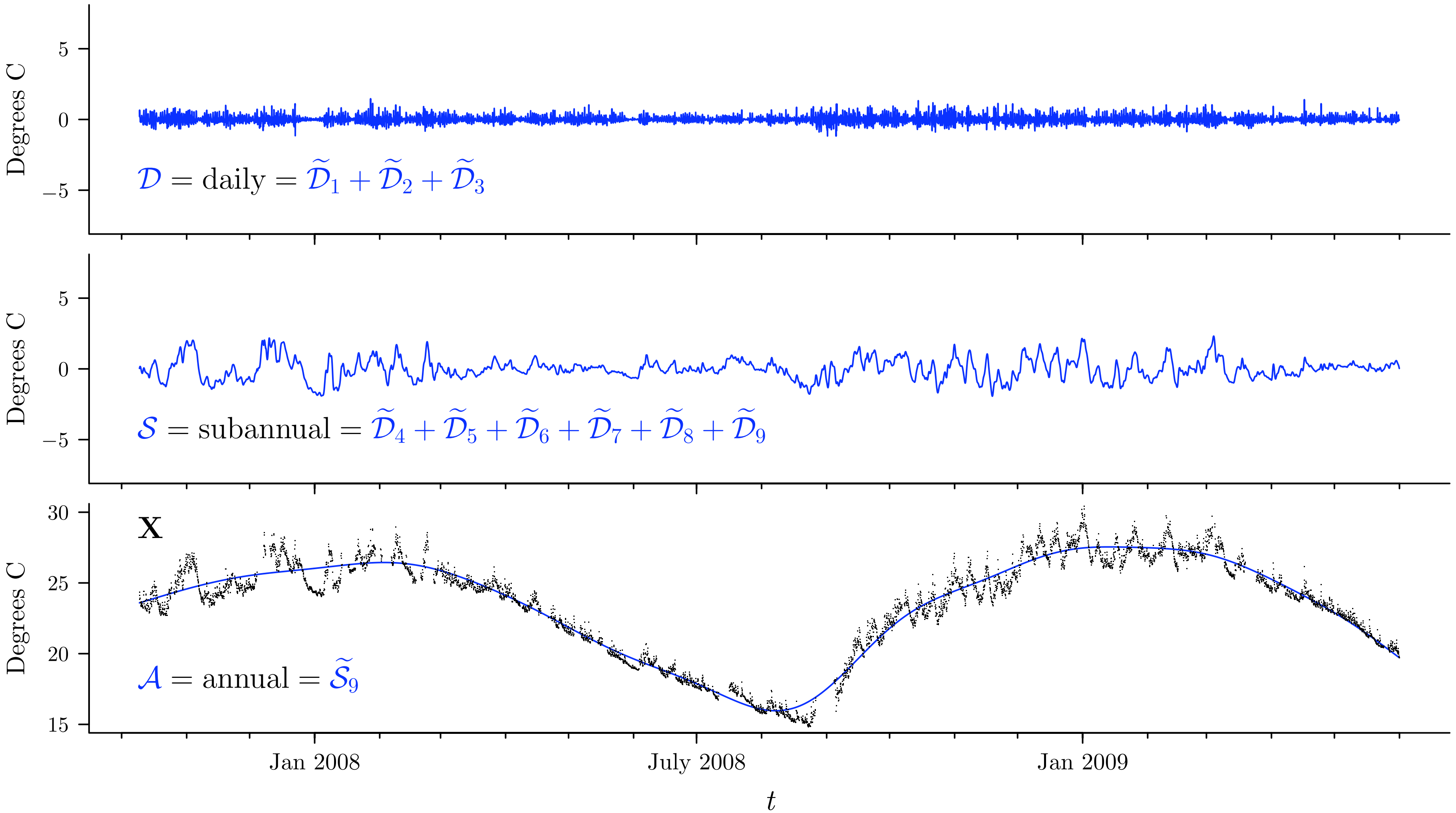
$$\mathcal{S} = \widetilde{\mathcal{D}}_4 + \widetilde{\mathcal{D}}_5 + \cdots + \widetilde{\mathcal{D}}_9,$$

  leading to the modified MRA

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

- will refer to this modified MRA as the DSA decomposition

DSA Decomposition for 1 m Water Temperatures



Adaptation of ANOVA for Water Temperatures: I

- can formulate ANOVA corresponding to DSA decomposition in two ways (one obvious, and the other, not so obvious)
- obvious way is to just add squared wavelet coefficients for each level involved in forming $\mathcal{D}$ and $\mathcal{S}$
  - elucidation of statistical properties of combination requires model to sort out relative influence of squared coefficients from different $\widetilde{\mathbf{W}}_j$'s (not easy to come by)
- not-so-obvious way is define a new transform, say $\mathbf{U} = \mathcal{U}\mathbf{X}$, with associated synthesis equation $\mathbf{X} = \mathcal{U}^T\mathbf{U}$
- $\mathbf{U}$ contains three types of coefficients $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$, each having N elements

Adaptation of ANOVA for Water Temperatures: II

- coefficients satisfy $\|\mathbf{D}\|^2 + \|\mathbf{S}\|^2 + \|\mathbf{A}\|^2 = \|\mathbf{X}\|^2$, with

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

- above leads to an ANOVA based upon the $\mathcal{U}$ transform:

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

where

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

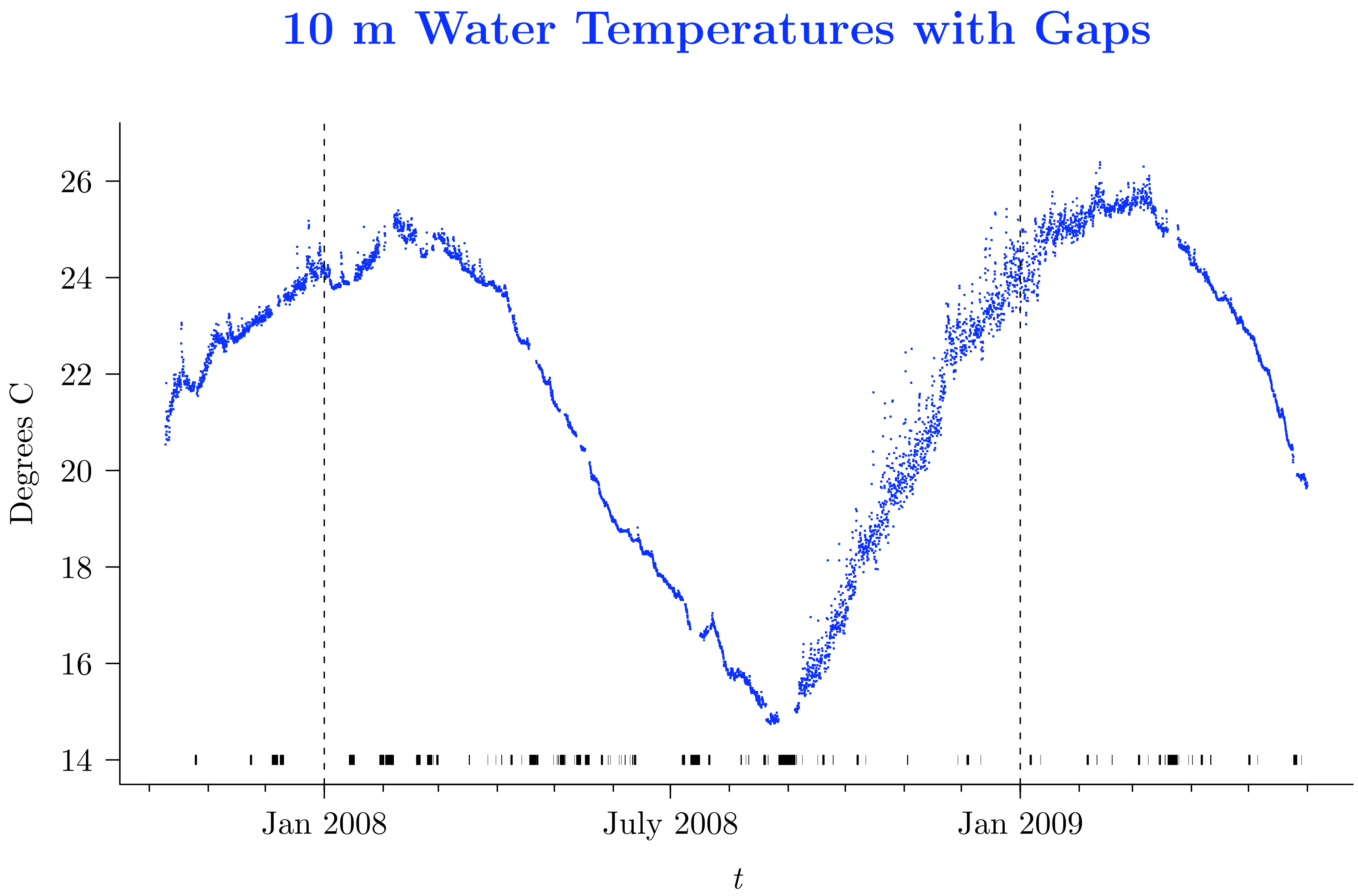
- manipulation of synthesis equation $\mathbf{X} = \mathcal{U}^T\mathbf{U}$ leads to DSA decomposition $\mathbf{X} = \mathcal{D} + \mathcal{S} + \mathcal{A}$

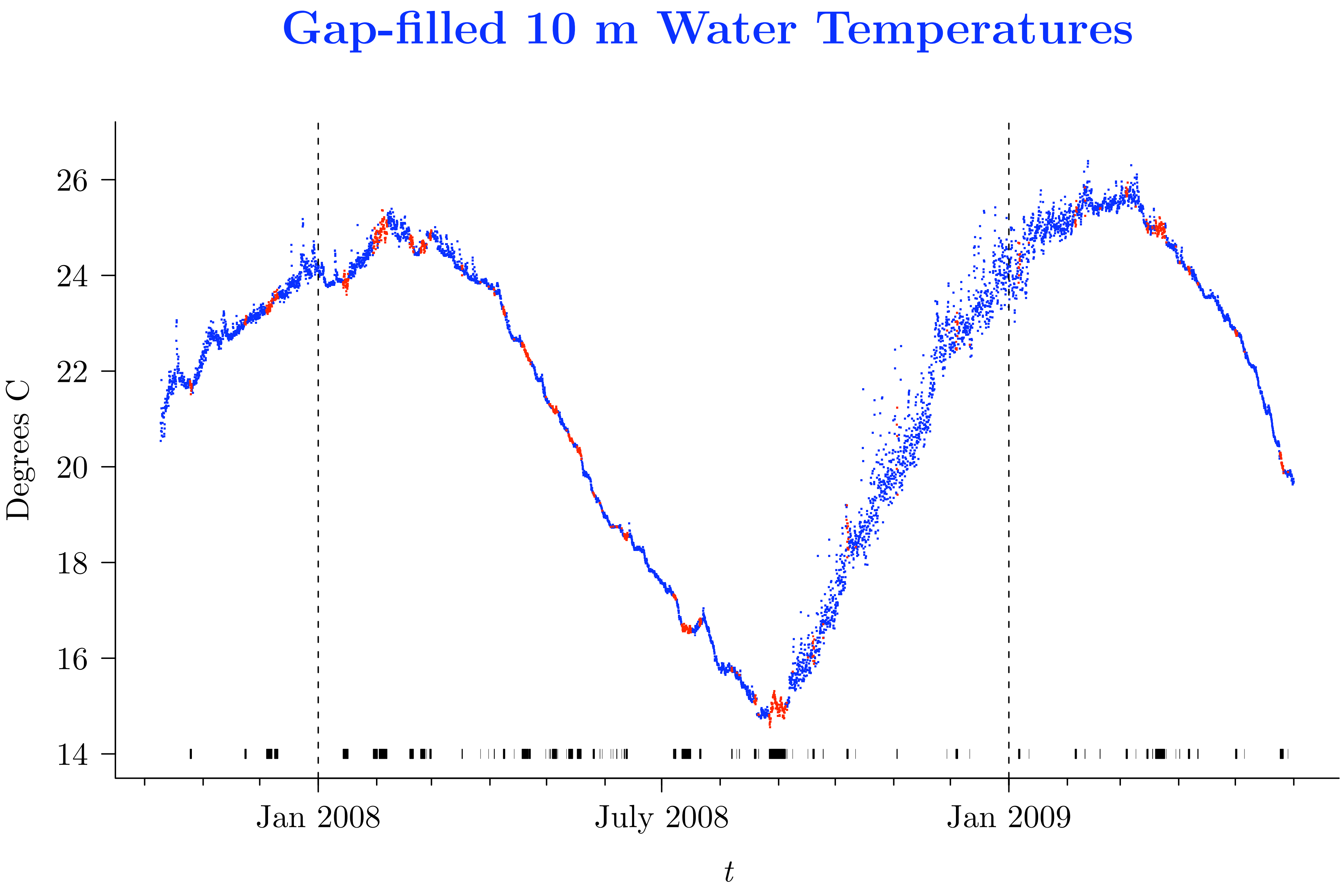
Adaptation of ANOVA for Water Temperatures: III

- have essentially 'collapsed' $3N$ wavelet coefficients in $\widetilde{\mathbf{W}}_1$, $\widetilde{\mathbf{W}}_2$ and $\widetilde{\mathbf{W}}_3$ into N coefficients $\mathbf{D}$, from which can determine $\mathcal{D}$
- likewise, have collapsed $6N$ wavelet coefficients in $\widetilde{\mathbf{W}}_4$, $\widetilde{\mathbf{W}}_5$, $\ldots, \widetilde{\mathbf{W}}_9$ into N coefficients $\mathbf{S}$, from which can determine $\mathcal{S}$
- will refer to
  - $\mathcal{U}$ as the DSA transform
  - $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$ collectively as DSA transform coefficients
  - $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$ alone as daily, subannual & annual coefficients
  - elements of $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$ by D_t, S_t and A_t

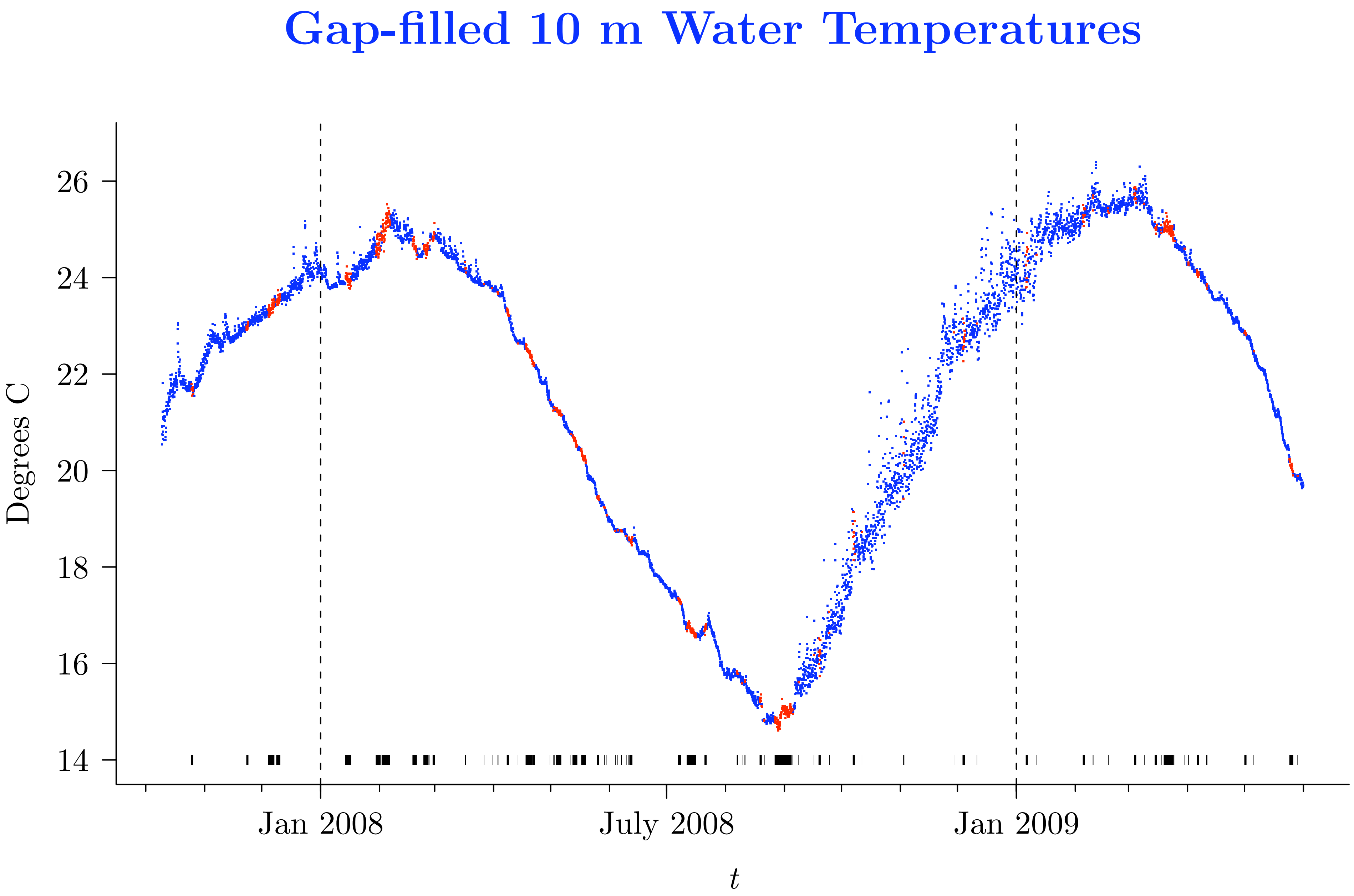
Data Preparation

- monitoring system at Wivenhoe Dam designed to measure water temperature and other variables at depths of $1, 2, \ldots, 20$ m every two hours (will concentrate on 1, 5, 10, 15 and 20 m as representative depths)
- protocol successfully adhered, with the exception of
  - some gaps in the data
  - some jitter in collection times (unlikely to impact analysis)
- can fill in gaps using a stochastic interpolation scheme
- also need to pay attention to how to handle boundary conditions for MODWT and DSA transforms

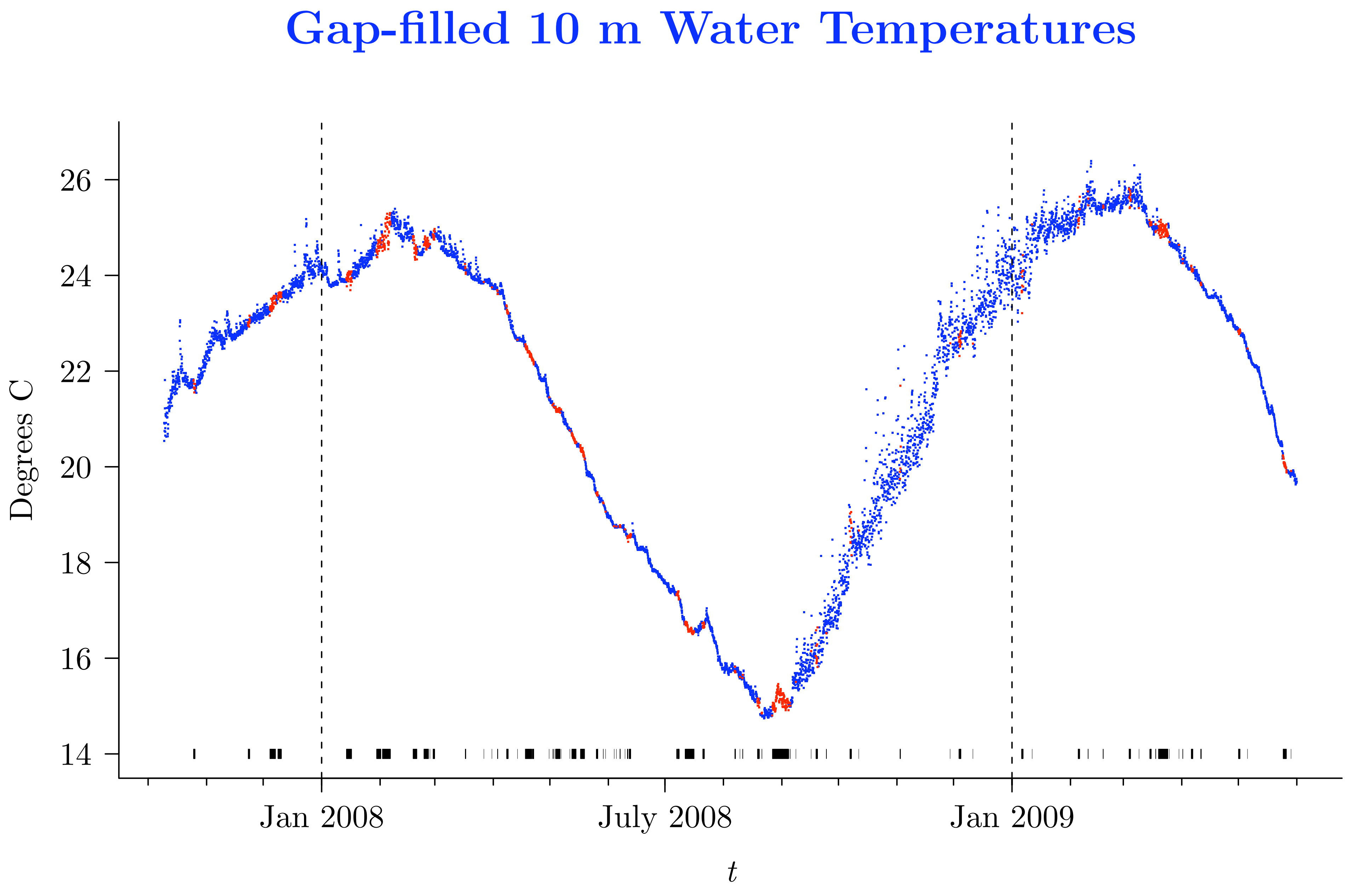

10 m Water Temperatures with Gaps
Degrees C
26
24
22
20
18
16
14
Jan 2008
July 2008
Jan 2009
t



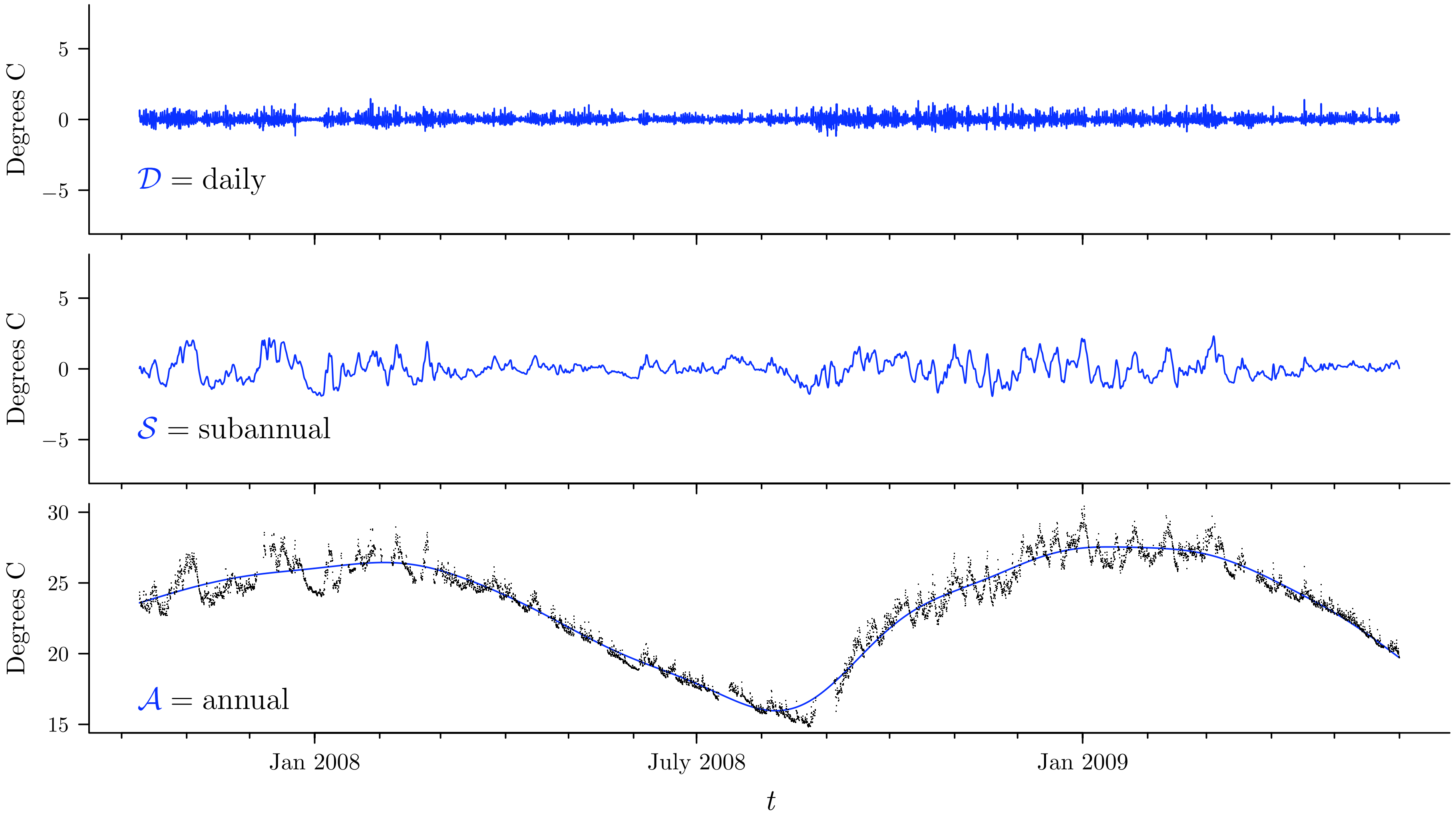
Gap-filled 10 m Water Temperatures
Degrees C
26
24
22
20
18
16
14
Jan 2008
July 2008
Jan 2009
t



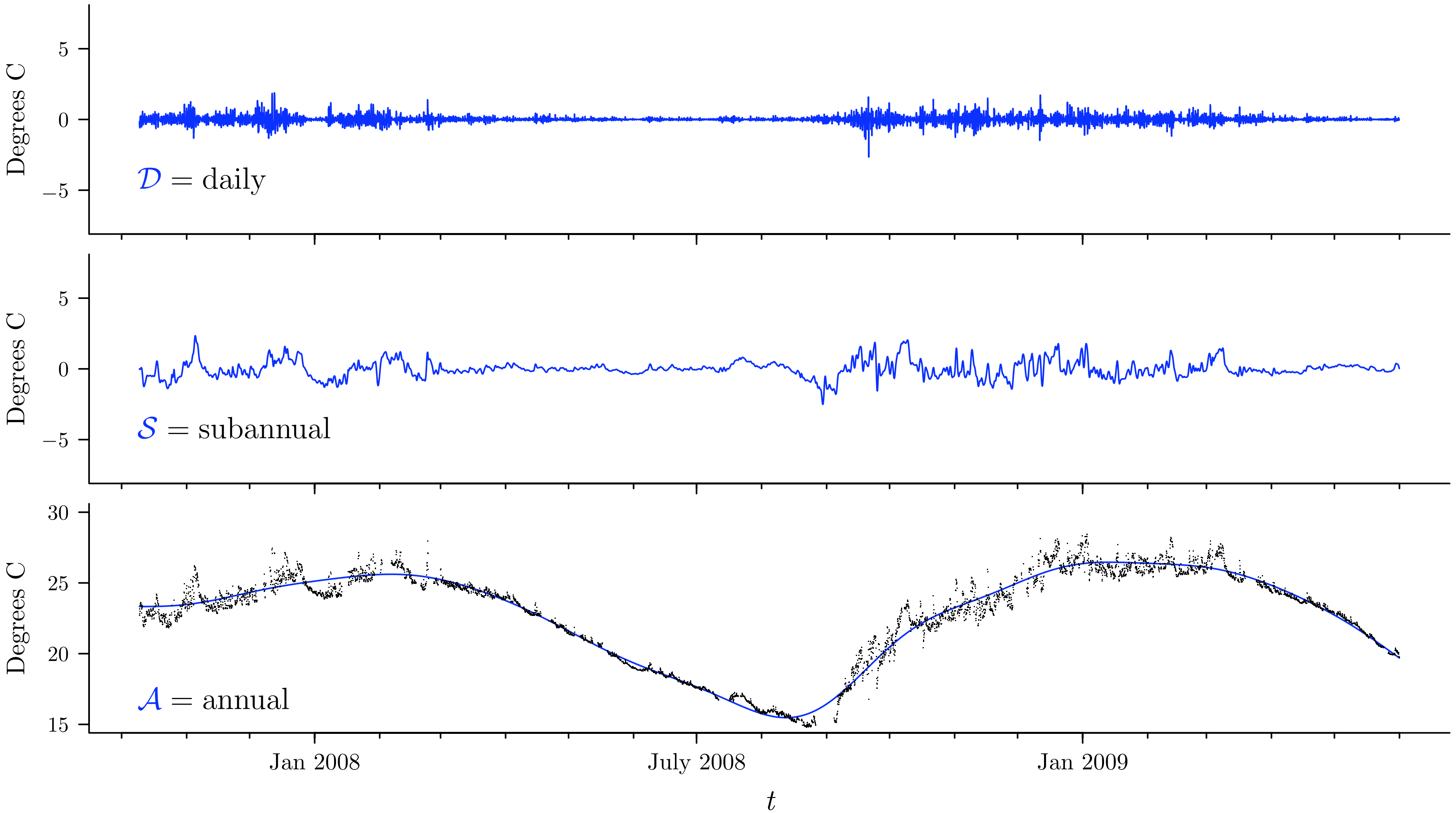
Gap-filled 10 m Water Temperatures
Degrees C
26
24
22
20
18
16
14
Jan 2008
July 2008
Jan 2009
t


Gap-filled 10 m Water Temperatures
Degrees C
26
24
22
20
18
16
14
Jan 2008
July 2008
Jan 2009
t

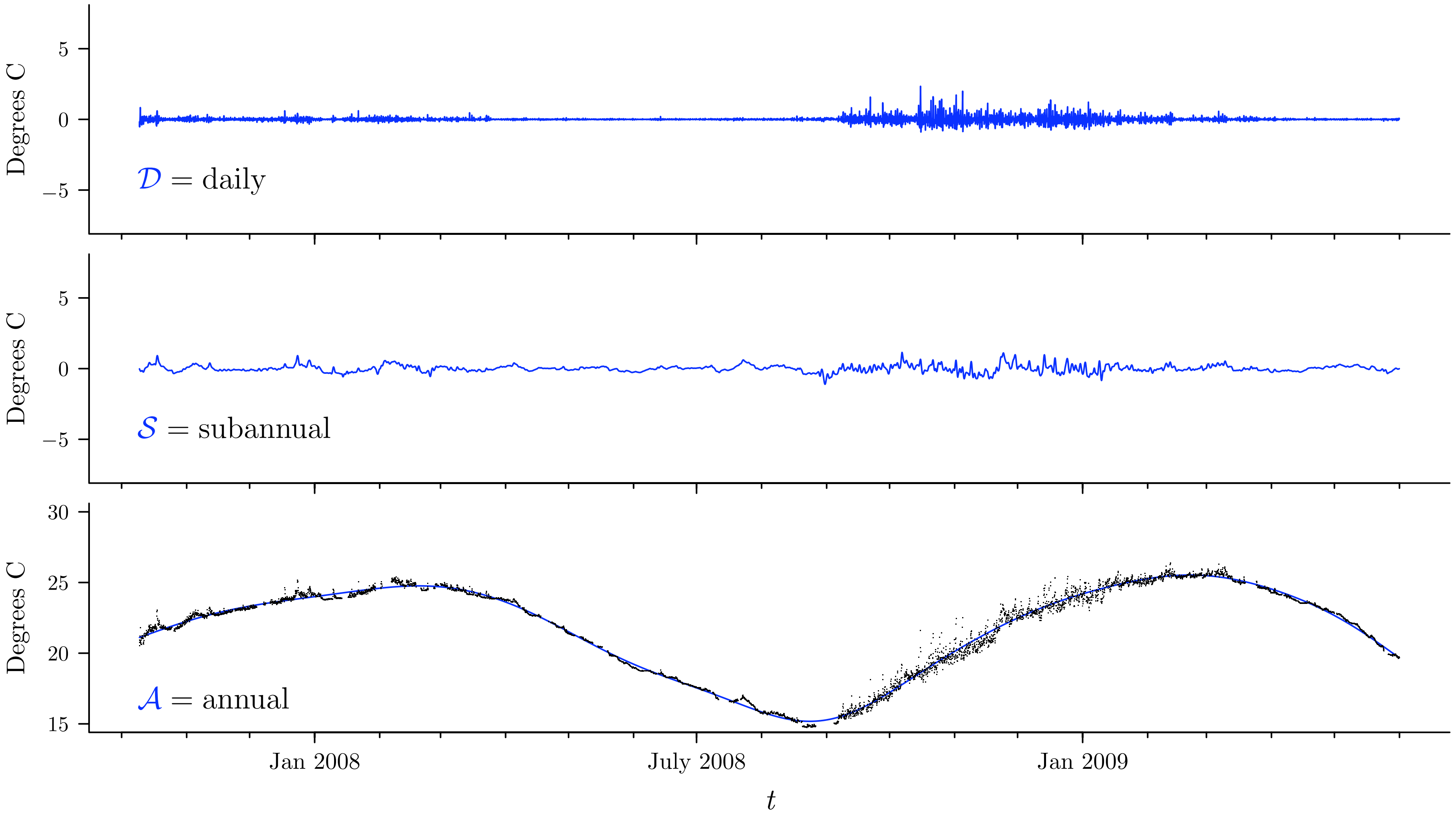
DSA Decomposition for 1 m Water Temperatures



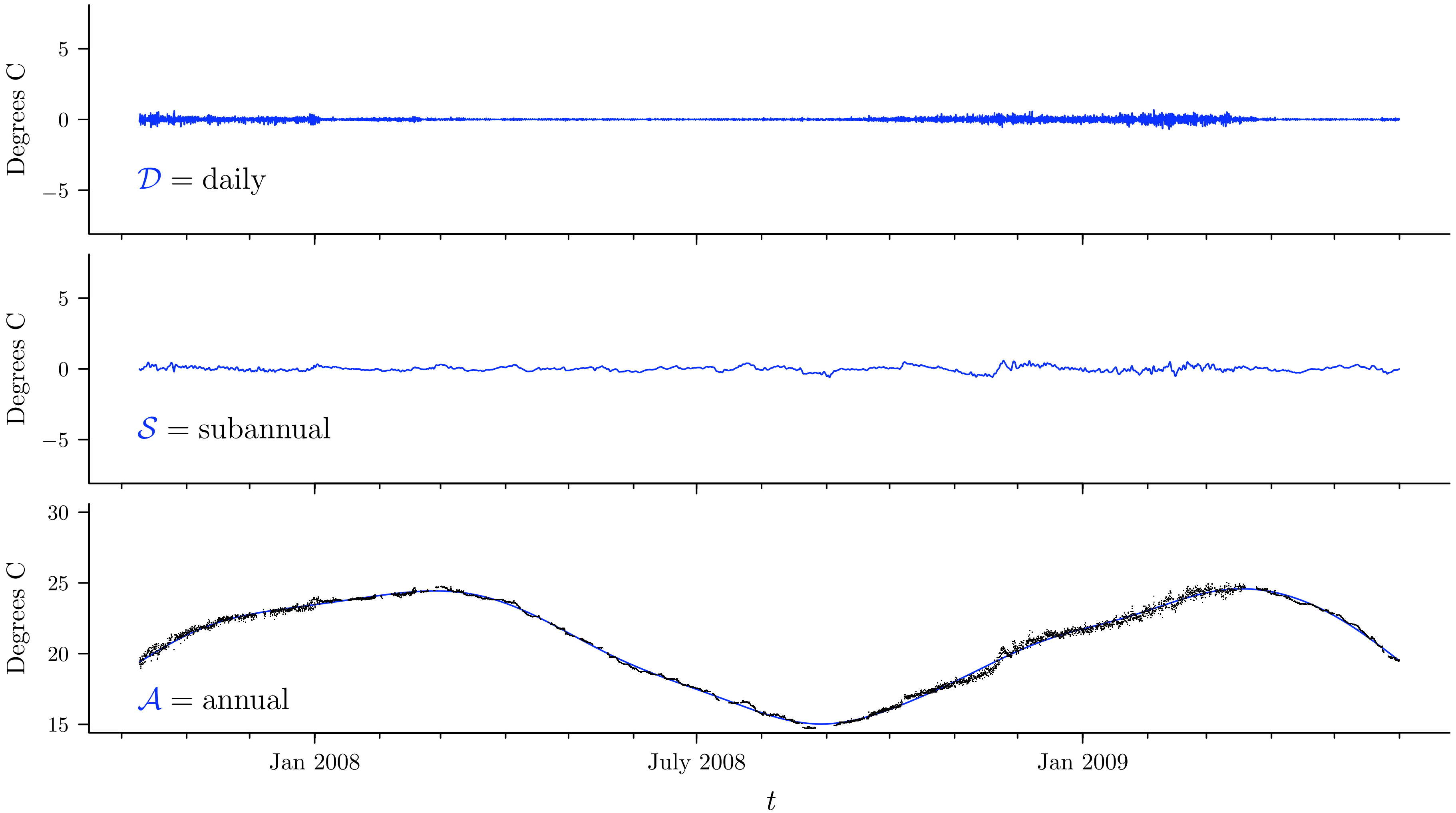
DSA Decomposition for 5 m Water Temperatures



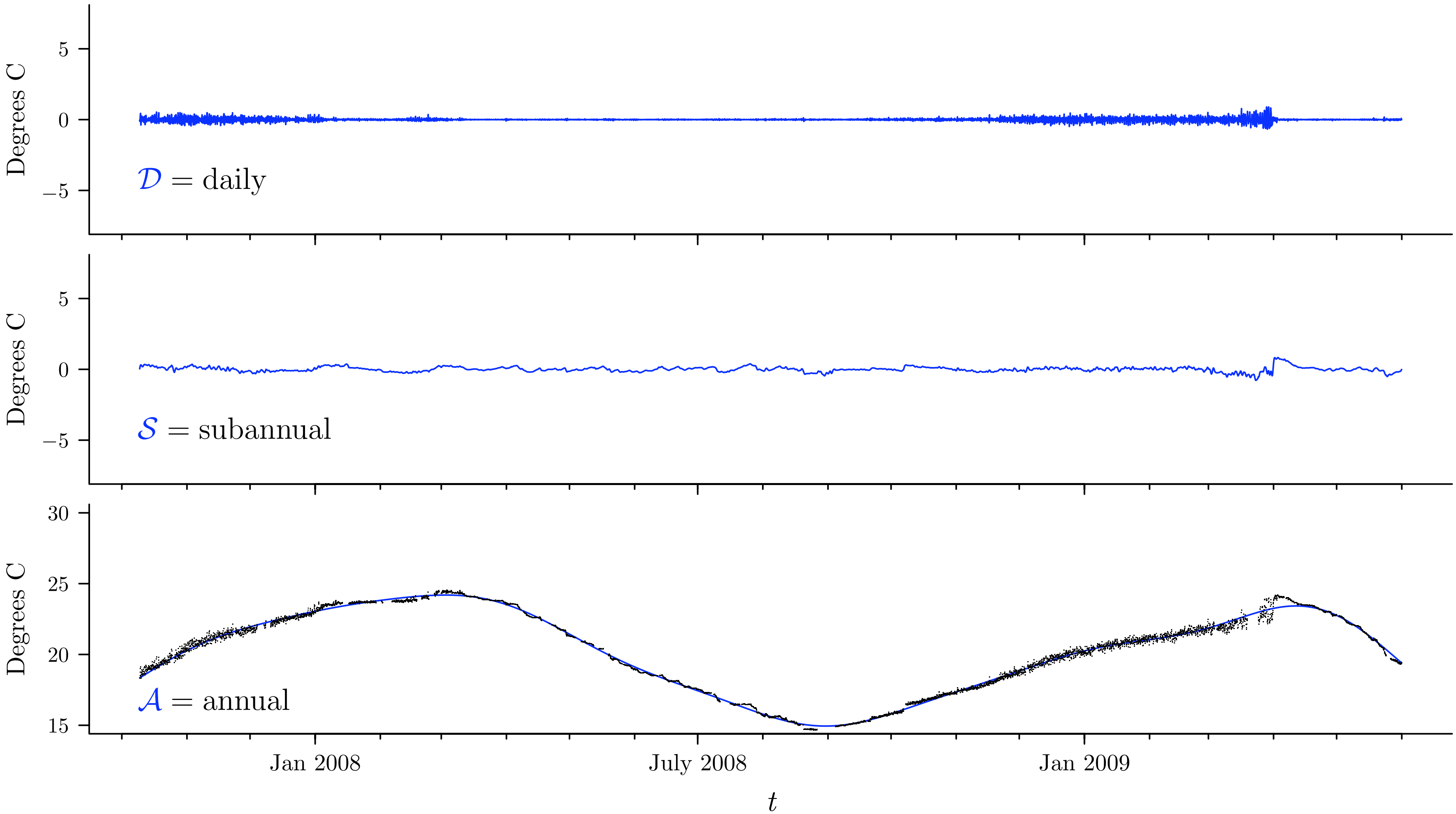
DSA Decomposition for 10 m Water Temperatures



DSA Decomposition for 15 m Water Temperatures



DSA Decomposition for 20 m Water Temperatures



Relative Importance of Three Components

- distance between adjacent vertical tick marks on all plots is 5 degrees Celsius
- in terms of overall (global) variability in each $\mathbf{X}$, annual component $\mathcal{A}$ is obviously dominant
- daily component $\mathcal{D}$ appears to contribute least, but there are local stretches over which it has greater variability than $\mathcal{S}$
- can quantify relative contribution of $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{A}$ coefficients to variance of each $\mathbf{X}$ globally using DSA-based ANOVA:

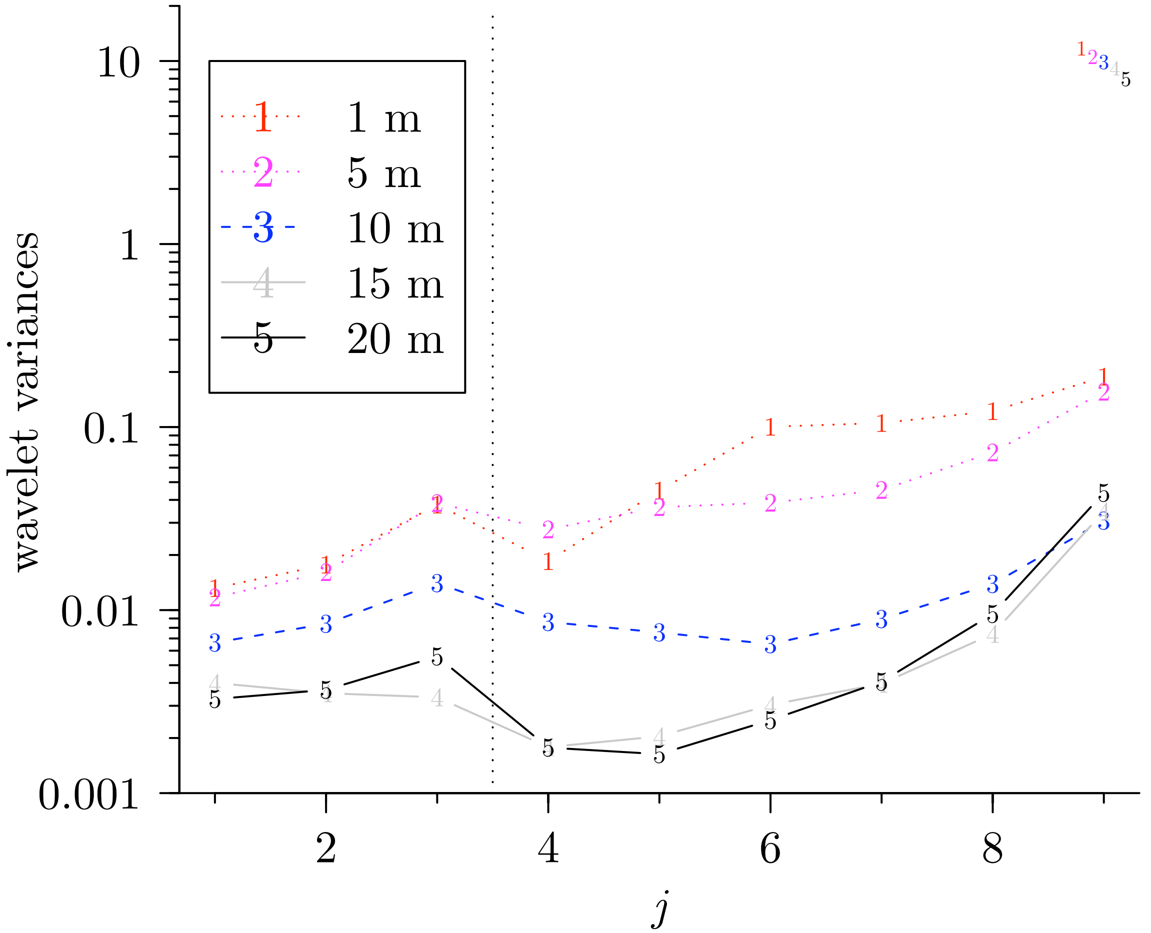
$$\hat{\sigma}_X^2 = \hat{\sigma}_D^2 + \hat{\sigma}_S^2 + \hat{\sigma}_A^2,$$

where

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

Wavelet-based ANOVA for Water Temperatures

- $\hat{\sigma}_j^2$'s to left (right) of vertical dashed line contribute to $\hat{\sigma}_D^2$ ($\hat{\sigma}_S^2$)



DSA-based ANOVA of Water Temperature

o

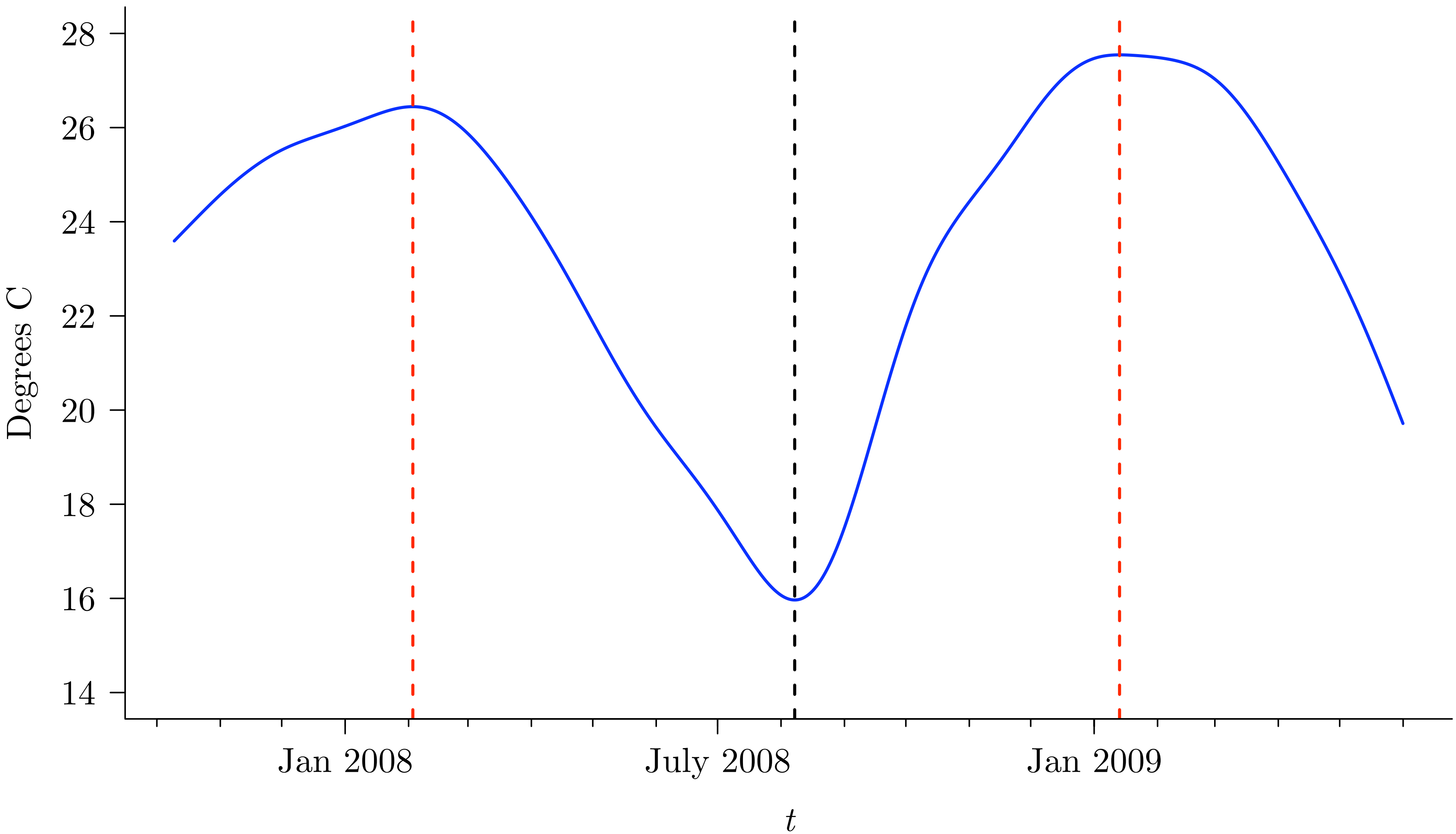
	1 m	5 m	10 m	15 m	20 m
$\hat{\sigma}_D^2$	0.07	0.07	0.03	0.01	0.01
$\hat{\sigma}_S^2$	0.58	0.38	0.08	0.05	0.06
$\hat{\sigma}_A^2$	11.74	10.56	9.88	9.14	7.98
$\hat{\sigma}_X^2$	12.39	11.00	9.99	9.20	8.06

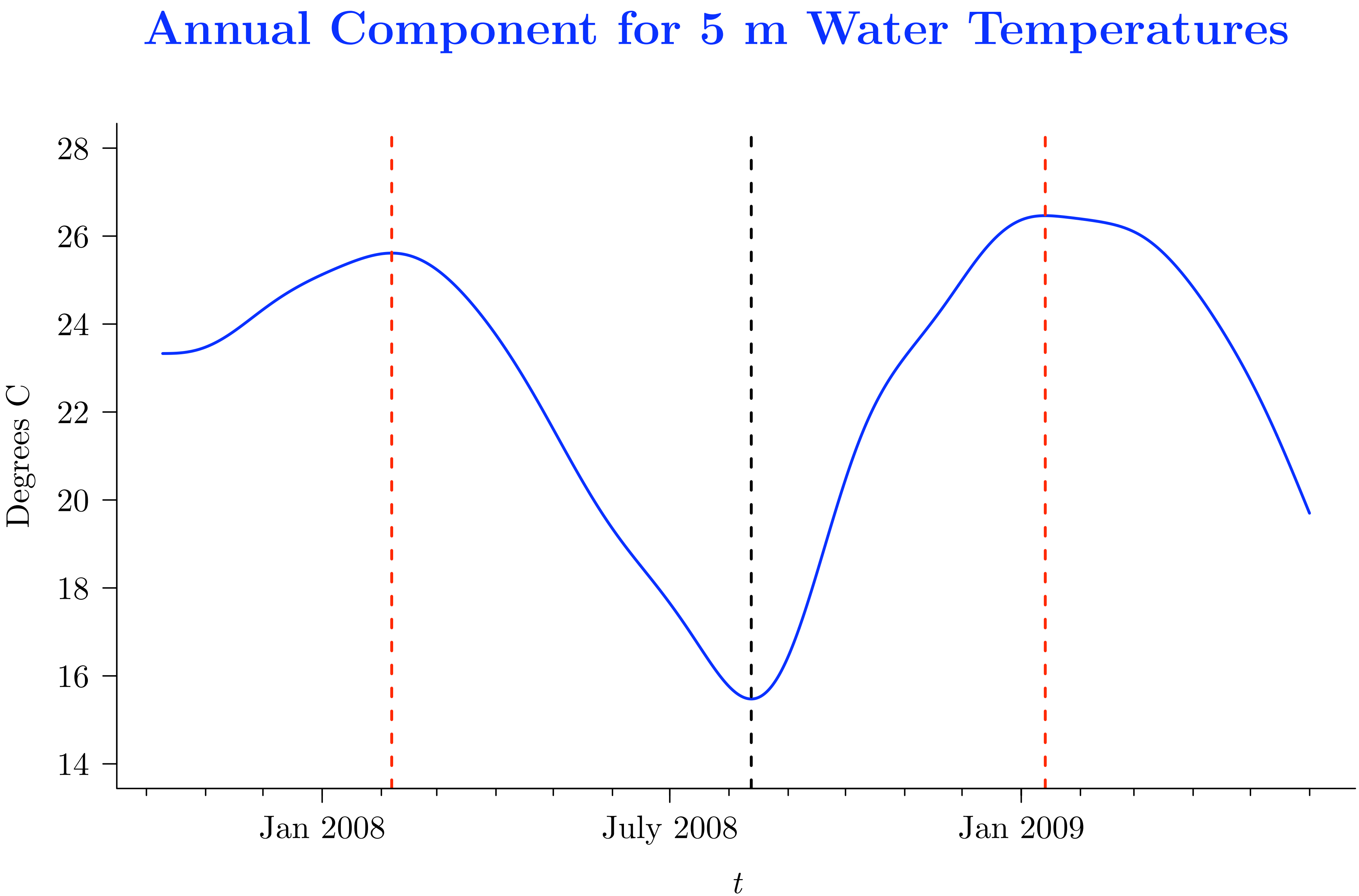
$\hat{\sigma}_D$	0.26	0.26	0.17	0.10	0.11
$\hat{\sigma}_S$	0.76	0.61	0.28	0.23	0.25
$\hat{\sigma}_A$	3.43	3.25	3.14	3.02	2.83
$\hat{\sigma}_X$	3.52	3.32	3.16	3.03	2.84

Notes on ANOVA

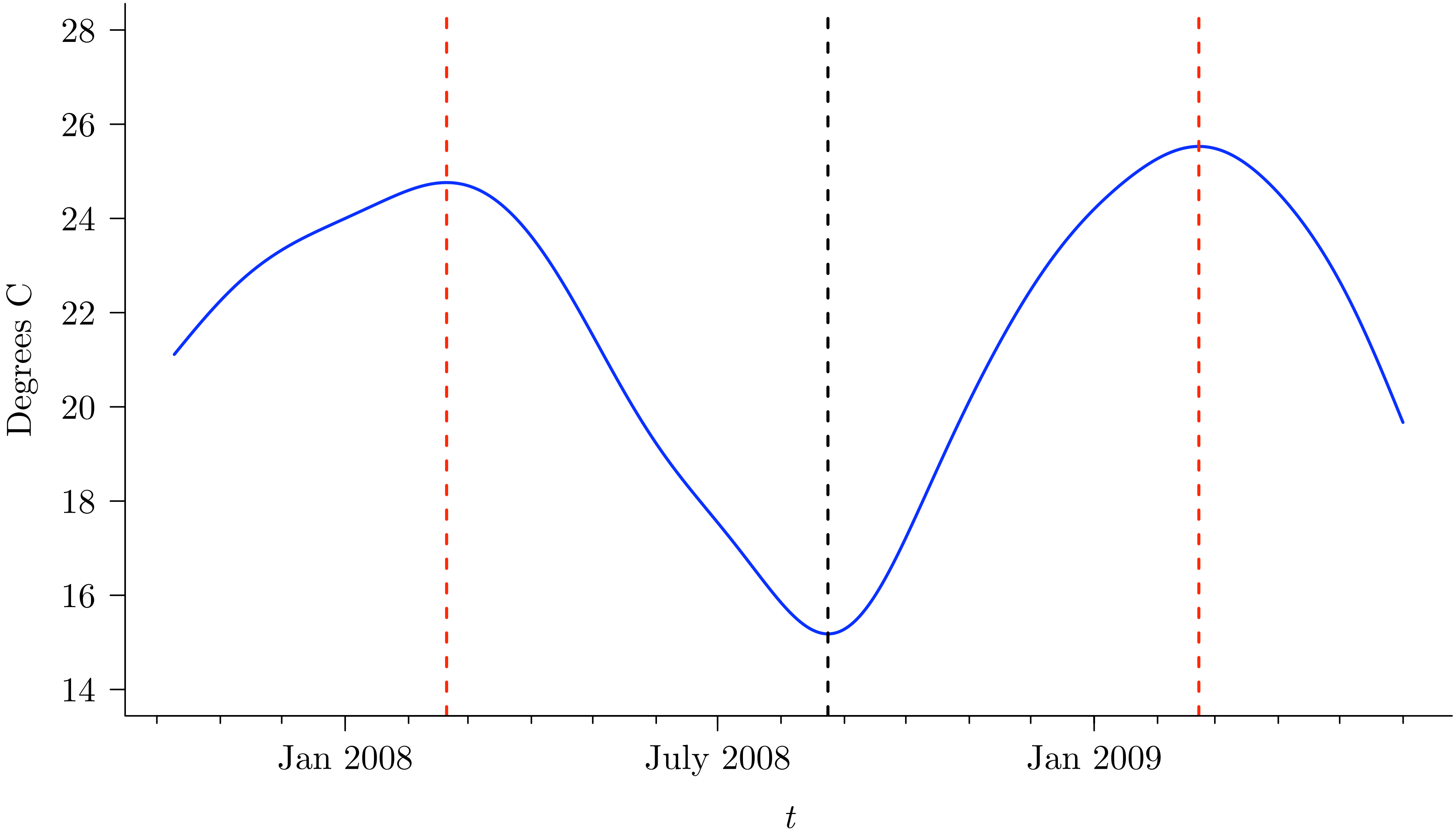
- pattern of $\hat{\sigma}_j^2$'s at 15 and 20 m quite similar, as are those for 1 and 5 m (with some divergence at $j = 6, 7$ & 8)
- pattern for 10 m represents a transition between shallower and deeper depths
- gross patterns are largely the same across all depths: increase from $j = 1$ to $j = 3$, followed by a drop between $j = 3$ & 4, and tendency to increase after that
- fundamental frequency of daily variations evidently more important than harmonics
- variance of annual coefficients $\mathbf{A}$ is one or two orders of magnitude greater than that of subannual coefficients $\mathbf{S}$
- variance of $\mathbf{S}$ is greater than that of $\mathbf{D}$ by at least a factor of 2

Annual Component for 1 m Water Temperatures

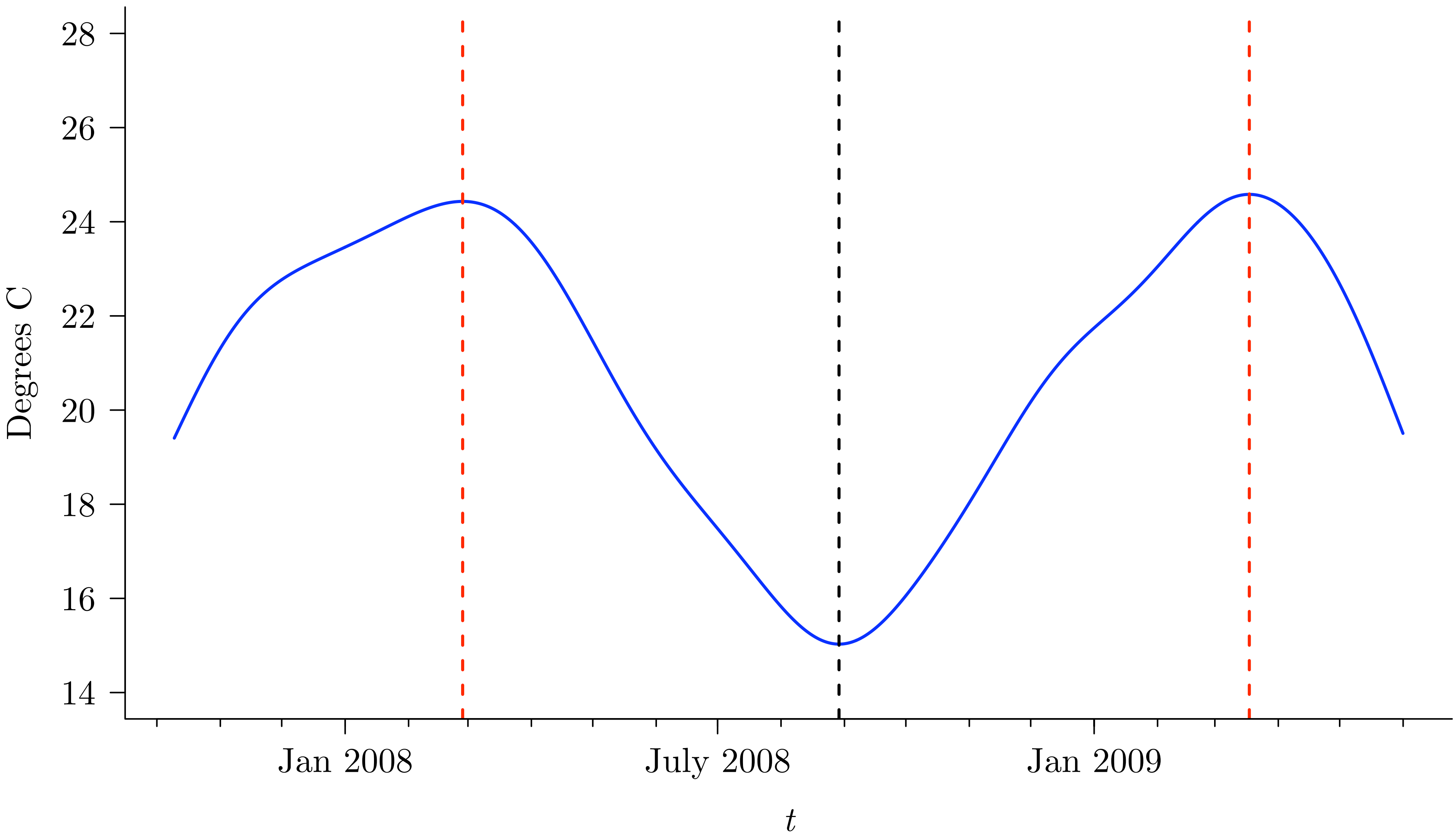



Annual Component for 5 m Water Temperatures
28
26
24
22
20
18
16
14
Degrees C
Jan 2008
July 2008
Jan 2009
t

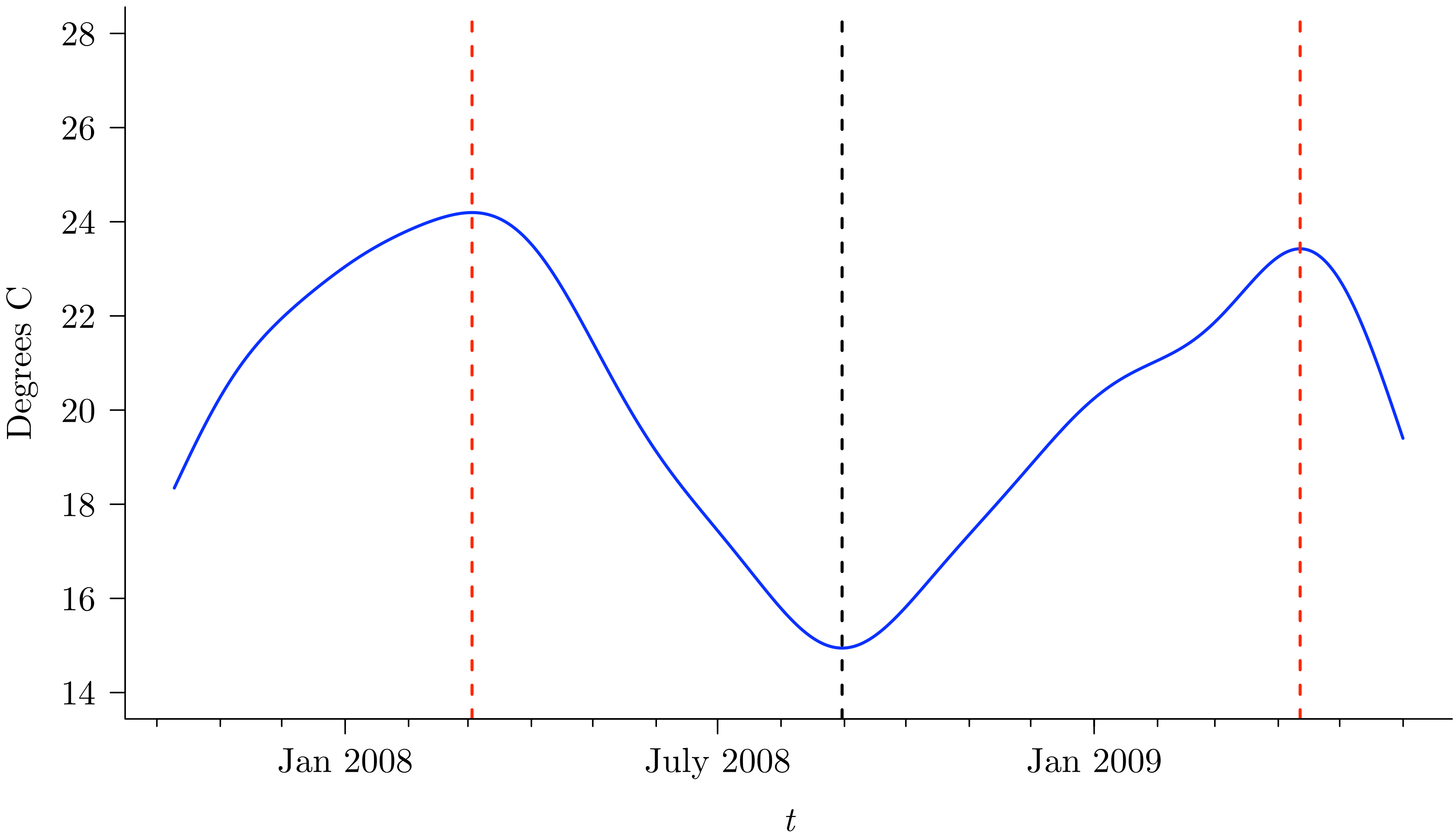
Annual Component for 10 m Water Temperatures



Annual Component for 15 m Water Temperatures

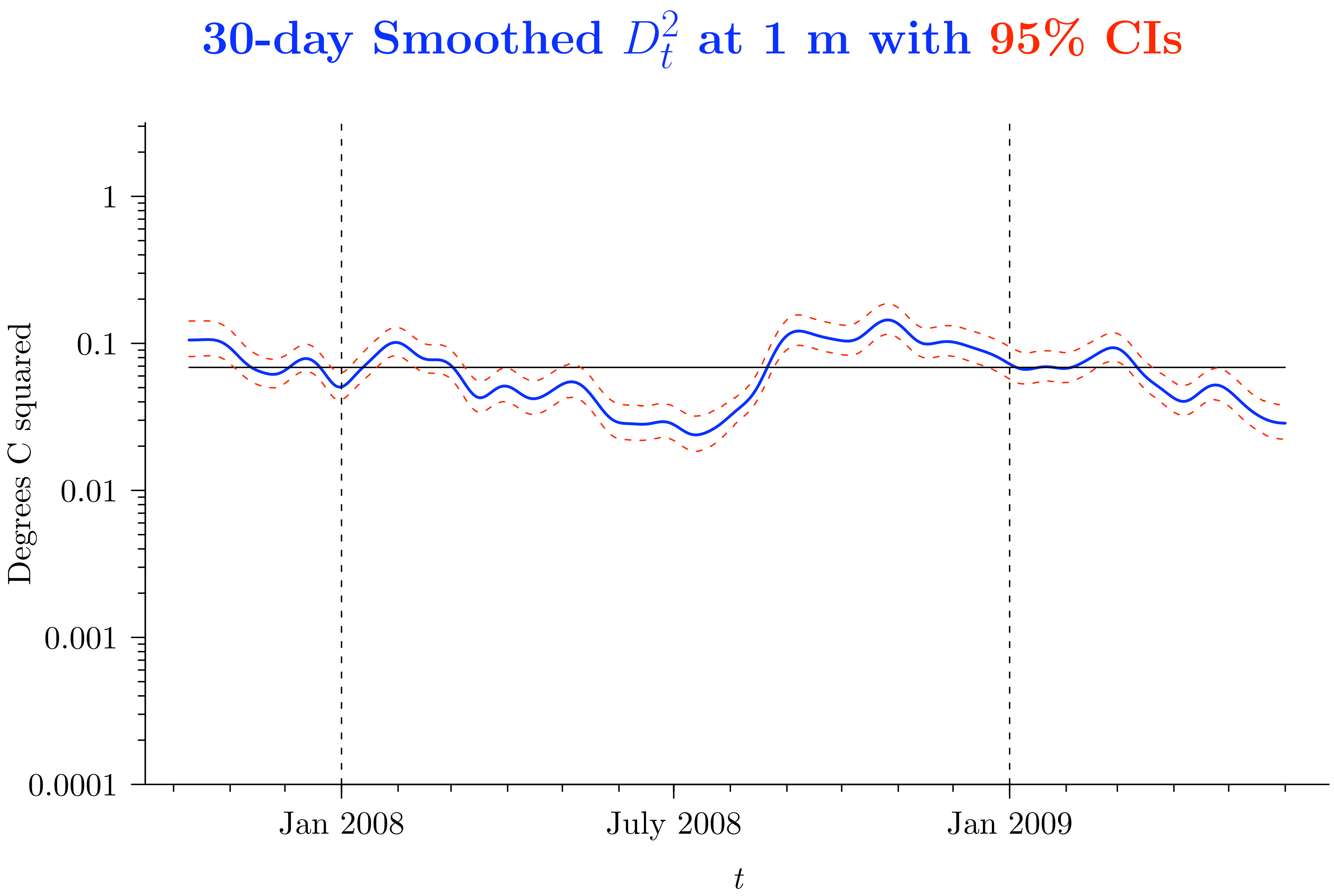


Annual Component for 20 m Water Temperatures

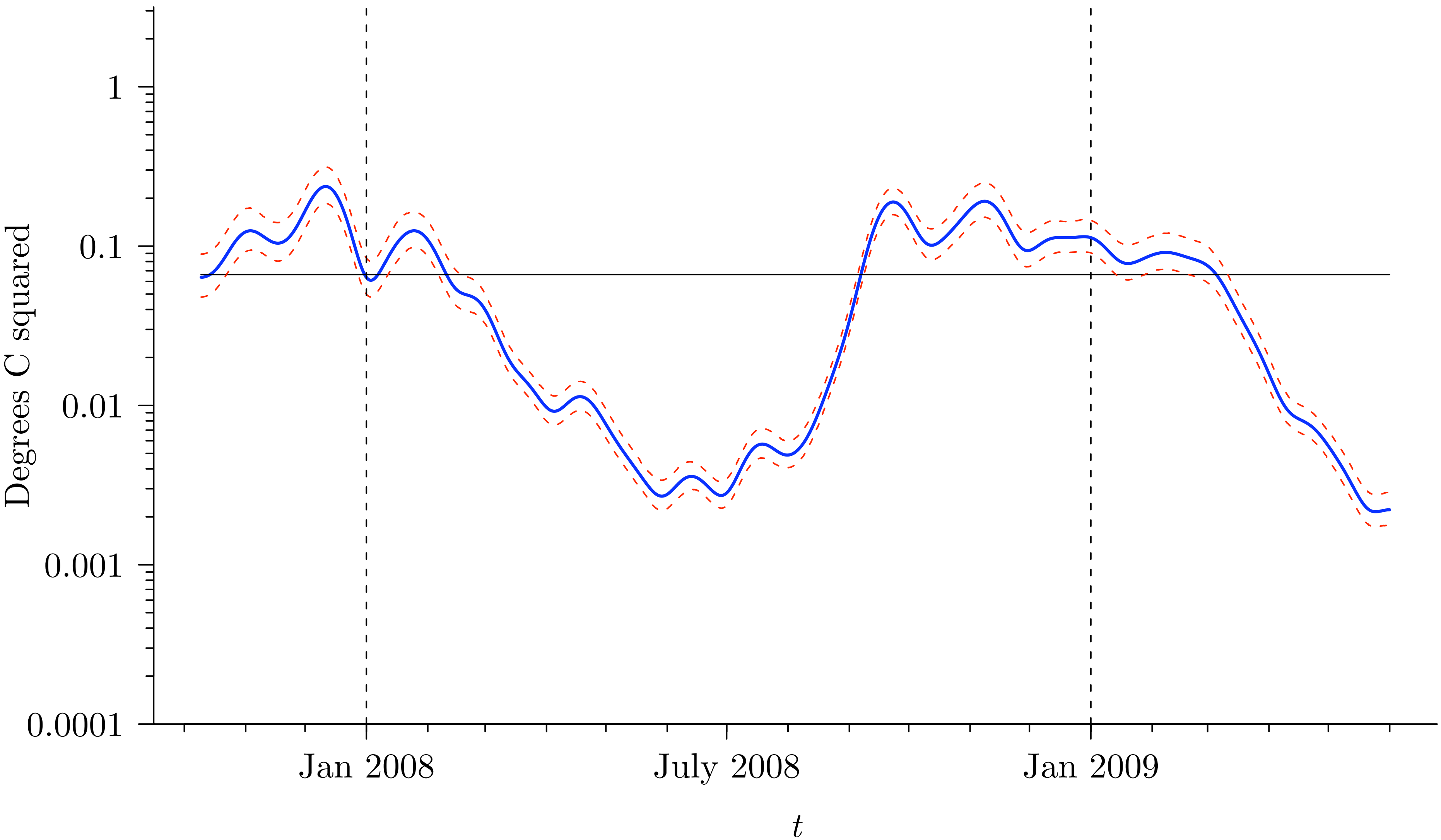


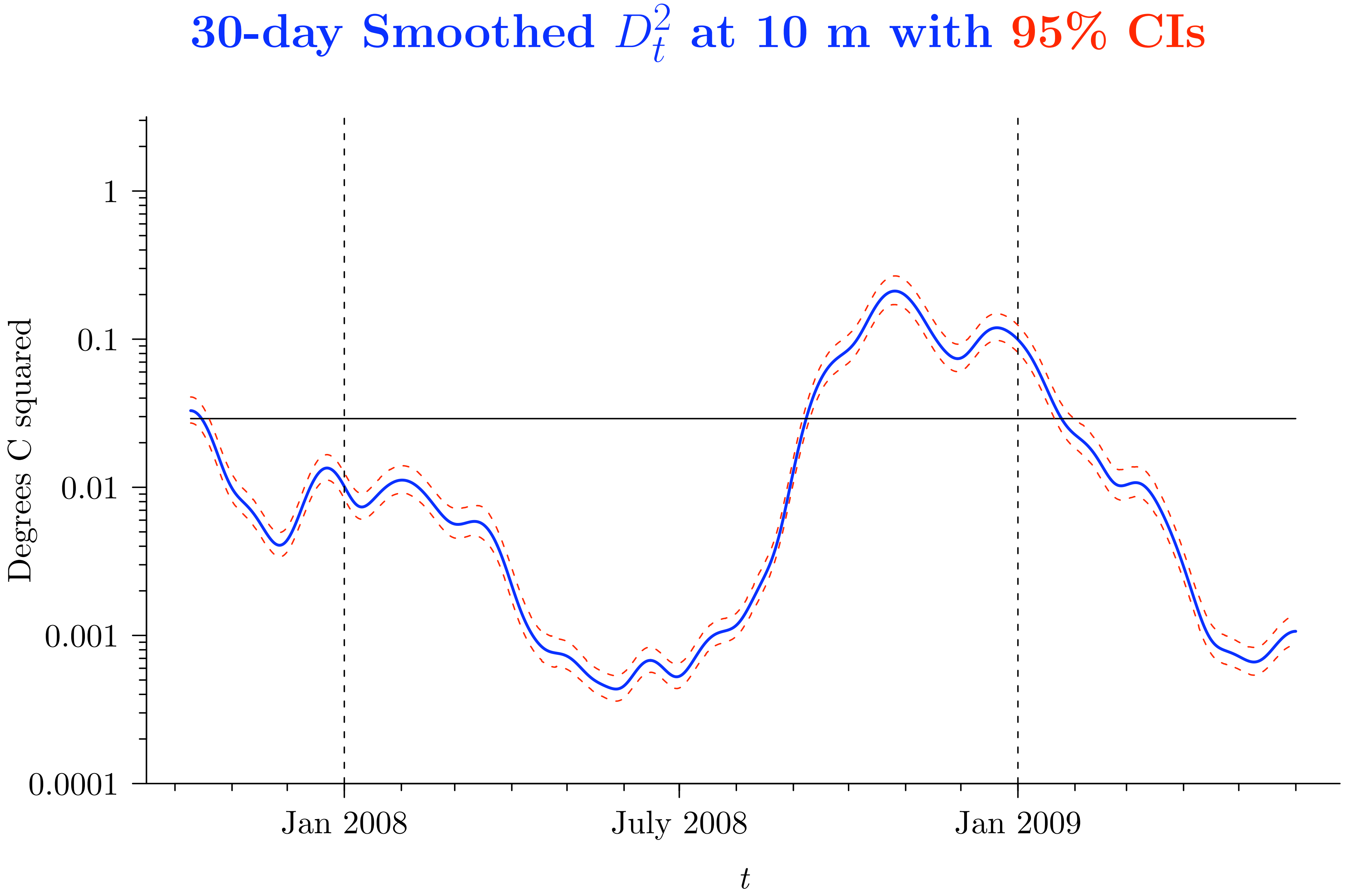
Notes on Annual Components

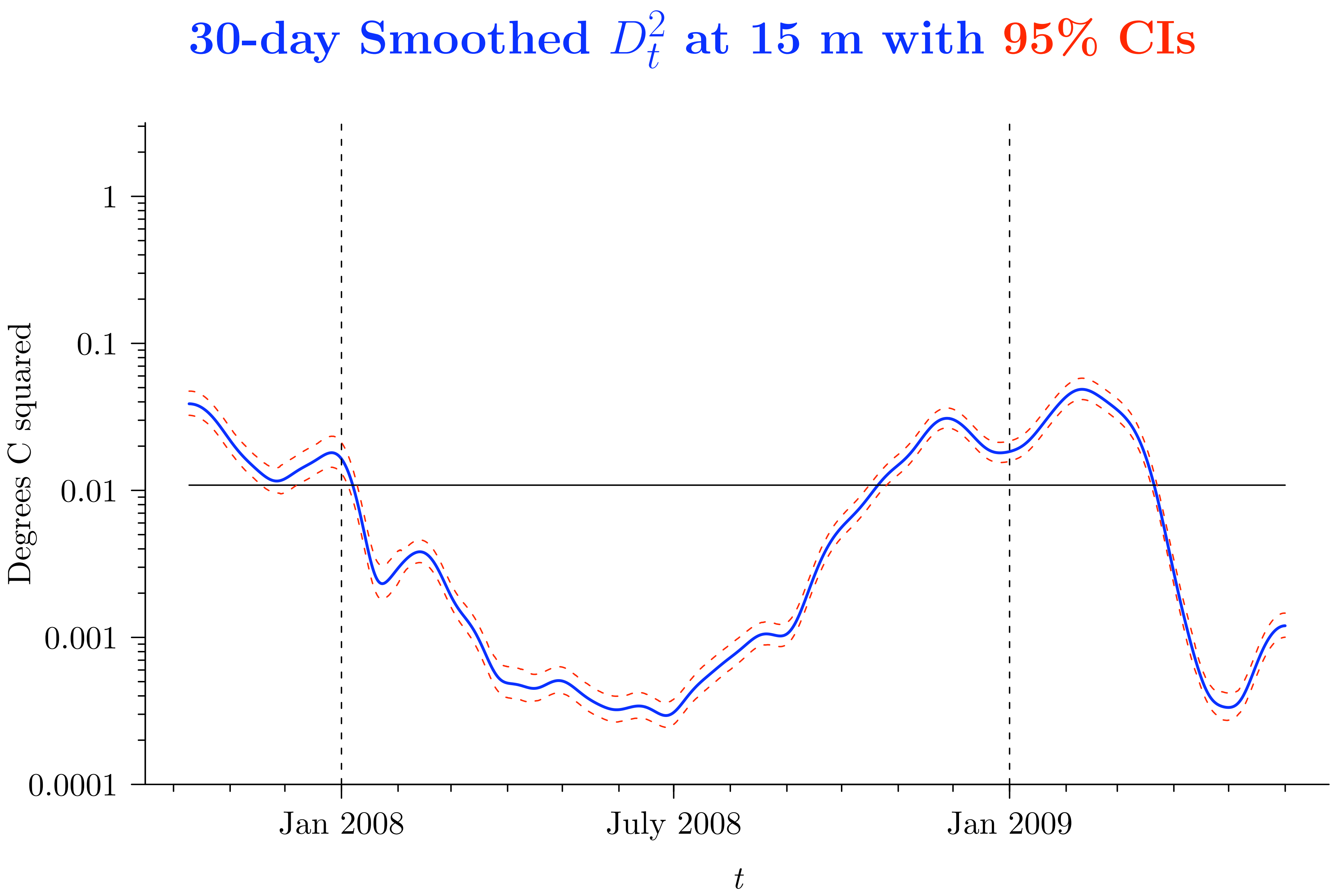
- overall height decreases with depth
- times of peaks in both 2008 and 2009 and of valley in 2008 arrive at later times the deeper the depth
- times of peaks in 2008 occur between start of February and start of March (span of one month)
- times of peaks in 2009 occur between start of January and start of April (span of three months)
- time span between 2008 & 2009 peaks increases with depth (11.5 months for 1 & 5 m, increasing to 13 months at 20 m)
- times of peaks for 1 & 5 m closely linked in both 2008 & 2009
- annual component in 2008 at each depth does not repeat itself in 2009

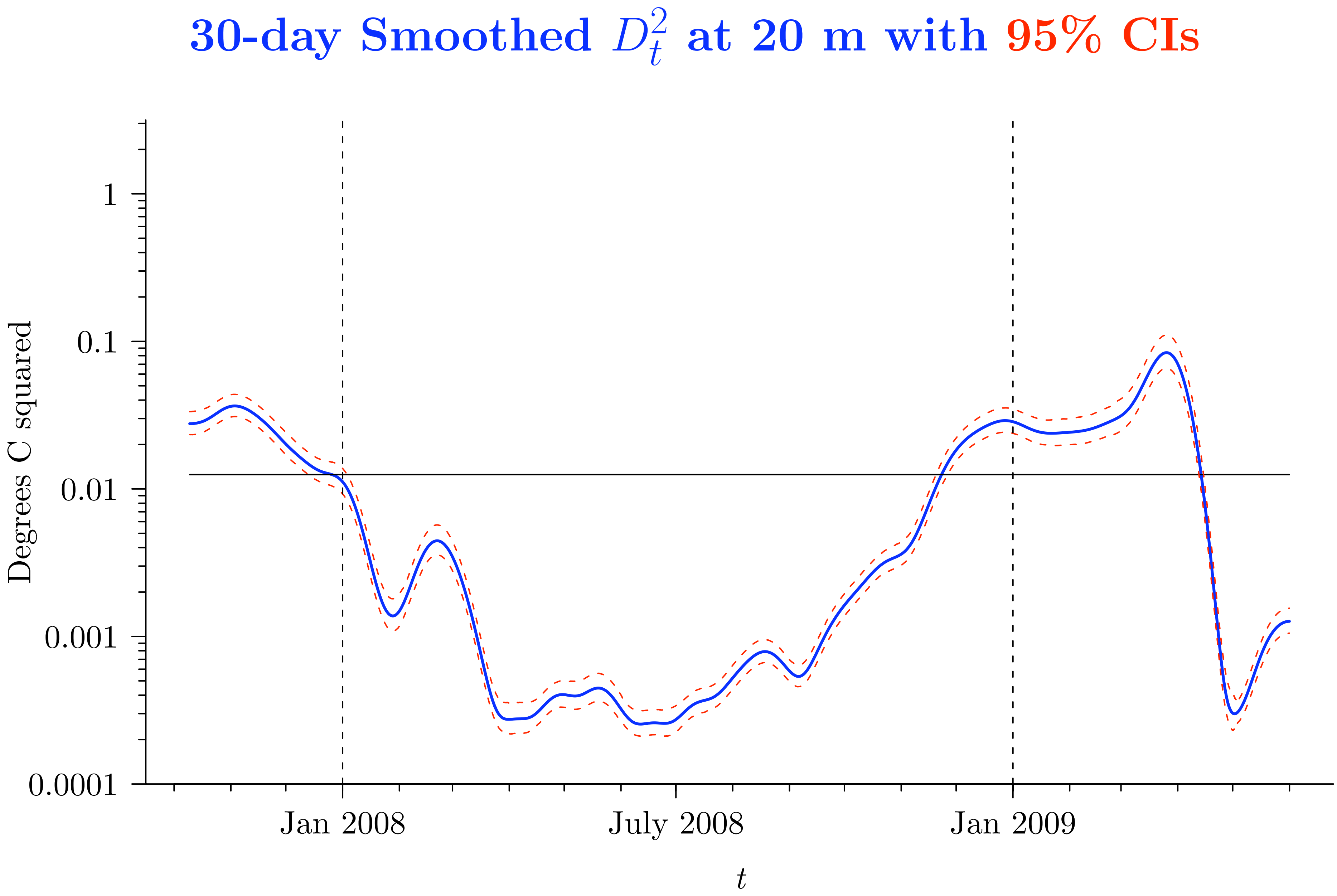

30-day Smoothed D_t^2 at 1 m with 95% CIs
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t

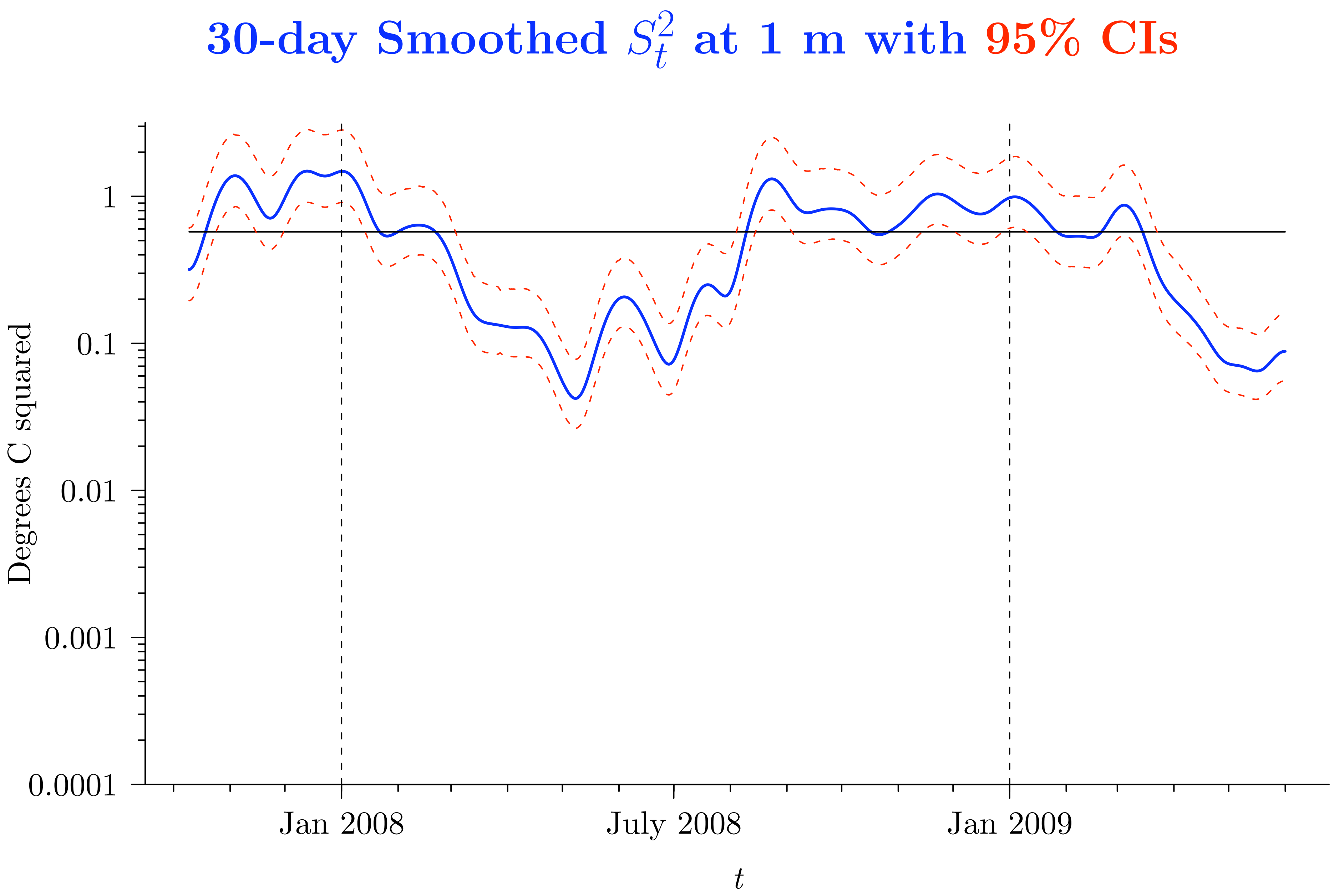
30-day Smoothed D_t^2 at 5 m with 95% CIs



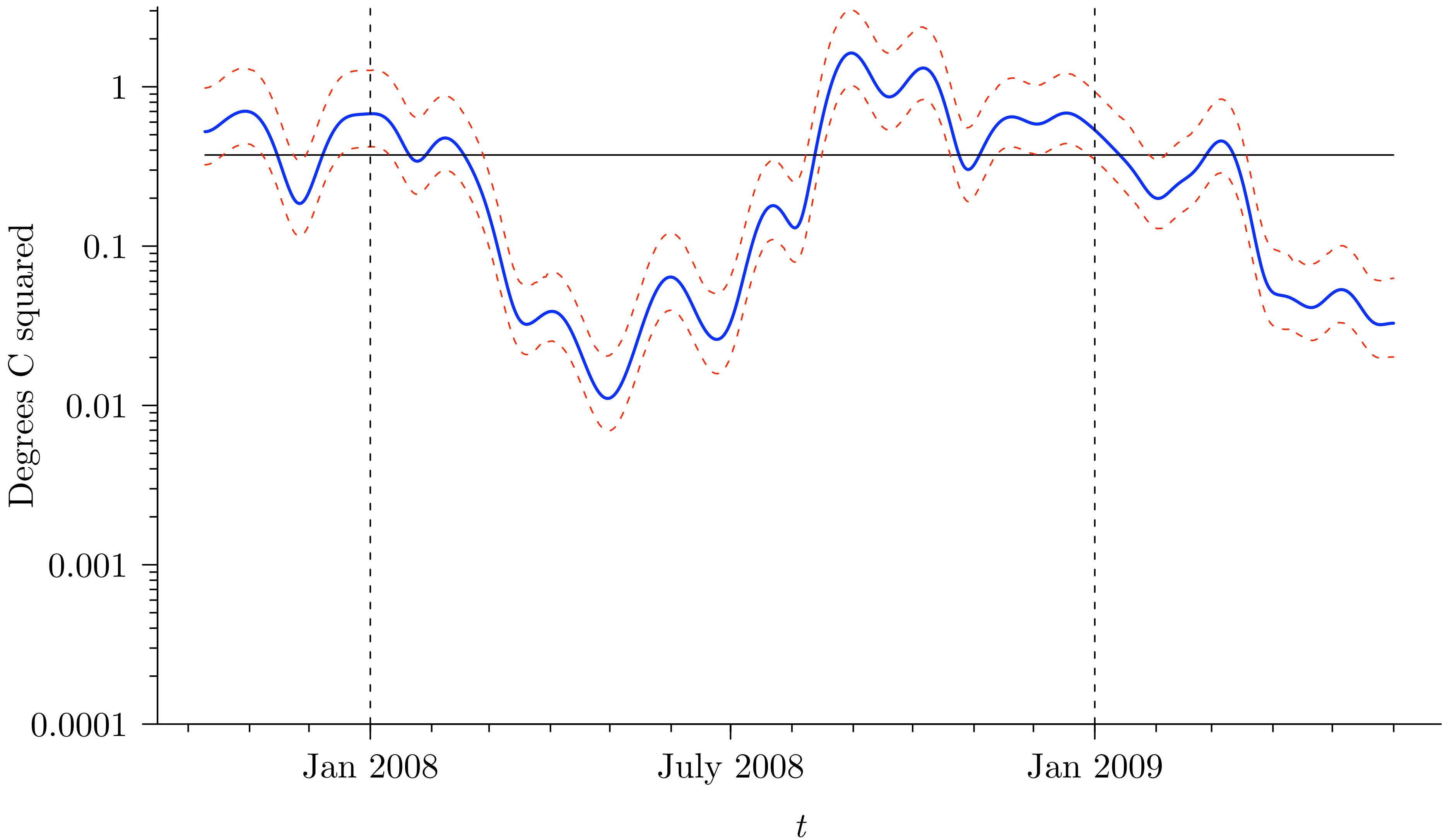

30-day Smoothed D_t^2 at 10 m with 95% CIs
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t

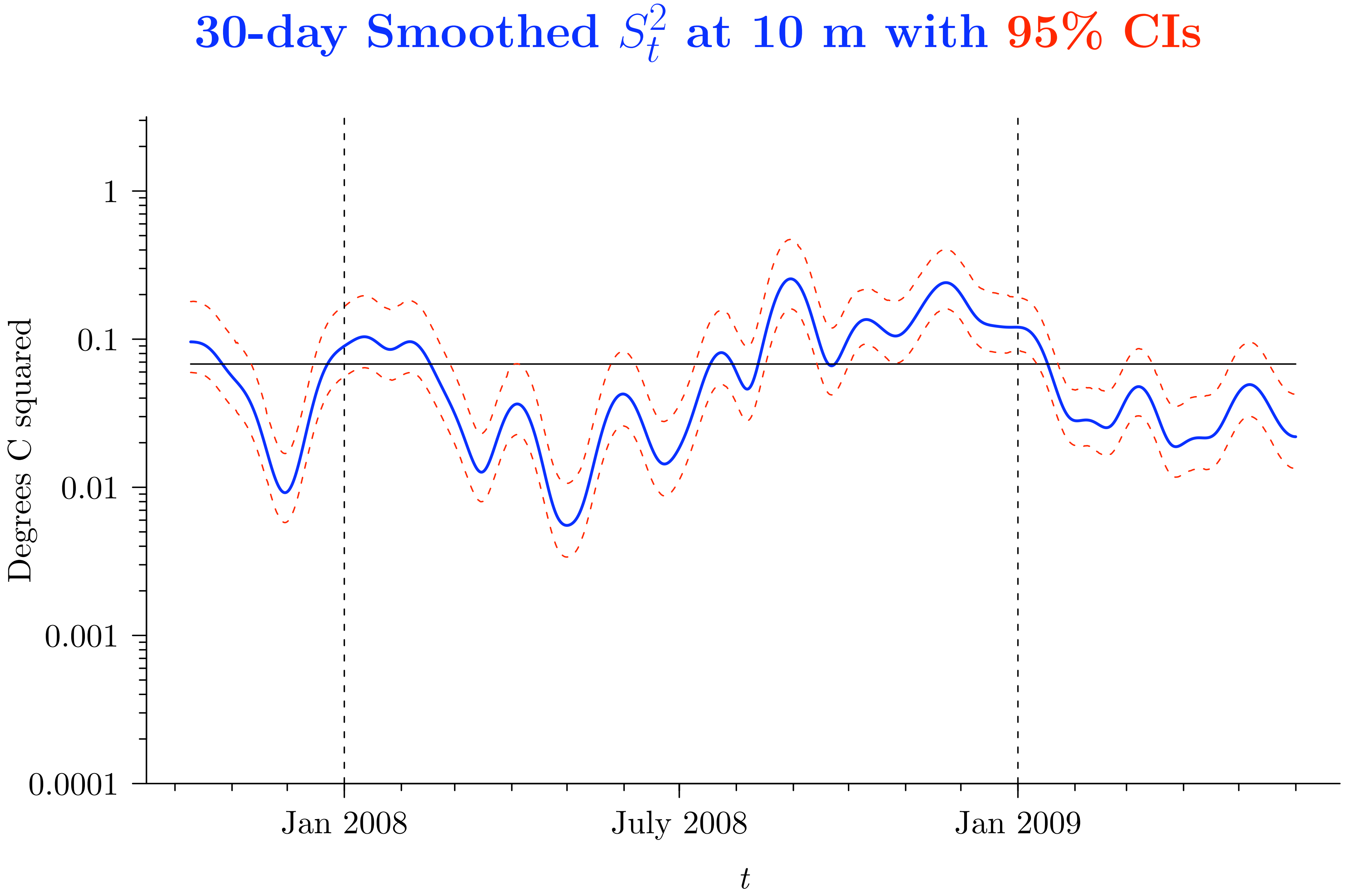

30-day Smoothed D_t^2 at 15 m with 95% CIs
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t

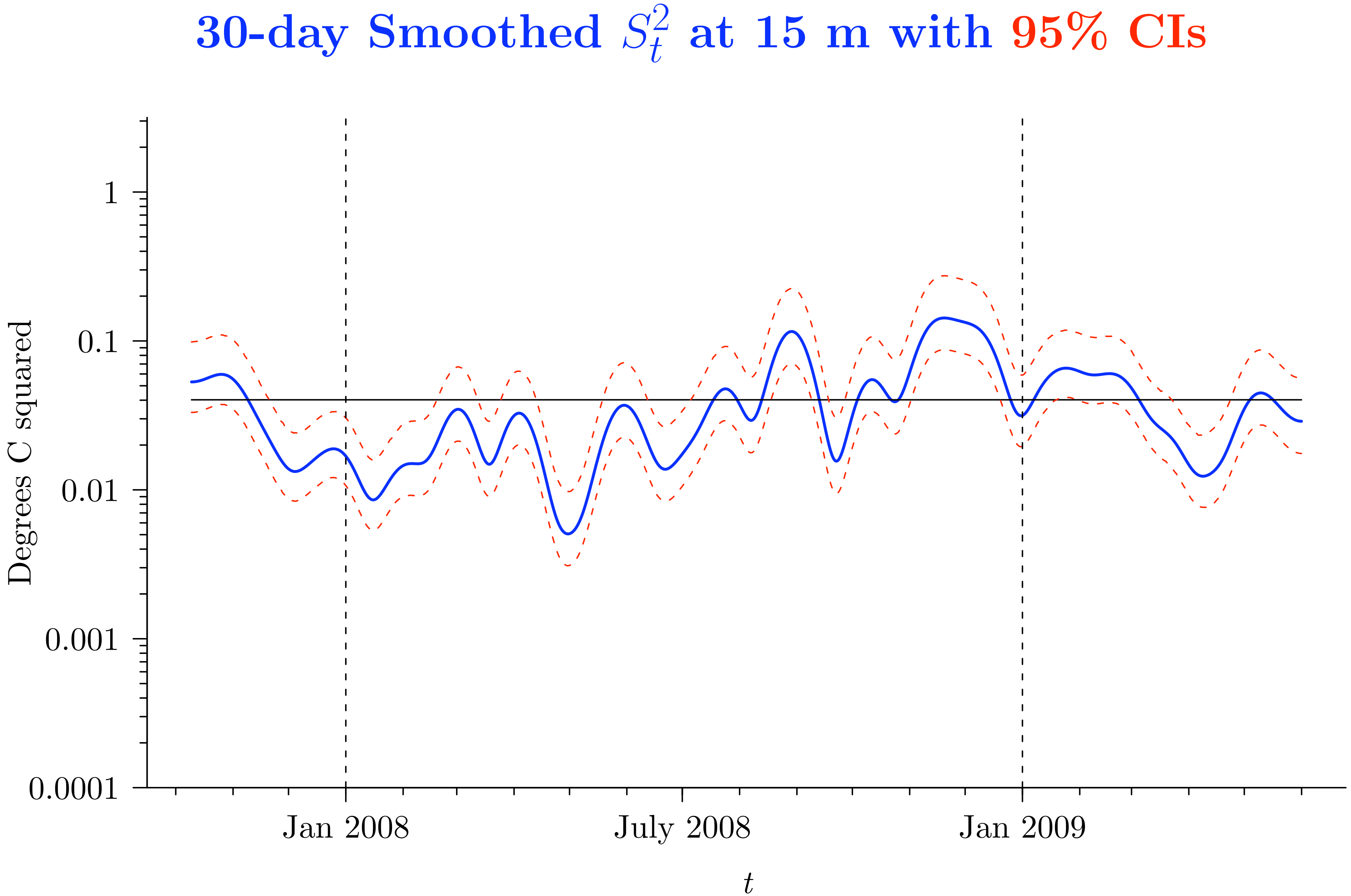

30-day Smoothed D_t^2 at 20 m with 95% CIs
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t


30-day Smoothed S_t^2 at 1 m with 95% CIs
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t

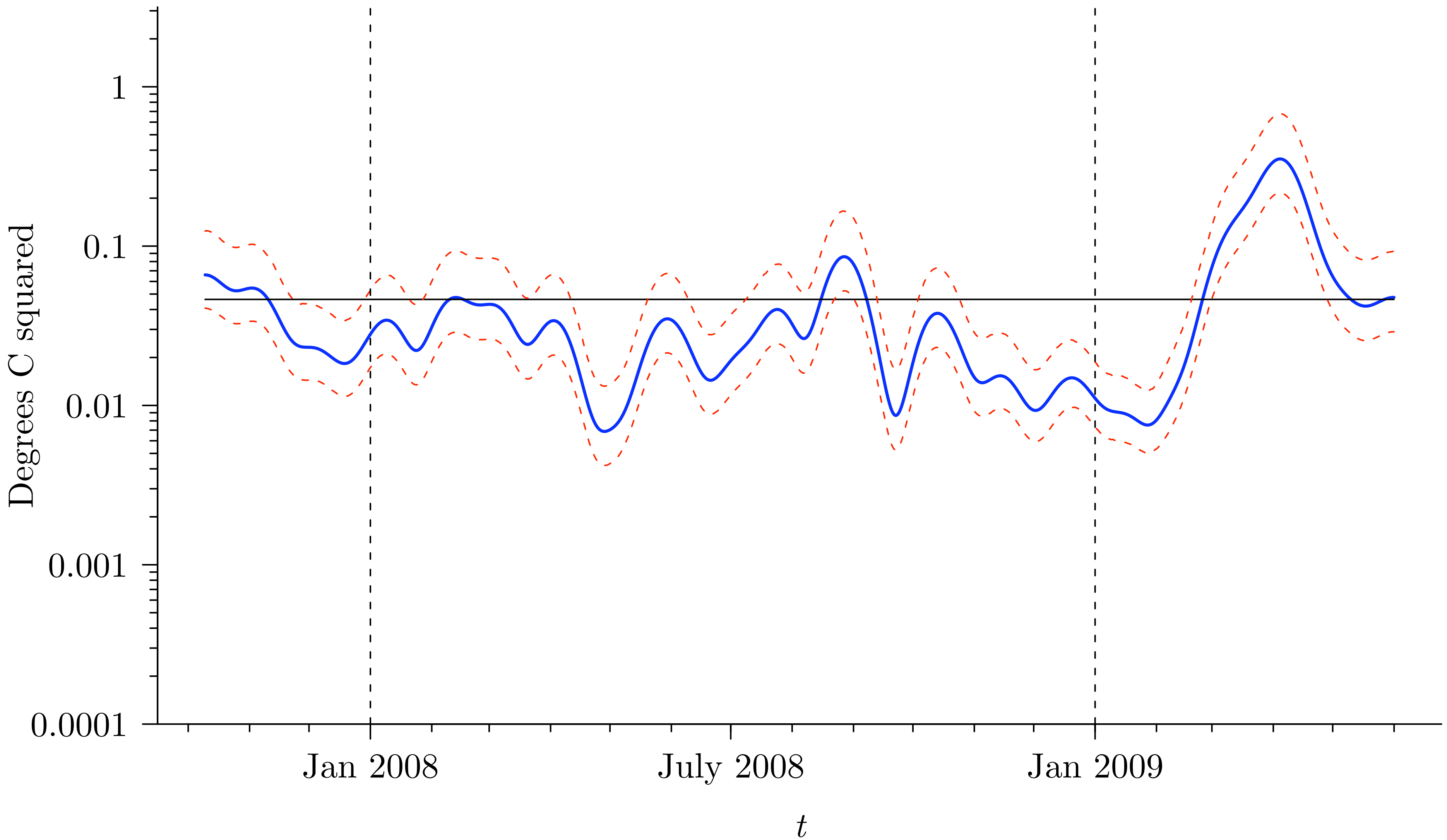
30-day Smoothed S_t^2 at 5 m with 95% CIs

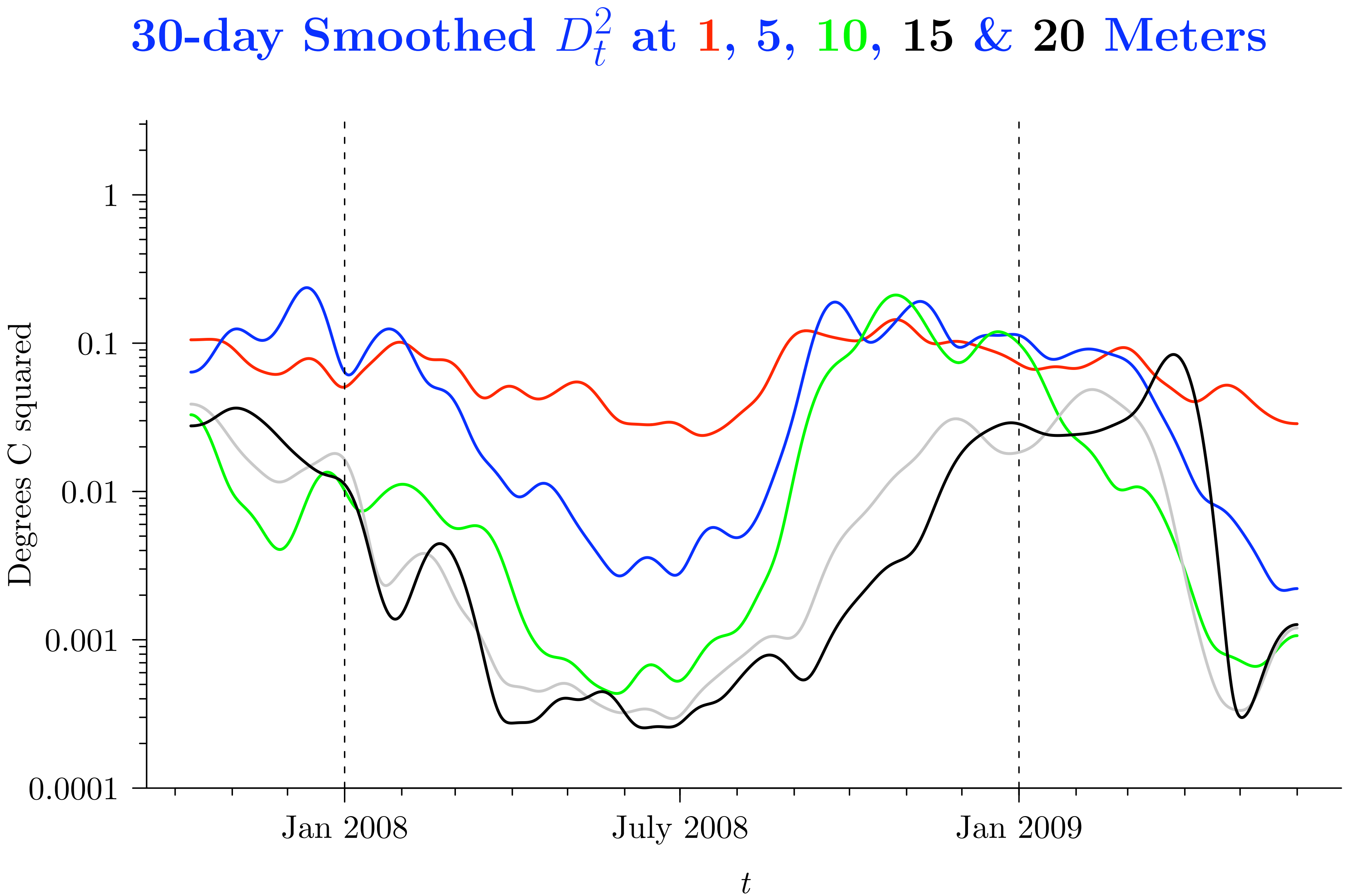



30-day Smoothed S_t^2 at 10 m with 95% CIs
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t

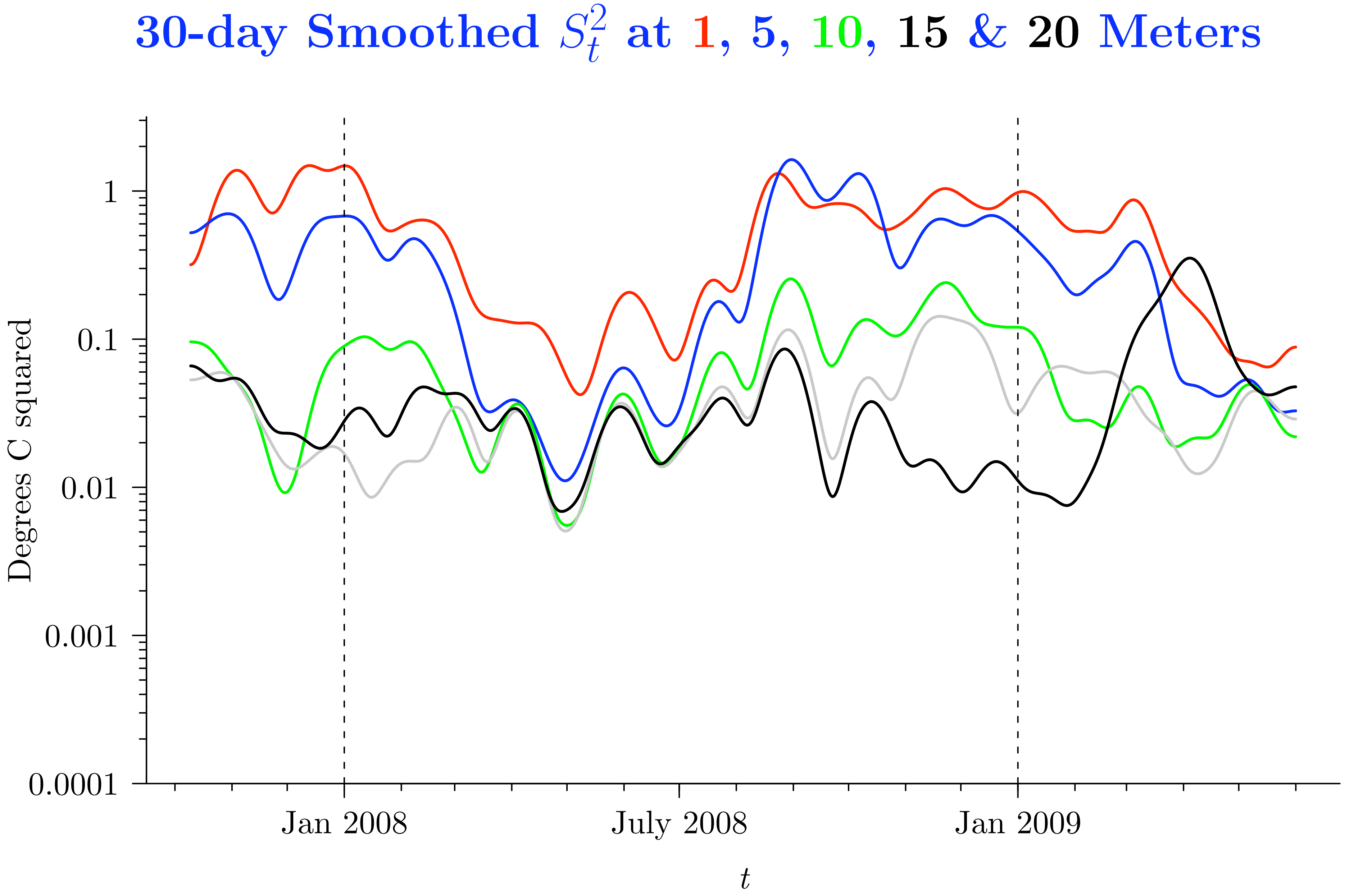

30-day Smoothed S_t^2 at 15 m with 95% CIs
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t

30-day Smoothed S_t^2 at 20 m with 95% CIs





30-day Smoothed D_t^2 at 1, 5, 10, 15 & 20 Meters
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t


30-day Smoothed S_t^2 at 1, 5, 10, 15 & 20 Meters
Degrees C squared
1
0.1
0.01
0.001
0.0001
Jan 2008
July 2008
Jan 2009
t

Notes on Time-Varying Variances of D and S

- variance of **D** at 1 m is relatively stable across time, but not at lower depths
- opposite pattern holds for **S**: three lower depths have more homogeneous variances than two shallower ones

Global Cross-Correlations Between DSA Coefficients

- 15 sets of coefficients in all (**D**, **S** and **A** at 5 depths)
- there are $\binom{15}{2} = 105$ cross-correlations to consider
- 75 are 'between-type' cross-correlations, i.e., involving different types of coefficients either at same depth or different depths
- between-type cross-correlations generally small: 6 are between 0.1 & 0.15, and remaining 69 are between -0.03 and 0.1
- lends credence to claim that DSA transform is separating $\mathbf{X}$ into different types of coefficients (**D**, **S** and **A**) that are approximately uncorrelated
- remaining 30 cross-correlations are 'within-type'

Global Within-Type Cross-Correlations

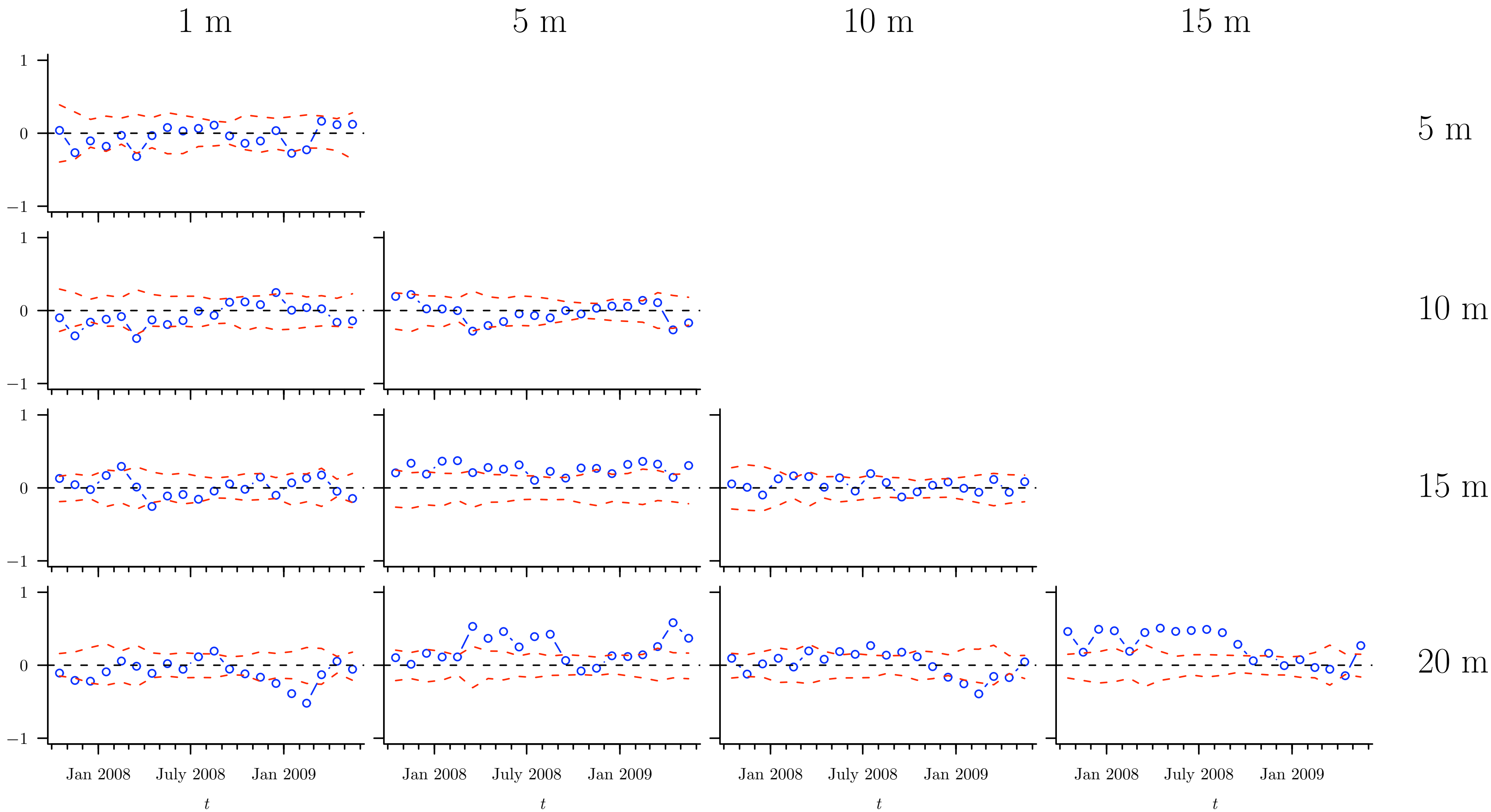
D	1 m	5 m	10 m	15 m
5 m	-0.09			
10 m	0.03	0.22		
15 m	0.05	0.01	0.06	
20 m	0.05	-0.18	-0.07	0.12

S	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

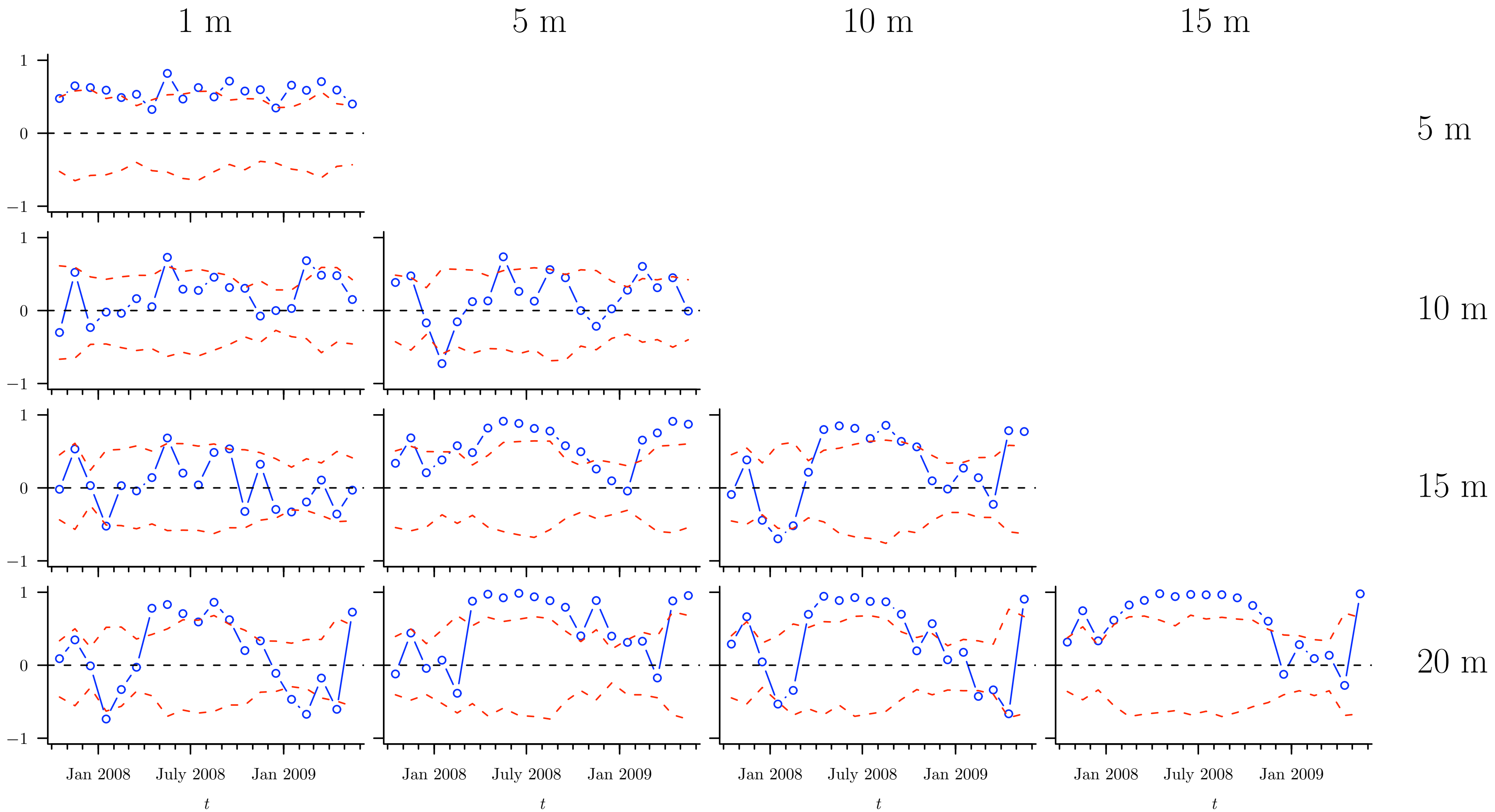
A	1 m	5 m	10 m	15 m
5 m	0.99			
10 m	0.92	0.95		
15 m	0.79	0.84	0.96	
20 m	0.68	0.74	0.89	0.97

X	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

Month-by-month Correlations for D with 95% CIs



Month-by-month Correlations for S with 95% CIs



Notes on Month-by-Month Cross-Correlations

- cross-correlations for **D** tend to be smaller and less time dependent than those for **S**
- in particular, small cross-correlations between **D** at 1 m and deeper depths indicate little direct daily co-temporaneous variations between temperatures near surface and at deeper levels
- cross-correlations for **S** have stretches of high correlation, e.g., between 15 and 20 m from Feb to Sept 2008, followed by gradual decline (period also associated with decreased variability at both depths)
- periods of high correlation well aligned with known periods of stratification

Concluding Remarks: I

- biggest contributor to variance of **X** is **A** (annual coefficients)
- evident from just 600 days of data that **A** can vary considerably from year to year (might be able to identify explanatory covariates from sparsely sampled historical data)
- next biggest contributor is **S** (subannual coefficients), while **D** (daily) is smallest
- **S** is more homogeneous in variability at deeper depths, whereas **D** is most homogeneous at 1 m (can interpret in terms of influence of atmospheric conditions)
- global statistics do not necessarily reflect localized patterns in various **X**, pointing to advantages of current sampling scheme and of localized measures such as the DSA transform

Concluding Remarks: II

- DSA approach is largely descriptive, but addresses some questions that could be answered also by formal statistical models
- *plenty* of opportunity for future work, including study of other water quality indicators (chlorophyll-a, turbidity, dissolved oxygen, specific conductivity) collected by the profiling system and their relationship to temperature

Thanks to . . .

- Sarah Lennox for organizing this seminar
- numerous folks at CSIRO who made my visit possible
- Seqwater for opportunity to analyze their data