Reply to "Editor comment" by Prof. Bettina Schaefli

The reviewers had only minor comments, to which the authors have responded in this public discussion. I look forward to the revised version, which will be a very interesting technical note for the readership of HESS

Response:

Prof. Schaefli, thank you very much for handling our manuscript and giving us a chance to improve this manuscript.

We have replied to the comments by two reviewers as follows. Meanwhile, we made some revisions in the manuscript.

Reply to "Interactive comment on "Technical Note: Multiple wavelet coherence for untangling scale-specific and localized multivariate relationships in geosciences" by W. Hu and B. C. Si " by Referee #1

The manuscript of Multiple wavelet coherence by Hu and Si presented an important topic. In characterizing scale specific variations, wavelet coherence has been used in many field but was restricted to only two variables. Presentation of wavelet coherence produces a step forward on the methodological development aspect. The method will support a lot of different fields including soil science and hydrology. The scientific content is suitable for the journal and the readers of this journal will be interested in this topic. Therefore, my suggestion is for acceptance of the manuscript with some minor corrections such as English, which could be improved. Another thing, authors used the artificial series to compare with other multi-variate analysis. Just wondering, how will you confirm about you claimed superior information of the new method compare to other methods. I mean to say, how will you say that this variations, what is shown by other methods are also showing the right information. The variations showing here could be spurious as identified by different methods.

Response:

Thank you for the positive comments.

In terms of language, we have tried our best to correct it, and we have asked an English editing company double check the language.

We are not very sure we understood your second comment, but we will try to explain a bit here. The two existing methods (i.e., multiple spectral coherence and multivariate empirical mode decomposition) are widely used for spatial or temporal series analysis in different disciplines. Actually we have known that these two methods cannot deal with localized relationships between variables. Therefore, the advantages of the new method over these two methods is demonstrated mainly in terms of relationships between response and predictor variables at various scales of the response variable. The reason for using the artificial data is that the major features (e.g., scale) are known. Then, the superiority of the new method over these two methods can be assessed by whether the known major features of the artificial data are demonstrated by these methods. Our results clearly show that localized multivariate relationships are not available by the two existing methods and both methods are likely to underestimate the degree of multivariate relationships for non-stationary processes. Because the cosine-like artificial datasets mimic many time series and spatial series in geosciences. Therefore, we conclude that the new method is superior.

All above mentioned information can be found in the revised copy. Please refer to them at Lines 82-84, 154-160, 183-187, and 381-384.

Reply to "Interactive comment on "Technical Note: Multiple wavelet coherence for untangling scale-specific and localized multivariate relationships in geosciences" by W. Hu and B. C. Si " by Referee #2

General Comments

The multiple wavelet coherence methodology presented in the manuscript by Hu and Si represents an important contribution to wavelet analysis. In particular, Hu and Si build upon the previous work of Ng and Chan (2012) to extend multiple wavelet coherence to case of more than two predictor variables. The authors further demonstrate that the new multiple wavelet coherence methodology is better suited for situations where the predictor variables are cross-correlated. The problems with the traditional formulation are clearly stated and consistent with the objective of the paper proposed in the introduction section. Theoretical examples were also presented to highlight the advantages of the new methodology relative to existing ones. I their recommend that the manuscript be accepted after the substantial correction of grammatical errors and the consideration of more specific comments presented below.

Response:

Thank you for the positive comments.

Specific comments

The conclusion section simply summarizes the results of the paper. The authors could consider expanding the conclusion section into a discussion section to comment on limitations of the method. After all, wavelet analysis, while useful, is not a scientific panacea. More specifically, the inclusion of more predictor variables may result in the statistical significance threshold at a particular wavelet scale and time to approach unity, which would impose a limit on how much statistical information can be gained. This phenomenon occurs with the traditional multiple wavelet coherence formulation, where the threshold for 5% significance, for example, is higher than that for bivariate wavelet coherence at a given wavelet scale.

Response:

We agree with you that one of the limitation is that the critical values increase with the number of predictor variables. This is also why the percentage area of significant coherence (PASC) for three predictor variables (z2, z4, and noised z4) are even lower than for only two predictor variables (z2 and z4) when the third predictor variable (noised z4) is not statistically significant to explain the variation of the response variable. Please see Lines 260-261.

We put this limitation in the conclusion part as " Theoretically, any number of predictor variables can be included in the multiple wavelet analysis. However, the statistical significance threshold usually increases with the number of predictor variables (Grinsted et al., 2004; Ng and Chan, 2012a).In addition, the inclusion of too many predictor variables may result in the statistical significance threshold at particular wavelet scales (e.g., the lowest and largest scales) to approach unity. This would restrict the availability of statistical information." (Lines389-395).

The author may also consider discussing at least briefly the problem of simultaneously testing multiple statistical hypothesis, as discussed in Maraun and Kurths (2004), Maraun et al. (2007), Schulte et al. (2015), and Schulte (2016). Multiple-testing problem is a major problem in wavelet analysis and therefore merits consideration in a discussion section. Presenting clearly the methodological limitations will better guide the likely interdisciplinary readership in making decisions regarding what analysis tools to implement.

Response:

The multiple-testing problem has been briefly discussed in the conclusion part. "Furthermore, similar to bivariate wavelet analysis, the new method also suffers from the multiple-testing problem (Maraun and Kurths, 2004; Maraun et al., 2007; Schulte et al., 2015; Schulte, 2016). Therefore, a more robust statistical significance testing method may be beneficial to the new method." (Lines395-399).

Throughout the manuscript, the authors mention how geoscience data are often nonstationary. Perhaps the term is used too loosely in some instances and is sometimes inconsistent with the strict time series analysis definition. Even white and red-noise processes contain time and scale-localized features in wavelet space, even though theirrespectivestatisticsarestationaryatallorders. Time-andscale-localizedfeatures are evident in the wavelet power spectrum of say, the North Atlantic Oscillation (NAO), even though the statistics of the NAO are consistent with a first-order Markov process (Feldstein,2000). Therefore,insomeinstances,Irecommendchangingtheword“nonstationary” to “transient” or “transitory”.

Response:

We agree. In the introduction, we made this more clear as " More often than not, geoscience data are transient, consisting of a variety of frequency regimes that may be localized in space or time (Torrence and Compo, 1998; Si and Zeleke, 2005; Graf et al., 2014). The transient characteristics exist widely in non-stationary processes, but also sometimes occur in stationary processes (Feldstein, 2000)." (Lines35-39).

At many instances, we changed the "non-stationary" to "transient" when suitable, such as Line 42, 58, 65 in the attached revision.

Some Technical Corrections

Page 2 Line 3536. Change “geoscience data is” to “geoscience data are”.

Response:

Yes, done at L36.

Page 2 Line 39. Is it better to say bivariate wavelet coherency rather than “simple wavelet coherency”

Response:

Yes, we changed all throughout the paper.

Page 5, Line 97. Add comma before “respectively”.

Response:

Yes, we did throughout the paper.

Page9,Line169-171. The sentence can be slightly simplified by changing“ white noise with a mean of 0” to “zero-mean white noise”. Perhaps it is redundant to write that the white noise processes were generated. Authors could consider just saying that white noise was added to the predictor variables.

Response:

We agree. Now, it changed to" zero-mean white noises with standard deviations of 0.3, 1, and 4 are added to the predictor variables of y2 (or z2) and y4 (or z4)".

Page 9, Lines 171-173. The sentence “The resulting noised series are termed weakly, moderately, 172 and highly noised series respectively, and have a correlation coefficient of 0.9, 0.5, 173 and 0.1 respectively, with their original predictor variable” needs to be rewritten and simplified. Consider breaking the sentence into two separate sentences.

Response:

We changed it to two sentences. Now, it is like "The resulting noised series have correlation coefficients of 0.9, 0.5, and 0.1, respectively, with their original predictor variable. Therefore, we will refer to them as weakly, moderately, and highly noised series, respectively." (Lines 175-177).

The authors should carefully check for grammatical errors and make similar changes throughout the manuscript.

Response:

Yes, done.

We have asked an English editing company double check the language.

References

Feldstein SB (2000) The timescale, power spectra, and climate noise properties of teleconnection patterns. J Clim 13:4430–4440. doi:10.1175/15200442(2000)0132.0.CO;2 Maraun, D. and Kurths, J.: Cross wavelet analysis: significance testingand pitfalls, Nonlin. Processes Geophys., 11, 505–514, 2004. Maraun, D., Kurths, J., and Holschneider, M.: Nonstationary Gaussian processes in wavelet domain: synthesis,estimation, and significance testing, Phys. Rev. E, 75,doi:10.1103/PhysRevE.75.016707, 2007. Ng, E. K. W. and Chan, J. C. L.: Geophysical applications of partial wavelet coherence and multiple wavelet coherence, J. Atmos. Ocean. Tech., 29, 1845–1853, doi: 10.1175/JTECH-D-12-00056.1, 2012. Schulte, J. A., Duffy, C., and Najjar, R. G.: Geometric and topological approaches to significance testing in wavelet analysis, Nonlin.Processes Geophys., 22, 139–156, doi:10.5194/npg-22-139-2015, 2015. Schulte, J. A.: Cumulative areawise testing in wavelet analysis and its application to geophysical time series, Nonlin. Processes Geophys., 23, 45-57, doi:10.5194/npg-2345-2016, 2016.

Response:

Appreciate for the good references. We cited them when we made relevant discussion.

Technical Note: Multiple wavelet coherence for untangling scale-specific and localized multivariate relationships in geosciences

Wei Hu[2,3] and Bing Cheng Si[1,3]

[1]College of Hydraulic and Architectural Engineering, Northwest A&F University, Yangling 712100, China

[2]New Zealand Institute for Plant & Food Research Limited, Private Bag 4704, ~~8140~~ Christchurch 8140, New Zealand

[3]University of Saskatchewan, Department of Soil Science, Saskatoon, SK S7N 5A8, Canada

Correspondence to: Wei Hu (wei.hu@plantandfood.co.nz) and Bing Cheng Si (bing.si@usask.ca)

Abstract

The scale-specific and localized bivariate relationships in geosciences can be revealed using bivariate~~simple~~ wavelet coherence. The objective of this study ~~is~~ was to develop a multiple wavelet coherence method for examining scale-specific and localized multivariate relationships. Stationary and non-stationary artificial datasets, generated with the response variable as the summation of five predictor variables (cosine waves) with different scales, were used to test the new method. Comparisons were also conducted using existing multivariate methods, including multiple spectral coherence and multivariate empirical mode decomposition (MEMD). Results show that multiple spectral coherence is unable to identify localized multivariate relationships, and underestimates the scale-specific multivariate relationships for non-stationary processes. The MEMD method was able to separate all variables into

components at the same set of scales, revealing scale-specific relationships when combined with multiple correlation coefficients, but has the same weakness as multiple spectral coherence. However, multiple wavelet coherences are able to identify scale-specific and localized multivariate relationships, as they are close to 1 at multiple scales and locations corresponding to those of predictor variables. Therefore, multiple wavelet coherence outperforms other common multivariate methods. Multiple wavelet coherence was applied to a real dataset and revealed the optimal combination of factors for explaining temporal variation of free water evaporation at Changwu site in China at multiple scale-location domains. Matlab codes for multiple wavelet coherence we~~a~~re developed and are provided in the supplement.

1. Introduction

Geoscience data such as topography, climate, and ocean waves usually present cyclic patterns, with high-frequency (small-scale) processes being superimposed on low-frequency (large-scale) processes (Si, 2008). More often than not, geoscience data ~~is~~ are ~~non-stationary~~transient, consisting of a variety of frequency regimes that may be localized in space or time (Torrence and Compo, 1998; Si and Zeleke, 2005; Graf et al., 2014). The transient characteristics exist~~s~~ widely in non-stationary processes, but also sometimes occur in stationary processes (Feldstein, 2000). The wavelet method is a common tool for detecting multi-scale and localized features of ~~non-stationary~~transient processes in geosciences. ~~Simple~~ Bivariate wavelet coherency has been widely used for untangling scale-specific and localized relationships for ~~non-stationary~~transient processes in areas including geophysics (Lakshmi et al., 2004; Müller et al., 2008), hydrology (Labat et al., 2005; Das and Mohanty, 2008; Tang and Piechota, 2009; Carey et al., 2013; Graf et al., 2014), soil science (Si and Zeleke, 2005; Biswas and Si, 2011), meteorology (Torrence and Compo, 1998), and ecology (Polansky et al., 2010). This method, however, is limited to two variables. Processes in geosciences are usually complex and may be affected by more than two environmental factors. A method is needed for analyzing multivariate (>2 variables) and localized relationships at multiple scales.

Several methods have been used for characterizing multivariate relationships. For example, multiple spectral coherence (MSC) has been used to explore the scale-specific relationships between soil saturated hydraulic conductivity (K_s) and multiple soil physical properties (Koopmans, 1974; Si, 2008), but requires a stationary data series, which is rare in geosciences. Multivariate empirical mode decomposition (MEMD), a data-driven method, ~~—~~decomposes each variable into different components (intrinsic mode functions (IMFs)) with each IMF corresponding to a "common scale" inherent in multiple variables (Rehman and Mandic, 2010). The MEMD method is meritorious due to its ability to deal with both ~~non-stationary~~transient and nonlinear systems. The combination of the squared multiple correlation coefficient and MEMD (MCC_{memd}) has been used to explore the multivariate control of soil water content ~~or~~ and K_s~~saturated hydraulic conductivity~~ at multiple scales (Hu and Si, 2013; She et al., 2013, 2015; Hu et al., 2014). However, the sum of variances from different components ~~typically does~~ did not typically equal the total variance of the original series, which may ~~result in~~produce misleading MCC$_{memd}$ results. In addition, ~~in geosciences,~~ multivariate relationships in geosciences are most likely to change with time or space due to the ~~non-stationarity~~transient nature of the processes involved. However, localized multivariate relationships are not available using any of the existing multivariate methods. Therefore, extending the wavelet coherence from two variables to multiple variables ~~it~~ is required ~~to extend the wavelet coherence from two variables to multiple variables~~.

An attempt to extend wavelet coherence from two to three variables has been made by Mihanović et al. (2009). Their method was also applied later in the marine sciences (Ng and Chan, 2012a, b). Limitations arise when using the trivariate~~three variable~~ wavelet coherence: first, only two predictor variables are considered; second, the two predictor variables must be orthogonal. Otherwise, extremely high or low (spurious) coherence (>>1 or <0) may be produced. This spuriousness is inconsistent with the definition of coherence, ~~and~~ which may limit the application of this method in geosciences~~,~~ where environmental variables are usually cross-correlated. Therefore, a robust method for calculating MWC, which produces coherence in the closed interval of [0, 1], is needed.

The objective of this paper is to develop an MWC that applies to cases where there are multiple environmental variables, of which may be cross-correlated. This method is first tested with artificial datasets to demonstrate its advantages over existing multivariate methods. The superiority of the new method over others can be assessed by determining whether the known major features of the artificial data are demonstrated by these methods. ~~It~~ The new method is then applied to a temporal series of evaporation (*E*) from free water surface and meteorological factors at Changwu site in Shaanxi, China.

2. Theory

Bivariate~~Simple~~ wavelet coherence can be understood as the traditional correlation coefficient localized in the scale-location domain (Grinsted et al., 2004). Just as correlation coefficients can be ~~extensions~~ extended from two variables to multiple (>2) variables, wavelet coherence between two variables may also be extended to multiple variables. Similar to bivariate~~simple~~ wavelet coherence, MWC is based on a series of auto- and cross-wavelet power spectra, at different scales and spatial (or temporal) locations, for the response variable and all predictor variables.

Following Koopman (1974), a matrix representation of the smoothed auto- and cross-wavelet power spectra for multiple predictor variables X ($X=\{X1, X2, \ldots, Xq\}$) can be written as

$$\overleftrightarrow{W}^{X,X}(s,\tau)=\begin{bmatrix} \overleftrightarrow{W}^{X1,X1}(s,\tau) & \overleftrightarrow{W}^{X1,X2}(s,\tau) & \cdots & \overleftrightarrow{W}^{X1,Xq}(s,\tau) \\ \overleftrightarrow{W}^{X2,X1}(s,\tau) & \overleftrightarrow{W}^{X2,X2}(s,\tau) & \cdots & \overleftrightarrow{W}^{X2,Xq}(s,\tau) \\ \vdots & \vdots & & \vdots \\ \overleftrightarrow{W}^{Xq,X1}(s,\tau) & \overleftrightarrow{W}^{Xq,X2}(s,\tau) & \cdots & \overleftrightarrow{W}^{Xq,Xq}(s,\tau) \end{bmatrix}, \quad (1)$$

where ~~—~~ $\overleftrightarrow{W}^{Xi,Xj}(s,\tau)$ is the smoothed auto-wavelet power spectra (when *i*=*j*) or cross-wavelet power spectra (when *i*≠*j*) at scale *s* and spatial (or temporal) location

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Font color: Auto

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Font color: Auto

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Font color: Auto

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Font color: Auto

τ, respectively. For the detailed calculation of smoothed auto- and cross-wavelet power spectra, see Supplement, Sect. S1.

The matrix of smoothed cross wavelet power spectra between response variable *Y* and predictor variables *Xi* can be defined as

$$\overleftrightarrow{W}^{Y,X}(s,\tau)=\left[\overleftrightarrow{W}^{Y,X1}(s,\tau)\ \ \overleftrightarrow{W}^{Y,X2}(s,\tau)\ \cdots\ \overleftrightarrow{W}^{Y,Xq}(s,\tau)\right], \qquad (2)$$

Field Code Changed

where $\overleftrightarrow{W}^{Y,Xi}(s,\tau)$ is the smoothed cross-wavelet power spectra between *Y* and *Xi* at scale *s* and spatial (or temporal) location τ.

The smoothed wavelet power spectrum of response variable *Y* is $\overleftrightarrow{W}^{Y,Y}(s,\tau)$.

Following Koopmans (1974), the MWC at scale *s* and location τ, $\rho_m^{2}(s,\tau)$, can be written as

$$\rho_m^{2}(s,\tau)=\frac{\overleftrightarrow{W}^{Y,X}(s,\tau)\overleftrightarrow{W}^{X,X}(s,\tau)^{-1}\overline{\overleftrightarrow{W}^{Y,X}(s,\tau)}}{\overleftrightarrow{W}^{Y,Y}(s,\tau)}. \qquad (3)$$

When only one predictor variable (e.g., *X1*) is included in *X*, Eq. (3) is the equation for bivariate~~simple~~ wavelet coherence, $\rho_b^2(s,\tau)$ ~~$\rho_s^{2}(s,\tau)$~~, ~~between two variables~~which can be expressed as (Torrence and Webster, 1999; Grinsted et al., 2004):

Field Code Changed

$$\rho_b^2(s,\tau)=\frac{\overleftrightarrow{W}^{Y,X1}(s,\tau)\overline{\overleftrightarrow{W}^{Y,X1}(s,\tau)}}{\overleftrightarrow{W}^{X1,X1}(s,\tau)\overleftrightarrow{W}^{Y,Y}(s,\tau)}. \qquad (4)$$

Field Code Changed

Therefore, bivariate~~simple~~ wavelet coherence is consistent with multiple wavelet coherence if only one predictor variable is included. In addition, the wavelet phase between a response variable (*Y*) and a predictor variable (*X1*) is

$$\phi(s,\tau) = \tan^{-1}\left(\mathrm{Im}\left(W^{Y,X1}(s,\tau)\right)/\mathrm{Re}\left(W^{Y,X1}(s,\tau)\right)\right), \tag{5}$$

where Im and Re denote the imaginary and real part of $W^{Y,X1}(s,\tau)$, respectively. Note that the phase information between a response variable Y and multiple predictor variables X cannot be obtained.

Multiple wavelet coherence at the 95% confidence level is calculated using the Monte Carlo method (Grinsted et al., 2004). Surrogate spatial series (i.e., red noise) of all variables are generated with a Monte Carlo simulation based on their first-order autocorrelation coefficient (AR1). The MWC at each scale and location is calculated using the simulated spatial series. This is repeated an adequate number of times (e.g., 1000) (Grinsted et al., 2004). At each scale, MWCs at all locations outside the cones of influence, from all simulations are ranked in ascending order. The value at the 95th percentile represents the 95% confidence level for the MWC at that scale. The Matlab codes and user manual document for calculating MWC and significance level are provided in the Supplement (Sect. S2–S4).

3. Data and analysis

3.1 Artificial data for method test

The method is tested using a stationary and non-stationary artificial dataset, generated following Yan and Gao (2007). The response variable (y for the stationary case and z for the non-stationary case) encompasses five cosine waves (y1 to y5 for the stationary case and z1 to z5 for the non-stationary case), with different dimensionless scales (Fig. 1). For the stationary case, y1=cos(2πx/4), y2=cos(2πx/8),

$y3=\cos(2\pi x/16)$, $y4=\cos(2\pi x/32)$, and $y5=\cos(2\pi x/64)$, where $x=0, 1, 2, \ldots, 255$. There is one regular cycle every 4, 8, 16, 32, and 64 locations, representing dimensionless scales of 4, 8, 16, 32, and 64 for y1, y2, y3, y4, and y5, respectively (Fig. 1a). The regular cycles make each predictor and response series stationary. For the non-stationary case, $z1=\cos(500\pi(x/1000)^{0.5})$, $z2=\cos(250\pi(x/1000)^{0.5})$, $z3=\cos(125\pi(x/1000)^{0.5})$, $z4=\cos(62.5\pi(x/1000)^{0.5})$, and $z5=\cos(31.25\pi(x/1000)^{0.5})$, where $x=0, 1, 2, \ldots, 255$. The equation ~~with~~ containing the square root of the location term results in the gradual change in frequency (scale), with the greatest dimensionless scales of 4, 8, 16, 32, and 64 at the right hand side for z1, z2, z3, z4, and z5, respectively (Fig. 1b). The average scales for these predictor variables are 3, 5, 9, 17, and 32, respectively. The location-varying scales make each predictor and response variable non-stationary.

For both the stationary and non-stationary series, the variance of the response variable is 2.5. The predictor variables, each with a variance of 0.5, are orthogonal to each other, and contribute equally to the total variance of the response variable. The cosine-like artificial datasets mimic many time series such as seismic signals, turbulence, air temperature, precipitation, hydrologic fluxes, and the El Niño-Southern Oscillation. They also mimic geoscientific spatial series such as ocean waves, seafloor bathymetry, land surface topography, and soil water content along a hummocky landscape ~~in geosciences~~. Therefore, they are representative of a geoscience data series and are suitable for testing the new method.

Multiple wavelet coherence between the response variable y (or z) and two (y2 and y4, or z2 and z4) or three (y2, y3, and y4, or z2, z3, and z4) predictor variables were calculated. The advantage of the artificial data is that the known scale- and localized features for all variables, and the known relationships between the response and each predictor variable, are exact. By definition, the coherence is 1 at scales corresponding to ~~that~~ those of the included predictor variables, and 0 at other scales.

To demonstrate the advantages of MWC in dealing with abrupt changes (a type of transient and localized feature), the second half of the original series of y2 (or z2) or y4 (or z4) ~~is~~ are replaced by 0, and MWC between the response variable and new set of predictor variables is calculated. We anticipate that the coherence changes from 1 to 0 at the location where the new predictor variable becomes 0.

Predictor variables may not be as regular as that shown in Fig. 1, and may also be cross-correlated to one another. For these reasons, zero-mean white noises with ~~a mean of 0 and a~~ standard deviations of 0.3, 1, and 4 are ~~generated and~~ added to the predictor variables of y2 (or z2) and y4 (or z4). The resulting noised series have correlation coefficients of 0.9, 0.5, and 0.1, respectively, with their original predictor variable. Therefore, we will refer to them ~~to~~as ~~are termed~~ weakly, moderately, and highly noised series, respectively~~,~~. ~~and have a correlation coefficient of 0.9, 0.5, and 0.1 respectively, with their original predictor variable.~~ Multiple wavelet coherences between the response variable and different predictor variables (original and noised series) are calculated to demonstrate the performance of MWC when noised or correlated predictor variables are involved. Only the non-stationary case will be demonstrated, because the performances of MWC for stationary and non-stationary cases are similar.

The MWC is compared to the MSC (Koopmans, 1974; Si, 2008) and MCC_{memd} (Hu and Si, 2013), which are widely used for spatial or temporal series analysis in different disciplines. The advantages of the new method over these two methods will be demonstrated mainly in terms of relationships between response and predictor variables at various scales of the response variable. The MSC is calculated based on the calculated auto- and cross- power spectra, using an equation similar to Eq. (3). The detailed introduction of this method can be found in Si (2008). For the calculation of MCC_{memd}, a set of response and predictor variables form a multivariate data series for MEMD. The MEMD is a data driven method and has the ability to align “common scales” present within multivariate data. Please refer to Rehman and Mandic (2010) and Hu and Si (2013) for the MEMD analysis, and the website (http://www.commsp.ee.ic.ac.uk/~mandic/research/emd.htm) for the related Matlab codes. The original series of response and predictor variables can be decomposed by the MEMD, into different components (IMFs) with ~~different~~ varying scales ~~by the MEMD~~. For IMFs at the same scale, multiple stepwise regressions are conducted between response and predictor variables, and the multiple correlation coefficients for each scale-specific IMF are calculated.

3.2 Real data for application

Daily evaporation (*E*) from free water surfaces ~~of~~ in an E601 evaporation pan (pan diameter of 61.8 cm), and other meteorological factors (i.e., relative humidity, mean temperature, sun hours, and wind speed) were collected from January 1, 1979 to December 31, 2013, at Changwu site in Shaanxi, China. The Changwu site is a transition area between semi-arid and subhumid climates, where agricultural productivity is mainly limited by water~~ limits agricultural productivity~~. Monthly averages of all variables were used in this study, because we are mainly interested in seasonal and inter-annual variability.

4. Results and discussion

4.1 MWC with orthogonal~~ly~~ predictor variables

For the stationary data, there are two narrow, horizontal bands (red color) representing an MWC value of around 1, at the respective scales of 8 and 32 for all locations (Fig. 2a). These two bands also correspond to the scales of 8 and 32, respectively, for the two predictor variables. When an additional predictor variable with the scale of 16 is introduced, a wide band appears from 6 to 40~~ appears~~, signifying that the MWC equals approximately 1 at all locations, at the scales of 8, 16, and 32. As anticipated, when all five predictor variables with scales ranging from 4 to 64 are included, coherence values of close to 1 are found in the whole scale-location domain (data not shown).

The application of MWC to the non-stationary datasets shows that the scales with significant MWC values gradually increase ~~with the increase in~~as distance increases. This increase in the scales is due to the non-stationarity of the variables (Fig. 2b). For example, when predictor variables of z2 and z4 are included, scales of the two bands corresponding to MWC around 1 increase from 4 to 8 and from 8 to 32, respectively. Furthermore, as expected, for only one predictor variable (stationary and non-stationary), MWC reduces to bivariate~~simple~~ wavelet coherence; there is only one band of coherence around 1, which corresponds to the scale of that predictor variable (data not shown). Note that the significant MWC values for both stationary and non-stationary cases are not exactly 1 at all scales or locations, due to the smoothing effect along both scales and locations. However, the mean MWC values of the significant bands are very high (i.e., 0.94~~-~~–1.00), and the MWC values at the centre of the significant band are 1, which corresponds to the exact scale of a predictor variable.

When the point values in the second half of the data series of a predictor variable are~~is~~ replaced by 0, the MWC values in that half of the data series ~~is~~ are almost 0 at scales corresponding to that predictor variable (Fig. 3). For the stationary case, when the point values in the second half of the data series of predictor variable y2 (or y4) ~~is~~ are replaced by 0, the MWC values ~~is~~ are around 1 at the scale of 8 (or 32) in the first half of the transect, and 0 in the second half (Fig. 3a). Similar results ~~were~~ are also found for the non-stationary case (Fig. 3b). This is expected because the constant series of 0 is not correlated to the response variables at any scale. Much like bivariate~~simple~~ wavelet coherence, the MWC method is able to detect abrupt changes in the data series, and has the advantages of dealing with localized multivariate relationships.

4.2 MWC with noised and correlated predictor variables

When z2 and a noised series derived from z2 are included as predictor variables, there is only one band of coherence close to 1 at scales corresponding to z2, irrespective of the correlation between z2 and a noised series of z2 (Fig. 4a). When z2 and a noised series of z4 are included as predictor variables, the coherence depends on the degree of the noise (Fig. 4b). For weakly noised series, there are two bands of coherence of around 1, corresponding to the scales of z2 and z4, respectively. The percentage area of significant coherence (PASC) is 23%, which equals that of when z2 and z4 are included. With the increasing magnitude~~e~~ of noise, the coherence and corresponding PASC at the scales corresponding to z4 decrease. When z2 and a strongly noised series of z4 are considered, the band of coherence around 1, at scales corresponding to z4, disappears.

The inclusion of a third noised z4 variable substantially increases the area with high coherence (in red) as compared to the case when only z2 and z4 are included (Fig. 4c). This indicates that MWC will increase ~~with the~~as ~~increase in~~ the number of predictor variables increases, with the highest coherence less or equal to 1, irrespective of the number of predictor variables. However, the area of significant coherence may not necessarily increase because of the simultaneously increased statistical significance threshold (Ng and Chan, 2012a). In fact, the PASC values for three predictor variables (19~~-~~–20%) are lower than ~~for only~~ those of the two predictor variables (23%). This indicates that, in this case, two predictor variables are better than three in terms of explaining the variations of the response variable. This ~~is~~ occurs because the variance of the response variable that is explained by the noised variable is already accounted for by other variables. Therefore, only an additional variable that can independently explain a fair amount of variance could contribute significantly to explaining variations of a response variable (Fig. 4b). This may~~can~~ also explain why there is only one band of coherence around 1 at scales corresponding to z2, when z2 and a noised series of z2 are included (Fig. 4a). This information is helpful in choosing predictor variables for developing scale-specific predictions, especially when predictor variables are correlated.

4.3 Comparison with other multivariate methods

4.3.1 MSC

The MSC as a function of scale is shown in Fig. 5a. For the stationary case, when y2 and y4 are included as predictor variables, there are two plateaus centered at the scales of 8 and 28, representing a coherence of 1. As expected, when an additional predictor variable y3 is added, the corresponding scale of 16 also shows coherence of 1. The MSC produces similar scale-specific relationships, as MWC does for a stationary dataset, with exception given to ~~that~~ the centered scale (i.e., 28) with a coherence of 1. Here, the scale with a unity MSC deviates from the expected value (i.e., 32) for predictor variable y4. For the non-stationary case,~~,~~ however, the MSC is much lower than 1 for the predictor variables of z2 and z4; an MSC of 1 is present only at the scale of 8 when an additional predictor variable z3 is added. Obviously, the MSC underestimates the multivariate relationships, and is not suitable ~~to~~ for non-stationary processes (Si, 2008) due to its inability to deal with localized features. The MSC at a specific scale provides the average of multivariate relationships, across all locations. ~~Because~~ Due to the change in scale of a predictor variable ~~changes~~ with location for the non-stationary case, the MSC deviates greatly from 1.

~~The inability of the MSC to deal with localized features is demonstrated further by the decrease of~~ The MSC decreases at scales when the second half of the included predictor variable series are replaced by 0 for both the stationary and non-stationary series (Fig. 5b). For example, when the second half of the y4 series in the stationary case ~~is~~ are replaced by 0, ~~for the stationary case,~~ the MSC at scales of around 32 decreases from 1 to 0.52. Although the MSC, throughout the second half of the series, can detect the decrease of coherence at the scales corresponding to the 0 values ~~throughout the second half of the series~~, the exact locations for the decrease cannot be identified. In fact, the coherence decreases only in the second half of the series, and does not change in the first half of the series. The location for the decrease can be easily identified by the MWC, but not by MSC. This further demonstrates the inability of the MSC to deal with localized features.

4.3.2 MCC_{memd}

Five intrinsic mode functions (IMFs) with non-negligible variance, are obtained for multivariate data series. While the obtained scales for the response variable y are in agreement with the true scales for the stationary case, the obtained scales (i.e., 3, 6, 11, 21, and 43) for the response variable z deviate slightly from the average scales for the non-stationary case. For the response variable, the contribution of IMFs to the total variance generally decreases (20% to 13% for stationary, and 27% to 11% for non-stationary) from IMF1 to IMF5.~~,~~ ~~which~~ This disagrees with the fact that each scale contributes equally (i.e., 20%) to the total variance. In addition, the sum of variances over all IMFs for each variable is less than 100% (ranging from 84% to 93%), indicating that MEMD cannot capture all the variances. For the detailed results of MEMD, see Supplement, Sect. S5.

The MCC_{memd} as a function of scale, is shown in Fig. 6a. For the stationary case, when predictor variables of y2 and y4 are included, the MCC_{memd} values are 0.98 and 0.93, respectively, at scales corresponding to ~~that~~ those of y2 and y4. When a predictor variable of y3 is included, the MCC_{memd} values are 1.00, 1.00, and 0.96, respectively, at scales corresponding to ~~that~~ those of y2, y3, and y4. For the non-stationary, two predictor variable case, the corresponding MCC_{memd} values are 0.80 and 0.85. ~~for the two predictor variable case, and~~For the non-stationary, three predictor variable case, the corresponding MCC_{memd} values are 0.95, 0.99, and 0.91, respectively~~, for the case of three predictor variables~~. Therefore, the MCC_{memd} can be used to determine the scale-specific multivariate relationships. Similar to MSC, however, the MCC_{memd} underestimates the multivariate relationships, especially for the non-stationary case with less predictor variables. On the contrary, the MCC_{memd} ~~can~~ also overestimates the multivariate relationships. For example, when considering only predictor variables corresponding to scales of 8, 16, and 32, ~~are considered,~~ the MCC_{memd} value for the stationary case is 0.47 at the scale of 64.~~,~~ ~~which~~ This deviates much from the expected MCC_{memd} value of 0 (Fig. 6a). The possible underestimation and overestimation by the MCC_{memd} may come from the decomposition errors inherent in the MEMD algorithm (Rehman and Mandic, 2010).

Similar to MSC, the localized multivariate relationships cannot be obtained from MCC_{memd}. This can be better explained by the decrease of MCC_{memd} when half of the series of the predictor variables are replaced by 0 (Fig. 6b). Take ~~For~~ the stationary case for example, the MCC_{memd} values at the scales corresponding to y2 ~~(or~~ and y4~~)~~ decrease from 0.98 to 0.49 ~~and from 0.93 to 0.62 when the second half of the y2 (or y4) series are replaced by 0.~~ , and from 0.93 to 0.62, respectively, when the second half of the y2 and y4 series are replaced by 0.

As explained above, the MWC has advantages in untangling localized multivariate relationships as compared to the common multivariate methods. It is important to reveal the multivariate relationships~~,~~ which vary with time or space, that are associated with different processes. For example, discharge usually ~~happens~~ occurs on knolls, while recharge usually ~~happens~~ occurs in neighboring depressions (Gates et al., 2011). Therefore, the controlling factors of soil water storage may vary with the land element characteristics of a location. ~~For example, l~~Local controls may be more important on knolls, while non-local controls may be more important in depressions (Grayson et al., 1997). In a temporal domain, vegetation transpiration contributes more to the evapotranspiration in the growing seasons, which may result in the changes of environmental factors explaining temporal variations of evapotranspiration in different seasons.

4.4 Application of the MWC

Each meteorological factor was significantly correlated to ~~the~~ *E*, but the dominant factors explaining variations in *E* differed with scale. For example, the relative humidity was the dominating factor at small (2–8 months) and large (>32 months) scales, while temperature was the dominating factor at the medium (8–32 months) scales. Overall, the relative humidity corresponded to the greatest mean MWC (0.62) and PASC value (40%) at multiple scale-location domains. For the detailed relationships between *E* and each factor, see Supplement, Sect. S6.

The MWC analysis shows that the combination of relative humidity and mean temperature produced the greatest mean MWC (0.82) and PASC (49%) among all two-factor cases.~~,~~ This ~~indicat~~suggested~~ing~~ that relative humidity and mean temperature ~~they are~~were the ~~best~~ most appropriate factors ~~to~~ for explaining variations in *E* at multiple scale-location domains (Fig. 7a). However, adding an additional factor such as sun hours, which was the best among all three-factor cases, increased the average coherence (0.91), but slightly decreased the PASC to 48% (Fig. 7b). This indicated that sun hours was not significantly different from red noise in explaining additional variation in *E*. Similar results were found when the wind speed was added. ~~The~~ This occurs because ~~reason for this was being that~~ most areas with significant coherence between *E* and sun hours or wind speed~~,~~ were a subset of areas with significant coherence between *E* and relative humidity or mean temperature (see Supplement, Sect. S3). Therefore, relative humidity and mean temperature were adequate ~~to~~ for explaining the temporal variation of *E* at various scales at this site.

This ~~is~~ was consistent with Li et al. (2012), who indicated that relative humidity and mean temperature ~~are~~ were the two main contributors to the temporal change of potential evapotranspiration on the Chinese Loess Plateau.

5. Conclusions

Multiple wavelet coherence ~~is~~ was developed to determine scale-specific and localized multivariate relationships in geosciences. The new method ~~is~~ was tested and compared with existing multivariate methods, using an artificial dataset. The new method can be used to determine the proportion of the variance of a response variable that is explained by predictor variables, at a specific scale and location (spatially or temporally). As compared with bivariate~~simple~~ wavelet coherence, more variation may be explained at multiple scale-location domains by the MWC. Including more variables is only beneficial if the variables are not strongly cross-correlated, and can independently explain a fair amount of variability in a response variable. Therefore, the best combinations of variables that explain multivariate, spatial or temporal variability at multiple scales can be determined. This is important for optimizing variables ~~for~~ to develop~~ing~~ scale-specific prediction.

The MSC and MCC_{memd} can determine multivariate relationships at multiple scales, but localized multivariate relationships are not available. Furthermore, ~~and~~ both MSC and MCC_{memd} are likely to underestimate the degree of multivariate relationships for non-stationary processes. In addition, the performance of MCC_{memd} relies on the performance of MEMD, which needs further development. Application of the MWC into the real dataset indicates that the combination of relative humidity and mean temperature are the optimal factors that can be used to explain temporal variations of *E* at the Changwu site in China.

Limitations of the new method also exist. Theoretically, any number of predictor variables can be included in the multiple wavelet analysis. However, the statistical significance threshold usually increases with the number of ~~the~~ predictor variables (Grinsted et al., 2004; Ng and Chan, 2012a). ~~, and i~~In addition, the inclusion of too many predictor variables may result in the statistical significance threshold at particular wavelet scales (e.g., the lowest and largest scales) to approach unity. This would restrict the availability of statistical information. ~~In addition~~Furthermore, similar to bivariate wavelet analysis, the new method also suffers from the multiple-testing problem (Maraun and Kurths, 2004; Maraun et al., 2007; Schulte et al., 2015; Schulte, 2016). Therefore, a more robust statistical significance testing method may be beneficial to the new method.

In summary, multiple wavelet coherence has advantages over existing multivariate methods, and provides an effective vehicle for untangling complex spatial or temporal variability for multiple controlling factors at multiple scales and locations. It may also be used as a data-driven tool for modeling and predicting various processes in the area of geosciences, such as precipitation, drought, soil water dynamics, stream flow, and atmospheric circulation.

Acknowledgements

The Matlab codes for calculating multiple wavelet coherence are developed based on the codes provided by A. Grinsted (http://noc.ac.uk/using-science/crosswavelet-wavelet-coherence) and, together with the user manual, are available in the Supplement (Sect. S2-S4). The project was partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Agriculture Development Fund of Saskatchewan. We thank the two anonymous reviewers for their constructive comments.

References

Biswas, A. and Si, B. C.: Identifying scale specific controls of soil water storage in a hummocky landscape using wavelet coherency, Geoderma, 165, 50–59, doi: 10.1016/j.geoderma.2011.07.002, 2011.

Carey, S. K., Tetzlaff, D., Buttle, J., Laudon, H., McDonnell, J., McGuire, K., Seibert, J., Soulsby, C., and Shanley, J.: Use of color maps and wavelet coherence to discern seasonal and interannual climate influences on streamflow variability in northern catchments, Water Resour. Res., 49, 6194–6207, doi: 10.1002/wrcr.20469, 2013.

Das, N.N. and Mohanty, B. P.: Temporal dynamics of PSR-based soil moisture across spatial scales in an agricultural landscape during SMEX02: A wavelet approach, Remote Sens. Environ., 112, 522–534, doi:10.1016/j.rse.2007.05.007, 2008.

Feldstein, S. B.: The timescale, power spectra, and climate noise properties of teleconnection patterns, J. Clim., 13, 4430–4440, doi:10.1175/15200442(2000)0132.0.CO, 2000.

Gates, J. B., Scanlon, B. R., Mu, X. M., and Zhang, L.: Impacts of soil conservation on groundwater recharge in the semi-arid Loess Plateau, China, Hydrogeol. J., 19, 865–875, 2011.

Graf, A., Bogena, H. R., Drüe, C., Hardelauf, H., Pütz, T., Heinemann, G., and Vereecken, H.: Spatiotemporal relations between water budget components and soil water content in a forested tributary catchment, Water Resour. Res., 50, 4837–4857, doi:10.1002/2013WR014516, 2014.

Grayson, R. B., Western, A. W., Chiew, F. H. S., and Blöschl, G.: Preferred states in spatial soil moisture patterns: local and nonlocal controls, Water Resour. Res., 33, 2897–2908, doi: 10.1029/97WR02174, 1997.

Grinsted, A., Moore, J. C., and Jevrejeva, S.: Application of the cross wavelet transform and wavelet coherence to geophysical time series, Nonlinear Proc. Geoph., 11, 561–566, 2004.

Hu, W., Biswas, A., and Si, B. C.: Application of multivariate empirical mode decomposition for revealing scale- and season-specific time stability of soil water storage, Catena, 113, 377–385, doi:10.1016/j.catena.2013.08.024, 2014.

Hu, W. and Si, B. C.: Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition, Geoderma, 193–194,180–188, doi: 10.1016/j.geoderma.2012.10.021, 2013.

Koopmans, L. H.: The spectral analysis of time series, Academic Press, New York, 1974.

Labat, D.: Recent advances in wavelet analyses: Part I. A review of concepts, J. Hydrol., 314, 275–288, doi: 10.1016/j.jhydrol.2005.04.003, 2005.

Lakshmi, V., Piechota, T., Narayan, U., and Tang, C. L.: Soil moisture as an indicator of weather extremes, Geophys. Res. Lett., 31, L11401, doi:10.1029/2004GL019930, 2004.

Li, Z., Zheng, F. L., and Liu, W. Z.: Spatiotemporal characteristics of reference evapotranspiration during 1961–2009 and its projected changes during 2011–2099 on the Loess Plateau of China, Agric. For. Meteorol., 154-155, 147–155, doi:10.1016/j.agrformet.2011.10.019, 2012.

Maraun, D. and Kurths, J.: Cross wavelet analysis: significance testing and pitfalls, Nonlin. Processes Geophys., 11, 505–514, doi: 10.5194/npg-11-505-2004, 2004.

Maraun, D., Kurths, J., and Holschneider, M.: Nonstationary Gaussian processes in wavelet domain: synthesis, estimation, and significance testing, Phys. Rev. E ., 75, 016707, doi:10.1103/PhysRevE.75.016707, 2007.

Mihanović, H., Qrlić, M., and Pasrić, Z.: Diurnal thermocline oscillations driven by tidal flow around an island in the Middle Adriatic, J. Marine Syst., 78, S157–S168, doi: 10.1016/j.jmarsys.2009.01.021, 2009.

Müller, W. A., Frankignoul, C., and Chouaib, N.: Observed decadal tropical Pacific–North Atlantic teleconnections, Geophys. Res. Lett., 35, L24810, doi:10.1029/2008GL035901, 2008.

Ng, E. K. W. and Chan, J. C. L.: Geophysical applications of partial wavelet coherence and multiple wavelet coherence, J. Atmos. Ocean. Tech., 29, 1845–1853, doi: 10.1175/JTECH-D-12-00056.1, 2012a.

Ng, E. K. W. and Chan, J. C. L.: Interannual variations of tropical cyclone activity over the north Indian Ocean, Int. J. Climatol., 32, 819–830, doi: 10.1002/joc.2304, 2012b.

Polansky, L., Wittemyer, G., Cross, P. C., Tambling, C. J., and Getz, W. M.: From moonlight to movement and synchronized randomness: Fourier and wavelet analyses of animal location time series data, Ecology, 91, 1506–1518, doi: 10.1890/08-2159.1, 2010.

Rehman, N. and Mandic, D. P.: Multivariate empirical mode decomposition, Proc. R. Soc. A., 466, 1291–1302, doi:10.1098/rspa.2009.0502, 2010.

Schulte, J. A.: Cumulative areawise testing in wavelet analysis and its application to geophysical time series, Nonlin. Processes Geophys., 23, 45–57, doi:10.5194/npg-2345-2016, 2016.

Schulte, J. A., Duffy, C., and Najjar, R. G.: Geometric and topological approaches to significance testing in wavelet analysis, Nonlin. Processes Geophys., 22, 139–156, doi:10.5194/npg-22-139-2015, 2015.

She, D. L., Liu, D. D., Peng, S. Z., and Shao ,M. A.: Multiscale influences of soil properties on soil water content distribution in a watershed on the Chinese Loess Plateau, Soil Sci., 178, 530–539, doi: 10.1097/SS.0000000000000021, 2013.

She, D. L., Zheng, J. X., Shao, M. A., Timm, L. C., and Xia, Y. Q.: Multivariate empirical mode decomposition derived multi-scale spatial relationships between saturated hydraulic conductivity and basic soil properties, Clean-Soil Air Water, doi:

10.1002/clen.201400143, 2015.

Si, B. C.: Spatial scaling analyses of soil physical properties: A review of spectral and wavelet methods, Vadose Zone J., 7, 547–562, doi: 10.2136/vzj2007.0040, 2008.

Si, B. C. and Zeleke, T. B.: Wavelet coherency analysis to relate saturated hydraulic properties to soil physical properties, Water Resour. Res., 41, W11424, doi:10.1029/2005WR004118, 2005.

Tang, C. L. and Piechota, T. C.: Spatial and temporal soil moisture and drought variability in the Upper Colorado River Basin, J. Hydrol., 379, 122–135, doi: 10.1016/j.jhydrol.2009.09.052, 2009.

Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc., 79, 61–78, doi: 10.1175/1520-0477(1998)079<0061:apgtwa>2.0.co;2, 1998.

Torrence, C. and Webster, P. J.: Interdecadal changes in the ENSO-monsoon system, J. Clim., 12, 2679–2690, doi: 10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2, 1999.

Yan, R. and Gao, R. X.: A tour of the Hilbert–Huang transform: an empirical tool for signal analysis, IEEE Instrum. Meas. Mag., 10, 40–45, doi: 10.1109/MIM.2007.4343566, 2007.

Figure captions

Figure 1. (a) Stationary and (b) non-stationary series of response variables (y for stationary and z for non-stationary case) encompassing five cosine waves (y1 to y5 for stationary and z1 to z5 for non-stationary case) with different dimensionless scales.

Figure 2. Multiple wavelet coherence (a) between response variable y and predictor variables y2 and y4; (b) between response y and predictors y2, y3, and y4; (c) between response z and predictors z2 and z4; and (d) between response z and predictors z2, z3, and z4. The artificial data series (y) encompasses five cosine waves (y1, y2, y3, y4, and y5) with different scales for the stationary case, and the artificial data series (z) encompasses five cosine waves (z1, z2, z3, z4, and z5) with different scales for the non-stationary case. The predictor variables, connected by a hyphen, are shown in the top right corner of each subplot. Thin solid lines demarcate the cones of influence, and thick solid lines show the 95% confidence levels.

Figure 3. Multiple wavelet coherence (a) between y and y2h0 and y4; (b) between y and y2 and y4h0; (c) between z and z2h0 and z4; and (d) between z and z2 and z4h0. The artificial data series (y) encompasses five cosine waves (y1, y2, y3, y4, and y5) with different scales for the stationary case and the artificial data series (z) encompasses five cosine waves (z1, z2, z3, z4, and z5) with different scales for the non-stationary case. The variables y2h0 (or z2h0) and y4h0 (or z4h0) refer to the new series of y2 (or z2) and y4 (or z4), in which the second half ~~is~~ are replaced by 0. The

predictor variables, connected by a hyphen, are shown in the top right corner of each subplot. Thin solid lines demarcate the cones of influence and thick solid lines show the 95% confidence levels.

Figure 4. Multiple wavelet coherence of an artificial data series (z) encompassing five cosine waves (z1, z2, z3, z4, and z5) with different scales and (a) z2 and noised z2, (b) z2 and noised z4, and (c) z2, z4, and noised z4 for the non-stationary case. The predictor variables are connected by a hyphen and shown in the top right corner of each subplot. z2wn (z4wn), z2mn (z4mn), and z2sn (z4sn) indicate weakly, moderately, and strongly noised z2 (z4) series, respectively. Weakly, moderately, and strongly noised series are correlated with original series, having correlation coefficients of 0.9, 0.5, and 0.1, respectively. Thin solid lines demarcate the cones of influence and thick solid lines show the 95% confidence levels.

Figure 5. Multiple spectral coherence (MSC) of an artificial data series (y or z) encompassing five cosine waves (y1 to y5; or z1 to z5) with different scales and (a) two (y2 and y4; or z2 and z4) or three (y2, y3, and y4; or z2, z3, and z4) data series, and (b) two (y2 and y4; or z2 and z4) data series when the second half of one data series ~~is~~ are replaced by 0. The variables y2h0 (or z2h0) and y4h0 (or z4h0) refer to the new series of y2 (or z2) and y4 (or z4) in which the second half ~~is~~ are replaced by 0. **Figure 6.** Multiple correlation coefficient between multivariate empirical mode decomposition (MCC_{memd}) of an artificial series (y or z) and (a) two (y2 and y4; or z2 and z4) or three (y2, y3, and y4; or z2, z3, and z4) data series, and (b) two (y2 and y4; or z2 and z4) data series when the second half of one data series ~~is~~ are replaced by 0.

The variables y2h0 (or z2h0) and y4h0 (or z4h0) refer to the new series of y2 (or z2) and y4 (or z4) in which the second half ~~is~~ are replaced by 0.

Figure 7. Multiple wavelet coherence between evaporation (*E*) from water surfaces and meteorological factors ((a) relative humidity and mean temperature and (b) relative humidity, mean temperature, and sun hours) at Changwu site in Shaanxi, China. Thin solid lines demarcate the cones of influence, and thick solid lines show the 95% confidence level.

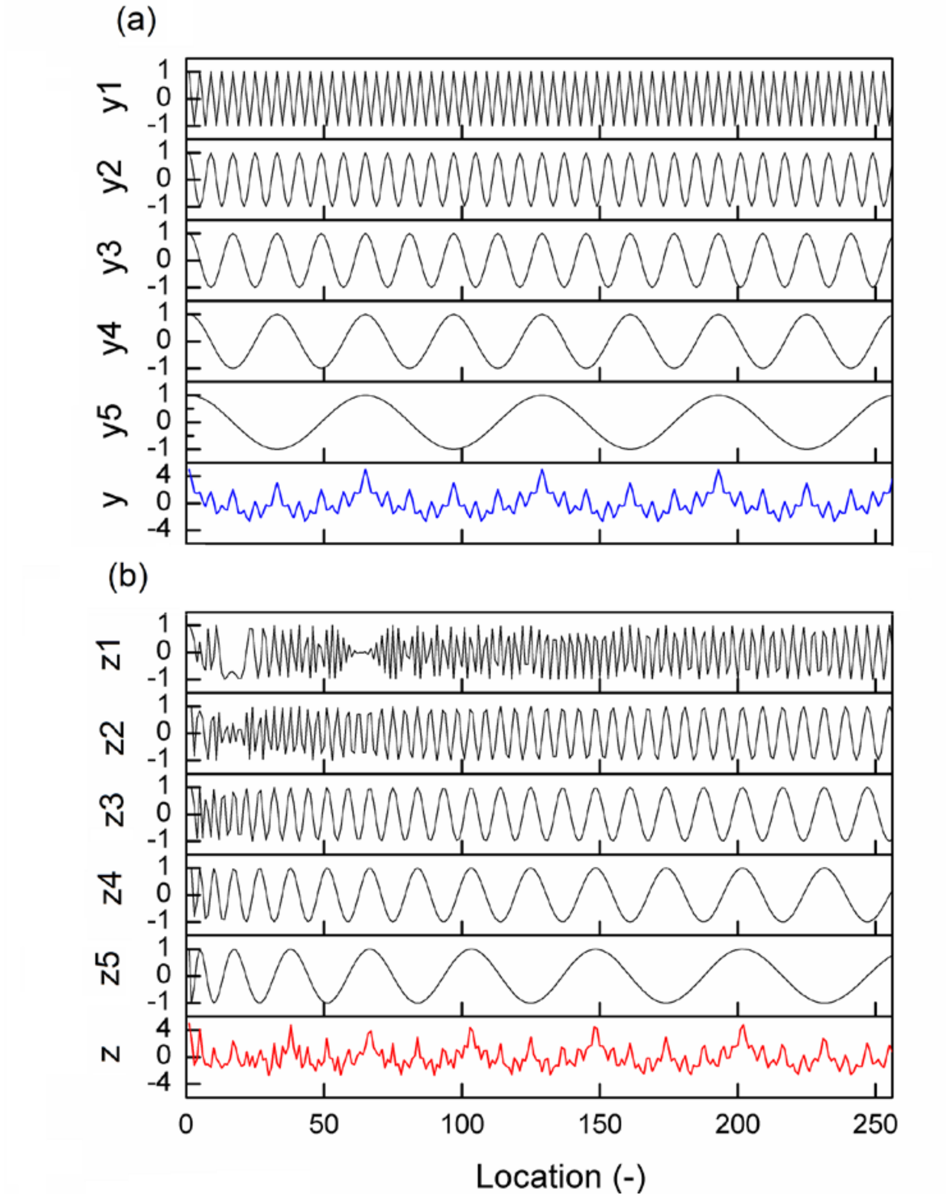

Figure 1. (a) Stationary and (b) non-stationary series of response variables (y for stationary and z for non-stationary case) encompassing five cosine waves (y1 to y5 for stationary and z1 to z5 for non-stationary case) with different dimensionless scales.

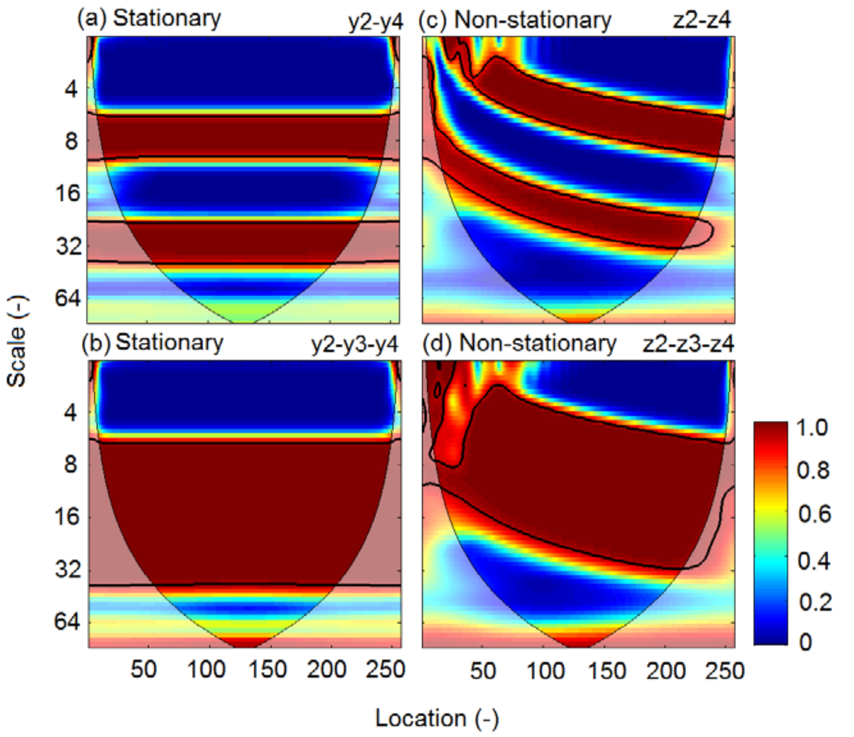

Figure 2. Multiple wavelet coherence (a) between response variable y and predictor variables y2 and y4; (b) between response y and predictors y2, y3, and y4; (c) between response z and predictors z2 and z4; and (d) between response z and predictors z2, z3, and z4. The artificial data series (y) encompasses five cosine waves (y1, y2, y3, y4, and y5) with different scales for the stationary case, and the artificial data series (z) encompasses five cosine waves (z1, z2, z3, z4, and z5) with different scales for the non-stationary case. The predictor variables, connected by a hyphen, are shown in the top right corner of each subplot. Thin solid lines demarcate the cones of influence, and thick solid lines show the 95% confidence levels.

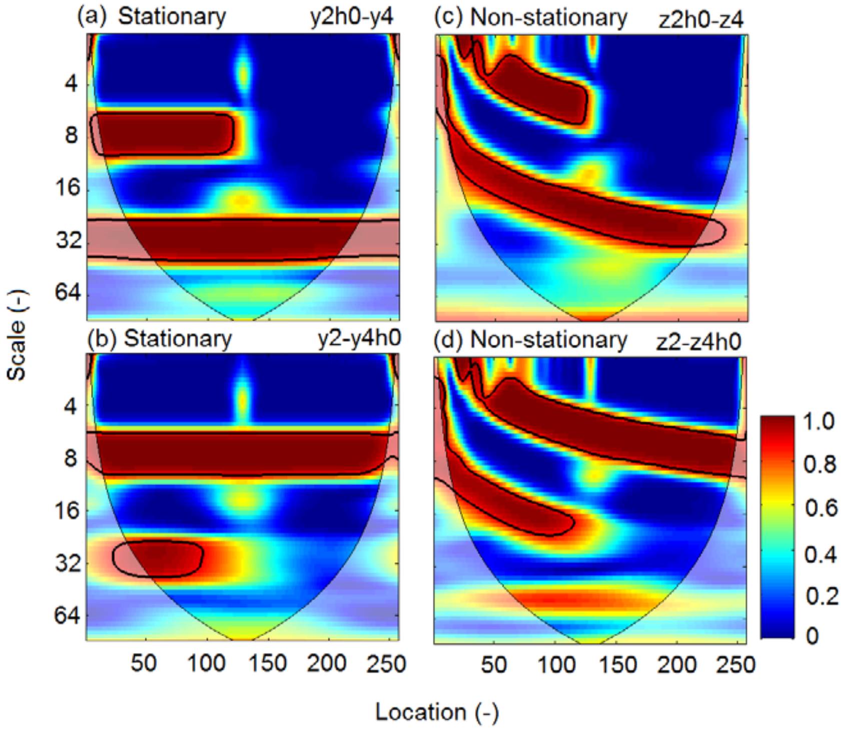

Figure 3. Multiple wavelet coherence (a) between y and y2h0 and y4; (b) between y and y2 and y4h0; (c) between z and z2h0 and z4; and (d) between z and z2 and z4h0. The artificial data series (y) encompasses five cosine waves (y1, y2, y3, y4, and y5) with different scales for the stationary case and the artificial data series (z) encompasses five cosine waves (z1, z2, z3, z4, and z5) with different scales for the non-stationary case. The variables y2h0 (or z2h0) and y4h0 (or z4h0) refer to the new series of y2 (or z2) and y4 (or z4), in which the second half ~~is~~ are replaced by 0. The predictor variables, connected by a hyphen, are shown in the top right corner of each subplot. Thin solid lines demarcate the cones of influence and thick solid lines show the 95% confidence levels.

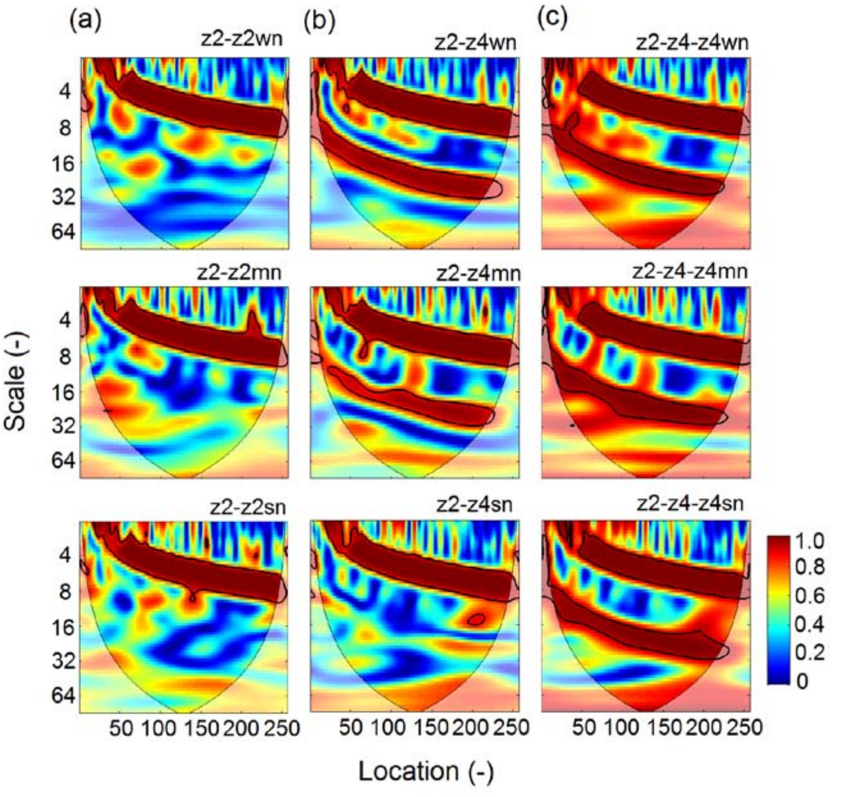

Figure 4. Multiple wavelet coherence of an artificial data series (z) encompassing five cosine waves (z1, z2, z3, z4, and z5) with different scales and (a) z2 and noised z2, (b) z2 and noised z4, and (c) z2, z4, and noised z4 for the non-stationary case. The predictor variables are connected by a hyphen and shown in the top right corner of each subplot. z2wn (z4wn), z2mn (z4mn), and z2sn (z4sn) indicate weakly, moderately, and strongly noised z2 (z4) series, respectively. Weakly, moderately, and strongly noised series are correlated with original series, having with correlation coefficients of 0.9, 0.5, and 0.1, respectively. Thin solid lines demarcate the cones of influence and thick solid lines show the 95% confidence levels.

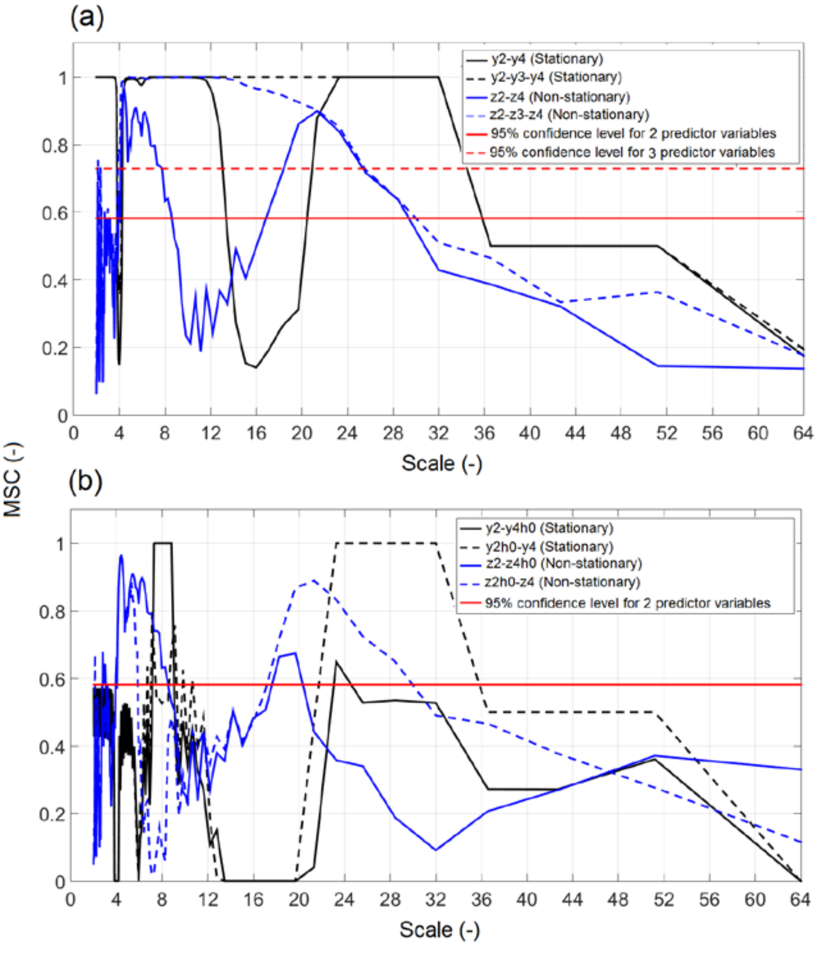

Figure 5. Multiple spectral coherence (MSC) of an artificial data series (y or z) encompassing five cosine waves (y1 to y5; or z1 to z5) with different scales and (a) two (y2 and y4; or z2 and z4) or three (y2, y3, and y4; or z2, z3, and z4) data series, and (b) two (y2 and y4; or z2 and z4) data series when the second half of one data series ~~is~~ are replaced by 0. The variables y2h0 (or z2h0) and y4h0 (or z4h0) refer to the new series of y2 (or z2) and y4 (or z4) in which the second half ~~is~~ are replaced by 0.

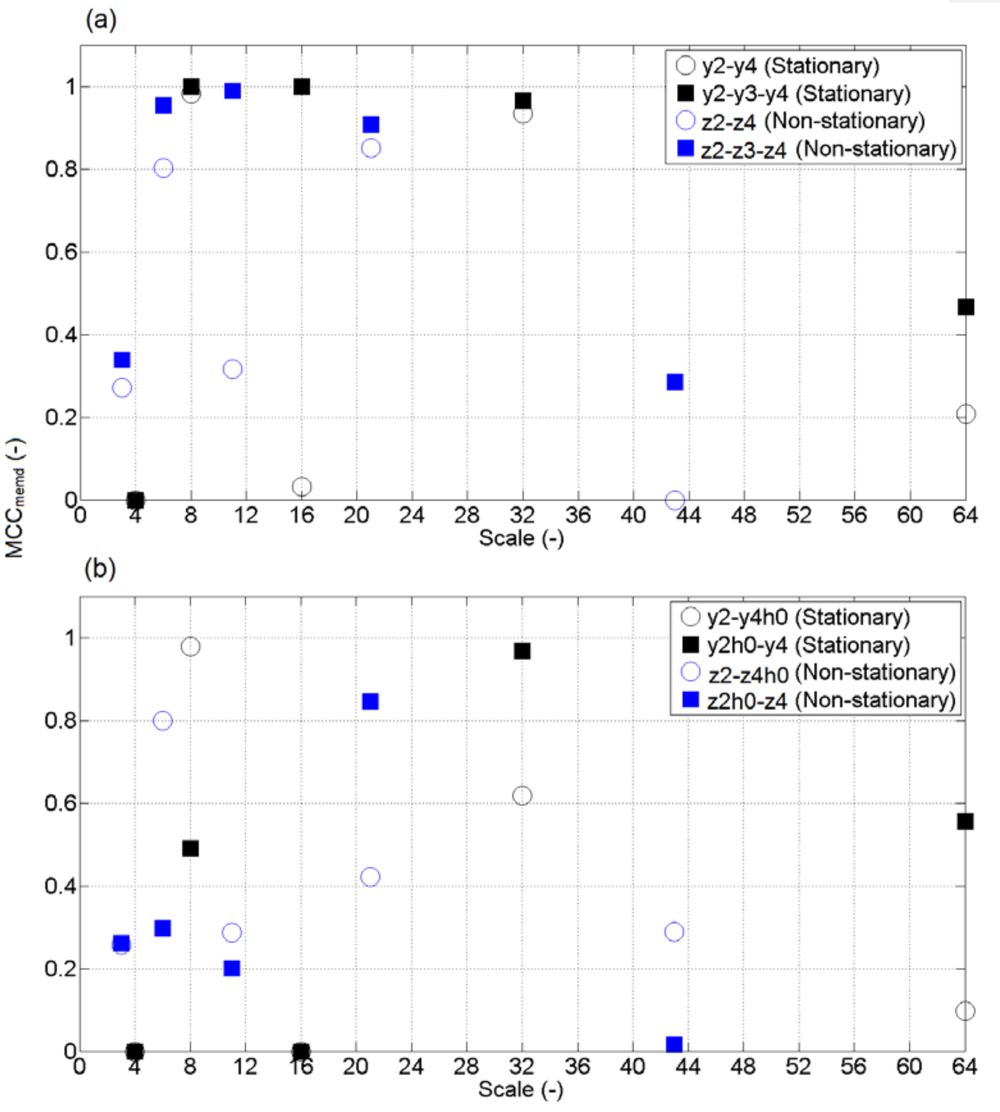

Figure 6. Multiple correlation coefficient between multivariate empirical mode decomposition (MCC_{memd}) of an artificial series (y or z) and (a) two (y2 and y4; or z2 and z4) or three (y2, y3, and y4; or z2, z3, and z4) data series, and (b) two (y2 and y4; or z2 and z4) data series when the second half of one data series ~~is~~ are replaced by 0. The variables y2h0 (or z2h0) and y4h0 (or z4h0) refer to the new series of y2 (or z2) and y4 (or z4) in which the second half ~~is~~ are replaced by 0.

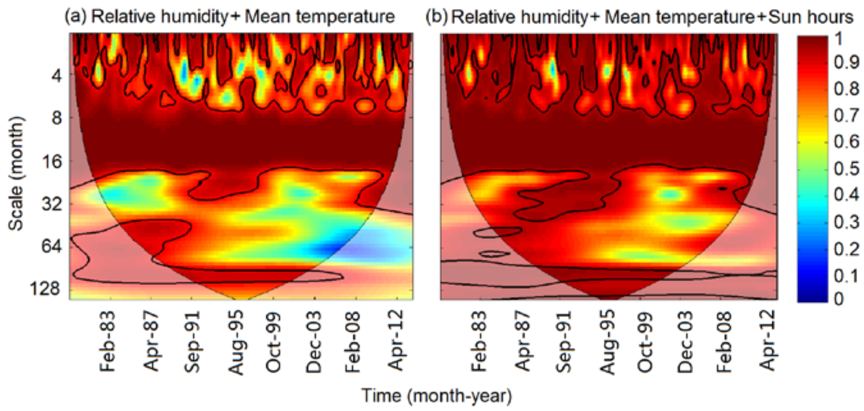

Figure 7. Multiple wavelet coherence between evaporation (E) from water surfaces and meteorological factors ((a) relative humidity and mean temperature and, (b) relative humidity, mean temperature, and sun hours) at Changwu site in Shaanxi, China. Thin solid lines demarcate the cones of influence, and thick solid lines show the 95% confidence level.

Supplement of

Technical Note: Multiple wavelet coherence for untangling scale-specific and localized multivariate relationships in geosciences

Wei Hu and Bing Cheng Si

Correspondence to: Wei Hu (wei.hu@plantandfood.co.nz) and Bing Cheng Si (bing.si@usask.ca)

Introduction

S1 Calculation of smoothed auto- and cross-wavelet power spectra

S2 Matlab code for calculating multiple wavelet coherence

S3 Matlab code for significance test on multiple wavelet coherence

S4 User manual for S2 (mwc.m) and S3 (mwcsignif.m)

S5 Results of MEMD

S6 Results of bivariate~~simple~~ wavelet coherency for *E*

S1 Calculation of smoothed auto- and cross-wavelet power spectra

In this section, we will only introduce the basics related to the calculation of smoothed auto- and cross-wavelet power spectra. Detailed information on the calculations of wavelet coefficients, cross-wavelet power spectra, and bivariate~~simple~~ wavelet coherence can be found elsewhere (Kumar and Foufoula-Georgiou, 1997; Torrence and Compo, 1998; Torrence and Webster, 1999; Grinsted et al., 2004; Das and Mohanty, 2008; Si, 2008). ~~Here, we will only introduce the basics related to the calculation of smoothed auto- and cross-wavelet power spectra.~~ The~~se~~ smoothed auto- and cross-wavelet power spectra require the calculation of wavelet coefficients, at different scales and spatial (or temporal) locations, for the response variable and all predictor variables. For convenience, only spatial variables will be referred to, as temporal variables can be similarly analyzed.

The continuous wavelet transform (CWT) of a spatial variable $X1$ of length N ($X1_h$, h=1, 2, …, N) with equal incremental distance δx, can be calculated as the convolution of $X1_h$ with the scaled and normalized wavelet (Torrence and Compo, 1998)

$$W^{X1}(s,\tau)=\sqrt{\frac{\delta x}{s}}\sum_{\tau=1}^{N}X1_h\psi\left[(h-\tau)\frac{\delta x}{s}\right], \tag{1}$$

where $W^{X1}(s,\tau)$ is the wavelet coefficient of spatial variable $X1$ at scale s and location τ, and $\psi[\ ]$ is the mother wavelet function. The Morlet wavelet is used in the CWT because it allows ~~us to~~for the identification~~y~~ of both location-specific amplitude and phase information at different scales in a spatial series (Torrence and Compo, 1998). The Morlet wavelet can be expressed as (Grinsted et al., 2004)

$$\psi(\eta) = \pi^{-1/4} e^{i\omega\eta - 0.5\eta^2}, \tag{2}$$

where ω and η are the dimensionless frequency and space ($\eta = s/x$), respectively.

The auto-wavelet power spectrum of spatial variable $X1$ can be expressed as

$$W^{X1,X1}(s,\tau) = W^{X1}(s,\tau)\overline{W^{X1}(s,\tau)}, \tag{3}$$

where $\overline{W^{X1}(s,\tau)}$ is a complex conjugate of $W^{X1}(s,\tau)$. Therefore, Eq. (3) can also be expressed as the squared amplitude of $W^{X1}(s,\tau)$, which is

$$W^{X1,X1}(s,\tau) = \left|W^{X1}(s,\tau)\right|^2. \tag{4}$$

The cross-wavelet spectrum between spatial variables of Y and $X1$ can be defined as

$$W^{Y,X1}(s,\tau) = W^{Y}(s,\tau)\overline{W^{X1}(s,\tau)}, \tag{5}$$

where $W^{Y}(s,\tau)$ is the wavelet coefficient of spatial variable Y.

Both the auto- and cross-wavelet spectra can be smoothed using the method suggested by Torrence and Compo (1998),

$$\overleftrightarrow{W}(s,\tau) = \mathrm{SM}_{scale}\left[\mathrm{SM}_{space}\left(W(s,\tau)\right)\right], \tag{6}$$

where $\overleftrightarrow{(\cdot)}$ is a smoothing operator. SM_{scale} and SM_{space} indicate the smoothing along the wavelet scale axis and spatial distance, respectively (Si, 2008). The $\overleftrightarrow{W}$ is the normalized real Morlet wavelet and has a similar footprint as the Morlet wavelet

$$\frac{1}{s\sqrt{2\pi}} e^{\left(-\tau^2/\left(2s^2\right)\right)}. \tag{7}$$

Therefore, the smoothing along spatial distance can be calculated as

$$\mathrm{SM}_{scale}\left(W\left(s,\tau\right)\right)=\sum_{k=1}^{N}\left(W\left(s,\tau\right)\frac{1}{s\sqrt{2\pi}}e^{\left(-\left(\tau-x_k\right)^2/\left(2s^2\right)\right)}\right)\Big|_s\ , \tag{8}$$

where $|_s$ ~~means at~~represents a fixed s value. The Fourier transform of Eq. (7) is $e^{\left(-2s^2\omega^2\right)}$. Therefore, Eq. (8) can be implemented using Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT) based on the convolution theorem, and is written as

$$\mathrm{SM}_{scale}\left(W\left(s,x\right)\right)=\mathrm{IFFT}\left(\mathrm{FFT}\left(W\left(s,x\right)\right)\left(e^{\left(-2s^2\omega^2\right)}\right)\right). \tag{9}$$

The smoothing along scales is then written as [*Torrence and Compo*, 1998]

$$\mathrm{SM}_{scale}\left(W\left(s_k,x\right)\right)=\frac{1}{2m+1}\sum_{l=k-m}^{k+m}\left(\mathrm{SM}_{space}\left(W\left(s_l,x\right)\right)\Pi\left(0.6s_l\right)\right)\Big|_x\ , \tag{10}$$

where Π is the rectangle function, $|_x$ indicates ~~at~~ a fixed x value, and l is the index for the scales. The coefficient of 0.6 is the empirically determined scale decorrelation length for the Morlet wavelet (Torrence and Compo, 1998).

S2 Matlab code for MWC (mwc.m)

```
% This is a Matlab code (mwc.m) for calculating multiple wavelet coherence.
% Please copy the following content into a txt file and rename it to “mwc.m” prior to running.

function varargout=mwc(X,varargin)
%  Multiple Wavelet coherence
% Creates a figure of multiple wavelet coherence
% USAGE: [Rsq,period,scale,coi,sig95]=mwc(X,[,settings])
%
% Input: X: a matrix of multiple variables equally distributed in space
%           or time. The first column corresponds to the dependent variable,
%        and the second and consequent columns are independent variables.
%
% Settings: Pad: pad the time series with zeros?
% .        Dj: Octaves per scale (default: '1/12')
% .        S0: Minimum scale
% .        J1: Total number of scales
% .        Mother: Mother wavelet (default 'morlet')
% .        MaxScale: An easier way of specifying J1
% .        MakeFigure: Make a figure or simply return the output.
% .        BlackandWhite: Create black and white figures
% .        AR1: the ar1 coefficients of the series
% .            (default='auto' using a naive ar1 estimator. See ar1nv.m)
% .        MonteCarloCount: Number of surrogate data sets in the significance calculation. (default=1000)

% Settings can also be specified using abbreviations. e.g. ms=MaxScale.
% For detailed help on some parameters type help wavelet.
% Example:
%    t=1:200;
%    mwc([sin(t),sin(t.*cos(t*.01)),cos(t.*sin(t*.01))])

% Please acknowledge the use of this software package in any publications,
% by including text such as:

%   "The software for the multiple wavelet coherence was  provided by W. Hu
%   and B. Si, and is available in the Supplement of Hu and Si (2016)
% (http://to be determined).~~(http://???).~~"
%   and cite the paper:
% "Hu, W., and B. Si (2016), Technical Note: Multiple wavelet coherence for untangling scale-specific and localized
%  multivariate relationships in geosciences, Hydrol. Earth Syst. Sci., to be determined~~??? (under review)~~"
```

Formatted: Font color: Red

Formatted: Font color: Red, Not Highlight

Formatted: Font color: Red, Not Highlight

Formatted: Font color: Red

```
%
%  (C) W. Hu and B. Si 2016
%
% ----------------------------------------------------------
%
%   Copyright (C) 2016, W. Hu and B. Si 2016
%   This software may be used, copied, or redistributed as long as it is not
%   sold and this copyright notice is reproduced on each copy made.  This
%   routine is provided as is without any express or implied warranties
%   whatsoever.
%
%    Wavelet software was provided by C. Torrence and G. Compo,
%       and is available at URL: http://paos.colorado.edu/research/wavelets/.
%
%    Crosswavelet and wavelet coherence software were provided by
%      A. Grinsted and is available at URL:
%    http://noc.ac.uk/using-science/crosswavelet-wavelet-coherence
%
%   We acknowledge Aslak Grinsted for his wavelet coherency code (wtc.m) on
% which this code builds.
%
%----------------parse function arguments----------------------------------------------------

[row,col]=size(X)
[y,dt]=formatts(X(:,1))
mm=y(1,1)
nn=y(end,1)

for i=2:col
[x,dtx]=formatts(X(:,i))

if (dt~=dtx)
    error('timestep must be equal between time series')
end

mm1=x(1,1)
nn1=x(end,1)

if mm1>mm
mm=mm1
end
```

```
if nn1<nn
nn=nn1
end

x1(:,(i-1))=x(:,1)
x2(:,(i-1))=x(:,2)

end

t=(mm:dt:nn)'


%common time period
if length(t)<4
    error('The three time series must overlap.')
end

n=length(t);

%----------default arguments for the wavelet transform-----------
Args=struct('Pad',1,...      % pad the time series with zeroes (recommended)
            'Dj',1/12, ...    % this will do 12 sub-octaves per octave
            'S0',2*dt,...    % this says start at a scale of 2 years
            'J1',[],...
            'Mother','Morlet', ...
            'MaxScale',[],...   %a more simple way to specify J1
            'MakeFigure',(nargout==0),...
            'MonteCarloCount',1000,...
            'BlackandWhite',0,...
            'AR1','auto',...
            'ArrowDensity',[30 30],...
            'ArrowSize',1,...
            'ArrowHeadSize',1);

Args=parseArgs(varargin,Args,{'BlackandWhite'});

if isempty(Args.J1)
    if isempty(Args.MaxScale)
        Args.MaxScale=(n*.17)*2*dt; %auto maxscale
    end
    Args.J1=round(log2(Args.MaxScale/Args.S0)/Args.Dj);
end
```

```
ad=mean(Args.ArrowDensity);
Args.ArrowSize=Args.ArrowSize*30*.03/ad;
%Args.ArrowHeadSize=Args.ArrowHeadSize*Args.ArrowSize*220;
Args.ArrowHeadSize=Args.ArrowHeadSize*120/ad;

if ~strcmpi(Args.Mother,'morlet')
   warning('MWC:InappropriateSmoothingOperator','Smoothing operator is designed for morlet wavelet.')
end

if strcmpi(Args.AR1,'auto')
     for i=1:col
arc(i)= ar1nv(X(:,i))
      end
Args.AR1=arc
   if any(isnan(Args.AR1))
       error('Automatic AR1 estimation failed. Specify it manually (use arcov or arburg).')
   end
end

%-----------:::::::::::::--------- ANALYZE ----------::::::::::::------------

%Calculate and smooth wavelet spectrum Y and X


[Y,period,scale,coiy] = wavelet(y(:,2),dt,Args.Pad,Args.Dj,Args.S0,Args.J1,Args.Mother);
sinv=1./(scale');
smY=smoothwavelet(sinv(:,ones(1,n)).*(abs(Y).^2),dt,period,Args.Dj,scale);


dte=dt*.01;
idx=find((y(:,1)>=(t(1)-dte))&(y(:,1)<=(t(end)+dte)));
Y=Y(:,idx);
smY=smY(:,idx)
coiy=coiy(idx);

coi=coiy

for  i=2:col
 [XS,period,scale,coix] = wavelet(x2(:,(i-1)),dt,Args.Pad,Args.Dj,Args.S0,Args.J1,Args.Mother);

idx=find((x1(:,(i-1))>=(t(1))-dte)&(x1(:,(i-1))<=(t(end)+dte)));
```

```
XS=XS(:,idx);
coix=coix(idx);

XS1(:,:,(i-1))=XS
coi=min(coi,coix)

end

% -------- Calculate Cross Wavelet Spectra----------------------------

% ---- between dependent variable and independent variables------

for i=1:(col-1)
Wyx=Y.*conj(XS1(:,:,i))
sWyx=smoothwavelet(sinv(:,ones(1,n)).*Wyx,dt,period,Args.Dj,scale)
sWyx1(:,:,i)=sWyx
end

% ----between independent variables and independent variables------
for i=1:(col-1);
for j=1:(col-1);
Wxx=XS1(:,:,i).*conj(XS1(:,:,j))
sWxx=smoothwavelet(sinv(:,ones(1,n)).*Wxx,dt,period,Args.Dj,scale)
sWxx1(:,:,i,j)=sWxx
end
end

% --------------- Mutiple wavelet coherence ---------------------------------
% calculate the multiple wavelet coherence
for i=1:length(scale)
  parfor j=1:n
a=transpose(squeeze(sWyx1(i,j,:)))
b=inv(squeeze(sWxx1(i,j,:,:)))
c=conj(squeeze(sWyx1(i,j,:)))
d=smY(i,j)
Rsq(i,j)=real(a*b*c/d)
  end
end

% --------------- make figure--------------------------------------------
if (nargout>0)||(Args.MakeFigure)
```

```
mwcsig=mwcsignif(Args.MonteCarloCount,Args.AR1,dt,length(t)*2,Args.Pad,Args.Dj,Args.S0,Args.J1,Args.Mother,.6);
    mwcsig=(mwcsig(:,2))*(ones(1,n));
    mwcsig=Rsq./mwcsig;
end

if Args.MakeFigure

    Yticks = 2.^(fix(log2(min(period))):fix(log2(max(period))));

    if Args.BlackandWhite
        levels = [0 0.5 0.7 0.8 0.9 1];
        [cout,H]=safecontourf(t,log2(period),Rsq,levels);

        colorbarf(cout,H)
        cmap=[0 1;.5 .9;.8 .8;.9 .6;1 .5];
        cmap=interp1(cmap(:,1),cmap(:,2),(0:.1:1)');
        cmap=cmap(:,[1 1 1]);
        colormap(cmap)
        set(gca,'YLim',log2([min(period),max(period)]), ...
            'YDir','reverse', 'layer','top', ...
            'YTick',log2(Yticks(:)), ...
            'YTickLabel',num2str(Yticks'), ...
            'layer','top')
        ylabel('Period')
        hold on

        if ~all(isnan(mwcsig))
            [c,h] = contour(t,log2(period),mwcsig,[1 1],'k');%#ok
            set(h,'linewidth',2)
        end
        %suptitle([sTitle ' coherence']);
        %plot(t,log2(coi),'k','linewidth',2)
                tt=[t([1 1])-dt*.5;t;t([end end])+dt*.5];
        %hcoi=fill(tt,log2([period([end 1]) coi period([1 end])]));
        %hatching- modified by Ng and Kwok
        hcoi=fill(tt,log2([period([end 1]) coi period([1 end])]),'w');

        hatch(hcoi,45,[0 0 0]);
        hatch(hcoi,135,[0 0 0]);
        set(hcoi,'alphadatamapping','direct','facealpha',.5)
```

```
        plot(t,log2(coi),'color','black','linewidth',1.5)
        hold off
    else
        H=imagesc(t,log2(period),Rsq);%#ok
        %[c,H]=safecontourf(t,log2(period),Rsq,[0:.05:1]);
        %set(H,'linestyle','none')

        set(gca,'clim',[0 1])

        HCB=safecolorbar;%#ok

        set(gca,'YLim',log2([min(period),max(period)]), ...
            'YDir','reverse', 'layer','top', ...
            'YTick',log2(Yticks(:)), ...
            'YTickLabel',num2str(Yticks'), ...
            'layer','top')
        ylabel('Period')
        hold on

        if ~all(isnan(mwcsig))
            [c,h] = contour(t,log2(period),mwcsig,[1 1],'k');%#ok
            set(h,'linewidth',2)
        end
        %suptitle([sTitle ' coherence']);
        tt=[t([1 1])-dt*.5;t;t([end end])+dt*.5];
        hcoi=fill(tt,log2([period([end 1]) coi period([1 end])]),'w');
        set(hcoi,'alphadatamapping','direct','facealpha',.5)
        hold off
    end
end
%---------------------------------------------------------------%

varargout={Rsq,period,scale,coi,mwcsig};
varargout=varargout(1:nargout);

function [cout,H]=safecontourf(varargin)
vv=sscanf(version,'%i.');
if (version('-release')<14)|(vv(1)<7)
    [cout,H]=contourf(varargin{:});
else
    [cout,H]=contourf('v6',varargin{:});
end
```

```
function hcb=safecolorbar(varargin)
vv=sscanf(version,'%i.');

if (version('-release')<14)|(vv(1)<7)
    hcb=colorbar(varargin{:});
else
    hcb=colorbar('v6',varargin{:});
end
```

S3 Matlab code for significance test on multiple wavelet coherence

% This is a Matlab file (mwcsignif.m) for calculating significance tests on multiple wavelet coherence.

%Please copy the following content into a txt file and rename this file to “mwcsignif.m” prior to running.

```
function mwcsig=mwcsignif(mccount,ar1,dt,n,pad,dj,s0,j1,mother,cutoff)
% Multiple Wavelet Coherence Significance Calculation (Monte Carlo)
%
% mwcsig=mwcsignif(mccount,ar1,dt,n,pad,dj,s0,j1,mother,cutoff)
%
% mccount: number of time series generations in the monte carlo run
%(the greater the better)
% ar1: a vector containing two (in case of calculating wavelet
% coherence between two variables) or
% multiple (≥3) (in case of calculating multiple wavelet coherence
% with three or more variables)
% AR1 coefficients.
% dt,pad,dj,s0,j1,mother: see wavelet help...
% n: length of each generated timeseries. (obsolete)
%
% cutoff: (obsolete)
%
% RETURNED
% mwcsig: the 95% significance level as a function of scale... (scale,sig95level)
% -----------------------------------------------------------
% Please acknowledge the use of this software package in any publications,
% by including text such as:
%
%   "The software for the multiple wavelet coherence was  provided by W. Hu
%   and B. Si, and is available in the supplement of Hu and Si (2016)
% (http://to be determined~~http://???~~)"
%   and cite the paper:
% "Hu, W., and B. Si (2016), Technical Note: Multiple wavelet coherence for untangling scale-specific and localized
%  multivariate relationships in geosciences, Hydrol. Earth Syst. Sci., to be determined~~??? (under review)~~"
%
%  (C) W. Hu and B. C. Si 2016
%
% -----------------------------------------------------------
%
%  Copyright (C) 2016, W. Hu and B. C. Si 2016
%   This software may be used, copied, or redistributed as long as it is not
%   sold and this copyright notice is reproduced on each copy made.  This
```

Formatted: Font: Italic, Font color: Red

Formatted: Font color: Red

Formatted: Font: Italic

Formatted: Font: Italic, Font color: Red

Formatted: Font color: Red

```
%   routine is provided as is without any express or implied warranties
%  whatsoever.
%
%  Wavelet software was provided by C. Torrence and G. Compo,
%   and is available at URL: http://paos.colorado.edu/research/wavelets/.
%
%  Crosswavelet and wavelet coherence software were provided by
%    A. Grinsted and is available at URL:
%    http://noc.ac.uk/using-science/crosswavelet-wavelet-coherence
%
%
% We acknowledge Aslak Grinsted for his code (wtcsignif.m) on
% which this code builds.
%
%---------------------------------------------------------------------
persistent mypath
if isempty(mypath)
    mypath=strrep(which('mwcsignif'),'mwcsignif.m','');
end

%we don't need to do the monte carlo if we have a cached
%siglevel for ar1s that are almost the same. (see fig4 in Grinsted et al., 2004)
aa=round(atanh(ar1(:)')*4); %this function increases the sensitivity near 1 & -1
aa=abs(aa)+.5*(aa<0); %only positive numbers are allowed in the checkvalues (because of log)

% do a check that it is not the same as last time... (for optimization purposes)
checkvalues=[aa dj s0/dt j1 double(mother)]; %n & pad are not important.
%also the resolution is not important.

checkhash=['' mod(sum(log(checkvalues+1)*127),25)+'a' mod(sum(log(checkvalues+1)*54321),25)+'a'];

cachefilename=[mypath 'mwcsignif-cached-' checkhash '.bnm'];

%the hash is used to distinguish cache files.
try
    [lastmccount,lastcheckvalues,lastmwcsig]=loadbnm(cachefilename);
    if (lastmccount>=mccount)&(isequal(single(checkvalues),lastcheckvalues))
%single is important because bnm is single precision.
        mwcsig=lastmwcsig;
         return
    end
catch
```

```
end

%choose a n so that largest scale have atleast some part outside the coi
ms=s0*(2^(j1*dj))/dt; %maxscale in units of samples
n=ceil(ms*6);

warned=0;
%precalculate stuff that's constant outside the loop
%d1=ar1noise(n,1,ar1(1),1);
d1=rednoise(n,ar1(1),1);
[W1,period,scale,coi] = wavelet(d1,dt,pad,dj,s0,j1,mother);
outsidecoi=zeros(size(W1));
for s=1:length(scale)
    outsidecoi(s,:)=(period(s)<=coi);
end
sinv=1./(scale');
sinv=sinv(:,ones(1,size(W1,2)));

if mccount<1
    mwcsig=scale';
    mwcsig(:,2)=.71; %pretty good
    return
end

sig95=zeros(size(scale));

maxscale=1;
for s=1:length(scale)
    if any(outsidecoi(s,:)>0)
        maxscale=s;
    else
        sig95(s)=NaN;
        if ~warned
warning('Long wavelengths completely influenced by COI. (suggestion: set n higher, or j1 lower)');
         warned=1;
        end
    end
end

%PAR1=1./ar1spectrum(ar1(1),period');
%PAR1=PAR1(:,ones(1,size(W1,2)));
%PAR2=1./ar1spectrum(ar1(2),period');
```

```
%PAR2=PAR2(:,ones(1,size(W1,2)));

nbins=1000;
wlc=zeros(length(scale),nbins);

wbh = waitbar(0,['Running Monte Carlo (significance)... (H=' checkhash ')'],'Name','Monte Carlo (MWC)');

for ii=1:mccount
    waitbar(ii/mccount,wbh);

dy=rednoise(n,ar1(1),1)
[Wdy,period,scale,coiy] = wavelet(dy,dt,pad,dj,s0,j1,mother);
sinv=1./(scale');
smdY=smoothwavelet(sinv(:,ones(1,n)).*(abs(Wdy).^2),dt,period,dj,scale);

col=size(ar1,2)

for  i=2:col
dx=rednoise(n,ar1(i),1)
 [Wdx,period,scale,coix] = wavelet(dx,dt,pad,dj,s0,j1,mother);
Wdx1(:,:,(i-1))=Wdx
end

% -------- Calculate Cross Wavelet Spectra----------------------------

% ----between dependent variable and independent variables------

parfor i=1:(col-1)
Wdyx=Wdy.*conj(Wdx1(:,:,i))
sWdyx=smoothwavelet(sinv(:,ones(1,n)).*Wdyx,dt,period, dj,scale)
sWdyx1(:,:,i)=sWdyx
end

% ----between independent variables and independent variables------
for i=1:(col-1);
parfor j=1:(col-1);
Wdxx=Wdx1(:,:,i).*conj(Wdx1(:,:,j))
sWdxx=smoothwavelet(sinv(:,ones(1,n)).*Wdxx,dt,period,dj,scale)
sWdxx1(:,:,i,j)=sWdxx
end
end
```

```
% calculate the multiple wavelet coherence
for i=1:length(scale)
  parfor j=1:n
a=transpose(squeeze(sWdyx1(i,j,:)))
b=inv(squeeze(sWdxx1(i,j,:,:)))
c=conj(squeeze(sWdyx1(i,j,:)))
d=smdY(i,j)
Rsq(i,j)=real(a*b*c/d)
  end
end

 for s=1:maxscale
        cd=Rsq(s,find(outsidecoi(s,:)));
        cd=max(min(cd,1),0);
        cd=floor(cd*(nbins-1))+1;
        for jj=1:length(cd)
            wlc(s,cd(jj))=wlc(s,cd(jj))+1;
        end
    end
end
close(wbh);

for s=1:maxscale
    rsqy=((1:nbins)-.5)/nbins;
    ptile=wlc(s,:);
    idx=find(ptile~=0);
    ptile=ptile(idx);
    rsqy=rsqy(idx);
    ptile=cumsum(ptile);
    ptile=(ptile-.5)/ptile(end);
    sig95(s)=interp1(ptile,rsqy,.95);
end
mwcsig=[scale' sig95'];

if any(isnan(sig95))&(~warned)
    warning(sprintf('Sig95 calculation failed. (Some NaNs)'))
else
    savebnm(cachefilename,mccount,checkvalues,mwcsig); %save to a cache....
end
```

S4 User manual for S2 (mwc.m) and S3 (mwcsignif.m)

Multiple wavelet- coherence package

by Wei Hu and Bingcheng Si

Release date: ~~xx~~ 27 April 2016

This software package is written for performing multiple wavelet coherence.

This software package includes mwc.m and mwcsignif.m, which

are written in the Matlab program based on wtc.m and wtcsignif.m provided by A.

Grinsted

(http://noc.ac.uk/using-science/crosswavelet-wavelet-coherence).

Users are, therefore, required to download his software package and

combine these two packages into one to run the multiple wavelet coherence analysis.

Please acknowledge the use of ~~of~~ this software package in any publications by including

text such as:

****

The software for the multiple wavelet coherence was provided by W. Hu and B. C. Si,
and is available in the supplement of Hu and Si (2016) (http://~~???~~*to be determined*).

****

and cite the paper:

%%%

Hu, W., and B.C. Si (2016), Technical Note: Multiple wavelet coherence for untangling
scale-specific and localized multivariate relationships in geosciences, Hydrol. Earth Syst.
Sci., *to be determined*~~??? (under review)~~.~~"~~

%%%

Formatted: Font color: Red

Formatted: Font: Italic, Font color: Red

Formatted: Font color: Red

Formatted: Font color: Red

Acknowledgements:

Wavelet software was provided by C. Torrence and G. Compo,
and is available at URL: http://paos.colorado.edu/research/wavelets/.

Crosswavelet and wavelet coherence software were provided by
A. Grinsted and is available at URL:
http://noc.ac.uk/using-science/crosswavelet-wavelet-coherence

Should there be any enquiriese, please feel free to contact:

Wei Hu
Email: wei.hu@plantandfood.co.nz

Bing Si
Email: bing.si@usask.ca

S5 Results of MEMD

Six or seven intrinsic mode functions (IMFs) corresponding to different scales are obtained for multivariate data series (i.e., a combination of the response variable with two (y2 and y4, or z2 and z4) or three (y2, y3, and y4, or z2, z3, and z4) predictor variables) by MEMD. ~~Because~~ Due to~~the~~ IMFs, with a number of 6 or greater, contributing~~ed~~ negligible variance to the total, only the first five IMFs are presented (Fig. S1). For each IMF, the scale is calculated as the total number of points (i.e., 256) divided by the number of ~~cycles~~ cycles for each IMF. The obtained scales and percentage (%) of variance explained by each IMF are shown in Table S1. While the obtained scales for the response variable y are in agreement with the true scales for the stationary case, the obtained scales (i.e., 3, 6, 11, 21, and 43) for the response variable z deviate slightly from the average scales for the non-stationary case. For the response variable, the contribution of IMFs to the total variance generally decreases (20% to 13% for stationary and 27% to 11% for non-stationary) from IMF1 to IMF5, which disagrees with the fact that each scale contributes equally (i.e., 20%) to the total variance. The scale of the dominant variance from each predictor variable can be obtained (Table S1). However, the sum of variances over all IMFs for each variable is less than 100% (ranging from 84% to 93%), indicating that MEMD cannot capture all the variances, as was also previously observed (Hu et al., 2013; She et al., 2014).

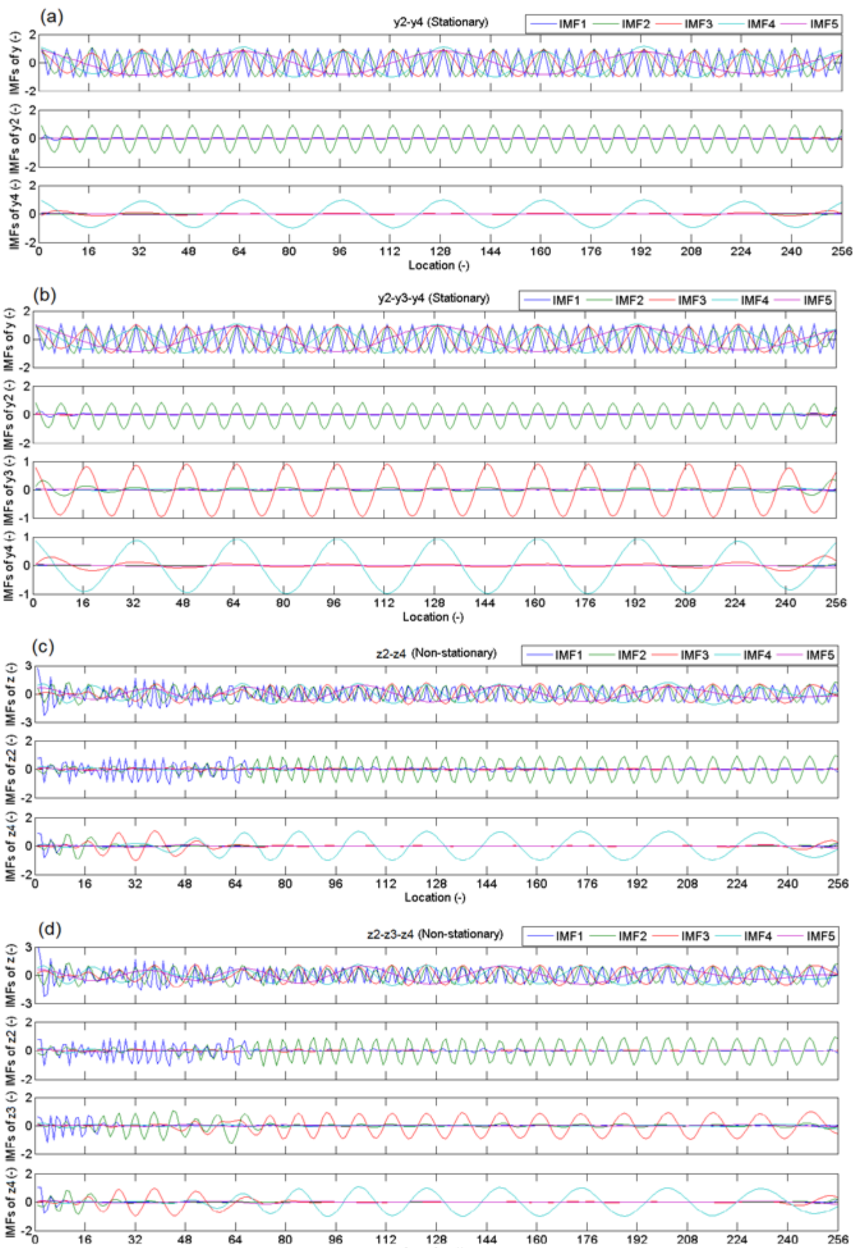

Figure S1. The first five intrinsic mode functions (IMFs) of response variable y (or z) and predictor variables (y2 and y4; y2 y3, and y4; z2 and z4; or z2, z3, and z4) obtained by multivariate empirical mode decomposition.

Table S1. Scales and percentage (%) of variance explained by each intrinsic mode function (IMF) of response variable y (or z) and predictor variables (y2 and y4; y2, y3 and y4; z2 and z4; or z2, z3, and z4) using the multivariate empirical mode decomposition method.

		Scale (-)	y (%)	y2 (%)	y3 (%)	y4 (%)
y2-y4 (Stationary)	IMF1	4	20	0		0
	IMF2	8	18	**90**		0
	IMF3	16	15	0		1
	IMF4	32	18	0		**88**
	IMF5	64	13	0		0
y2-y3-y4 (Stationary)	IMF1	4	20	1	0	0
	IMF2	8	17	**85**	1	0
	IMF3	16	16	0	**82**	2
	IMF4	32	16	0	0	**82**
	IMF5	64	15	0	0	0
z2-z4 (Non-stationary)	IMF1	3	27	22		2
	IMF2	6	17	**68**		4
	IMF3	11	17	0		11
	IMF4	21	17	0		**75**
	IMF5	43	11	0		0
z2-z3-z4 (Non-stationary)	IMF1	3	27	22	7	3
	IMF2	6	18	**69**	17	4
	IMF3	11	17	0	**61**	14
	IMF4	21	16	0	1	**68**
	IMF5	43	11	0	0	0

S6 Results of bivariate~~simple~~ wavelet coherency for *E*

The evaporation from free water surface was significantly correlated to each meteorological factor at scales of around 1 year, at all times, with exception ~~of~~ to a certain period for relative humidity and sun hours (Fig. S2). Each of mean temperature, sun hours, and wind speed was positively correlated to *E* at different scales. For relative humidity~~,~~ however, its influences on *E* changed with scale. For example, at scales of around 1 year, relative humidity was positively correlated to *E* during the period of 1979 to 1997. This is ~~because~~ due to high relative humidity ~~is~~ usually being associated with high summer temperatures~~s in summer~~, when high evaporation occurs. At other scales (e.g., 2–6 months or 5–10 years), the relative humidity was negatively correlated to the *E*, which was expected. The dominant factors explaining variation in *E* differed with scale. For example, the relative humidity was the dominating factor at small (2–8 months) and large (>32 months) scales, while temperature was the dominating factor at the medium (8–32 months) scales (Fig. S2). The relative humidity corresponded to the greatest mean MWC (0.62) and PASC value (40%) at multiple scale-location domains.

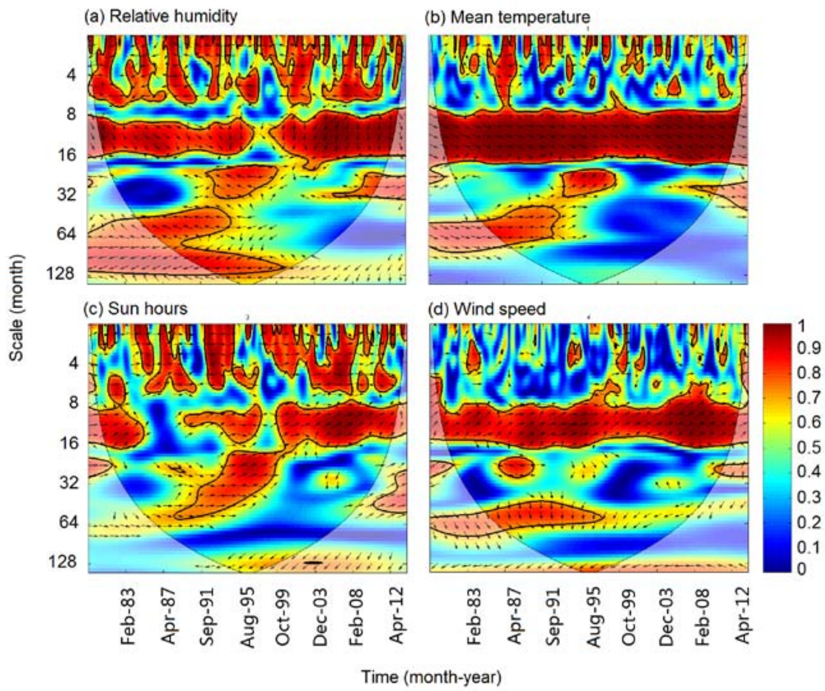

Figure S2. ~~Simple~~Bivariate wavelet coherency between evaporation (*E*) from water surfaces and each of the meteorological factors (relative humidity, mean temperature, sun hours, and wind speed) at Changwu site in Shaanxi, China. Arrows show the correlation type with the right ~~hand~~pointing arrows being positive and left~~-hand~~ pointing arrows being negative. Thin solid lines demarcate the cones of influence and thick solid lines show the 95% confidence levels.

References

Das, N.N. and Mohanty, B. P.: Temporal dynamics of PSR-based soil moisture across spatial scales in an agricultural landscape during SMEX02: A wavelet approach, Remote Sens. Environ., 112, 522–534, doi:10.1016/j.rse.2007.05.007, 2008.

Grinsted, A., Moore, J. C., and Jevrejeva, S.: Application of the cross wavelet transform and wavelet coherence to geophysical time series, Nonlinear Proc. Geoph., 11, 561–566, 2004.

Hu, W. and Si, B. C.: Soil water prediction based on its scale-specific control using multivariate empirical mode decomposition, Geoderma, 193–194,180–188, doi: 10.1016/j.geoderma.2012.10.021, 2013.

Kumar, P. and Foufoula-Georgiou, E.: Wavelet analysis for geophysical applications, *Rev. Geophys*., 35, 385–412, doi: 10.1029/97RG00427, 1997.

She, D. L., Tang, S. Q., Shao, M. A., Yu, S. E., and Xia, Y. Q.: Characterizing scale specific depth persistence of soil water content along two landscape transects, *J. Hydrol*., *519*, 1149–1161, doi:10.1016/j.jhydrol.2014.08.034, 2014.

Si, B. C.: Spatial scaling analyses of soil physical properties: A review of spectral and wavelet methods, Vadose Zone J., 7, 547–562, doi: 10.2136/vzj2007.0040, 2008.

Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc., 79, 61–78, doi: 10.1175/1520-0477(1998)079<0061:apgtwa>2.0.co;2, 1998.

Torrence, C. and Webster, P. J.: Interdecadal changes in the ENSO-monsoon system, J. Clim., 12, 2679–2690, doi: 10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2, 1999.