“Modeling Energy Prices using Neural Network and Spectral Analysis”

M.Panella¹, F.Barcellona², R.L. D’Ecclesia²
massimo.panella@uniroma1.it
francesco.barcellona@uniroma1.it

¹Dpt. DIET - Dipartimento di Ingegneria dell’Informazione, Elettronica e Telecomunicazioni
²Dpt. DAES - Dipartimento di Scienze Sociali
Outline

- Aim
- Some literature
- The dataset
- Tests
- Some results
- Conclusion and future developments
Motivation

- To provide a powerful tool to replicate economic and financial series dynamics
- To apply the filter bank for subband decomposition to describe energy prices dynamics
- To generate simulated prices using a neural model trained on each subband according to specific training and prediction techniques
- To backtest the results
Some Recent Literature:


The idea

Key points

- Use of NN to provide reliable forecasts for the price dynamics.
- Find interactions between pairs of variables in short and long term.
- Sort by relevance of input variables
- Find drivers of prices volatility among the chosen key variables (heterogeneity respect to equilibrium condition)
- Use of MSE to evaluate training and testing phases

- Signal to noise ratio to assess the goodness of fit
- Comparison between the simulated series and the observed ones
- Statistical analysis of the simulated series
The Spectral Analysis

- *Filter banks* represent a signal processing tool limiting the spectral distribution of time series to be predicted in every subchannel. They are adopted in association with neural networks to predict the subsequences regardless of the specific filter bank implementation.

- FB Generate maximally decimated subsequences, each subsampled (sample reduction) by a factor $M$ that is equal to the number of channels;

- Maintain the number of samples of each sequence but translating (modulation) in the baseband its spectrum.
The decomposition

The decomposition in $M$ subbands of a scalar sequence $s(n)$, on the sampled time, $n$, is performed in the AFB; each subband is associated with a scalar sequence $s(i) a(n)$, $i = 0, \ldots, (M-1)$. The AFB’s output can be arranged in a column vector

$$sa(n): sa(n) = s(0) a(n) s(1) a(n) \ldots s(M-1) a(n)t$$
The procedure

Given the complexity of the daily price series, the prediction model in each subchannel must be a nonlinear function inferred by a regularized learning paradigm and estimated through a given set of observed samples.

We consider well-known neural network models: the Radial Basis Function (RBF) neural network and the Adaptive Neuro Fuzzy Inference System (ANFIS),

A comparison is also performed with respect to a linear model whose parameters are estimated by a least-squares (LSE) technique. These models can be re-trained at every sample, to follow the intrinsic non stationarity or seasonality
Radial Basis Function Neural Network

A hidden layer of radial kernels:
Performs a non-linear transformation of input space. The resulting hidden space is typically of higher dimensionality than the input space.

An output layer of linear neurons:
The output layer performs linear regression to predict the desired targets.

\[ y = \sum_{j=1}^{M} w_j F_j(x) \]

\( F_j(x) \): nonlinear (vector) function
Mixture of Gaussian (MoG) Neural Network

Parameters related to the \( j \)-th Gaussian component \( p(x, y \mid j) \) in the joint input-output space:

\[
\begin{bmatrix}
K_{xx, j} & K_{xy, j} \\
K_{yx, j} & K_{yy, j}
\end{bmatrix} = \text{cov}[G^{(j)}_{y \mid x}]
\]
\[G^{(j)}_{y \mid x} \equiv p(y \mid x, j)\]

\[
m_{y \mid x, j} = E[G^{(j)}_{y \mid x}] = m_{y, j} + K_{yx, j} K_{xx, j}^{-1} (x - m_{x, j})
\]

Least–Squares Approximation (LSE):

\[
y = \sum_{j=1}^{M} h_j(x) m_{y \mid x, j}
\]
**ANFIS (Adaptive Neuro-Fuzzy Inference System) Network Architecture**

- **Sugeno 1st-order type rules:** IF $x_1$ is $B_1^{(k)}$, ..., AND $x_n$ is $B_n^{(k)}$ THEN $y^{(k)} = \sum_{j=1}^{n} a_j^{(k)} x_j + a_0^{(k)}$

- **Approximate reasoning:**
  - **Fuzzification** of crisp inputs: $\underline{x} = [x_1 \ x_2 \ ... \ x_n]$;
  - **Membership functions**: $\mu^{(k)}(x) \Rightarrow$ reliability of linguistic (fuzzy) rules
  - Soft (fuzzy) decision: $y(x) = \sum_{k=1}^{M} \alpha^{(k)}(x)y^{(k)}(x)$
Data set

EUR:
- Brent
- European Energy Exchange (EEX)
- National Balancing Point (NBP)

US:
- West Texas Intermediate (WTI)
- PJM Interconnection (Eastern Interconnection)
- Henry Hub (HH)

Input-output series:
- Input series (training phase): 8 two-year period from 2001-2002 to 2008-2011;
- Output series (testing phase): 8 years from 2003 to 2012 (each follows the previous two years)

Unique series for both phases (training and testing), described by data of N-year period (from 2001-2002-2003 to 2010-2011-2012).
The tests

We have approximately 250 prices per year;

The predictors are trained on 500 samples (about two years) and tested on the successive 250 samples (about one year).

- 2001-2002 (training) and 2003 (test);
- 2004-2005 (training) and 2006 (test);
- 2008-2009 (training) and 2010 (test).

The performance is measured by the Normalized Mean Squared Error (NMSE)

The subband prediction using maximally decimated filter banks allows an average NMSE reduction of about 5% for almost all the sequences under investigation and all the adopted prediction models: linear, RBF and ANFIS.
American electricity output (PJM) of 2007-2008. Starting sample of 150 elements to forecast the next 50.
American electricity output (PJM) of 2007-2008. Starting sample of 200 elements to forecast the next 100.
The training routines are:

- **LSE** (*Least Squares Estimation*)
- **RBF** (*Radial Basis Function*)
- **MOG**: developed by the Dept. DIET (gaussian model)
- **ANFIS2**: matlab plug-in
Three different criteria, based on a comparison of some series characteristics, are used:

- Reproduce the same characteristics of trend, seasonality, volatility and jumps
- Easy calculation and numerical implementation
- **Match of statistical properties of the two series**

Summarized in the four moments:

- **Mean**
- **Std Dv**
- **Skewness**
- **Kurtosis**

Features of the series probability distribution

RELIABLE FORECAST
Results analysis 2/7

2003 forecast for (Brent) using MOG predictor (alpha = 0.9, lambda = 0.5)

<table>
<thead>
<tr>
<th></th>
<th>Actual</th>
<th>Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>2.7645</td>
<td>2.7663</td>
</tr>
<tr>
<td>STD DV</td>
<td>0.0093</td>
<td>0.0091</td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>0.5547</td>
<td>0.1204</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>2.4811</td>
<td>2.2023</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.0016</td>
</tr>
<tr>
<td>NMSE</td>
<td>0.1715</td>
</tr>
<tr>
<td>SNR (dB)</td>
<td>36.8176</td>
</tr>
<tr>
<td>MAPE (%)</td>
<td>1.0926</td>
</tr>
</tbody>
</table>

Venice, September 10th 2012
2003 forecast for (Brent) using ANFIS2 predictor (T_OPT[1]=50)
2003 forecast for American oil (WTI) using MOG predictor (alfa=0.9, lambda=0.5)
Results analysis 5/7

2003 forecast for American oil (WTI) using ANFIS2 predictor (T_OPT[1]=50)

Venice, September 10th 2012

WTI 2003 Set Prediction

<table>
<thead>
<tr>
<th></th>
<th>ACTUAL</th>
<th>FORECAST</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>2.9718</td>
<td>2.9669</td>
</tr>
<tr>
<td>STD DV</td>
<td>0.0069</td>
<td>0.0067</td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>0.2596</td>
<td>0.2538</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>3.1255</td>
<td>3.1028</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.0006</td>
</tr>
<tr>
<td>NMSE</td>
<td>0.0943</td>
</tr>
<tr>
<td>SNR(dB)</td>
<td>41.3292</td>
</tr>
<tr>
<td>MAPE(%)</td>
<td>0.6462</td>
</tr>
</tbody>
</table>
2003 forecast for Brent using linear model (LSE) whose statistical analysis returned a perfect coincidence of all 4 moments.

**Actual**      | **Forecast**
---|---
**MEAN** | 2.7645 | 2.7648 |
**STD DV** | 0.0093 | 0.0092 |
**SKEWNESS** | 0.5547 | 0.5446 |
**KURTOSIS** | 2.4811 | 2.4653 |

**MSE** | 0.0005 |
**NMSE** | 0.0519 |
**SNR(dB)** | 42.0041 |
**MAPE(%)** | 0.6023 |

Perfect match
2003 forecast for **Brent** using **RBF** model (alfa=0.5), whose statistical analysis return a good match only for first and second moments.
Conclusions and Future developments

Prediction accuracy of NN on training and testing for Brent, EEX, NBP, WTI, PJM and HH.

Structuring data:
- Different energy assets
- Reference to current (actuality)
- Extend forecast time range

Reduction of computational cost

Correlation between NN and moments up to fourth order. This introduces new forecasting assumptions for energy commodities

Modeling:
- Different forecasting methods:
  - Embedded techniques
  - Prediction in subbands (serial decomposition)
- Other financial tools

Venice, September 10th 2012