Project Draft:
The Wavelet project

H. Niederau, B. Nikolic
sponsored by: Dr. Niederau Consulting&Research

September 11, 2008

1 Project Draft

1.1 Aim

The aim of the proposed work is to investigate if wavelet decomposition of a time series of financial data, such as prices of an index of stocks provides a way of approximately predicting the future values of the time series.

More specifically, given a time series, x_i, representing the (possibly transformed) value of a security/index at points in time, the aim of this report is to investigate if a wavelet transform of data with indices $j-n \leq i < j$ provides any information about the likely value of x_j and possibly some further points in range $j \leq i < j+m$, where m is a few. In terms of explanation, n is termed the rolling window size, while m denotes the forecasting horizon.

1.2 Data to be analyzed

The analysis will be based on the daily adjusted close value of the Dow Jones Industrial average in the period 1960-present as supplied by the Yahoo Finance website (`http://finance.yahoo.com/q/hp?s=^DJI`). The analysis will be based on the historical log-return of the index, i.e., the transformation $x_i = \ln(d_i/d_{i-1})$ will be carried out where d_i are the adjusted

index closing values. Adjustment removes the undesired artifacts from splitting and dividends.

We will also consider making use of the trading volume information by weighting either the log-returns or the wavelet coefficients in a 'suitable' way. While it has not been agreed upon how such a weighting would be realized, consensus is that we will proceed with full analysis of these trading volume data only if there is indication up-front that it may substantially improve our findings.

1.3 Method

The task logically separates into two main parts: computation of the wavelet transforms, and testing their predictive power. The entire analysis described below will be carried out with two different wavelet families (the baseline choice are the Haar and Daubechies families) to investigate if this has a significant impact on the final results.

1.3.1 Wavelet transform

The first step will be to calculate the discrete wavelet transform on a rolling data window with a number of different window lengths. The result of this step of the analysis will be a series of coefficients which will be a of function time, i.e., the index j above[1].

We will also calculate modified wavelet transforms using the 'cycle-spinning' approach to obtain translational symmetry. The number of shifts used for the cycle-spinning will be 8 or 16. These transforms will be used in the extrapolation method described below and compared to the results of the normal transform. Up-front, we do not expect that these will make much of an improvement as the sliding window approach naturally eliminates the artifacts from the structure of the wavelets.

1.3.2 Testing the predictive power

Given wavelet coefficients as a function of time, the question that arises is: can we predict with any sort of confidence the future value of the security being modelled?

[1] More precisely the coefficients fill take the form $c_l(j)$, where l specifies the wavelet coefficient index, while j specifies the rolling window position in which the transform was carried out.

There are two approaches we currently see to the way this prediction could be done, and which will be examined during the project.

1. Some of the individual coefficients of the wavelet transform may correlate with *future* pattern of security price, e.g., a turn from a declining to generally increasing market.

2. The wavelet transform, together with suitable 'thresholding', can be used to remove non-fundamental trends in the time series data allowing extrapolation of the time series into the near future.

We propose to examine the item (1) above by making suitable charts that will allow visual inspection of the values of wavelet coefficients together with the actual security value time series.

The approach for item (2) will be to calculate the extrapolation from the wavelet de-noised time series and then calculate the frequency with which this extrapolation predicts price movement in the same direction as what actually happened. Thresholding will be selectable as either 'hard' or 'soft' types and the threshold level will be manually selected. Future coefficients will be calculated by linear regression of a variable number of coefficients at previous times. This extrapolation will be calculated for each point in the available data and the overall *hit-rate* calculated. Obviously, a hit-rate significantly above 50% indicates a potentially viable trading strategy.

2 Results by end of 2008

The results of this study are expected by the end of December 2008 and will contain the following components.

2.1 Report

The project report will include the following elements:

- a short section stating the goal of the study;
- a complete description of the analysis method, including justification for the approach taken;

- a description of each stage numerical processing;
- results and conclusions based on both of the approaches described above;
- a discussion of how the methods used in the analysis compare to the methods already described in the literature, highlighting the original parts of the analysis;
- suggestions for a further analyses;
- a reference list.

2.2 Slides

A set of presentation slides will be delivered containing primarily graphical illustrations of the results of the investigations. An on-line presentation is planned if the results of the study justify the effort to design it.

2.3 The source code

The analysis program will use the Gnu Scientific Library (`http://www.gnu.org/software/gsl/`) implementation of the wavelet transform. This means that the entire program and the source will be licensed under the terms of the Gnu Public Licence V3 (`http://www.gnu.org/copyleft/gpl.html`).