

**Impact case study (REF3b)**

<p><b>Institution:</b> Imperial College London</p>
<p><b>Unit of Assessment:</b> 10 Mathematical Sciences</p>
<p><b>Title of case study:</b> C6 - Wavelet analysis techniques developed into multiple software packages and widely used internationally including in the biomedical, conservation and financial sectors</p>
<p><b>1. Summary of the impact</b> (indicative maximum 100 words)</p> <p>Methodological, algorithmic and interpretational advances in wavelet techniques for time series analysis are encapsulated in the research monograph by Percival and Walden (2000): “Wavelet Methods for Time Series Analysis” (WMTSA). Multiple language software packages have been developed from the book’s contents, including the Spotfire S+ package from the major commercial software company TIBCO (2008-present). TIBCO Spotfire clients span many sectors and include major companies such as GE, Chevron, GlaxoSmithKline and Cisco. Further applications of the wavelet techniques developed in WMTSA include in the biomedical, conservation and financial sectors. WMTSA is used, for example, in functional Magnetic Resonance Imaging by GlaxoSmithKline, to monitor cracks in the dome of the UNESCO world heritage site Santa Maria del Fiore Cathedral in Florence, and by the Reserve Bank of New Zealand in its analysis of measuring core inflation.</p>
<p><b>2. Underpinning research</b> (indicative maximum 500 words)</p> <p>Professor Andrew Walden started working on wavelets initially via the linkage between wavelets and long-memory time series [G1], resulting in publication [1] with Dr Emma McCoy in 1996. With Alberto Contreras Cristan, he then studied the maximal overlap (undecimated) discrete wavelet transform (MODWT) and maximal overlap wavelet packet transforms (MODWPT). In particular the optimum time shifts to apply to ensure approximate zero phase filtering at every level of the transform were derived, and applied to the wavelet packet coefficients to give phase corrections which ensure alignment with the original time series, making the analysis methodology useful to physical scientists [2]. In addition, in the context of matching pursuit, they showed how to expand the MODWT dictionary in a physically sensible way while maintaining key theoretical properties.</p> <p>A wavelet thresholding scheme for multitaper spectral estimators was derived with McCoy and Don Percival of the University of Washington [3] and is much simpler and preferable to the previously proposed scheme involving the periodogram. The ‘complete’ pilot spectrum estimator of Blackman and Tukey was updated and extended by using the MODWPT in work with Eva Tsakiroglou. The statistical properties of the wavelet variance estimator for the scale analysis of time series were studied with Abdeslam Serroukh and Percival [4]. The asymptotic distribution of the MODWT-based wavelet variance estimator was derived for a wide class of stochastic processes, not necessarily Gaussian or linear. It was shown how to estimate the variance of the estimator using spectral methods.</p> <p>In 1998 Percival and Walden largely completed the research monograph WMTSA [5]; this incorporated the aforementioned underpinning research. Percival contributed introductory, qualitative and graphical explanatory material on transforms, and his own research results, particularly applications to long-memory processes. Both authors took a novel filtering approach to the mathematical exposition (rendering results more easily used in a statistical context) including a proper treatment of the boundary wavelet coefficients, details and smooths showing how to delineate reliable values free of end-effects, of great importance to practitioners. The impact of the book led to demand for a Chinese translation, published by the China Machine Press (2006).</p> <p>Walden’s research continued with a comparison of wavelets versus wavelet packets for power spectrum estimation with Cristan and the new area of wavelet analysis of matrix-valued times series, closely connected to multiwavelets, with Abdeslam Serroukh (grant [G2]). Sofia Olhede joined the research group in 2000 as focus moved to continuous wavelet transforms, and particularly analysis of Morse wavelets, resulting in articles showing applications to polarization in</p>

## Impact case study (REF3b)

earthquakes and Doppler ultrasound. Analysis of multicomponent signals and the development of wavelet methodologies as an alternative to empirical mode decomposition followed (e.g., [6]). Further work took place on the use of analytic signals methodology for time series analysis, resulting in analytic thresholding and novel directional denoising schemes.

Statistics Section contributors:

- Andrew Walden, Professor of Statistics, Imperial College (1990-present)
- Emma McCoy, initially an RA supported on grant [G1], now Senior Lecturer in Statistics, Imperial College (1996-present)
- Alberto Contreras Cristan, PhD student, Imperial College (1994-1998)
- Evangelia Tsakiroglou, PhD student, Imperial College (1997-2000)
- Abdeslam Serroukh, RA supported on grants [G1, G2], Imperial College (1996-99)
- Sofia Olhede, Senior Lecturer in Statistics, Imperial College (2002-2007), now Pearson Professor of Statistics, UCL (2007-present).

### 3. References to the research (\* References that best indicate quality of underpinning research)

- [1] [E. J. McCoy](#) and [A. T. Walden](#), 'Wavelet analysis and synthesis of stationary long-memory processes', Journal of Computational and Graphical Statistics, 5, 26-56 (1996). [DOI](#)
- [2] [Walden, A.T.](#) and [Contreras Cristan, A.](#), 'The phase-corrected undecimated discrete wavelet packet transform and its application to interpreting the timing of events', Proc. R. Soc. Lond. A, 454, 2243-2266 (1998). [DOI](#)
- [3] \*[Walden, A.T.](#), [Percival, D.B.](#) and [McCoy, E.J.](#), 'Spectrum estimation by wavelet thresholding of multitaper estimators', IEEE Transactions on Signal Processing, 46, 3153-3165 (1998). [DOI](#)
- [4] \*[Serroukh A.](#), [Walden, A.T.](#) and [Percival, D.B.](#), 'Statistical properties and uses of the wavelet variance estimator for the scale analysis of time series'. J. Am. Stat. Assoc., 95, 184-196 (2000). [DOI](#)
- [5] \*[D. B. Percival](#) and [A. T. Walden](#), '[Wavelet Methods for Time Series Analysis](#)', Cambridge University Press, 594pp, Hb, (2006).
- [6] [S. Olhede](#) and [A. T. Walden](#), 'The Hilbert spectrum via wavelet projections', Proc. R. Soc. Lond. A, 460, 955-975 (2004). [DOI](#)

Grants:

- [G1] EPSRC, [GR/J62715/01](#), 'Time Series Analysis using the Discrete Wavelet Transform', PI: A.T. Walden, 01/10/94-31/03/98, £124,878
- [G2] EPSRC, [GR/L11182/01](#), 'Advances in the theory and practice of multiwavelets', PI: A.T. Walden, 20/01/97-19/07/00, £130,646

### 4. Details of the impact (indicative maximum 750 words)

The above research, in particular the WMTSA book [5], has had widespread impact in a variety of sectors, as is detailed below.

There have been four **software** implementations of the highly cited (>2000 citations, [A]) WMTSA book that have had widespread use:

- (1) **TIBCO Spotfire S+ Wavelets**: This software was developed primarily by Bill Constantine at Insightful (owners of S+). It is based very heavily on WMTSA. This is acknowledged in [B]: "The new methodology implemented in S+Wavelets 2.0 stems almost entirely from Don Percival's book entitled, Wavelet Methods for Time Series Analysis, co-authored by Andrew Walden and published by Cambridge University Press in 2001" [B].

Subsequently the S+ software was sold to TIBCO [C] and v8.1 incorporated into their Spotfire system in 2008 (see under Powerful New Statistics) [D]. The press release announcing the release of TIBCO Spotfire S+® version 8.1 states "*Spotfire S+® is the only statistical programming platform that delivers a fully integrated development environment, a commercially supported analytic packaging system, and the ability to scale a desktop to manipulate gigabyte*

*class data sets. Spotfire S+ enables statisticians and business analysts to prototype, test, and deploy analytics much faster than with alternative statistical modeling environments. It delivers a wider range of robust statistics tools, as well as improved deployment and integration capabilities that help business analysts and researchers make informed and reliable decisions at critical points across the organization*” [D]. Specifically the Wavelets Package is described as providing “*advanced signal and image analysis, time series analysis, statistical signal estimation, and data compression analysis*” [E]. Spotfire customers span the Life Sciences, Financial Services, Energy, Government, Consumer Goods, Manufacturing and Telecommunications sectors [F]. Companies include GE, Chevron, PerkinElmer Inc., GlaxoSmithKline, Cisco, the BNP Paribas, Aberdeen Group, Salesforce, and BUCS Analytics. 65 US Government agencies including the CIA, NSA, Defense Advanced Research Projects Agency (DARPA) and the Defense Threat Reduction Agency (DTRA) use Spotfire S+ [F].

- (2) **WMTSA:** a MATLAB toolkit developed by Charlie Cornish (Department of Atmospheric Sciences, University of Washington). This is widely used in fMRI (functional Magnetic Resonance Imaging). For example, it is used in the processing pipeline by the Brain Mapping Unit at Cambridge [G] and its Neuroscience use is frequently acknowledged (e.g., three papers in 2010 and 2012 are given in [H]). It is recommended software (2010) used by the international company Disha Life Sciences Ltd for Metabolomics (regulation and fluxes in cells) [p. 88 of I]).
- (3) **Waveslim:** R package for “basic wavelet routines for time series (1D), image (2D) and array (3D) analysis” [J] developed by Brandon Whitcher. The waveslim code is based in part “*on wavelet methodology developed in Percival and Walden (2000)...*” [J]. This has been extensively used with applications ranging from a Reserve Bank of New Zealand document on measuring core inflation in New Zealand [K] to studies of copy number alterations in breast cancer led by Fred Hutchinson Cancer Research Centre [L].
- (4) **wmtsa:** a package developed by Bill Constantine and Don Percival in the R programming language [M]. As with the MATLAB distribution, this software has also found use in fMRI (e.g. [N]).

As a further specific example of the use of WMTSA in the biomedical sector, in an email Dr. Brandon Whitcher (now Senior R Consultant, Mango Solutions) confirms the use of methodology from WMTSA to analyse univariate and bivariate time series in support of early-phase drug development at GlaxoSmithKline during the period 2005-2009. The methodology was applied “*in the quantitative analysis of both functional and pharmacological MRI (magnetic resonance imaging) experiments in a variety of pre-clinical models for the neurology and psychiatry therapeutic areas. It facilitated the rapid and efficient processing of time series data, and produced easily interpretable results to the imaging scientists that supported our internal decision-making process.*” [O]

### **Monitoring of Santa Maria del Fiore Dome in Florence**

Based on the MODWPT algorithm introduced by Walden & Contreras Cristan, and included in WMTSA, Gabbanini, Vannucci et al [P] used wavelet packet variances to analyse crack widths in the famous dome of the Santa Maria del Fiore Cathedral in Florence. Their analysis revealed “*interesting aspects regarding the dynamics of crack evolutions and the structural functions of the different elements of the dome*”. The influence of this work is confirmed in the OPA Workshop on Monitoring of Great Historical Structures (Florence, January 2012) discussion “60 Years Results of the Monitoring System on Santa Maria del Fiore Dome in Florence” by Blasi & Ottoni which cites Gabbanini, Vannucci et al as one of the “essential references for Santa Maria del Fiore monitoring issue” [Q].

### **Ontario Ministry of Natural Resources**

Methodology from WMTSA was used in the modelling of tree taper of jack pine (*Pinus banksiana*) trees grown in the Canadian boreal forest region in a study carried out by the Ontario Ministry of Natural Resources [R]. A key point in their model fitting analysis is the use of reflecting boundary

## Impact case study (REF3b)

conditions: “*Following Percival and Walden (2000, p. 140)...*”The results of the study “*opens new possibilities for analysing longitudinal or taper data collected across time or space.*”

### Finance

Marco J. van der Burgt [S] of Atradius Group Risk Management (Amsterdam) used the MODWT algorithm and wavelet variance results as given in WMTSA to analyse monthly observed default rates to answer the question “*how long is a business cycle and where are we in the business cycle?*” This enabled the inclusion of business cycle effects in default probability validation.

### 5. Sources to corroborate the impact (indicative maximum of 10 references)

- [A] WMTSA citation count, <http://scholar.google.com/citations?user=Ki2lg9YAAAAJ&hl=en> (archived at <https://www.imperial.ac.uk/ref/webarchive/ktf>)
- [B] S+WAVELETS Version 2 information page, <http://www.msi.co.jp/splus/addon/wave2.html> (archived at <https://www.imperial.ac.uk/ref/webarchive/ltf> on 9/10/13)
- [C] TIBCO press release, 3/9/08, <http://www.tibco.com/company/news/releases/2008/press924.jsp> (archived at <https://www.imperial.ac.uk/ref/webarchive/ttf> on 9/10/13)
- [D] ‘TIBCO Reveals Industry’s Most Flexible Statistics-Driven Analytic Platform’ press release, 8/12/08, <http://www.tibco.com/company/news/releases/2008/press939.jsp> (archived at <https://www.imperial.ac.uk/ref/webarchive/mtf> on 9/10/13)
- [E] TIBCO Spotfire, ‘What’s new in v8.1’, <http://spotfire.tibco.com/~-/media/content-center/datasheets/whats-new-splus-8-1.ashx> (available [here](#))
- [F] Who uses Spotfire? webpage, <http://spotfire.tibco.com/en/discover-spotfire/who-uses-spotfire.aspx> (archived at <https://www.imperial.ac.uk/ref/webarchive/hwf> on 15/10/13) and Spotfire Case Studies webpage, <http://spotfire.tibco.com/en/resources/content-center.aspx?Content%20Type=Case%20Studies#content-center.aspx?Content%20Type=Case%20Studies%2C> (archived at <https://www.imperial.ac.uk/ref/webarchive/jwf> on 15/10/13)
- [G] Brain Mapping Unit webpage, University of Cambridge, <https://wiki.cam.ac.uk/bmuwiki/FMRI> (archived at <https://www.imperial.ac.uk/ref/webarchive/ntf> on 9/10/13)
- [H] Three neuroscience papers using WTMSA for fMRI analysis: [DOI\(1\)](#), [DOI\(2\)](#), [DOI\(3\)](#)
- [I] Disha Life Sciences presentation, [www.scribd.com/doc/27187414/In-Silico-Analysis-to-Metabolomics](http://www.scribd.com/doc/27187414/In-Silico-Analysis-to-Metabolomics) (pages 1 & 88 available [here](#))
- [J] ‘waveslim’ R software information page, <http://cran.r-project.org/web/packages/waveslim/index.html> (archived [here](#) on 9/10/13)
- [K] ‘Using wavelets to measure core inflation: the case of New Zealand’, Reserve Bank of New Zealand, May 2009, [http://www.rbnz.govt.nz/research\\_and\\_publications/discussion\\_papers/2009/dp09\\_05.pdf](http://www.rbnz.govt.nz/research_and_publications/discussion_papers/2009/dp09_05.pdf) (archived [here](#))
- [L] Breast cancer paper using Waveslim, May 2011: [DOI](#)
- [M] ‘wmtsa’ R software information page, <http://cran.r-project.org/web/packages/wmtsa/index.html> (archived at <https://www.imperial.ac.uk/ref/webarchive/stf> on 9/10/13)
- [N] Paper using WMTSA for fMRI analysis: [DOI](#)
- [O] Email from Senior R Consultant, Mango Solutions, formerly of GlaxoSmithKline, 17/10/12 (available from Imperial on request)
- [P] ‘Wavelet Packet Methods for the Analysis of Variance of Time Series With Application to Crack Widths on the Brunelleschi Dome’, [DOI](#)
- [Q] OPA Workshop abstract, (available [here](#))
- [R] ‘Applying wavelet-based functional approach in modelling tree taper’, 2/8/11, [DOI](#)
- [S] M. van der Burgt, ‘*Wavelet analysis of business cycles for validation of probability of default: what is the influence of the current credit crisis on model validation?*’, The Journal of Risk Model Validation, Vol 3, No 1, 3-22 (2009) (available [here](#))