Testing and interpreting coherence in biological fluctuations

Lawrence W. Sheppard[1], James R. Bell[2], Richard Harrington[2], Emma Defriez[3], Philip C. Reid[4], Vesna Vuksanovic[5], Peter V.E. McClintock[6], Aneta Stefanovska[6], and Daniel C. Reuman[1,7]

[1]University of Kansas, Department of Ecology and Evolutionary Biology and Kansas Biological Survey, Lawrence, USA

[2]Rothamsted Research, Rothamsted Insect Survey, Harpenden, UK

[3]Silwood Park, Imperial College London, UK

[4]SAHFOS, The Laboratory, Plymouth, UK

[5]Bernstein Center for Computational Neuroscience, Berlin, Germany

[6]Lancaster University, Physics Department, Lancaster, UK

[7]Rockefeller University, Laboratory of Populations, New York, USA

Motivation

In biophysical systems multivariate data is often autocorrelated and cross correlated and incorporates a number of different spectral components.

Wavelet coherence methods can be used to evaluate the strength of statistical relationships between the fluctuations found in these signals.

What do coherence methods measure?

A wavelet coherence measures the tendency for a particular phase relationship to exist between signals, and can illustrate how this relationship varies with frequency.

Different kinds of dependencies can produce different patterns of frequency dependent phase relationship:

Correlation and anticorrelation

Rate dependencies

AR models

Synchrony of periodic components

Combination of different periodic processes in the same time series etc.

For illustration, we generate two artificial time series each including two sinusoids with variable frequency (0.3 and 0.6 Hz) hidden in detrended brown noise. The 0.6Hz component is the same in both cases, the 0.3Hz components and noise are different.

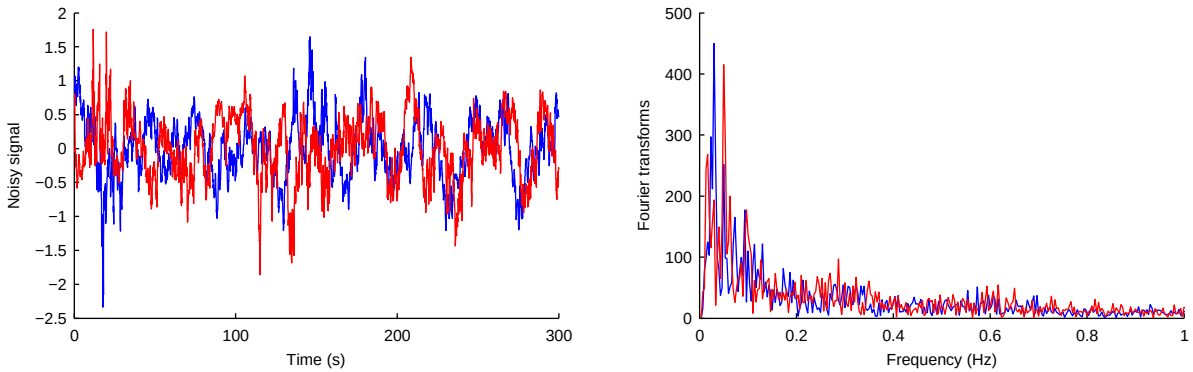


Figure : Example artificial data.

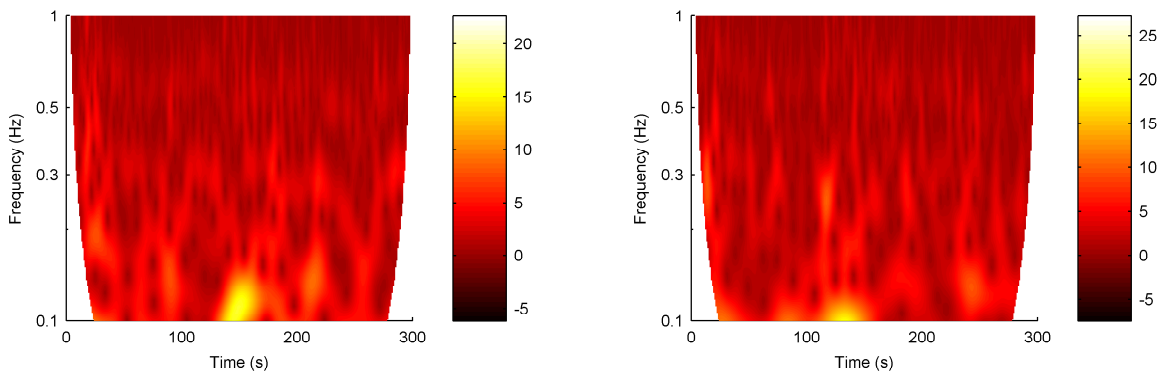


Figure : Wavelet transforms for the artificial data.

Wavelet phase coherence

Wavelet phase coherence is obtained at each frequency by extracting only the phase values, $\phi_k(t_n)$, from the wavelet transform [7]. The phase difference is $\phi(t_n) = \phi_1(t_n) - \phi_2(t_n)$.

$$\Pi_\Phi = \sqrt{\left(\frac{1}{N}\sum_{n=1}^{N} cos\phi(t_n)\right)^2 + \left(\frac{1}{N}\sum_{n=1}^{N} sin\phi(t_n)\right)^2} \tag{1}$$

$$= \left|\left(\frac{1}{N}\sum_{n=1}^{N} e^{i(\phi_1(t_n)-\phi_2(t_n))}\right)\right|. \tag{2}$$

This magnitude is the coherence and the phase of the summation tells us the phase difference that is maintained between these components.

Example coherence

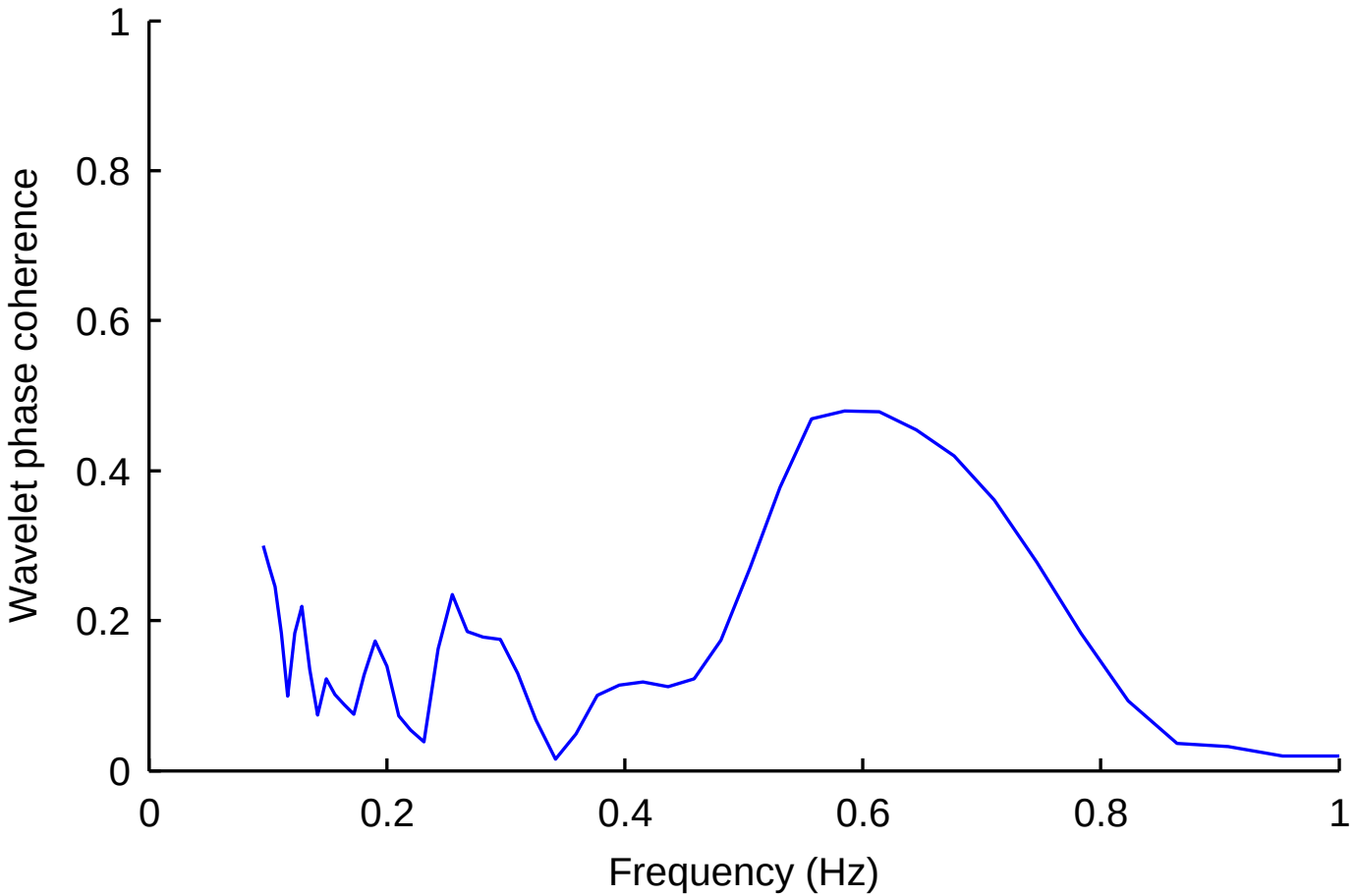


Figure : The wavelet phase coherence has a peak at 0.6Hz, where the same variability is found in both signals.

Phase differences as vectors

We consider a mean phasor $\left(\frac{1}{N}\sum_{n=1}^{N} e^{i(\phi_1(t_n)-\phi_2(t_n))}\right)$ as a walk in the complex plane. Non-interacting subsystems produce random walks; interacting subsystems produce directed walks. The nature and net length of the walk will depend on the self-correlations in the subsystems for both cases.

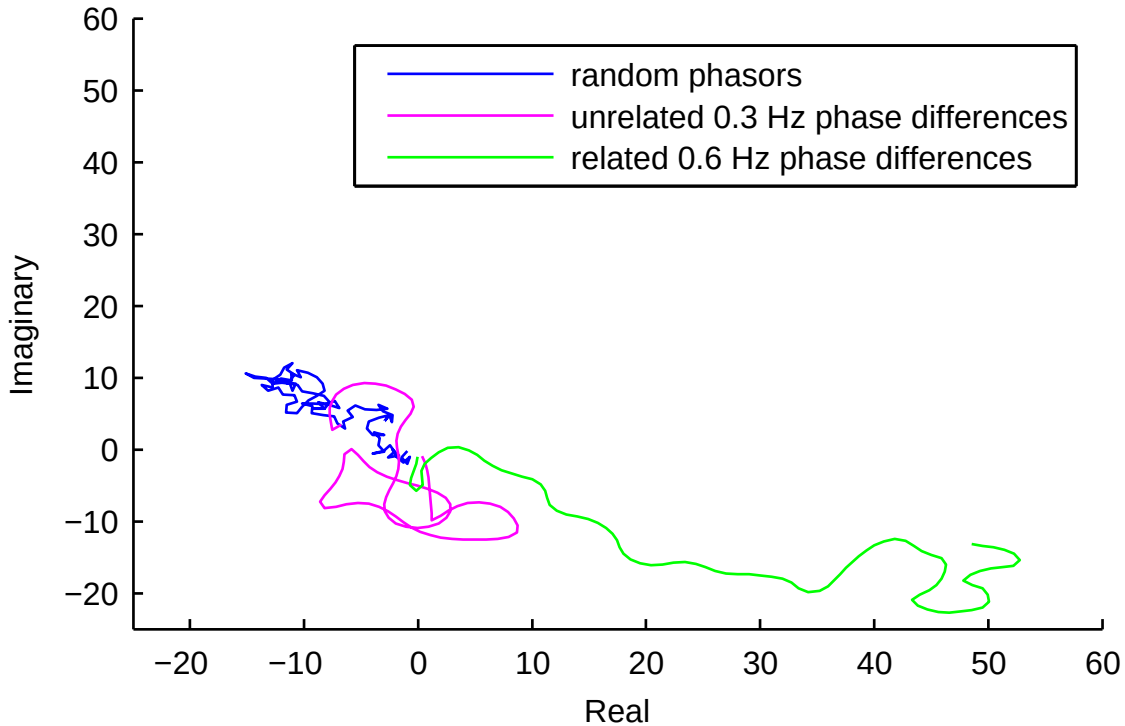


Figure : Walks in the complex plane.

Applied to biophysics

In biophysical systems multivariate data is often autocorrelated and cross correlated and incorporates a number of different spectral components.

Wavelet coherence methods can be used to evaluate the strength of statistical relationships between the fluctuations found in these signals.

Examples include blood flow and temperature fluctuations in human skin subject to heating and cooling [1], fluctuations in human blood pressure and skin conductivity during anaesthesia [2], and the intracranial pressure (ICP) and arterial blood pressure (ABP) in subjects with a traumatic brain injury [6]

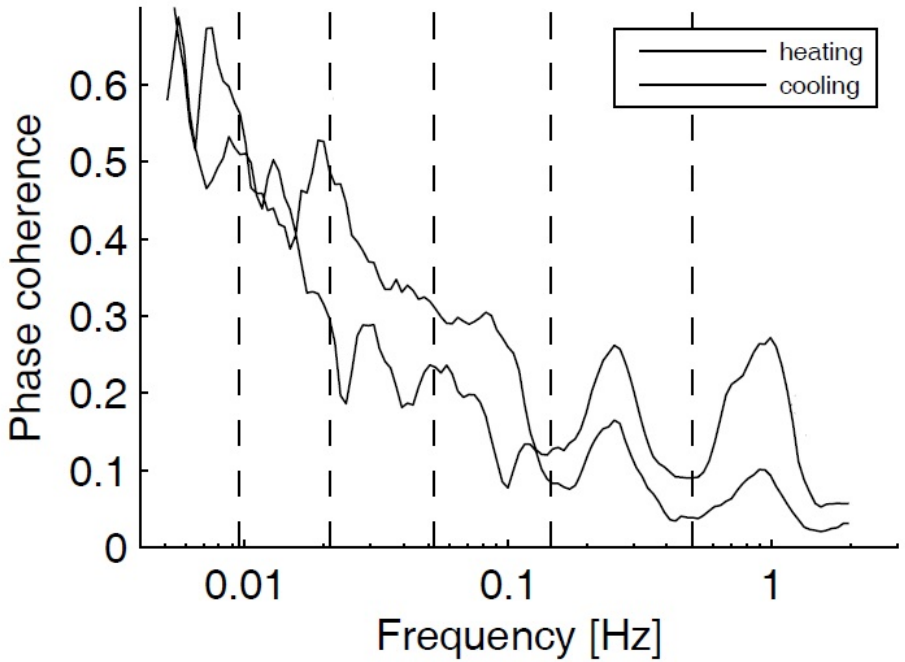


Figure : The phase coherence between skin blood flow and temperature is stronger in the heated (vasodilated) state

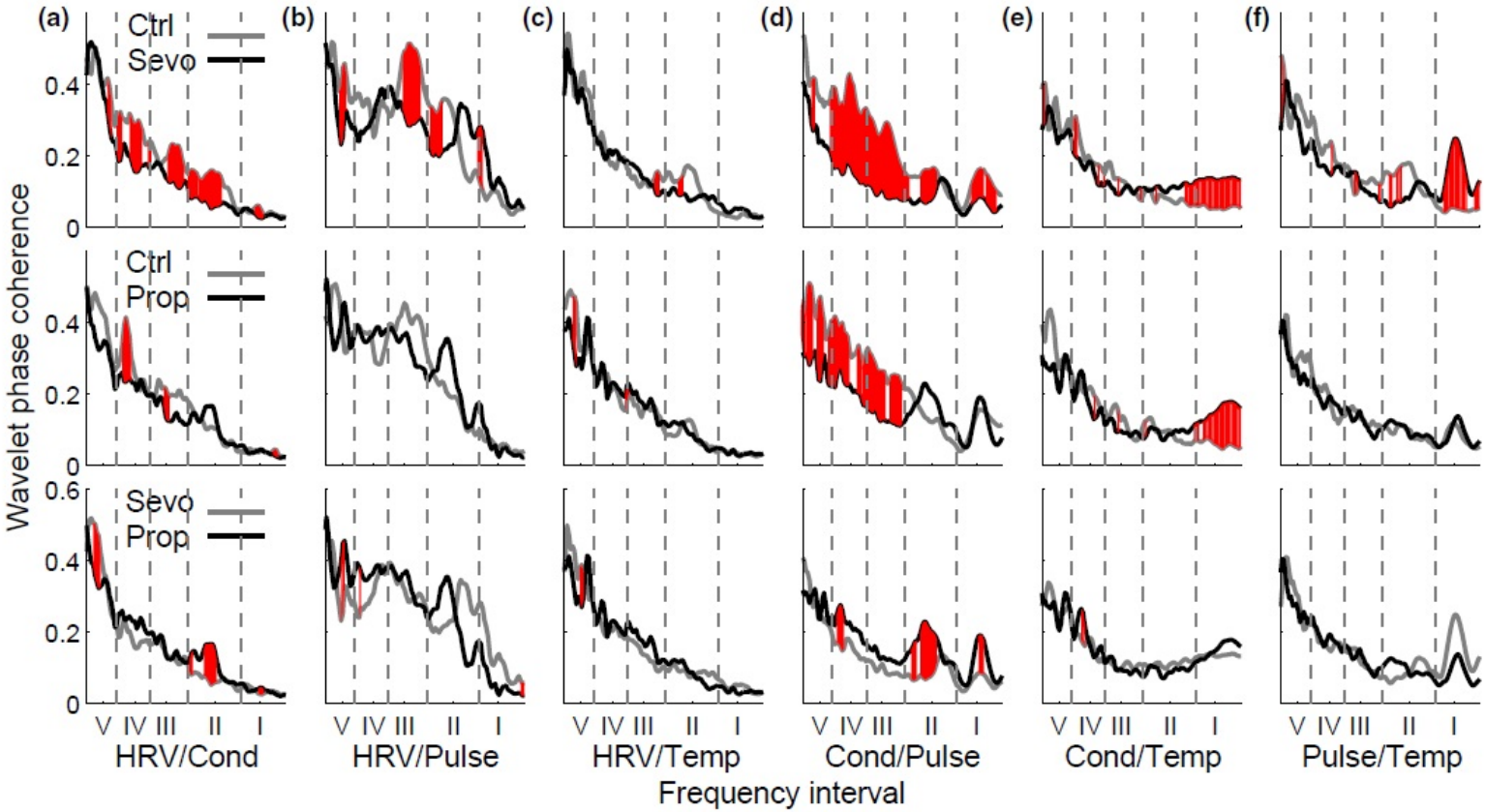


Figure : Average phase coherences for different groups of subjects compared. Top row: controls (grey) compared with sevoflurane-induced anaesthesia (black). Middle: controls (grey) and propofol-induced anaesthesia (black). Bottom: the two anaesthetic agents compared with each other. Red-shaded regions denote statistically significant differences ($p < 0.05$) between the two plots at the frequencies indicated.

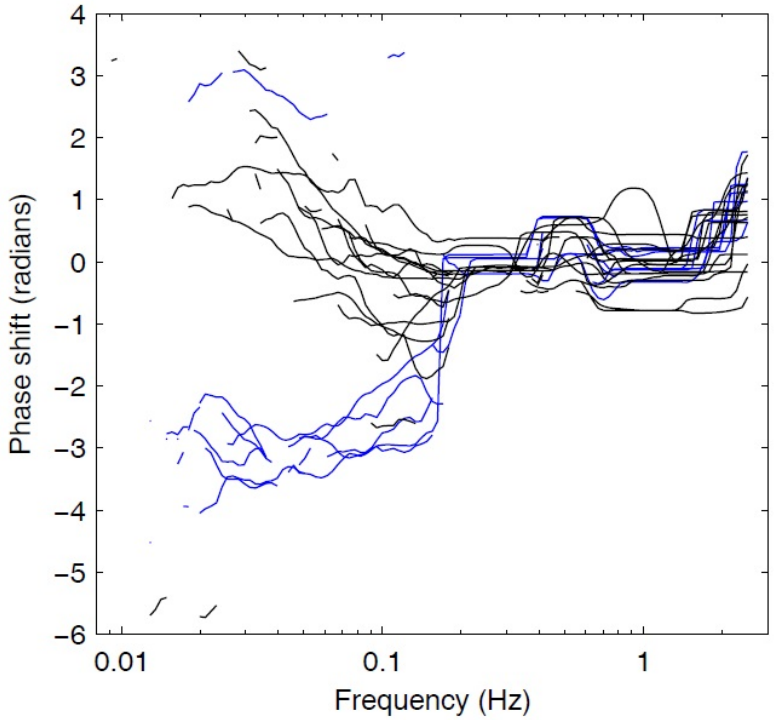


Figure : Good cerebrovascular reactivity in blue, poor cerebrovascular reactivity in black. Phase differences of ICP and ABP wavelet transforms plotted where coherent according to mismatches: two standard deviations or more above the mean of the mismatch distribution.

Wavelet coherence

Wavelet phase coherence can be written

$$\Pi_\Phi = \left| \left(\frac{1}{T} \sum_{t=1}^{T} e^{i(\phi_1(t) e^{-\phi_2(t))}} \right) \right|. \tag{3}$$

We can define a wavelet coherence to include the time-varying magnitudes of the wavelet components ($W_i(\sigma, t) = A_i(\sigma, t) e^{i\phi_i(\sigma,t)}$), with normalisation by wavelet power:

$$w_i(\sigma, t) = W_i(\sigma, t) / \sqrt{\left(\frac{1}{T} \sum_{t=1}^{T} W_i(\sigma, t) W_i(\sigma, t)^* \right)}, \tag{4}$$

$$\Pi_\Phi = \left| \left(\frac{1}{T} \sum_{t=1}^{T} w_1 w_2^* \right) \right|. \tag{5}$$

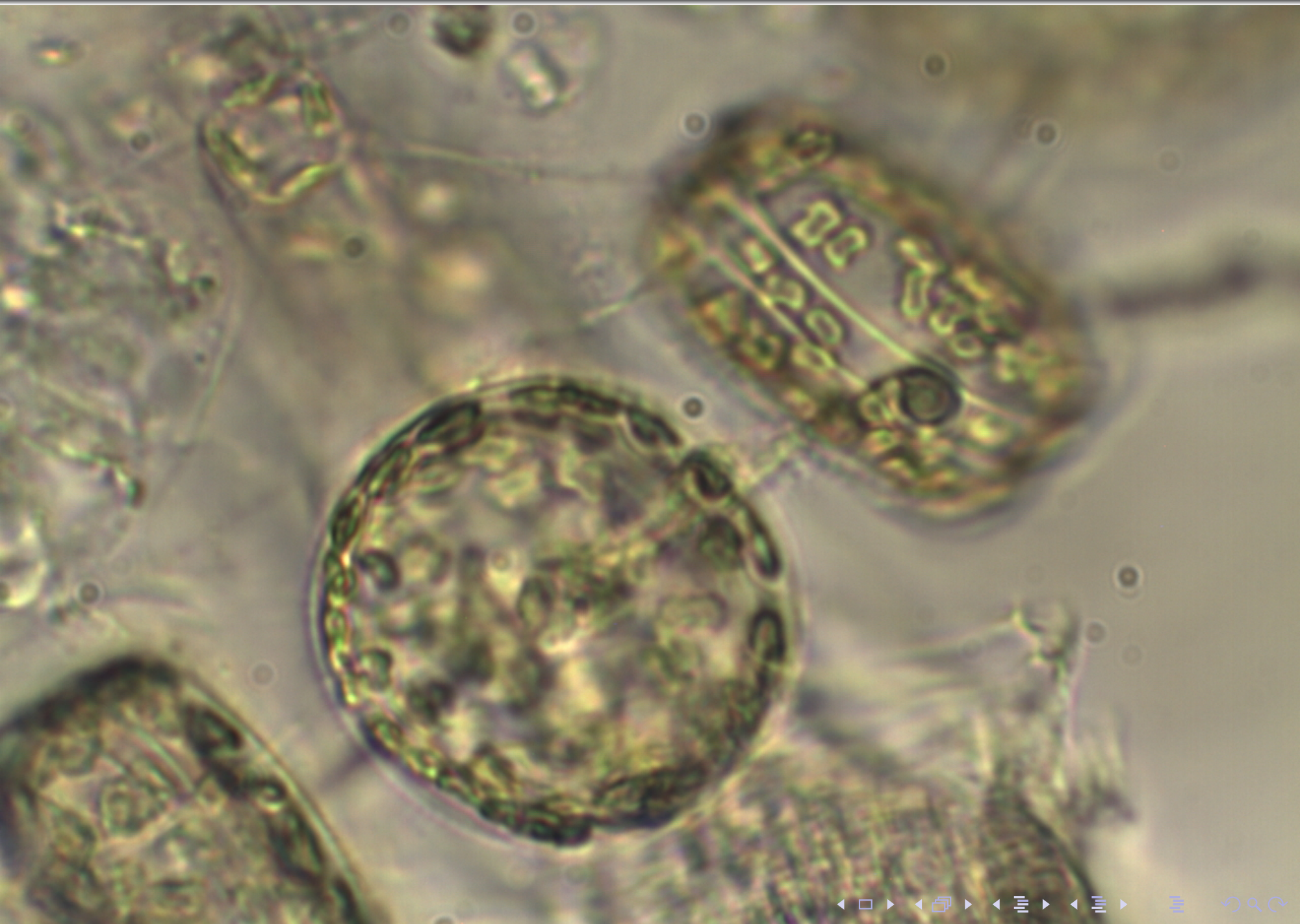
Analogous to a correlation coefficient, but maintaining the advantages of frequency specificity and sensitivity to phase.

Spatial coherence

Sometimes we have bivariate data drawn from a number (N) of locations simultaneously, in biomedical applications and in ecology. With appropriate normalisation, we can define a Spatial Coherence to measure the aggregate tendency for all the data to exhibit a given phase relationship [9]

$$w_i(\sigma, t) = W_i(\sigma, t) / \sqrt{\left(\frac{1}{N} \sum_{n=1}^{N} \frac{1}{T} \sum_{t=1}^{T} W_{n,i}(\sigma, t) W_{n,i}(\sigma, t)^* \right)}. \qquad (6)$$

$$\Pi_\Phi = \left| \left(\frac{1}{N} \sum_{n=1}^{N} \frac{1}{T} \sum_{t=1}^{T} w_{n,1} w_{n,2}^* \right) \right|. \qquad (7)$$



Plankton around the UK

The Continuous Plankton Recorder (CPR) survey at SAHFOS monitors near surface plankton in the sea around the British Isles.

We examine time series of annualised data representing changes in primary productivity (phytoplankton colour index or PCI) and the abundances of many individual plankton species, for 26 areas of the North Sea.

Using our wavelet approach we compare the fluctuations in relevant environmental and ecological variables.

Testing

We have a limited number of phase difference samples to average over, so determining the statistical significance of apparent phase relationships in spatio-temporal data requires careful testing. We must test against a properly constructed distribution of possible coherence values consistent with the null hypothesis that no actual phase relationship exists between variables, but including the properties of spatial and temporal autocorrelation that are present in the data. This distribution depends on the frequency of the wavelet component under consideration, and on how its rate of phase growth varies over time [4].

So-called Fourier surrogates are artificial timeseries that can be generated to estimate the null distribution [5,8]. These surrogates preserve the Fourier transform power spectrum, and thus the autocorrelation function, of the data, but they do not preserve the temporal information.

Fourier surrogates

To generate a Fourier surrogate timeseries we first use the efficient FFT algorithm to find the Fourier transform of the data. The FFT of the data has phase and amplitude information for each frequency. For real data it is symmetric about the Nyquist frequency.

We scramble the phase information by adding a vector of random phases with the same symmetry properties, then Inverse FFT to obtain a timeseries in which internal correlations are preserved but correlations with other timeseries are lost.

If we have multivariate data with known self- and cross- correlations, and we wish to test their relationship with further data, we may wish to use the same vector of random phases to produce multivariate surrogates.

Normalisation and Scalloping

The Fourier surrogate procedure does not preserve the distribution of the values in the data, it yields values with a Gaussian distribution. It may be necessary to detrend, normalise, and box-cox transform data during pre-processing for good comparability with the surrogates

The wavelet transform is typically scalloped to remove edge values. This makes estimation of the actual phase coherence more accurate, but also discards some useful information from the ends of the timeseries. We can include edge values provided we are comparing the actual and surrogate coherences including these edge values. Numerical experimentation with artificial data indicates this test has the same false positive rate and lower false negative rate than testing with scalloped data.

A simpler algorithm

We apply Fourier-phase-shuffled-surrogate based methods [5] to testing large spatio-temporal data sets for wavelet coherence. A method is described for testing wavelet coherence with a high degree of computational efficiency, exploiting the relationship between the Fourier and Morlet wavelet transforms.

In this way we can robustly demonstrate relationships between paired timeseries (for example, fluctuations in abundance of two plankton species [3]) drawn from a large number of locations. The methods could also be applied to purely physical variables or biomedical timeseries drawn from multiple measurement sites on the human body.

Efficient computation

We can make use of a number of mathematical relationships to make the calculation of a null coherence distribution highly efficient.

By the convolution theorem the Morlet wavelet transform can be efficiently calculated by FFT'ing the data, multiplying by the FFT of the Morlet wavelet at each scale, and inverse FFT'ing.

In other words, the FFT of a wavelet component at given scale is the FFT of the data multiplied by a weighting vector.

Using Parseval's theorem, the wavelet coherence (a product summed over time) can be found from the FFT's of the two wavelet components (a product summed over frequency).

The coherence is thus a weighted sum of the Fourier cross spectrum (the FFT of one timeseries multiplied by the complex conjugate FFT of the other).

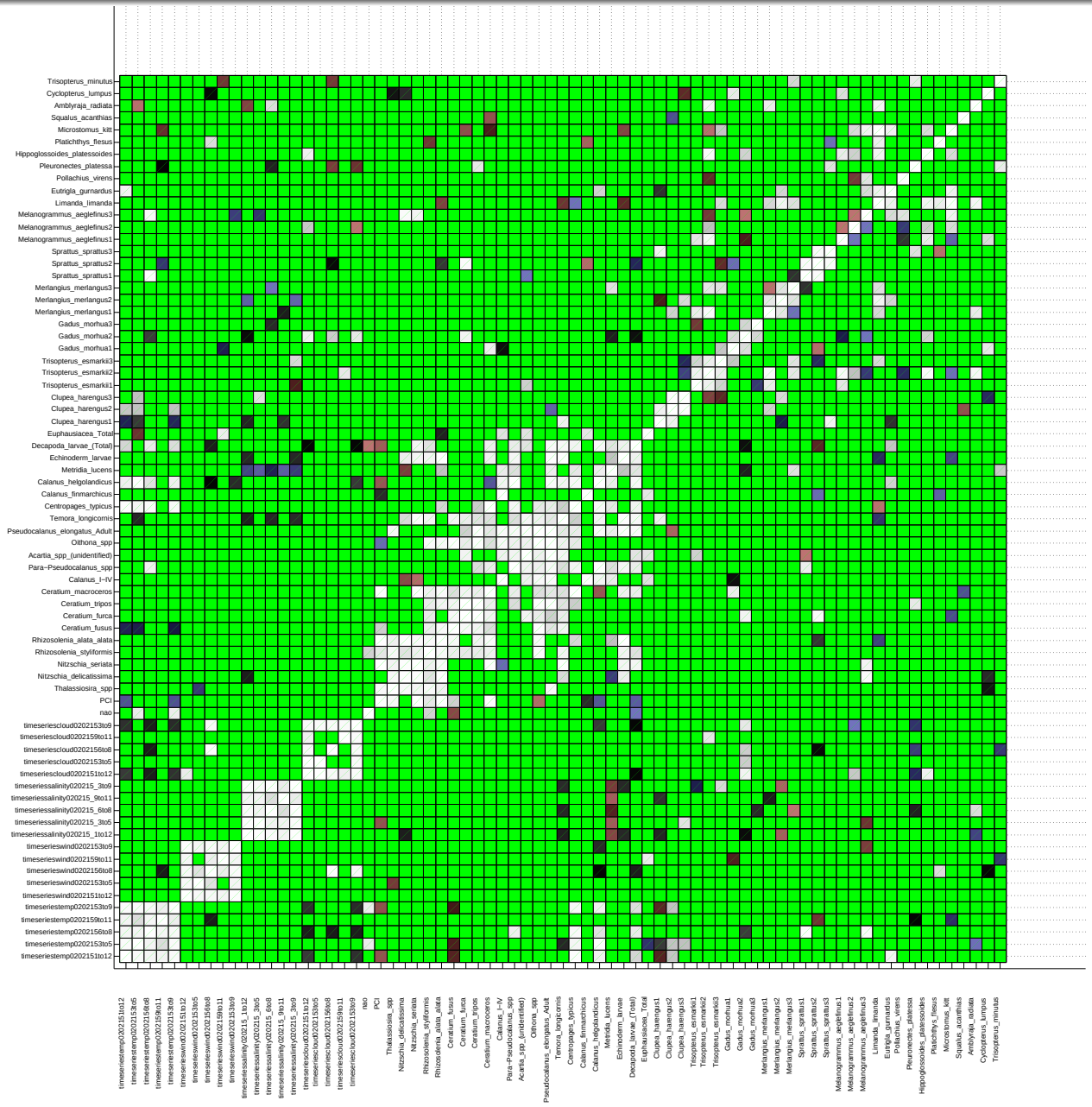
The sum of the FFT's of multivariate data is the FFT of their sum. This equality is preserved if all are subject to the same phase randomisation

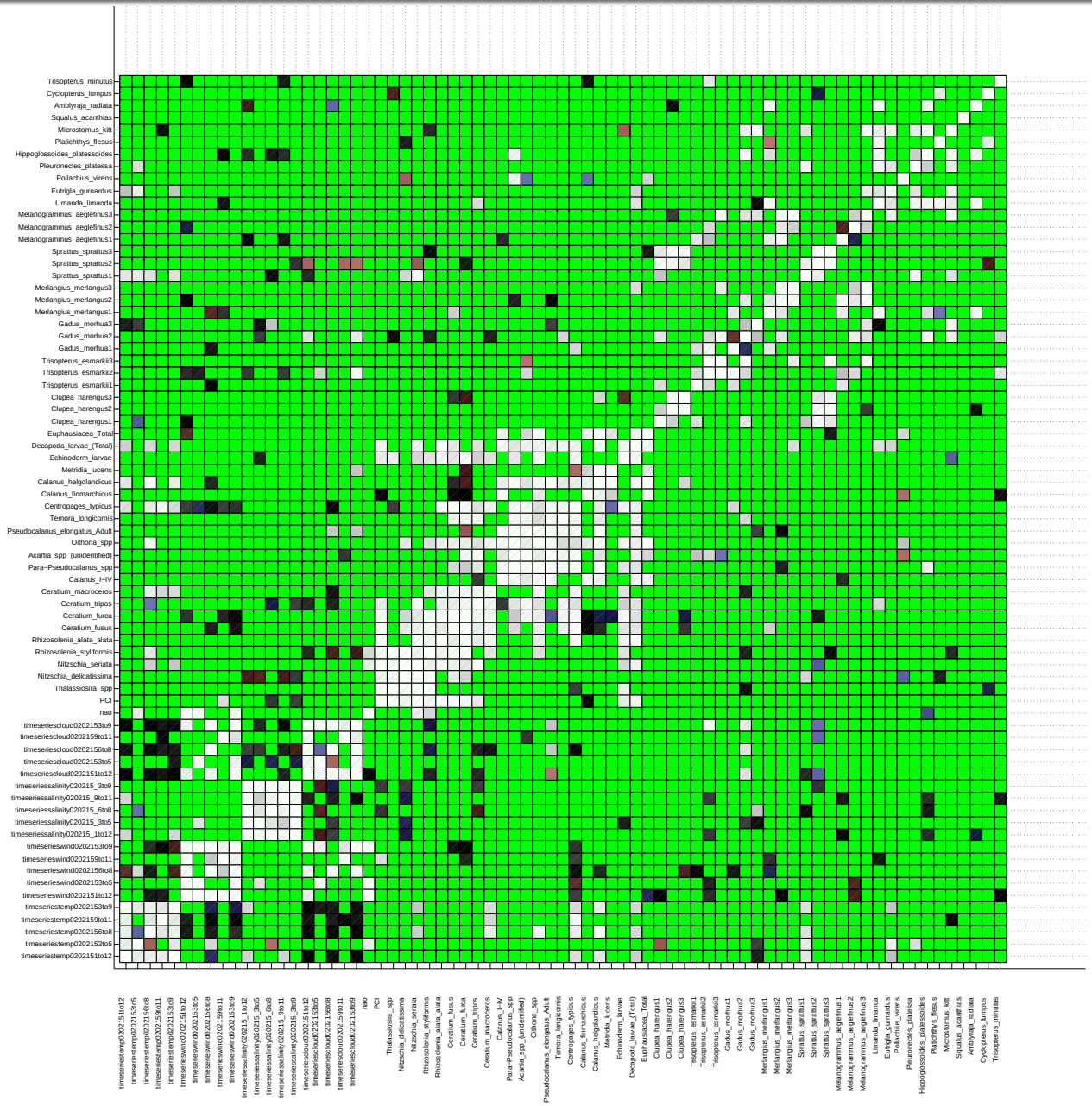
Fourier surrogates without Fourier transforms

The last step of generating a Fourier surrogate timeseries is an IFFT. The first step of caluculating a wavelet transform is an FFT. Thus it is un-necessary to evaluate the surrogate timeseries given by a particular phase randomisation to find its wavelet transform.

Coherence can be determined using Parseval's theorem. Thus it is un-necessary to evaluate the wavelet transform in the time domain.

This makes computation of surrogate coherence values, including spatial surrogate values, highly efficient. **No FFT operations are required after finding the FFT's of the actual timeseries**. Only generating random phase values, vector multiplications and summations are required to find null coherence values associated with Fourier surrogates.


Trisopterus_minutus
Cyclopterus_lumpus
Amblyraja_radiata
Squalus_acanthias
Microstomus_kitt
Platichthys_flesus
Hippoglossoides_platessoides
Pleuronectes_platessa
Pollachius_virens
Eutrigla_gurnardus
Limanda_limanda
Melanogrammus_aeglefinus3
Melanogrammus_aeglefinus2
Melanogrammus_aeglefinus1
Sprattus_sprattus3
Sprattus_sprattus2
Sprattus_sprattus1
Merlangius_merlangus3
Merlangius_merlangus2
Merlangius_merlangus1
Gadus_morhua3
Gadus_morhua2
Gadus_morhua1
Trisopterus_esmarkii3
Trisopterus_esmarkii2
Trisopterus_esmarkii1
Clupea_harengus3
Clupea_harengus2
Clupea_harengus1
Euphausiacea_Total
Decapoda_larvae_(Total)
Echinoderm_larvae
Metridia_lucens
Calanus_helgolandicus
Calanus_finmarchicus
Centropages_typicus
Temora_longicornis
Pseudocalanus_elongatus_Adult
Oithona_spp
Acartia_spp_(unidentified)
Para−Pseudocalanus_spp
Calanus_I−IV
Ceratium_macroceros
Ceratium_tripos
Ceratium_furca
Ceratium_fusus
Rhizosolenia_alata_alata
Rhizosolenia_styliformis
Nitzschia_seriata
Nitzschia_delicatissima
Thalassiosira_spp
PCI
nao
timeseriescloud0202153to9
timeseriescloud0202159to11
timeseriescloud0202156to8
timeseriescloud0202153to5
timeseriescloud0202151to12
timeseriessalinity020215_3to9
timeseriessalinity020215_9to11
timeseriessalinity020215_6to8
timeseriessalinity020215_3to5
timeseriessalinity020215_1to12
timeserieswind0202153to9
timeserieswind0202159to11
timeserieswind0202156to8
timeserieswind0202153to5
timeserieswind0202151to12
timeseriestemp0202153to9
timeseriestemp0202159to11
timeseriestemp0202156to8
timeseriestemp0202153to5
timeseriestemp0202151to12
nao
PCI
Thalassiosira_spp
Nitzschia_delicatissima
Nitzschia_seriata
Rhizosolenia_styliformis
Rhizosolenia_alata_alata
Ceratium_fusus
Ceratium_furca
Ceratium_tripos
Ceratium_macroceros
Calanus_I−IV
Para−Pseudocalanus_spp
Acartia_spp_(unidentified)
Oithona_spp
Pseudocalanus_elongatus_Adult
Temora_longicornis
Centropages_typicus
Calanus_finmarchicus
Calanus_helgolandicus
Metridia_lucens
Echinoderm_larvae
Decapoda_larvae_(Total)
Euphausiacea_Total
Clupea_harengus1
Clupea_harengus2
Clupea_harengus3
Trisopterus_esmarkii1
Trisopterus_esmarkii2
Trisopterus_esmarkii3
Gadus_morhua1
Gadus_morhua2
Gadus_morhua3
Merlangius_merlangus1
Merlangius_merlangus2
Merlangius_merlangus3
Sprattus_sprattus1
Sprattus_sprattus2
Sprattus_sprattus3
Melanogrammus_aeglefinus1
Melanogrammus_aeglefinus2
Melanogrammus_aeglefinus3
Limanda_limanda
Eutrigla_gurnardus
Pollachius_virens
Pleuronectes_platessa
Hippoglossoides_platessoides
Platichthys_flesus
Microstomus_kitt
Squalus_acanthias
Amblyraja_radiata
Cyclopterus_lumpus
Trisopterus_minutus


Trisopterus_minutus
Cyclopterus_lumpus
Amblyraja_radiata
Squalus_acanthias
Microstomus_kitt
Platichthys_flesus
Hippoglossoides_platessoides
Pleuronectes_platessa
Pollachius_virens
Eutrigla_gurnardus
Limanda_limanda
Melanogrammus_aeglefinus3
Melanogrammus_aeglefinus2
Melanogrammus_aeglefinus1
Sprattus_sprattus3
Sprattus_sprattus2
Sprattus_sprattus1
Merlangius_merlangus3
Merlangius_merlangus2
Merlangius_merlangus1
Gadus_morhua3
Gadus_morhua2
Gadus_morhua1
Trisopterus_esmarkii3
Trisopterus_esmarkii2
Trisopterus_esmarkii1
Clupea_harengus3
Clupea_harengus2
Clupea_harengus1
Euphausiacea_Total
Decapoda_larvae_(Total)
Echinoderm_larvae
Metridia_lucens
Calanus_helgolandicus
Calanus_finmarchicus
Centropages_typicus
Temora_longicornis
Pseudocalanus_elongatus_Adult
Oithona_spp
Acartia_spp_(unidentified)
Para−Pseudocalanus_spp
Calanus_I−IV
Ceratium_macroceros
Ceratium_tripos
Ceratium_furca
Ceratium_fusus
Rhizosolenia_alata_alata
Rhizosolenia_styliformis
Nitzschia_seriata
Nitzschia_delicatissima
Thalassiosira_spp
PCI
nao
timeseriescloud0202153to9
timeseriescloud0202159to11
timeseriescloud0202156to8
timeseriescloud0202153to5
timeseriescloud0202151to12
timeseriessalinity020215_3to9
timeseriessalinity020215_9to11
timeseriessalinity020215_6to8
timeseriessalinity020215_3to5
timeseriessalinity020215_1to12
timeserieswind0202153to9
timeserieswind0202159to11
timeserieswind0202156to8
timeserieswind0202153to5
timeserieswind0202151to12
timeseriestemp0202153to9
timeseriestemp0202159to11
timeseriestemp0202156to8
timeseriestemp0202153to5
timeseriestemp0202151to12

Acknowledgments

We thank Joel E. Cohen, Robert Desharnais and the BRACCIA team for helpful discussions. LS was supported and DCR was partly supported by United Kingdom Natural Environment Research Council grants NE/H020705/1, NE/I010963/1, and NE/I011889/1. Travel was facilitated by US National Science Foundation grant DMS-1225529.

References

1. Sheppard LW, Vuksanović V, McClintock PVE, Stefanovska A 2011 *Phys Med Biol* **56** 3583–3601

2. Kenwright DA, Bernjak A, Drægni T, Dzeroski M, Entwistle M, Horvat M, Kvandal P, Landsverk S, McClintock PVE, Musizza B, Petrovčič J, Ræder J, Sheppard LW, Smith AF, Stankovski T, Stefanovska A 2015 *Anaesthesia* **70** 1356-1368

3. Sheppard LW, Defriez E, Reid PC, Reuman, DC 2016 *In preparation*

4. Sheppard LW, Stefanovska A, McClintock PVE 2013 *Phys. Rev. E* **85** 046205

5. Prichard D, Theiler J 1994 *Phys. Rev. Lett.* **7** 951-954

References

6. Wavelet phase dynamics of intracranial pressure monitoring, Lawrence William Sheppard (BA, M.Sci), Department of Physics, Lancaster University, Lancaster, UK. PhD Thesis. September 16, 2010

7. Paul S. Addison, The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance, 2002.

8.Theiler, J. and Eubank, S. and Longtin, A. and Galdrikian, B and Farmer, J.D., Testing for nonlinearity in time series: the method of surrogate data, Physica D, 77–94,58, 1992.

9. Changes in large-scale climate alter spatial synchrony of aphid pests. Lawrence W. Sheppard, James R. Bell, Richard Harrington and Daniel C. Reuman Nature Climate Change (2015) doi:10.1038/nclimate2881