■ Methods

Solar spectrum through the wavelet lens

Agnieszka Jach, ORCID: 0000-0001-6806-5685
Department of Finance and Statistics, Hanken School of Economics, 00101 Helsinki, Finland
DOI: 10.19232/uv4pb.2016.2.14 © 2016 The Author, licensed under

Measurements of spectral irradiance, which are used to describe the solar spectrum, can be viewed as signals made up of rough and smooth components. Wavelet analysis (Percival and Walden 2000), which we present in this article, allows one to extract these components. Wavelet analysis was developed in the 1980's and 90's (Daubechies 1992; Mallat 1998) and is primarily used as a tool for exploring time series and image data, in fields ranging from physics to finance. It breaks up the signal into components associated with scale and time (in the context of spectral irradiance, with scale and wavelength), where small scales correspond to the rough part of the signal, while large scales correspond to the smooth. This gives rise to the so-called Multi-Resolution Decomposition (equation (7.1) below). In the current paper we emphasize the smoothing application of the methodology, though we discuss the decomposition for all scales, and explain how to perform and interpret wavelet analysis of spectral irradiance data.

Why would we want to smooth out the spectral irradiance? Smoothing is not needed when calculating typical indicators of light quality such as red:far-red or UVB:PAR photon ratios as these are based on integrals of spectral irradiance over particular wavelength bands. However, if we want to compare entire spectra (functions) collected under different experimental conditions rather than integrals (numbers), such a prerequisite step might be necessary. This is precisely the case with functional data analysis (Horváth and Kokoszka 2012; Ramsay and Silverman 2005). Within this framework signals are treated as functions, but they have to be smooth at the outset. A comparison of functions, in our case (smoothed) spectral irradiance measurements, can then be facilitated by, for example, functional multi-way ANOVA (Albertos and Bande 2010).

Let $X = [X_{\lambda_1}, X_{\lambda_2}, \dots, X_{\lambda_n}]$ represent the signal (e.g., one of the measurements from the top-right panel of Figure 7.1 or the average of those measurements depicted in the top-left, given on a grid of equi-spaced wavelengths $\lambda_1, \lambda_2, \dots, \lambda_n$, with the sampling interval $\Delta\lambda = \lambda_i - \lambda_{i-1}$), then the Multi-Resolution Decomposition, MRD, states that

$$X = D_1 + D_2 + \dots + D_{J_0} + S_{J_0}, \qquad (7.1)$$

where $D_j = [D_{j,\lambda_1}, D_{j,\lambda_2}, \dots, D_{j,\lambda_n}]$, $1 \le j \le J_0$ and $S_{J_0} = [S_{J_0,\lambda_1}, S_{J_0,\lambda_2}, \dots, S_{J_0,\lambda_n}]$. '$D$' stands for Detail (there are several of them), while 'S' stands for Smooth (there is only one smooth). The number of components on the right-hand side of equation (7.1), governed by the positive integer J_0 ($J_0 \le J = \lfloor \log_2(n) \rfloor$), depends on the task in hand. In the example of Figure 7.1 we chose $J_0 = 3$ because for such J_0 we were able to separate the smooth part of the measurement of the spectral irradiance (bottom-left panel of Figure 7.1 for the average and bottom-right for the individual ones) from the rough one captured by the details D_1, D_2 and D_3 (middle panels of Figure 7.1, left for the average and right for the individual). Had we have stopped at $J_0 = 2$,

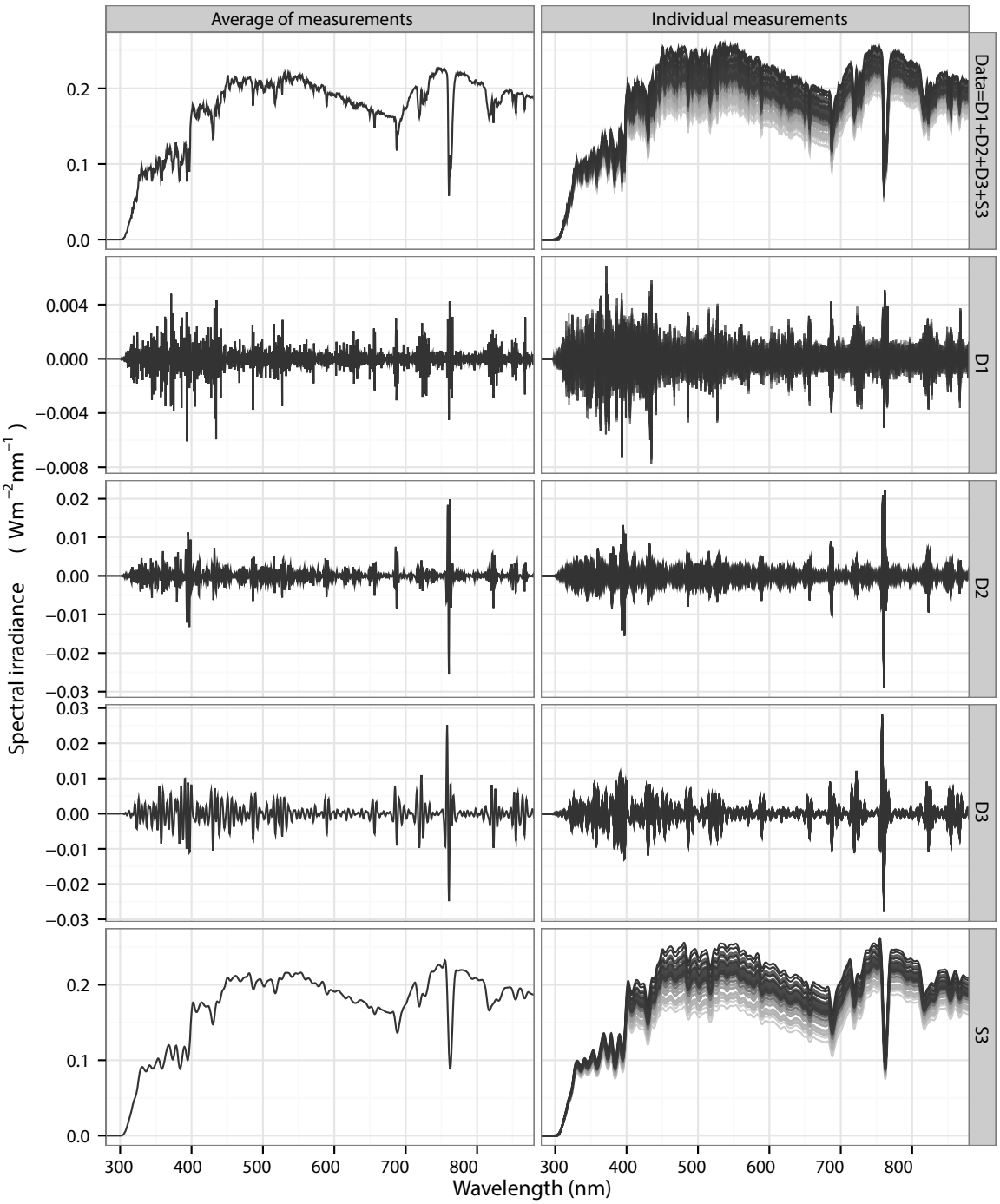

Figure 7.1: *Top:* spectral irradiance: average of measurements in the left and individual measurements in the right. *Middle and bottom:* Elements of the Multi-Resolution Decomposition (D_1, D_2, D_3, S_3) of the spectral irradiance based on wavelet analysis. Data collected with an array spectrometer (*Maya2000Pro*, Ocean Optics, 100 measurements over 50 seconds) on the 22.05.2015 around solar noon at Lammi Biological Station, Lammi, Finland (61°03'N, 25°02'E) in a leaf semi-shade position of a *Betula pendula* stand.

© 2016-2017 by the authors

our final element of the decomposition, S_2, would have consisted of the sum of the two bottom rows of Figure 7.1, and as such would have been affected by the roughness coming from D_3. The elements of the details and of the smooth, which essentially reflect variations in local averages of the signal, are interpreted as follows: D_{j,λ_i} is associated with the change/difference in weighted averages of X over two adjacent wavelength bands of width $2^{j-1}\Delta\lambda$, one band ending at λ_i and the other starting at λ_i ($2^{j-1}\Delta\lambda$ is referred to as scale); S_{J_0,λ_i} is associated with averages of X over wavelength bands of width $\geq 2^{J_0}\Delta\lambda$.

Let's have a closer look at the elements of the decomposition (left panels for the average of measurements and right panels for the individual measurements, at wavelengths (in nm) $\lambda_1 = 280, \lambda_2 = 280.5, \ldots, \lambda_n = 880$, with $\Delta\lambda = 0.5$). The measurements, recorded at nearly equi-spaced wavelengths with gaps from 0.43 to 0.48 nm, were interpolated to equi-spaced wavelengths (constant, left-continuous interpolation). The smooth (bottom panels), tracks its respective signal very well and indeed is smoother than the data itself. Hence smoothing of the spectral irradiance is achieved by dropping all the elements of the decomposition apart from S_3. What about the details? First of all, their y-ranges are much smaller than the y-range of the corresponding smooth. Detail D_1 (capturing the changes in the signal over the smallest wavelength band of size $2^{j-1}\Delta\lambda = 2^0\Delta\lambda$, hence 0.5 nm) based on the averaged measurements exhibits less variation compared to the details D_1 for the individual measurements, because we essentially averaged out the coarse part of the data in the former. Also, since in the right panels the different measurements (and their MRDs) are shown in different shades of grey, from lighter to darker, we notice that the details are largely consistent between the measurements (darker shades are superimposed on the lighter shades), which is not the case for the smooth. If we compare D_2 and D_3 in the left panels to those in the right ones, we also observe a high degree of overlap, hence the impact of averaging the measurements on these two components is far smaller than on D_1.

Several spikes in D_1, D_2 and D_3 around wavelengths (in nm) 370, 390, 430, 490, and 760 can be spotted. They imply that the changes in local averages of the spectral irradiance over scales of, respectively, 0.5, 1, 2 nm ($= 2^{j-1}\Delta\lambda$ for $j = 1, 2, 3$) at those λ_i's were much bigger compared to others. The inspection of the signal corroborates this, but it would be harder to detect such features without the wavelet lens. Another interesting quality of the MRD is an apparent increase in the variability of D_1 for wavelengths of up to 450 nm (it is also possible to perform a test of homogeneity of variance for the wavelet coefficients from which the details are derived). This suggests that the changes in local averages over scales of 0.5 nm at the beginning of the spectrum are much higher than for the remaining part. We leave the physical interpretation of these phenomena to the reader.

One can obtain MRD and plot it in `R` in three lines of code thanks to the package `wmtsa`, which accompanies the book of (Percival and Walden 2000). We demonstrate it below with the `sunspots` data set included in that package. Of course we skipped the technical details of the analysis, e.g., the choice of a particular wavelet transform, wavelet filter, handling boundary conditions, and refer the reader to Section 5.11 of (Percival and Walden 2000) for the details. The data and the `R` code (exceeding three lines) utilized in the current paper are available upon request.

```
#install.packages("wmtsa")
library(wmtsa)
x.modwt<-wavMODWT(x=sunspots,
                  wavelet="s8",
                  n.levels=3)
x.MRD<-wavMRD(x.modwt)
plot(x.MRD)
```

© 2016-2017 by the authors

Acknowledgements: The data was supplied by T.M. Robson and pre-processed using `R` package `MayaCalc`.

References

Albertos, J. C. and M. F. Bande (2010). "A simple multiway ANOVA for functional data". In: *TEST* 19, pp. 537–557. DOI: 10.1007/s11749-010-0185-3.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. ISBN: 9780898712742. Philadelphia: SIAM. DOI: 10.1137/1.9781611970104.fm.

Horváth, L. and P. Kokoszka (2012). *Inference for Functional Data with Applications*. ISBN: 9781461436546. Berlin: Springer. DOI: 10.1007/978-1-4614-3655-3.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing*. ISBN: 9780124666054. San Diego: Academic Press.

Percival, D. and A. Walden (2000). *Wavelet Methods for Time Series Analysis*. ISBN: 9780521685085. Cambridge: Cambridge University Press.

Ramsay, J. and B. Silverman (2005). *Functional Data Analysis*. ISBN: 9780387400808. Berlin: Springer. DOI: 10.1007/b98888.

Peer reviewed article.
Published on-line on 2016-12-28.
Edited by: P. J. Aphalo.

 © 2016-2017 by the authors