Modelling and predicting flow regimes using wavelet representations of ERT data

Alex Goodwin*, Stuart Barber and Robert G. Aykroyd

University of Leeds

1 Tomography

Tomographic techniques use measurements taken outside or on the boundary of a region in an attempt to describe what is happening within the region. Such techniques are widely used in geophysical, industrial and medical investigations. In some industrial applications, multiple voltages are recorded between electrodes attached to the boundary of, for example, a pipe. The usual first step of the analysis is then to reconstruct the conductivity distribution within the pipe. Usually, such problems are ill posed because they have multiple solutions which do not depend continuously on the observed data. Stable solution then requires regularization (West *et al.* 2004). Even if reliable reconstruction is possible it only provides an image representing the conditions within the pipe. Although this is useful for process visualisation, automatic control of a process does not require such an image, which will require further post-processing to allow control parameters to be obtained.

There is a growing sentiment (Stitt and James, 2003) that in a field application simple questions such as “Is there, or is there not a problem?” are more important than obtaining high quality flow and phase patterns within the vessel. Therefore control parameter estimation, rather than process visualization, is the more appropriate output of a data analysis in many real situations. In particular, there is a need for sensing systems, models and algorithms that are simple, fast and can operate largely unsupervised, to detect problems.

2 Wavelets

A signal can be expressed as the sum of a series of sines and cosines, known as a Fourier expansion. The disadvantage of a Fourier expansion is that it has only frequency resolution and no time resolution. This means that although we might be able to determine all the frequencies present in a signal, we do not know when they are present. To overcome this problem wavelets can be used which are able to represent a signal in the time and frequency domain simultaneously. Wavelets are basis functions which can be used to approximate an underlying signal, in a similar way to Fourier transforms and since they are localised in frequency and time, can handle a wider range of signals than Fourier analysis. The disadvantage of wavelets is that the transform of a dataset of length $n = 2^J$ only has representations of the data at $J = \log_2(n)$ “resolution levels”, each resolution level having a representation at approximately twice the frequency of the previous level. We use the non-decimated wavelet representation of a zero mean time series $X_t \ : \ t = 0, \ldots, n-1$, which is given by

$$X_t = \sum_{j=0}^{J-1} \sum_{k=0}^{n-1} d_{jk} \psi_{jk}(t/n),$$

where ψ_{jk} is the wavelet at scale j and location k derived from a mother wavelet ψ via $\psi_{jk}(t) = 2^{1/2}\psi(2^j(t-k))$. The set of coefficients $\{d_{jk}\}$ is referred to as the wavelet transform of $\{X_t\}$.

The main use of wavelets in statistics is in wavelet shrinkage, a smoothing technique used to remove noise from a corrupted signal. This is done by discarding or modifying coefficients whose magnitude is smaller than a pre-specified threshold λ. The idea is that any noise contained in the signal will be exhibited as small coefficients and so, by removing them, some of the noise component can be isolated from the signal. There are many variations on how to choose this threshold; Vidakovic (1999) gives a discussion of some of these. In this paper, we use the universal threshold proposed by Donoho and Johnstone (1994), $\lambda_u = s\sqrt{2\ln(n)}$, where s is an estimate of the noise standard deviation, proposed by Donoho and Johnstone (1994).

3 Data simulation

Consider the flow of a gas through a liquid in a section of vertical pipe. The gas enters at the bottom of the section of pipe under pressure and travels rapidly up the length of the pipe. The gas fraction and bubble size are controlled by the inlet size and by the input pressure. To control process efficiency it is important to monitor the flow regime, and to adjust the input parameters accordingly. In the simulation study two types of flow will be considered— small, fast and frequent pockets of gas, known as “bubble” flow and large, slow and less frequent pockets of gas, known as “churn” flow.

In electrical tomography, for given conductivity distributions the boundary voltages are found using Maxwell’s equations, and appropriate boundary conditions. This is the forward or direct problem. The forward problem is solved numerically using the finite-element method (FEM). For examples of FEM-based approaches see West *et al.* (2004). The key part of the simulation is to generate spatial patterns for the bubbles that evolve through time, which will define the conductivity distributions.

The process is allowed to evolve for 256 time points. Once noise-free voltages are obtained uncorrelated Gaussian noise is added, representing measurement error, to yield the simulated dataset. Hence the full dataset forms 49 electrode-pair time series each of length 256.

4 Predicting flow regimes using wavelet representation

The model was used to predict flow regime in a set of independently generated flow patterns. This consisted of 168 time points of bubble flow, followed by 256 time points of churn flow, followed by 88 time points of bubble, with added Gaussian white noise with the same standard deviation as in the training dataset. The predictions generated by the model are probabilities of the time point being in bubble type flow. In this case the thresholded version is greatly superior, with the predictions being much more clearly separated into bubble and churn regimes. If this probability was greater than or equal to 0.5, the model was regarded as predicting bubble flow, otherwise the prediction was churn flow. The correct classification rates (i.e. the percentage of time points which are classified correctly) for the models built using both thresholded and non-thresholded activity measures are shown in Figure 1.

In our example we have a dichotomous response; at each time point the flow is either bubble (labelled 1) or churn (labelled 0). We use logistic regression to predict the probability that the flow is in the “bubble” state, using summary statistics of the wavelet representation of the electrode-pair time series as predictors.

For training the logistic regression model, a dataset of 256 time points of churn flow, followed by 256 time points of bubble flow was used. Let $Y_t = 0$ if the process is in bubble flow at time t and $Y_t = 1$ otherwise. We have 49 sensor time series $\{X_t^s : t = 0, \ldots, n-1\ ;\ s =$

$1, \ldots, 49\}$ and using the Haar basis, we obtain the non-decimated wavelet transform of each to obtain the coefficients $\{d_{jks} : j = 0, \ldots, J-1\ ;\ k = 0, \ldots, n-1\ ;\ s = 1, \ldots, 49\}$. These transforms are aggregated over sensors to produce a single "activity" value for each level/location pair, (j, k), which we define as $a_{jk} = \sum_s |d_{jks}|$. Relabelling the location index as t for clarity, we can consider the $\{a_{jk}\}$ as J time series $\{a_{jt} : t = 0, \ldots, n-1\}$ for levels $j = 0, \ldots, J$. We then model $\{Y_t\}$ in terms of the $\{a_{jt}\}$ by fitting the logistic regression model

$$\frac{Y_t}{1 - Y_t} = \alpha + \sum_{j=0}^{J-1} \beta_j a_{jt} + \varepsilon_t, \quad \varepsilon_t \text{ iid } N(0, \sigma^2)$$

To remove the effects of noise in the data, the coefficients $\{d_{jks}\}$ can be thresholded before aggregation; any coefficient which is smaller in magnitude than $\lambda_u = s\sqrt{2\ln(n)}$ is replaced by zero. We shall see below (Figure 1) that this radically improves the performance of our predictors in some cases.

We also consider an alternative approach based on linear discriminant analysis (see for example Manly, 2005), which aims to map objects into one of several groups by means of the recorded measurements. The activity measure used in the logistic regression modelling in the previous section resulted in a nine dimensional response for each time point, each number representing activity at a different frequency. The training dataset from above was decomposed using the Daubechies Least Asymmetric (8) wavelet to generate activity measures as described above. A linear discriminant rule was then trained on these activity measures; this rule was then used to classify each time point in the test dataset to either bubble or churn flow.

5 Results and discussion

Figure 1 shows the correct classification rates for each of the three methods (logistic regression with thresholding; logistic regression without thresholding; discriminant analysis), averaged over 100 replications. Also shown are point-wise 95% confidence intervals calculated using the standard errors from the replication.

For all noise standard deviations below 0.1, it can be seen that the logistic regression method outperforms the linear discriminant analysis, with the activity measures calculated after thresholding performing the best. Here, the performance of each method is approximately constant. The correct classification rate for the thresholded logistic regression was approximately 90% for all noise standard deviations below 0.1. A standard deviation of 0.1 corresponds to a signal to noise ratio of 1; i.e. the "strength" of the signal and noise are roughly equal. For larger noise standard deviations, the performance of all methods degrades, but the degradation is much swifter for the logistic regression method with thresholding.

Wavelets break down activity into a representation localised both in time (so we can analyse data at a fixed time point) and scale (so we can distinguish between high frequency bubble flow and low frequency churn flow). Statistical techniques of logistic regression and linear discriminant analysis have been used to classify flow at a given time into the "bubble" or "churn" regimes.

The linear discriminant analysis and the logistic regression without thresholding methods both show a slow but steady decrease in performance as the noise standard deviation increases beyond 0.1. In contrast, the logistic regression with thresholding decreases rapidly. This is due to the nature of the thresholding algorithm, which discards any wavelet coefficient whose magnitude is smaller than the universal threshold, $\lambda_u = s\sqrt{2\ln(n)}$. Thus, as the noise standard deviation increases, so does s causing the threshold to increase. Hence as the data becomes

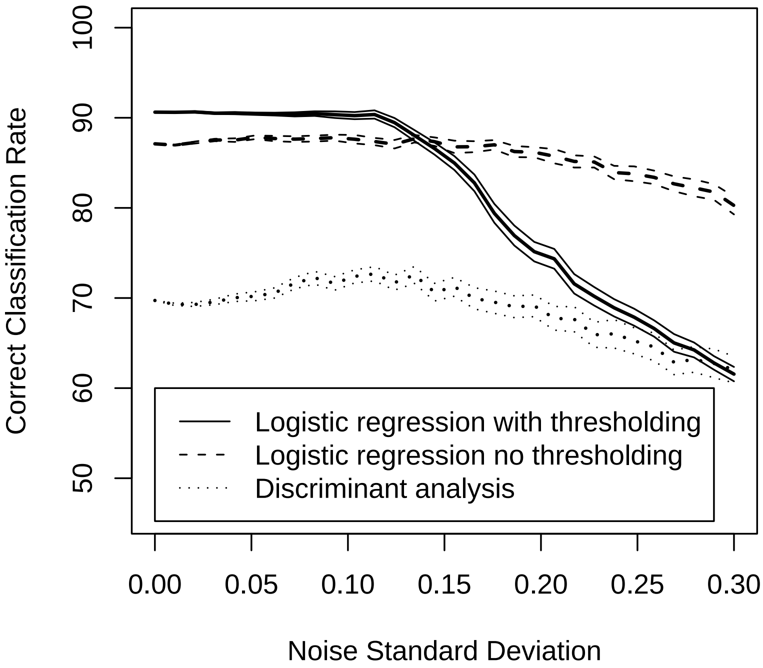

Figure 1: The correct classification rates for the different modelling methods, for a given noise standard deviation

noisier the thresholding step will eliminate much of the information in the signal along with the noise. In contrast, when no thresholding is done, the information is retained.

References

Donoho, D.L. and Johnstone, I.M., (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81, pp. 425-455.

Manly, B.F.G., (2005). *Multivariate Statistical Methods: A Primer*. Chapman & Hall/CRC, Boca Raton.

Nason, G.P., (1998). *WaveThresh3 Software*. Department of Mathematics, University of Bristol, Bristol, UK. http://www.stats.bris.ac.uk/∼wavethresh

Stitt, E.H. and James, K., (2003). Process Tomography and Particle Tracking: Research tool and commercial diagnostic tool. *Proc, 3rd World Congress on Industrial Process Tomography*, 2-10.

Vidakovic, B., (1999). *Statistical Modeling by Wavelets.* New York, John Wiley & Sons, Inc.

West, R. M., Aykroyd, R. G., Meng, S. and Williams, R. A., (2004). Markov Chain Monte Carlo Techniques and Spatial-Temporal Modelling for Medical EIT. *Physiological Measurements*, **25**, 181-194.