Intelligent Data Engineering and Automated Learning - IDEAL 2004
5th International Conference, Exeter, UK
Series: Lecture Notes in Computer Science , Vol. 3177

Wavelet Multi-scale Analysis of High Frequency FX Rates

Saif Ahmad

Department of Computing

University of Surrey, Guildford, UK

August 27, 2004

Talk Outline

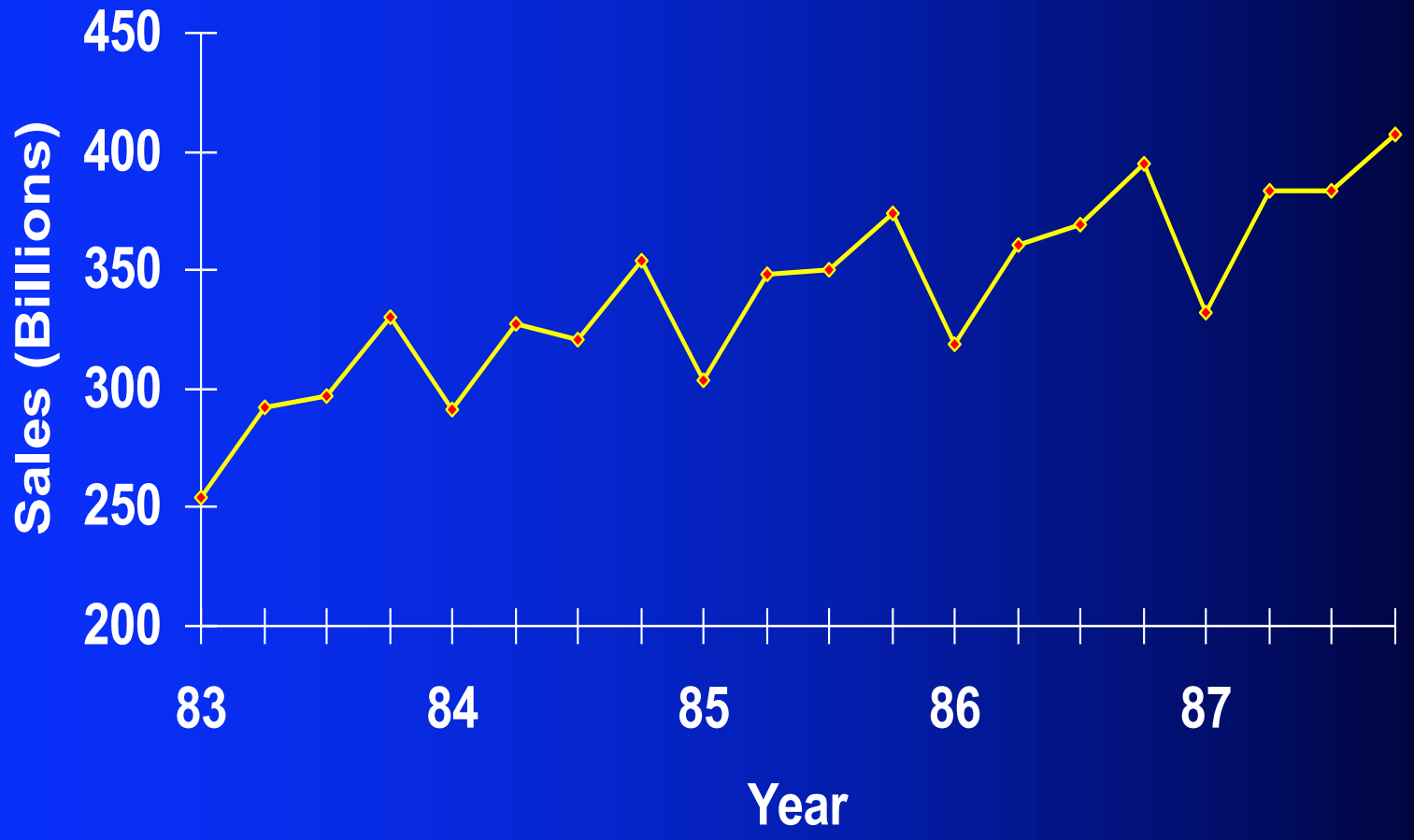
- Describing Time Series Data
- Financial Time Series Data Characteristics
- Wavelet Multiscale Analysis
- Our Time Series Analysis Approach
  - Algorithms
  - Prototype System
  - Case Study
  - Conclusions
- Questions

What Is a Time Series?

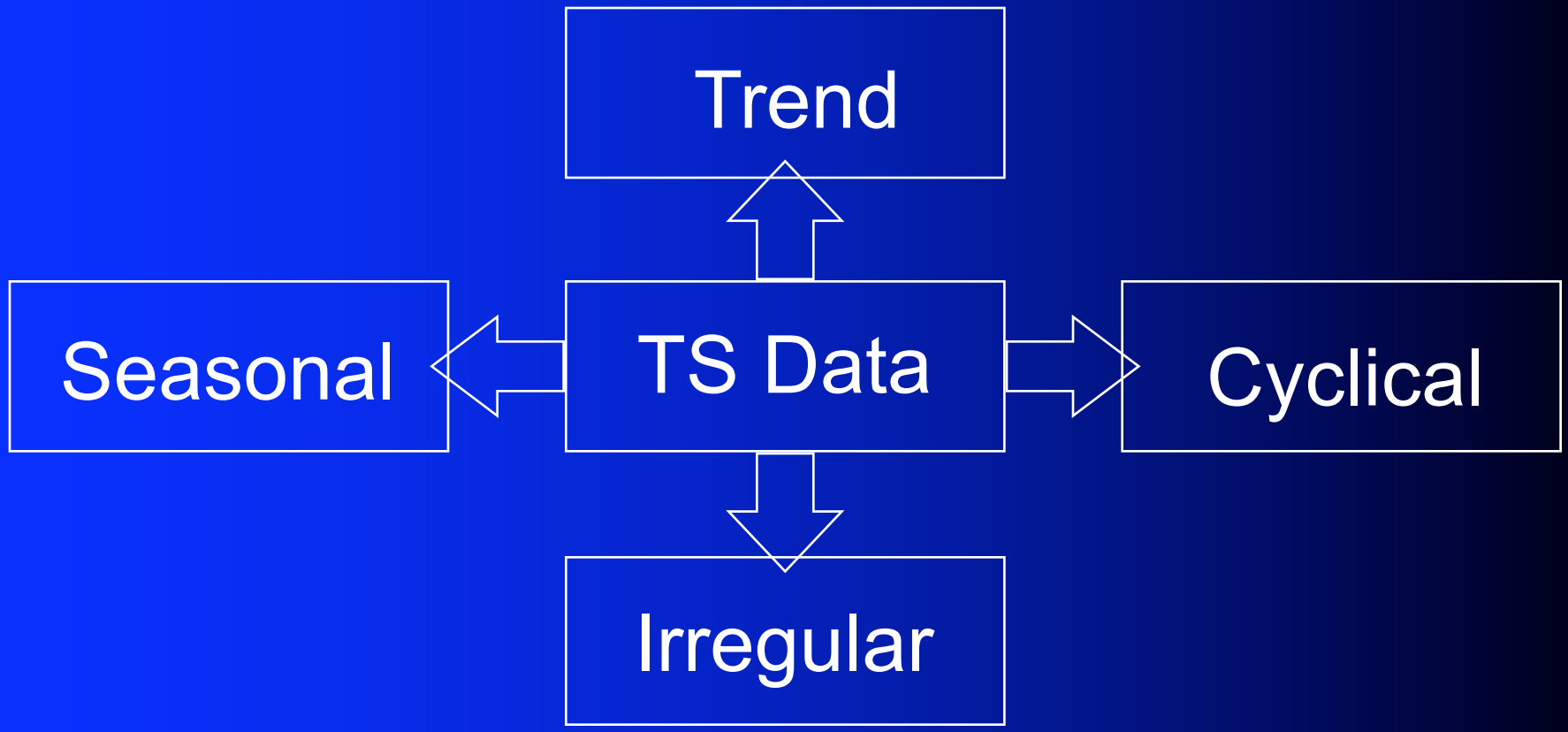
- A chronologically arranged sequence of data on a particular variable
- Obtained at regular time interval
- Assumes that factors influencing past and present will continue

U.S. Retail Sales

Quarterly Data



Time Series Components

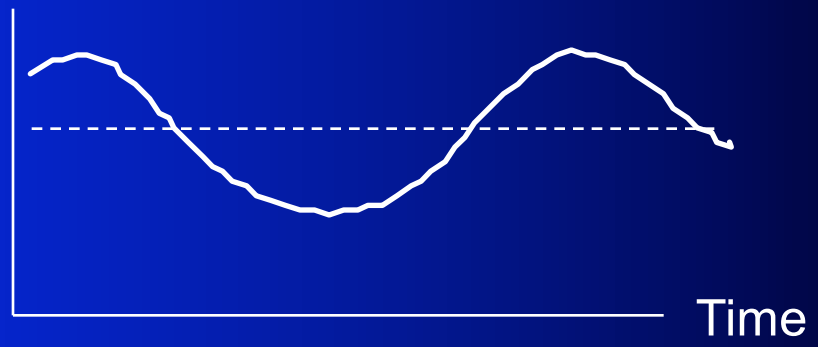


Trend Component

- Indicates the very long-term behavior of the time series
- Typically as a straight line or an exponential curve
- This is useful in seeing the overall picture

Cyclical Component

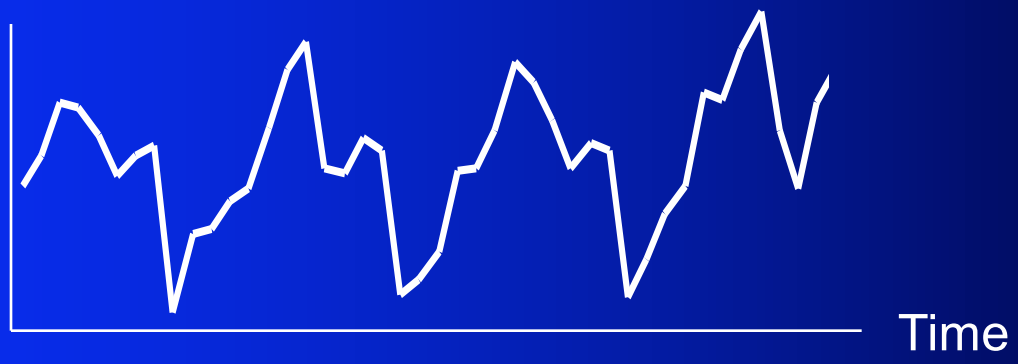
- A non-seasonal component which varies in a recognizable period
  - Peak
  - Contraction
  - Trough
  - Expansion



- Due to interactions of economic factors
- The cyclic variation is especially difficult to forecast beyond the immediate future → more of a local phenomenon

Seasonal Component

- Regular pattern of up and down fluctuations within a fixed time

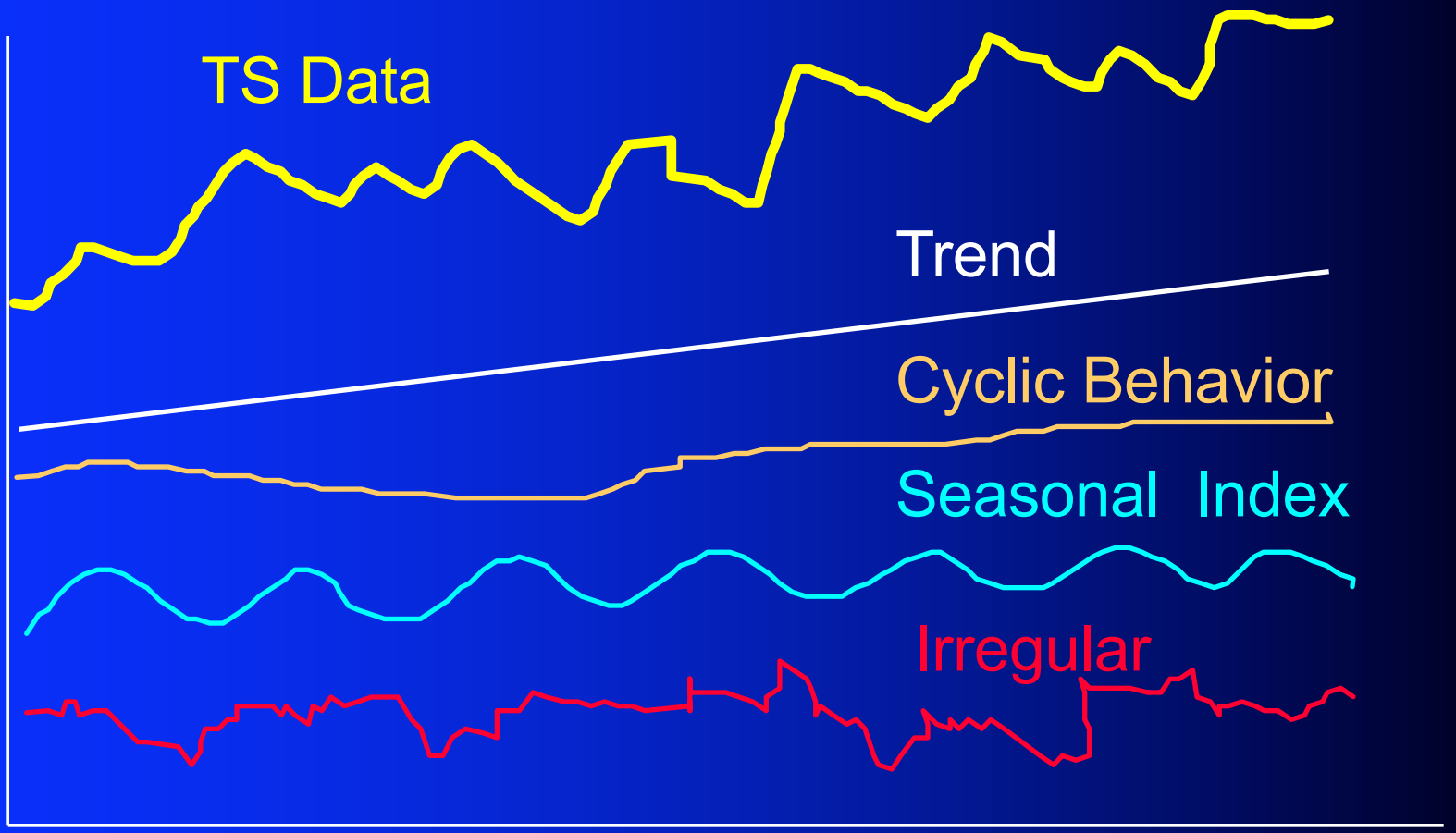


- Due to weather, customs etc.
- Periods of fluctuations more regular, hence more profitable for forecasting

Irregular Component

- Random, unsystematic, “residual” fluctuations
- Due to random variation or unforeseen events
- Short duration and non-repeating
- A forecast, even in the best situation, can be no closer (on average) than the typical size of the irregular variation

Time Series Data Broken-Down*



*For illustration purposes only.

Financial Time Series Data Characteristics

- Evolve in a **nonlinear** fashion over time
- Exhibit quite complicated patterns, like **trends**, **abrupt changes**, and **volatility clustering**, which appear, disappear, and re-appear over time → **nonstationary**
- There may be purely **local changes** in **time** domain, **global changes** in **frequency** domain, and there may be **changes** in the **variance parameters**

Financial Time Series Data Characteristics

Nonstationary

Time Varying Volatility

Financial Time Series Data Characteristics

Having said that…

- The nonlinearities and nostationarities do contain certain regularities or patterns
- Therefore, an analysis of nonlinear time series data would involve quantitatively capturing such regularities or patterns effectively

How and Why?

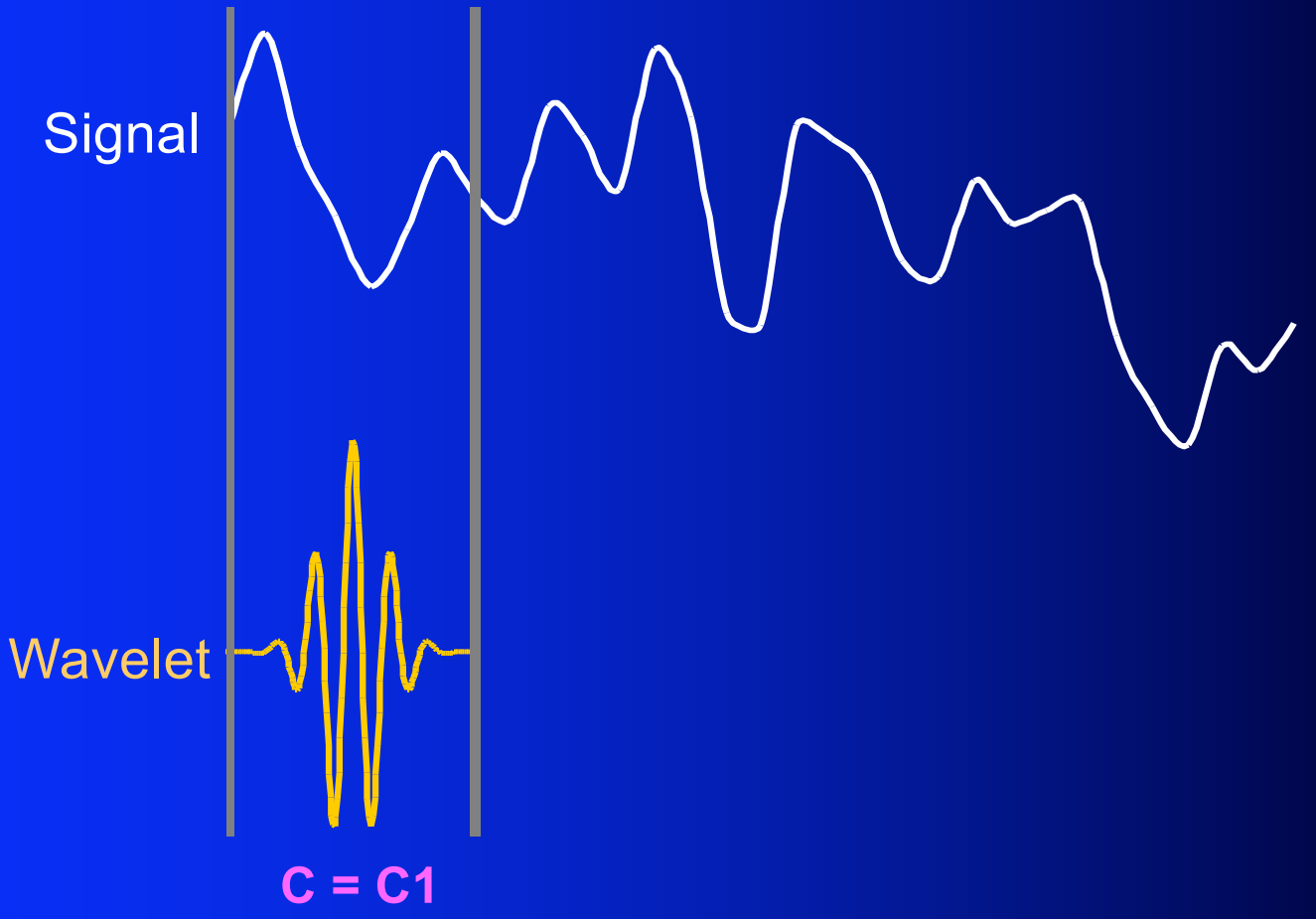
Wavelet Multiscale Analysis

Overview

- Wavelets are mathematical functions that cut up data into different frequency components and then study each component with a resolution matched to its scale
- Wavelets are treated as a ‘lens’ that enables the researcher to explore relationships that were previously unobservable
- Provides a unique decomposition (deconstruction) of a time series in ways that are potentially revealing

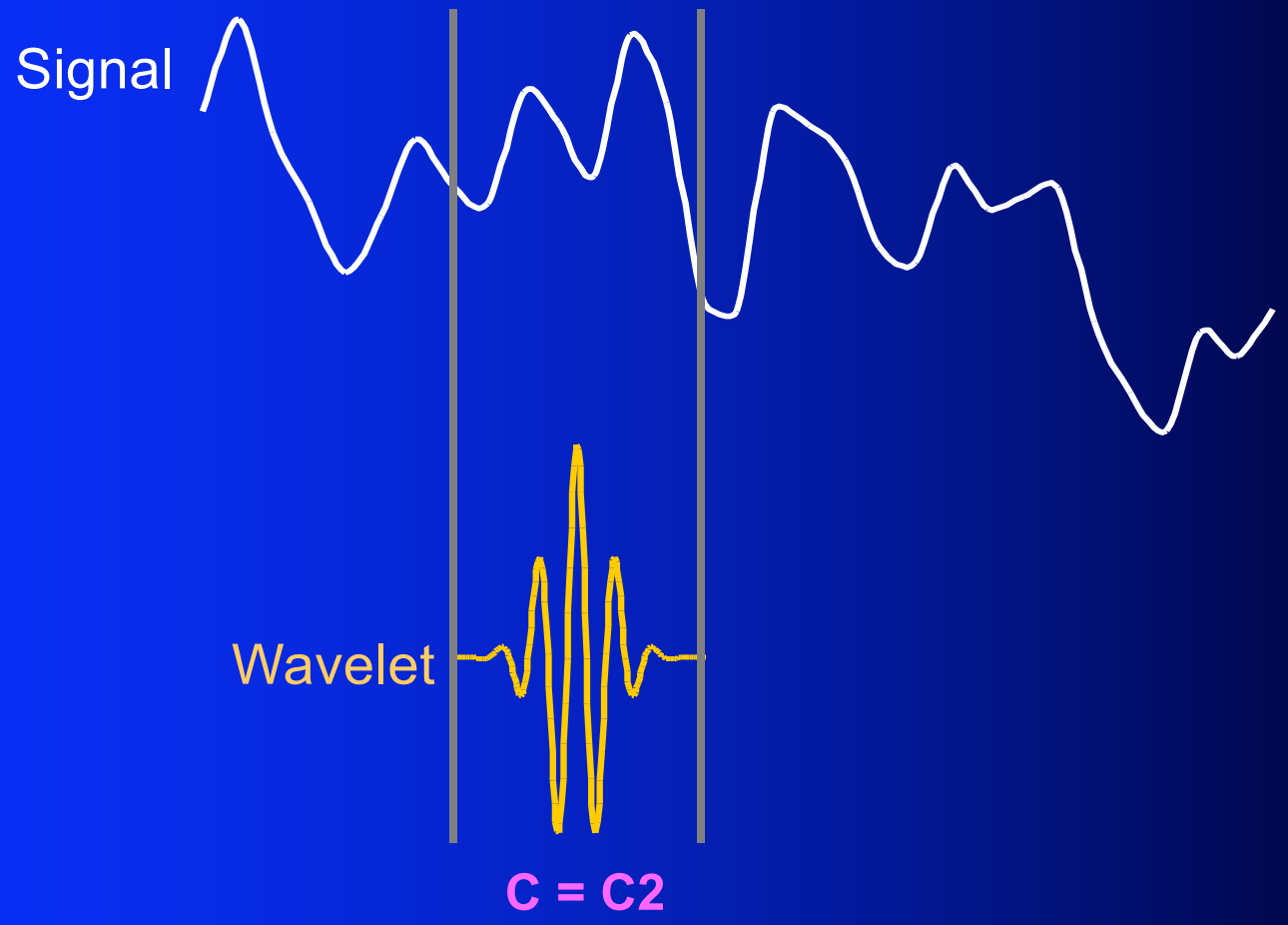
Wavelet Multiscale Analysis

Step I: Take a wavelet and compare it to a section at the start of the original signal. Calculate **C** to measure closeness (**correlation**) of wavelet with signal



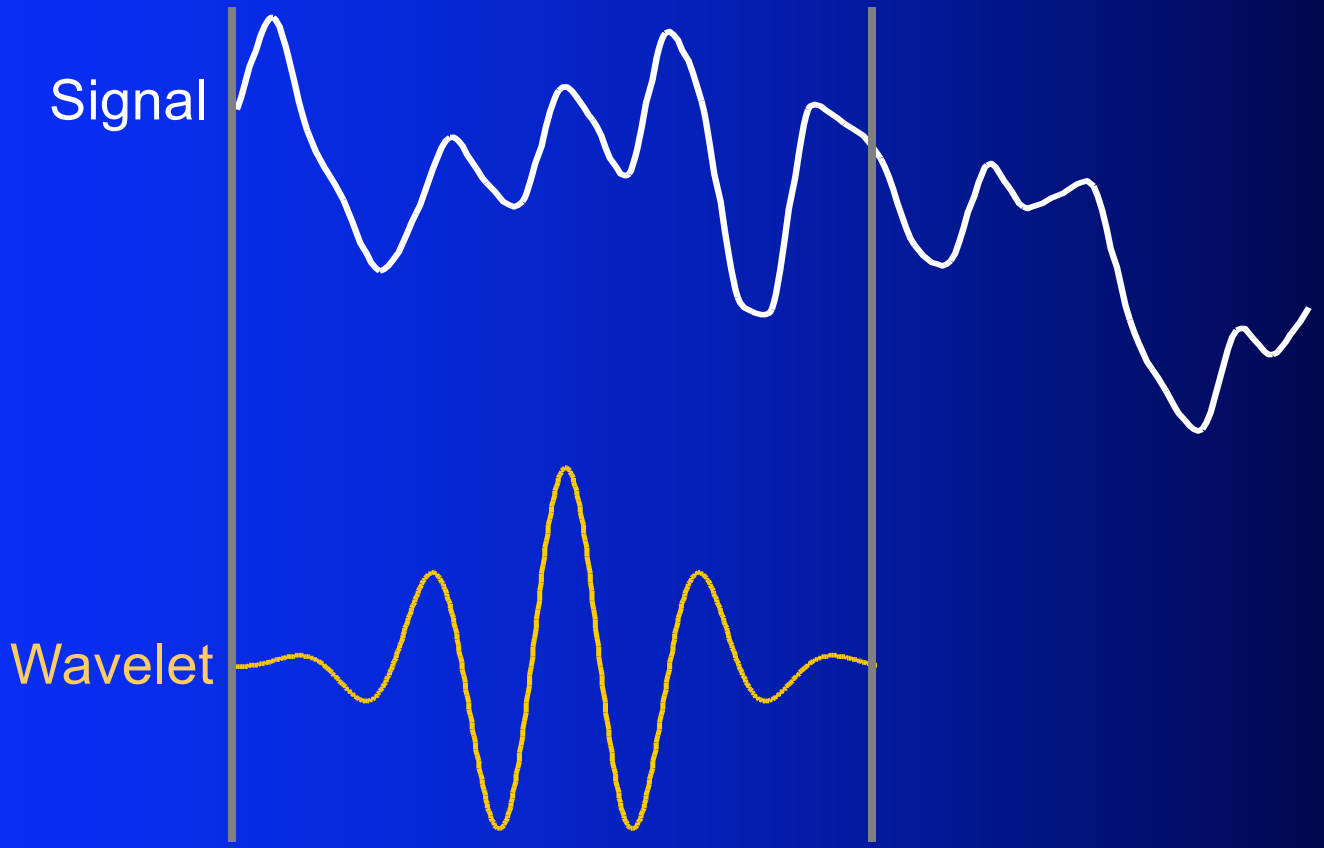
Wavelet Multiscale Analysis

Step II: Keep shifting the wavelet to the right and repeating Step I until whole signal is covered



Wavelet Multiscale Analysis

Step III: Scale (**stretch**) the wavelet and repeat Steps I & II



Step IV: Repeat Steps I to III for all scales

Wavelet Multiscale Analysis

Filter Bank Approach

Discrete Convolution: The original signal is convolved with a set of high or low pass filters corresponding to the prototype wavelet

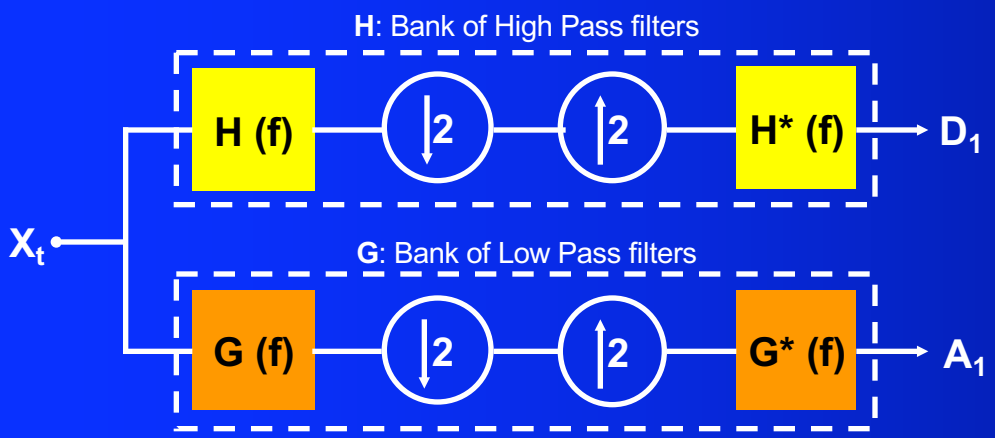
$$x[t] * \mathrm{W}[\mathrm{t}] = \sum_{k=-\infty}^{\infty} x[k] \bullet W[t-k]$$

X_t → Original Signal

W → High or low pass filters

Wavelet Multiscale Analysis

Filter Bank Approach



H (f) – high-pass decomposition filter

H* (f) – high-pass reconstruction filter

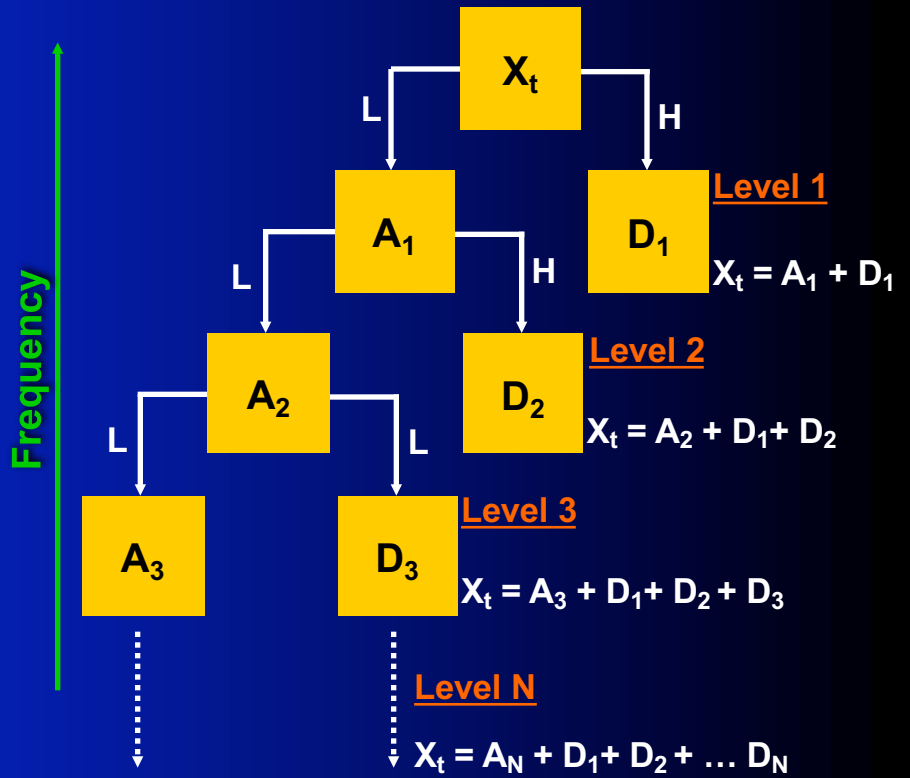
G (f) – low-pass decomposition filter

G* (f) – low-pass reconstruction filter

Up arrow with 2 – upsampling by 2

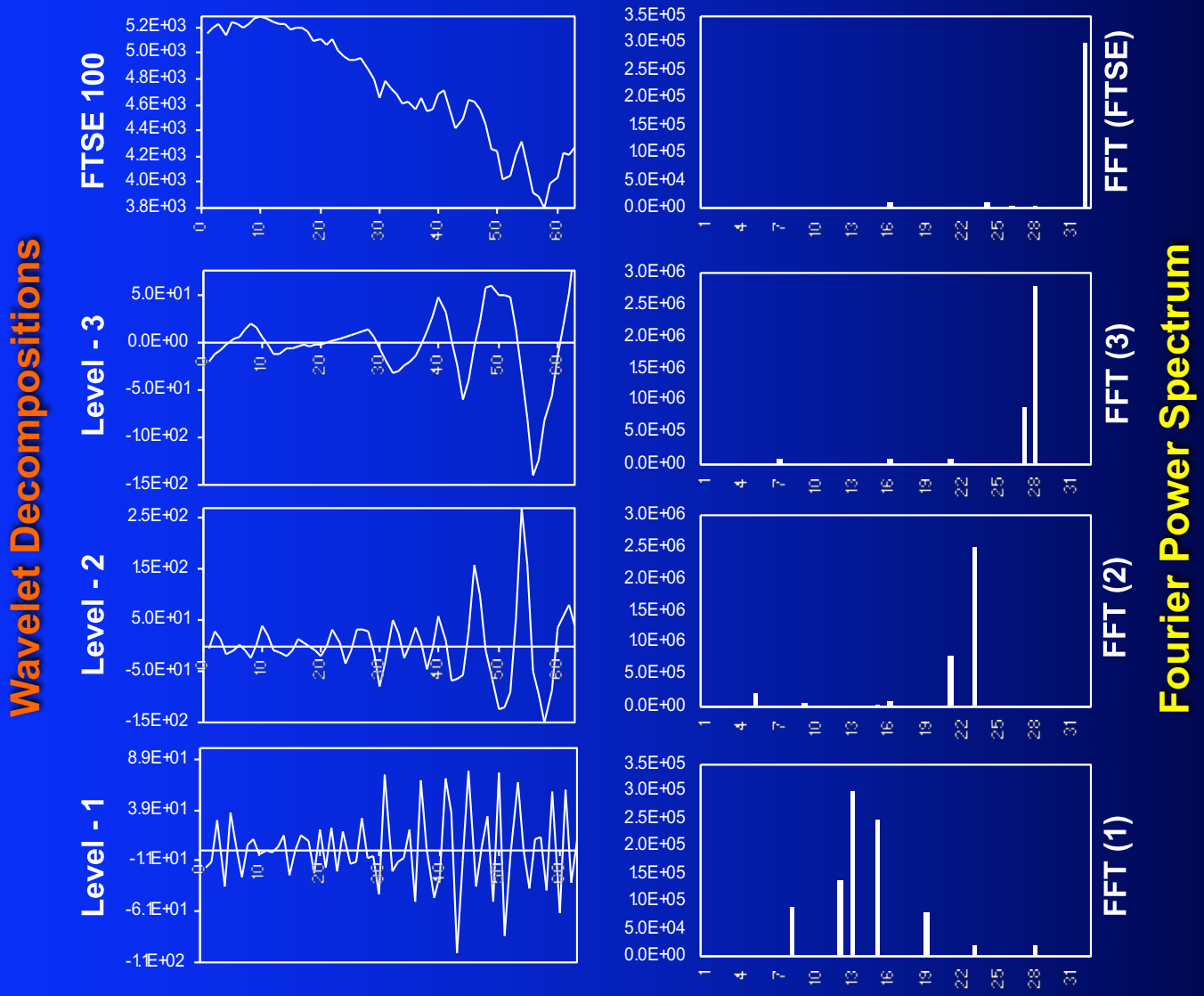
Down arrow with 2 – downsampling by 2

Iteration gives scaling effect at each level

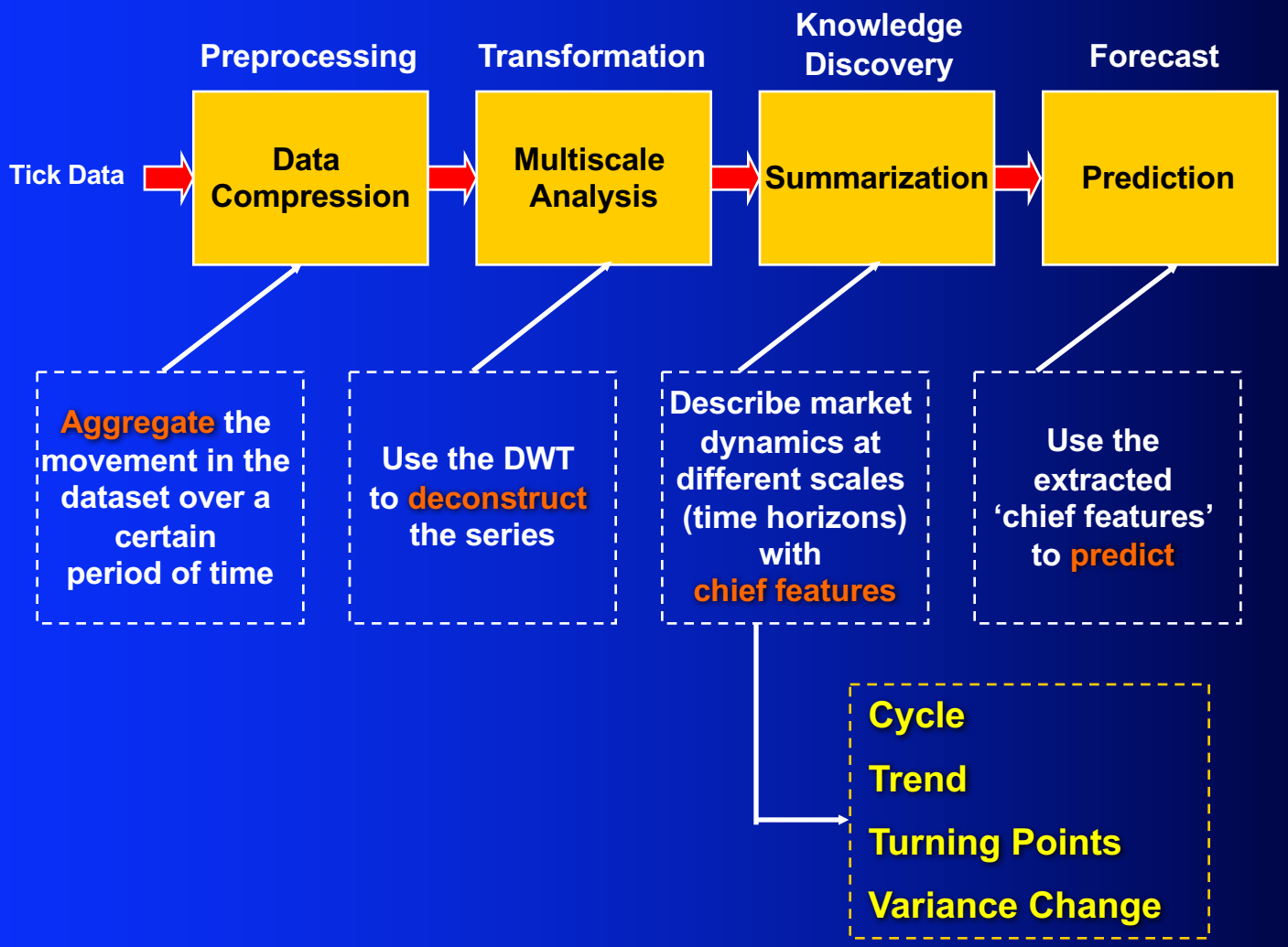


Mallat's Pyramidal Filtering Approach

Wavelet Multiscale Analysis



Analyzing High-frequency Financial Data: Our Approach



Analyzing High-frequency Financial Data: Our Approach

Generalized Algorithm: Summarization

I. **Compress** the tick data to get Open (O), High (H), Low (L) and Close (C) value for a given compression period (for example, one minute or five minutes).

II. Calculate the level L of the DWT needed based on number of samples N in C of Step I,

L = floor [log (N)/log (2)].

III. Perform a **level-L DWT** on **C** based on results of Step I and Step II to get, $\mathbf{D}_i$, i = 1, . . ., L, and $\mathbf{A}_L$.

III-1. Compute **trend** by performing linear regression on $\mathbf{A}_L$.

III-2. Extract **cycle** (seasonality) by performing a Fourier power spectrum analysis on each $\mathbf{D}_i$ and choosing the $\mathbf{D}_i$ with maximum power as $\mathbf{D}_S$.

III-3. Extract **turning points** by choosing extremas of each $\mathbf{D}_i$.

IV. Locate a single **variance change** in the series by using the NCSS index on **C**.

V. Generate a graphical and verbal **summary** for results of Steps III-1 to III-3 and IV.

Analyzing High-frequency Financial Data: Our Approach

Generalized Algorithm: Prediction

I. **Summarize** the tick data using the time series summarization algorithm.

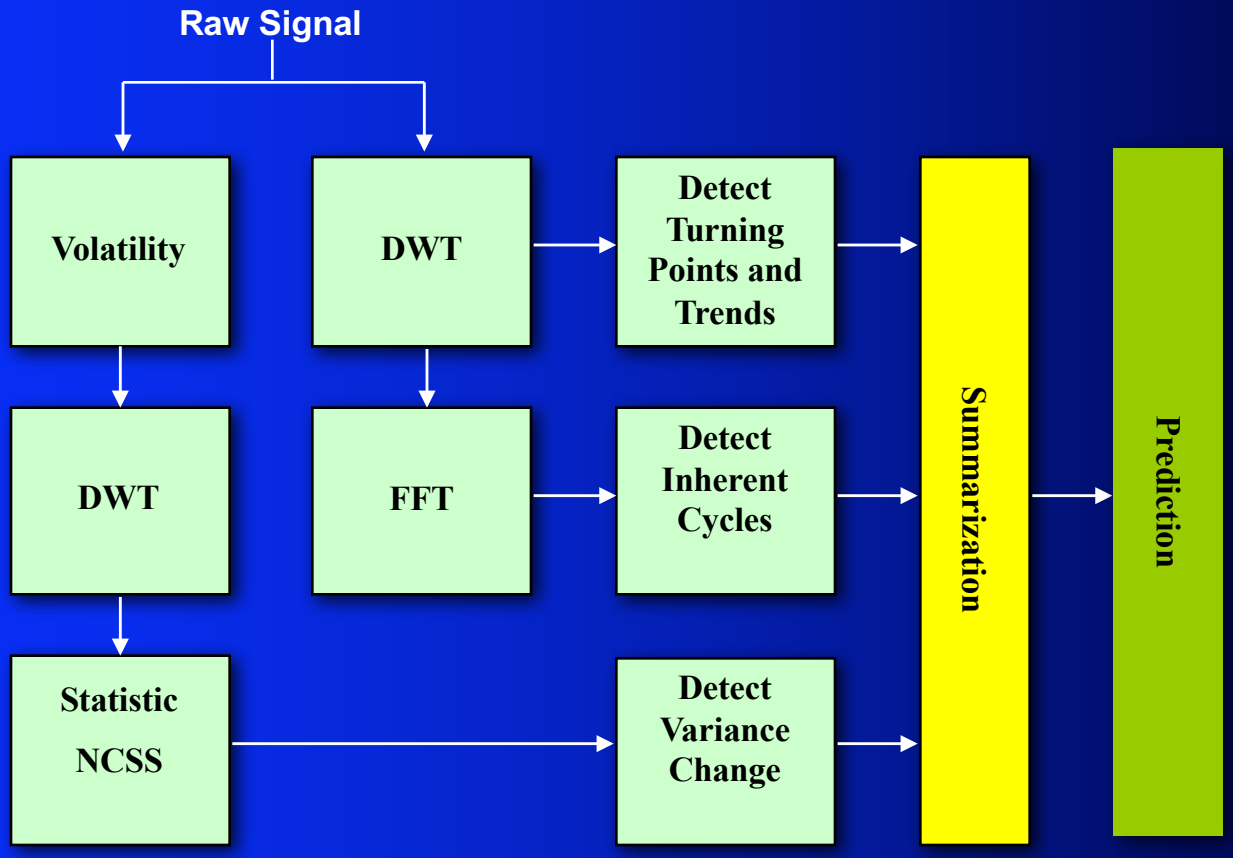
II. For a N-step ahead **forecast**, **extend the seasonal** component $\mathbf{D_S}$ *symmetrically* N points to the right to get $\mathbf{D_{S,\ forecast}}$.

III. For a N-step ahead **forecast**, **extend the trend component** $\mathbf{A_N}$ *linearly* N points to the right to get $\mathbf{A_{N,\ forecast}}$.

IV. Add the results of Steps II and III to get an **aggregate** N-step ahead **forecast**, $\boldsymbol{Forecast = D_{S,\ forecast} + A_{N,\ forecast}}$.

Analyzing High-frequency Financial Data: Our Approach

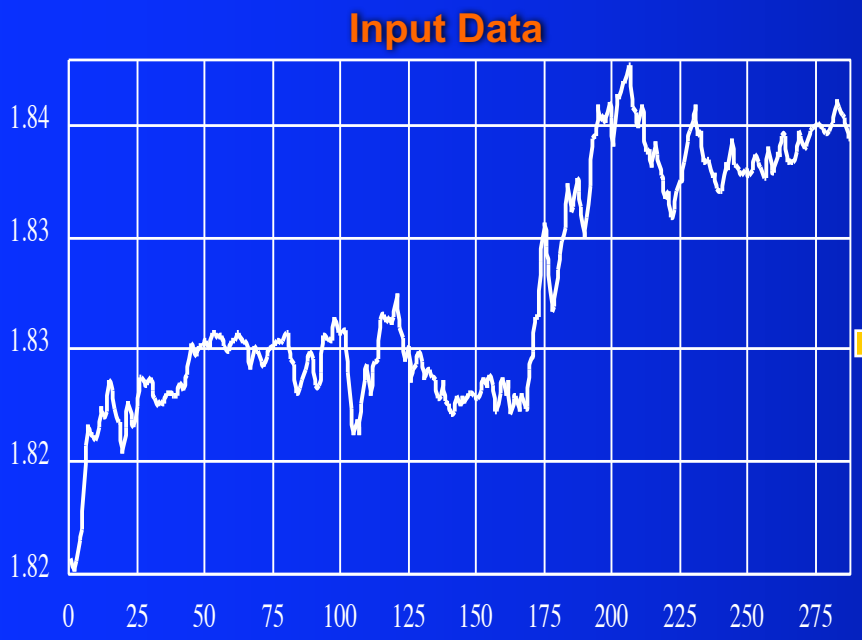
A **prototype system** has been implemented that automatically extracts "chief features" from a time series and give a prediction based on the extracted features, namely trend and seasonality

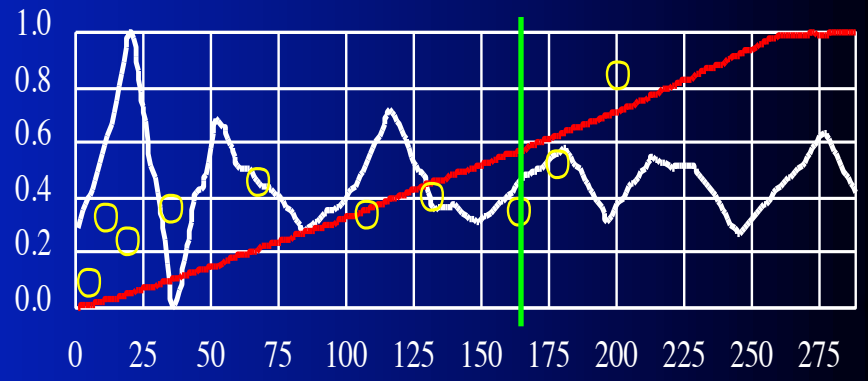


Analyzing High-frequency Financial Data: Our Approach

A Case Study

Consider the five minutes compressed tick data for the £/$ exchange rate on March 18, 2004





Feature	Phrases	Details
Trend	*1st Phase*	$x_1^{Trend} = 6.36e\text{-}5\,t + 1.81, t < 260$
	2nd Phase	$x_2^{Trend} = 3.65e\text{-}6\,t + 1.83, 261 < t < 288$
Turning Points	*Downturns*	108, 132, 164, and 178
	Upturns	5, 12, 20 36, 68, and 201
Variance Change	*Location*	164
Cycle	*Period*	42
	Peaks at	21, 54, 117, 181, 215, and 278

Analyzing High-frequency Financial Data: Our Approach

A Case Study

For **prediction**, we use the 'chief features' of the previous day (March 18, 2004), information about the dominant cycle and trend (summarization), to reproduce the elements of the series for the following day (March 19, 2004):

Root Means Square Error = 0.0000381

Correlation = + 62.4 %

Analyzing High-frequency Financial Data: Our Approach

Conclusions

I. We have **presented** a time series summarization, annotation, and prediction framework based on the multiscale wavelet analysis to deal with nonstationary, volatile and high frequency financial data

II. **Multiscale analysis** can effectively **deconstruct** the total series into its constituent time scales: specific forecasting techniques can be applied to each timescale series to gain **efficiency in forecast**

III. **Results** of experiments performed on Intraday exchange data **show promise** for summarizing and predicting highly volatile time series

IV. **Continuously evolving** and randomly shocked **economic systems** demand for a **more rigorous** and extended **analysis**, which is being planned

V. Successful analysis of **agents operating** on **several scales** simultaneously and of **modeling these components** could result in more **exact forecasts**

Questions / Comments