Alexandridis, Antonis and Zapranis, Achilleas (2011) *Wind Derivatives: Modeling and Pricing.* In: 1st International Conference of the Financial Engineering and Banking Society (F.E.B.S), 10-12 June 2011, Chania, Greece.

Presented in 1st International Conference of Financial Engineering and Banking Society (FEBS). Chania, Greece (2011)

Wind Derivatives: Modeling and Pricing

A. Alexandridis[1], A. Zapranis[1],

[1]*Department of Accounting and Finance, University of Macedonia of Economics and Social Studies, 156 Egnatia St., P.O. 54006, Thessaloniki, Greece.*

aalex@uom.gr, zapranis@uom.gr

Abstract. Wind is considered to be a free, renewable and environmentally friendly source of energy. However, wind farms are exposed to excessive weather risk since the power production depends on the wind speed and the wind direction. This risk can be successfully hedged using a financial instrument called weather derivatives. In this study the dynamics of the wind generating process are modeled using a non-parametric non-linear wavelet network. Our model is validated in New York. The proposed methodology is compared against alternative methods, proposed in prior studies. We find that wavelet networks can model the wind process very well and consequently they constitute an accurate and efficient tool for wind derivatives pricing. Finally, we provide the pricing equations for wind futures.

Keywords: *Wind Derivatives, Weather Derivatives, Pricing, Forecasting, Wavelet Networks*

1. Introduction

Weather derivatives are financial tools that can help organizations or individuals to reduce risk associated with adverse or unexpected weather conditions. Weather derivatives can be used as part of a risk management strategy. Weather derivatives linked to various weather indices, such as rainfall, temperature or wind, are extensively traded in Chicago Mercantile Exchange (CME) as well as on Over-The-Counter (OTC) market. According to (Challis 1999; Hanley 1999) nearly $1 trillion of the US economy is directly exposed to weather risk. It is estimated that nearly 30% of the US economy and 70% of the US companies are affected by weather, (CME 2005). The electricity sector is especially sensitive to the temperature and wind since. Hence, it is logical that energy companies are the main investors of the weather market. In (Zapranis and Alexandridis 2008, 2009; Alexandridis 2010; Zapranis and Alexandridis 2010) a detailed framework for modeling and pricing temperature derivatives was developed. In this study we focus on wind derivatives.

The notional value of the traded wind-linked securities is around $36 million indicating a large and growing market, (WRMA 2010). Wind is free, renewable and environmentally friendly source of energy, (Billinton et al. 1996). While the demand for electricity is closely related to the temperature, the electricity produced by a wind farm is dependent on the wind conditions. The risk exposure of the wind farm depends on the wind speed and the wind direction, (F. E. Benth and Saltyte-Benth 2009). However, modern wind turbines include mechanisms that allow turbines to rotate on in the appropriate wind direction, (Caporin and Pres 2010). Hence, the risk exposure of a wind farm can be analyzed by quantifying only the wind speed.

Many different approaches have been proposed so far for modeling the dynamics of the wind speed process. The most common is the generalized autoregressive moving average (ARMA) approach. There have been a number of studies on the use of linear ARMA models to simulate and forecast wind speed in various locations (Daniel and Chen 1991; Caporin and Pres 2010; Huang and Chalabi 1995; Torres et al. 2005; Billinton et al. 1996; Tol 1997; Kamal and Jafri 1997; Martin et al. 1999; Castino et al. 1998; J. r. a. Benth and Benth 2010). In (Kavasseri and Seetharaman 2009) a more sophisticated fractional integrated ARMA (ARFIMA) model was used. Most of these studies did not consider in detail the accuracy of the wind speed forecasts, (Huang and Chalabi 1995). On the other hand, (Ailliot et al. 2006) apply an autoregressive model (AR) with time-varying coefficients to describe the space-time evolution of wind fields. In (F. E. Benth and Saltyte-Benth 2009) a stochastic process, called Continuous AR (CAR) model is introduced in order to model and forecast daily wind speeds. Finally, in (Nielsen et al. 2006) various statistical methods were presented for short-term wind speed forecasting. (Sfetsos 2002) argues about the use of linear or meteorological models since their prediction error is not significantly lower than the elementary persistent method. Alternatively, some studies use space-state models to simultaneously fit the speed and the direction of the wind, (Castino et al. 1998; Cripps et al. 2005; Martin et al. 1999; Tolman and Booij 1998; Tuller and Brett 1984; Haslett and Raftery 1989).

Alternatively to the linear models, artificial intelligence was applied in wind speed modeling and forecasting. In (Sfetsos 2000, 2002; More and Deo 2003; Barbounis et al. 2006; Beyer et al. 1994; Mohamed A. Mohandes et al. 1998; Alexiadis et al. 1998) neural networks were applied in order to model the dynamics of the wind speed process. In (M. A. Mohandes et al. 2004) support vector machines were used while in (Pinson and Kariniotakis 2003) fuzzy neural networks were applied.

Depending on the application, wind modeling is based on hourly, (Yamada 2008; Ailliot et al. 2006; Torres et al. 2005; Sfetsos 2000, 2002; Castino et al. 1998; Martin et al. 1999; Kamal and Jafri 1997; Daniel and Chen 1991), daily, (Caporin and Pres 2010; F. E. Benth and Saltyte-Benth 2009; More and Deo 2003; Tol 1997; Billinton et al. 1996; Huang and Chalabi 1995), weekly or monthly basis, (More and Deo 2003). When the objective is to hedge against electricity demand and production, hourly modeling is used while for weather derivative pricing the daily method is used. More rarely, weekly or monthly modeling is used in order to estimate monthly wind indexes. Since, we want to focus on weather derivative pricing the daily modeling approach is followed; however, the proposed method can be easily adapted in hourly modeling too.

Wind speed modeling is much more complicated than temperature modeling since wind has a direction and is greatly affected by the surrounding terrain such as building, trees, etc. (Jewson et al. 2005). However, in (F. E. Benth and Saltyte-Benth 2009) it is shown that wind speeds dynamics share a lot of common characteristics with the dynamics of temperature derivatives as it was found on (F. E. Benth and Saltyte-Benth 2007; Zapranis and Alexandridis 2008, 2009; Alexandridis 2010; Zapranis and Alexandridis 2010). In this context we use a mean reverting Ornstein-Uhlenbeck stochastic process to model the dynamics of the wind speed dynamics were the innovations are driven by a Brownian motion. The statistical analysis reveals seasonality in the mean and variance. In addition we use a novel approach to model the autocorrelation of the wind speeds. More precisely, a wavelet network (WN) is applied in order to capture accurately the autoregressive characteristics of the wind speeds.

The evaluation of the proposed methodology against alternative modeling procedures proposed in prior studies indicates that WNs can accurately model and forecast the dynamics and the evolution of the speed of the wind. The performance of each method was evaluated in-sample as well as out-of-sample and for different time periods.

The rest of the paper is organized as follows. In section 2 a statistical analysis of the wind speed dynamics is presented. In section 3 a linear model is fitted to the data while in section 4 a nonlinear nonparametric WN is applied. The evaluation of the studied models is presented in section 5. In section 6 we derive the pricing formulas for futures derivatives written on the wind index. Finally, in section 7 we conclude.

2. Wind Speed Modeling

In this section we derive empirically the characteristics of the daily average wind speed (DAWS) dynamics in New York, USA. The data were collected from NOAA[1] and correspond to DAWSs. The wind speeds are measured in 0.1 Knots. The measurement period is between 1st of January 1988 and 28th of February 2008. The first 20 years are used for the estimation of the parameters while the remaining two months are used for the evaluation of the performance of the proposed model. In order for each year to have the same number of observations the 29th of February is removed from the data resulting to 7,359 data points. The time-series is complete without any missing values.

In the Fig. 1 the DAWSs for the first 20 years are presented. The descriptive statistics of the in-sample data are presented in Table 1. The values of the data are always positive and range from 1.8 to 32.8 with mean around 9.91. Also, a closer inspection of Fig. 1 reveals seasonality.

The descriptive statistics of the DAWSs indicate that there is a strong positive kurtosis and skewness while the normality hypothesis is rejected based on the Jarque-Bera statistic. The same conclusions can be reached observing the first part of Fig. 2 where the histogram of the DAWSs is represented. Hence, the distribution of DAWSs deviates significantly from the normal and it is not symmetrical. In literature the Weibull or the Rayleigh (which is a special case of the Weibull) distributions were proposed, (F. E. Benth and Saltyte-Benth 2009; Brown et al. 1984; Daniel and Chen 1991; Garcia et al. 1998; Justus et al. 1978; Kavak Akpinar and Akpinar 2005; Nfaoui et al. 1996; Torres et al. 2005; J. r. a. Benth and Benth 2010; Celik 2004; Tuller and Brett 1984). In addition, some studies propose the use of the lognormal distribution, (F. E. Benth and Saltyte-Benth 2009; Garcia et al. 1998), or the Chi-square, (Dorvlo 2002). Finally, in (Jaramillo and Borja 2004) a bimodal Weibull and Weibull distribution is used.

[1] http://www.noaa.gov/

However, empirical studies favor the use of the Weibull distribution, (Celik 2004; Tuller and Brett 1984).

A closer inspection of part (a) of Fig. 2 reveals that the DAWSs in New York follow a Weibull distribution with scale parameter $\lambda = 11.07$ and shape parameter $k = 3.04$. Following (Brown et al. 1984; F. E. Benth and Saltyte-Benth 2009; Daniel and Chen 1991) in order to symmetrize the data, the Box-Cox transform is applied. The Box-Cox transformation is given by:

$$W^{(l)} = \begin{cases} \dfrac{W^{l} - 1}{l} & l \neq 0 \\ \ln(W) & l = 0 \end{cases} \tag{1}$$

where $W^{(l)}$ is the transformed data. The parameter l is estimated by maximizing the log-likelihood function. Note that the Log-transform is a special case of the Box-Cox transform with $l = 0$. The optimal l of the Box-Cox transform for the DAWS in New York is estimated to be 0.014. In the second part of Fig. 2 the histogram of the transformed data can be found while the second row of Table 1 shows the descriptive statistics of the transformed data.

The DAWSs exhibit a clear seasonal pattern which is preserved in the transformed data. The same conclusion can be reached by examining the autocorrelation function (ACF) of the DAWS in the first part of Fig. 3. In (F. E. Benth and Saltyte-Benth 2009; J. r. a. Benth and Benth 2010; Caporin and Pres 2010) the seasonality was captured by series of sinusoids. As in (Zapranis and Alexandridis 2008, 2009, 2010) for the case of temperature process, the seasonal effects are modeled by a truncated Fourier series given by:

$$S(t) = a_0 + b_0 t + \sum_{i=1}^{I_1} a_i \sin\left(2\pi i (t - f_i)/365\right) + \sum_{j=1}^{J_1} b_i \sin\left(2\pi j (t - g_j)/365\right) \tag{2}$$

In addition we examine the data for a linear trend representing the global warming or the urbanization around the meteorological station. First, we quantify the trend by fitting a linear regression to the DAWS data. The regression is statistically significant with intercept $a_0 = 2.3632$ and slope $b_0 = -0.000024$ indicating a slightly decrease in the DAWSs. Next, the seasonal periodicities are removed from the detrended data. The remaining statistically significant estimated parameters of equation (2) with $I_1 = J_1 = 1$ are presented in Table 2. As it is shown on the second part of Fig. **3** the seasonal mean was successfully removed. The same conclusion was reached in previous studies for daily models for both temperature and wind, (Alexandridis 2010; Zapranis and Alexandridis 2008, 2009, 2010; F. E. Benth and Saltyte-Benth 2005, 2007; F. E. Benth et al. 2007; F. E. Benth and Saltyte-Benth 2009; F. E. Benth et al. 2009)

3. The linear ARMA model

In literature, various methods for studying the statistical characteristics of the wind speed, in daily or hourly measurements, were proposed. However the majority of the studies utilize variations of the general ARMA model, (Ailliot et al. 2006; Billinton et al. 1996; Brett and Tuller 1991; Daniel and Chen 1991; Huang and Chalabi 1995; Kamal and Jafri 1997; Lei et al. 2009; Nfaoui et al. 1996; Rehman and Halawani 1994; Torres et al. 2005). In this paper we will first estimate the dynamics of the detrended and deseasonalized DAWSs process using a general ARMA model and then we will compare our results with a WN.

First, in order to select the correct ARMA model, we examine the ACF of the detrended and deseasonalized DAWS. A closer inspection of the second part of Fig. 3 reveals that the 1st, 2nd and the 4th lags are significant. On the other hand by examining the Partial Autocorrelation Function (PACF) in Fig. 4 we conclude that the first 4 lags are necessary to model the autoregressive effects of the winds speed dynamics.

In order to find the correct model we estimate the Log Likelihood function (LLF) and the Akaike Information Criterion (AIC). Consistent with the PACF, both criteria suggest that an AR(4) model is adequate for modeling the wind process since they were minimized when a model with 4 lags was used. The estimated parameters and the corresponding *p*-values are presented in Table 3. It is clear that the three first parameters are statistically very significant since their *p*-value is less than 0.05. The parameter of the 4th lag is statistically significant with *p*-value 0.0657. The AIC for this model is 0.46852 while the LLF is -1705.14.

Observing the residuals of the AR model in the first part of Fig. 5 we conclude that the autocorrelation was successfully removed. However, the ACF of the squared residuals indicates a strong seasonal effect in the variance of the wind speed as it is shown in Fig. 6. The same conclusion was reached in previous studies for daily models for both temperature and wind, (Alexandridis 2010; Zapranis and Alexandridis 2008, 2009, 2010; F. E. Benth and Saltyte-Benth 2005, 2007; F. E. Benth et al. 2007; F. E. Benth and Saltyte-Benth 2009; F. E. Benth et al. 2009). Following (Alexandridis 2010) we model the seasonal variance with a truncated Fourier series:

$$\sigma^2(t) = c_0 + \sum_{i=1}^{I_2} c_i \sin\left(2i\pi t/365\right) + \sum_{j=1}^{J_2} d_j \sin\left(2\pi jt/365\right) \tag{3}$$

Note that we assume that the seasonal variance is periodic and repeated every year, i.e. $\sigma^2(t+365) = \sigma^2(t)$ where $t = 1,...,7359$. The empirical and the fitted seasonal variance are presented in Fig. 7 while in Table 4 the estimated parameters of equation (3) are presented. Non-surprisingly, the variance exhibits the same characteristics as in the case of temperature, (Alexandridis 2010; Zapranis and Alexandridis 2008; F. E. Benth and Saltyte-Benth 2007). More precisely the seasonal variance is higher in the winter and early summer while it reaches its lower values during the summer period.

Finally, the descriptive statistics of the final residuals are examined. A closer inspection of Table 5 shows that the autocorrelation has successfully removed as indicated by the Ljung-Box Q-statistic. In addition the distribution of the residuals is very close to the normal distribution as it is shown on the first part of Fig. 8; however, small negative skewness exists. More precisely, the residuals have mean 0 and standard deviation of 1. In addition, the kurtosis is 3.03 and the skewness is -0.09.

Concluding, the previous analysis indicates that an AR(4) model provides a good fit for the wind process while the final residuals are very close to the normal distribution.

4. Wavelet Networks for Wind Speed Modeling

In this section WNs are used in the transformed, detrended and deseasonalized wind speed data in order to model the daily dynamics of wind speeds in New York. Motivated by the waveform of the data we expect a wavelet function to better fit the wind speed. In addition, it is expected that the non-linear form of the

WN will provide more accurate representation of the dynamics of the wind speed process both in-sample and out-of-sample.

In Fig. 9 the structure and the mathematical expressions of a WN are presented. In (Alexandridis 2010) detailed explanation of how to use WNs in model identification problems is described. Model identification can be separated in two parts, in model selection and in variable significance testing. Since WNs are nonlinear tools, criteria like AIC or LLF cannot be used. Hence, in this section WNs will be used in order to select the significant lags, to select the appropriate network structure, to train a WN in order to learn the dynamics of the wind speeds and, finally, to forecast the future evolution of the wind speeds.

The algorithm developed by (Alexandridis 2010) simultaneously estimates the correct number of lags that must be used in order to model the wind speed dynamics and the architecture of the WN by using a recurrent algorithm. An illustration of the model identification algorithm is presented in Fig. 10. In (Alexandridis 2010) we give a concise treatment of the WN theory. Here the emphasis is in presenting the basic notions of the model selection algorithm. For a more detailed exposition on the mathematical aspects of WN we refer to (Alexandridis 2010).

Our backward elimination algorithm examines the contribution of each available explanatory variable to the predictive power of the WN. First, the prediction risk of the WN is estimated as well as the statistical significance of each variable. If a variable is statistically insignificant it is removed from the training set and the prediction risk and the new statistical measures are estimated. The algorithm stops if all explanatory variables are significant. In this study the selected statistical measure is the Sensitivity Based Pruning (SBP) proposed by (Moody and Utans 1992). Previous analysis in (Alexandridis 2010) indicates that the SBP fitness criterion was found to significantly outperform alternative criteria in the variable selection algorithm. The SBP quantifies the effect on the empirical loss of replacing a variable by its mean. Analytical description of the SBP is given in (Alexandridis 2010; Zapranis and Refenes 1999; Moody and Utans 1992). In each step the SBP and the corresponding p-value are calculated. For analytical explanation of each step of the algorithm we refer to (Alexandridis 2010).

The proposed variable selection framework will be applied on the transformed, detrended and deseasonalized wind speeds in New York in order to select the length of the lag series. The target values of the WN are the DAWSs. The explanatory variables are lagged versions of the target variable. The relevance of a variable to the model is quantified by the SBP criterion which was introduced in (Moody and Utans 1992). Initially the training set contains the dependent variable and 7 lags. The analysis in the previous section indicates that a training set with 7 lags will provide all the necessary information of the ACF of the detrended and deseasonalized DAWSs. Hence, the training set consists of 7 inputs, 1 output and 7293 training pairs.

Table 6 summarizes the results of the model identification algorithm for the New York. Both the model selection and the variable selection algorithms are included in Table 6. The algorithm concluded in 4 steps and the final model contains only 3 variables, i.e 3 lags. The prediction risk for the reduced model is 0.0937 while for the original model was 0.0938. On the other hand the empirical loss slightly increased from 0.0467 for the initial model to 0.0468 for the reduced model indicating that the explained variability (unadjusted) slightly decreased. Finally, the complexity of the network structure and number of parameters were significantly reduced in the final model. The initial model needed 1 hidden unit (HU) and 7 inputs. Hence, 23 parameters were adjusted during the training phase. Hence the ratio of the number of training pairs n to the number of parameters p was 317.4. In the final model only 2 HU and 3 inputs were used. Hence only 18 parameters were adjusted during the training phase and the ratio of the number of training pairs n to the number of parameters p was 405.6.

The proposed algorithm suggests that a WN needs only 3 lags to extract the autocorrelation from the data while the linear model needed 4 lags. A closer inspection of Table 6 reveals that the WN with 3 and 4 lags have the same predictive power in-sample and out-of-sample. Hence, we chose the simpler model. Our model is similar to an AR(3) model with time-varying parameters.

Examining the second part of Fig. 5 we conclude that the autocorrelation was successfully removed from the data; however, the seasonal autocorrelation in the squared residuals is still present as it is shown in Fig. 6. We will remove the seasonal autocorrelation using equation (3). The estimated parameters are presented in Table 7 and as it was expected their values are similar to those of the case of the linear model. In Fig. 7 the empirical and the fitted seasonal variance is presented. Again, the same conclusions are reached for the seasonal variance. The variance is higher at winter period while it reaches its minimum during the summer period.

Finally, examining the final residuals of the WN model, we observe that the distribution of the residuals is very close to the normal distribution as it is shown in Fig. 8 while the autocorrelation was successfully removed from the data. In addition we observe an improvement in the distributional statistics in contrast to the case of the linear model. The distributional statistics of the residuals are presented in Table 8.

Concluding, the distributional statistics of the residuals indicate that in-sample the two models can accurate represent the dynamics of the DAWSs however an improvement is evident when a nonlinear nonparametric WN is used.

5. Forecasting daily average wind speeds

In this section our proposed model will be validated out-of-sample. In addition the performance of our model will be tested against two models, first, against the linear model previously described and second, against the simple persistent method usually referred as benchmark. The linear model is the AR(4) model described in the previous section. The persistent method assumes that the today's and tomorrow's DAWSs will be equal, i.e. $W^{*}(t+1)=W(t)$ where the W^{*} indicates the forecasted value.

The three models will be used for forecasting DAWSs for two different periods. Usually wind derivatives are written for a period of a month. Hence, DAWSs for 1 and 2 months will be forecasted. The out-of-sample dataset correspond to the period from January 1st to February 28th 2008 and were not used for the estimation of the linear and nonlinear models. Note that our previous analysis reveals that the variance in higher in the winter period indicating that it is more difficult to forecast accurately DAWS for these two months.

In Table 9 the performance of the three methods when the forecast window is one month is presented. Various error criteria are estimated like the Mean, Median and Maximum Absolute Error (Max. AE), the Mean Square Error (MSE), the Position of Change in Direction (POCID) and the Independent POCID. As it is shown on Table 9 our proposed method outperforms both the persistent and the AR(4) model. The AR(4) model performs better than the naïve persistent method however all error criteria are improved when a nonlinear WN is used. The MSE is 16.3848 for the persistent method, 10.6127 for the AR(4) model and 10.3309 for the WN. In addition our model can predict more accurately the movement of the wind speed since the POCID is 80% for the WN and the AR(4) models while it is only 47% for the persistent method. Moreover the IPOCID is 37% for the proposed model while it is only 33% for the other two methods.

In order to compare our model directly with the linear method, we estimate a linear AR(3) model. However, our proposed methodology still outperforms the linear method.

Next, the three forecasting methods are evaluated in two months day-ahead forecasts. The results are similar and presented in Table 10. The proposed WN outperforms the other two methods. Only the Max. AE and the POCID are slightly smaller when the AR(4) model is used. However the IPOCID is 38% for both methods. Also, our results indicate that the persistent method produces significantly worse forecasts. Finally, the WN and the linear AR(3) model are compared with first to show better forecasting ability.

Our results indicate that the WN can forecast the evolution of the dynamics of the DAWSs and hence they constitute an accurate tool for wind derivatives pricing.

6. Pricing

In this section the pricing formulas for wind derivatives are presented under the assumption of a normal driving noise process. The analysis that performed in the previous section indicates the assumption that the final residuals, after dividing out the seasonal variance, follow a normal distribution is justified.

When the market is complete, a unique risk-neutral probability measure $Q \sim P$ can be obtained, where P is the real world probability measure. This change of measure turns the stochastic process into a martingale. Hence, financial derivatives can be priced under the risk-neutral measure by the discounted expectation of the derivative payoff.

The weather market is an incomplete market in the sense that the underlying weather derivative cannot be stored or traded. Moreover the market is relatively illiquid. In principle, (extended) risk-neutral valuation can be still carried out in incomplete markets, (Xu et al. 2008). However, in incomplete markets a unique price cannot be obtained using the no-arbitrage assumption. In other words, under every measure *Q* all assets are martingales after discounting.

The change of measure from the real world to the risk-neutral world under the dynamics of a BM can be performed using the Girsanov theorem. The Girsanov theorem tells us how changes in the measure affect a stochastic process. Then the discounted expected payoff of the various weather contracts can be estimated. However, in order to estimate the expected payoff of each derivative, the solution of the stochastic differential equation that describes the wind speed dynamics must be solved. This can be done by applying the Itô's Lemma when a BM is considered.

Following (F. E. Benth and Saltyte-Benth 2009) we focus on the Nordix Wind Speed Index which is the index that the US Future Exchange used to settle wind derivatives. The Nordix Index is given by:

$$I(\tau_1,\tau_2) = 100 + \sum_{s=\tau_1}^{\tau_2}\left(W(s) - w_{20}(s)\right) \tag{4}$$

and measures the daily wind speed deviations from the mean of the past 20 years over a period $[\tau_1,\tau_2]$.

The statistical analysis indicates that the transformed DAWSs can be modeled by a mean reverting Ornstein-Uhlenbeck process where the speed of mean reversion variable is a function of time:

$$dW_t^{(l)} = S(t) + a(t)\left(W_{t-1}^{(l)} - S(t-1)\right)dt + \sigma(t)dB_t \tag{5}$$

where *S(t)* is the seasonal function, $\sigma(t)$ is the seasonal variance which is bounded by zero, $a(t)$ is the speed of mean reversion and B_t is the driving noise process.

Using the Girsanov's theorem, under the risk neutral measure *Q*, we have that:

$$dB_t^{\theta} = dB_t - \theta(t) \tag{6}$$

where $\theta(t)$ is the market price of risk and

$$\int_0^T \theta^2(t)dt < \infty \tag{7}$$

Hence, applying Ito Formula on equations (5) and (6) the solution of the transformed DAWS under the risk neutral measure *Q* is given by:

$$W_t^{(l)} = S(t) + e^{\int_0^t a(z)dz}\left(W_0^{(l)} - S(0)\right) + e^{\int_0^t a(z)dz}\int_0^t \sigma(s)\theta(s)e^{-\int_0^s a(z)dz}ds + e^{\int_0^t a(z)dz}\int_0^t \sigma(s)e^{-\int_0^s a(z)dz}dB_t^{\theta} \quad (8)$$

The proposed model is an extension of the CAR(p) introduced by (Brockwell and Marquardt 2005) and applied by (F. E. Benth and Saltyte-Benth 2009) in wind derivative pricing. Hence, we follow a similar pricing approach presented in (F. E. Benth and Saltyte-Benth 2009).

The transformed, detrended and deseasonalized DAWS $\tilde{W}_t^{(l)} = W_t^{(l)} - S(t)$ are normally distributed with mean:

$$\mu_\theta(t,s,\tilde{W}_t^{(l)}) = e^{\int_t^s a(z)dz}\tilde{W}_s^{(l)} + e^{\int_t^s a(z)dz}\int_t^s \sigma(u)\theta(u)e^{-\int_t^u a(z)dz}du \quad (9)$$

And variance

$$\Sigma^2(t,s) = e^{2\int_t^s a(z)dz}\int_t^s \sigma^2(s)e^{-2\int_t^u a(z)dz}du \quad (10)$$

If Q is the risk-neutral probability r is the constant compounding interest rate, the arbitrage-free price of Nordix Wind Future index at time $t \le \tau_1 < \tau_2$ is given by:

$$e^{-r(\tau_2 - t)}\mathrm{E}_Q\left[100 - \sum_{s=\tau_1}^{\tau_2} w_{20}(s) + \sum_{s=\tau_1}^{\tau_2} W(s) - F_{NWI}(t,\tau_1,\tau_2) \mid \mathcal{F}_t\right] = 0 \quad (11)$$

and since $F_{NWI}(t,\tau_1,\tau_2)$ is $\mathcal{F}_t$ adapted we derive the price of a Nordix Wind Future Index to be:

$$F_{NWI}(t,\tau_1,\tau_2) = 100 - \sum_{s=\tau_1}^{\tau_2} w_{20}(s) + \mathrm{E}_Q\left[\sum_{s=\tau_1}^{\tau_2} W(s) \mid \mathcal{F}_t\right] \quad (12)$$

Applying the Lemma 4.1 from (F. E. Benth and Saltyte-Benth 2009) we find the explicit solution for the price of the Nordix Wind Future Index:

$$F_{NWI}(t,\tau_1,\tau_2) = 100 - \sum_{s=\tau_1}^{\tau_2} w_{20}(s) + \sum_{s=\tau_1}^{\tau_2} M_{1/l}\left(1 + l\left(S(s) + \mu_\theta(t,s,\tilde{W}_t^{(l)}\right), l^2\Sigma^2(t,s)\right) \tag{13}$$

where $M_k(a,b^2)$ is the k^{th} moment of a normal random variable with mean a and variance b^2.

7. Conclusions

In this study DAWSs from New York were studied. Our analysis revealed strong seasonality in the mean and variance. The DAWSs were modeled by a mean reverting Ornstein-Uhlenbeck process in the context of wind derivatives pricing. In this study the dynamics of the wind generating process are modeled using a non-parametric non-linear WN. Our proposed methodology was compared in-sample and out-of-sample against two methods often used in prior studies. The characteristics of the wind speed process are very similar to the process of daily average temperatures.

Our method is validated in a two month ahead out of sample forecast period. Moreover, the various error criteria produced by the WN are compared against the linear AR model and the persistent method. Results show that the WN outperforms the other two methods, indicating that WNs constitute an accurate model for forecasting DAWSs. More precisely the WN's forecasting ability is stronger in both samples. Testing the fitted residuals of the WN we observe that the distribution of the residuals is very close to the normal. Also, the WN needed only the information of the past 3 days while the linear method suggested a model with 4 lags. Finally, we provided the pricing equations for wind futures of the Nordix index. Although we focused on DAWSs our model can be easily adapted in hourly modeling.

The results in this study are preliminary and can be further analyzed. More precisely alternative methods for estimating the seasonality in the mean and in the variance can be developed. Alternative methods could improve the fitting to the original data as well as the training of the WN.

Also, it is important to test the largest forecasting window of each method. Since meteorological forecasts of a window larger than few days are considered inaccurate this analysis will suggest the best model according to the desired forecasting interval.

Finally, a large scale comparison must be conducted. Testing the proposed methods as well as more sophisticated models, like general ARFIMA or GARCH, in various meteorological stations will provide a better insight in the dynamics of the DAWS as well as in the predictive ability of each method.

References

Ailliot, P., Monbet, V., & Prevosto, M. (2006). An autoregressive model with time-varying coefficients for wind fields. *Envirometrics, 17*, 107-117.

Alexandridis, A. (2010). *Modelling and Pricing Temperature Derivatives Using Wavelet Networks and Wavelet Analysis*. University of Macedonia, Thessaloniki, Greece.

Alexiadis, M. C., Dokopoulos, P. S., Sahsamanoglou, H. S., & Manousaridis, I. M. (1998). Short-term forecasting of wind speed and related electrical power *Solar Energy, 63*(1), 61-68.

Barbounis, T. G., Theocharis, J. B., Alexiadis, M. C., & Dokopoulos, P. S. (2006). Long-term wind speed and power forecasting using local recurrent neural network models. *Energy Conversion, IEEE Transactions on, 21*(1), 273-284.

Benth, F. E., Hardle, W. K., & Lopez Cabrera, B. (2009). Pricing of Asian temperature risk. *SFB649 Working Paper*. Berlin: Humboldt-Universitat zu Berlin.

Benth, F. E., & Saltyte-Benth, J. (2005). Stochastic Modelling of Temperature Variations With a View Towards Weather Derivatives. *Applied Mathematical Finance, 12*(1), 53-85.

Benth, F. E., & Saltyte-Benth, J. (2007). The volatility of temperature and pricing of weather derivatives. *Quantitative Finance, 7*(5), 553-561.

Benth, F. E., & Saltyte-Benth, J. (2009). Dynamic Pricing of Wind Futures. *Energy Economics, 31*, 16-24.

Benth, F. E., Saltyte-Benth, J., & Koekebakker, S. (2007). Putting a price on temperature. *Scandinavian Journal of Statistics,, 34*, 746-767.

Benth, J. r. a., & Benth, F. E. (2010). Analysis and modelling of wind speed in New York. *Journal of Applied Statistics, 37*(6), 893 - 909.

Beyer, H. G., Degner, T., Hausmann, J., Hoffmann, M., & Rujan, P. Short-term prediction of wind speed and power outputof a wind turbine with neural netwroks. In *2nd European Congress on Intelligent Techniques and Soft Computing, Aachen, Germany, 20-23 September 1994*

Billinton, R., Chen, H., & Ghajar, R. (1996). Time-series Models for Reliability Evaluation of Power Systems Including Wind Energy. *Microelectronics and Reliability, 36*(9), 1253-1261.

Brett, A. C., & Tuller, S. E. (1991). The Autocorrelation of Hourly Wind Speed Observations. *Journal of Applied Meteorology, 30*(6), 823-833, doi:doi:10.1175/1520-0450(1991)030<0823:TAOHWS>2.0.CO;2.

Brockwell, P. J., & Marquardt, T. (2005). Levy-driven and fractionality integrated ARMA process with continuous time parameter. *Statistica Sinica, 15*, 477-494.

Brown, B. G., Katz, R. W., & Murphy, A. H. (1984). Time-series models to simulate and forecast wind speed and wind power, . *Journal of Climate and Applied Meteorology, 23*, 1184-1195.

Caporin, M., & Pres, J. (2010). Modelling and forecasting wind speed intensity for weather risk management. *Computational Statistics & Data Analysis*, doi:doi:10.1016/j.csda.2010.06.019.

Castino, F., Festa, R., & Ratto, C. F. (1998). Stochastic modelling of wind velocities time-series. *Journal of Wind Engineering and Industrial Aerodynamics, 74-76*, 141-151.

Celik, A. N. (2004). A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey. *Renewable Energy, 29*(4), 593-604, doi:10.1016/j.renene.2003.07.002.

Challis, S. (1999). Bright Forecast for Profits. *Reactions, June edition*.

CME (2005). An Introduction to CME Weather Products. http://www.cme.com/edu/res/bro/cmeweather. Accessed January 2007.

Cripps, E., Nott, D., Dunsmuir, W. T. M., & Wikle, C. (2005). Space-Time Modelling of Sydney Harbour Winds. *Australian & New Zealand Journal of Statistics, 47*(1), 3-17, doi:10.1111/j.1467-842X.2005.00368.x.

Daniel, A. R., & Chen, A. A. (1991). Stochastic simulation and forecasting of hourly average wind speed sequences in Jamaica. *Solar Energy, 46*(1), 1-11.

Dorvlo, A. S. S. (2002). Estimating wind speed distribution. *Energy Conversion and Management, 43*(17), 2311-2318, doi:Doi: 10.1016/s0196-8904(01)00182-0.

Garcia, A., Torres, J. L., Prieto, E., & de Francisco, A. (1998). Fitting wind speed distributions: a case study. *Solar Energy, 62*(2), 139-144, doi:Doi: 10.1016/s0038-092x(97)00116-3.

Hanley, M. (1999). Hedging the Force of Nature. *Risk Professional, 1*, 21-25.

Haslett, J., & Raftery, A. E. (1989). Space-Time Modelling with Long-Memory Dependence: Assessing Ireland's Wind Power Resource. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 38*(1), 1-50.

Huang, Z., & Chalabi, Z. S. (1995). Use of time-series analysis to model and forecast wind speed. *Journal of Wind Engineering and Industrial Aerodynamics, 56*, 311-322.

Jaramillo, O. A., & Borja, M. A. (2004). Wind speed analysis in La Ventosa, Mexico: a bimodal probability distribution case. *Renewable Energy, 29*(10), 1613-1630, doi:DOI: 10.1016/j.renene.2004.02.001.

Jewson, S., Brix, A., & Ziehmann, C. (2005). *Weather Derivative Valuation: The Meteorological, Statistical, Financial and Mathematical Foundations*. Cambridge, UK: Campbridge University Press.

Justus, C. G., Hargraves, W. R., Mikhail, A., & Graber, D. (1978). Methods for estimating wind speed frequency distributions. *Journal Name: J. Appl. Meteorol.; (United States); Journal Volume: 17:3*, Medium: X; Size: Pages: 350-385.

Kamal, L., & Jafri, Y. Z. (1997). Time-series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. *Solar Energy, 61*(1), 23-32.

Kavak Akpinar, E., & Akpinar, S. (2005). A statistical analysis of wind speed data used in installation of wind energy conversion systems. *Energy Conversion and Management, 46*(4), 515-532, doi:DOI: 10.1016/j.enconman.2004.05.002.

Kavasseri, R. G., & Seetharaman, K. (2009). Day-ahead wind speed forecasting using f-ARIMA models. *Renewable Energy, 34*(5), 1388-1393, doi:DOI: 10.1016/j.renene.2008.09.006.

Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., & Yan, Z. (2009). A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews, 13*(4), 915-920, doi:DOI: 10.1016/j.rser.2008.02.002.

Martin, M., Cremades, L. V., & Santabarbara, J. M. (1999). Analysis and modelling of time-series of surface wind speed and direction. *International Journal of Climatology, 19*, 197-209.

Mohandes, M. A., Halawani, T. O., Rehman, S., & Hussain, A. A. (2004). Support vector machines for wind speed prediction. *Renewable Energy, 29*(6), 939-947, doi:DOI: 10.1016/j.renene.2003.11.009.

Mohandes, M. A., Rehman, S., & Halawani, T. O. (1998). A neural networks approach for wind speed prediction. *Renewable Energy, 13*(3), 345-354, doi:Doi: 10.1016/s0960-1481(98)00001-9.

Moody, J. E., & Utans, J. (1992). Principled Architecture Selection for Neural Networks: Applications to Corporate Bond Rating Prediction. In A. P. Refenes (Ed.), *Neural Networks in the Capital Markets*: John Wiley & Sons.

More, A., & Deo, M. C. (2003). Forecasting wind with neural networks. *Marine Structures, 16*, 35-49.

Nfaoui, H., Buret, J., & Sayigh, A. A. M. (1996). Stochastic simulation of hourly average wind speed sequences in Tangiers (Morocco). *Solar Energy, 56*(3), 301-314, doi:Doi: 10.1016/0038-092x(95)00103-x.

Nielsen, T. S., Madsen, H., Nielsen, H. A., Pinson, P., Kariniotakis, G., Siebert, N., et al. (2006). *Short-term Wind Power Forecasting Using Advanced Statistical Methods.* Paper presented at the The European Wind Energy Conference, EWEC, Athens, Greece,

Pinson, P., & Kariniotakis, G. N. Wind power forecasting using fuzzy neural networks enhanced with on-line prediction risk assessment. In *Power Tech Conference Proceedings, 2003 IEEE Bologna, 23-26 June 2003 2003* (Vol. 2, pp. 8 pp. Vol.2)

Rehman, S., & Halawani, T. O. (1994). Statistical characteristics of wind in Saudi Arabia. *Renewable Energy, 4*(8), 949-956, doi:Doi: 10.1016/0960-1481(94)90229-1.

Sfetsos, A. (2000). A comaprison of various forecasting techniques applied to mean hourly wind speed time-series. *Renewable Energy, 21*, 23-35.

Sfetsos, A. (2002). A novel approach for the forecasting of mean hourly wind speed time series. *Renewable Energy, 27*, 163-174.

Tol, R. S. J. (1997). Autoregressive Conditional Heteroscedasticity in Daily Wind Speed Measurements. *Theoretical and Applied Climatology*, *56*, 113-122.
Tolman, H. L., & Booij, N. (1998). Modeling wind waves using wavenumber-direction spectra and a variable wavenumber grid. *The Global Atmosphere and Ocean System*, *6*, 295-309.
Torres, J. L., Garcia, A., De Blas, M., & De Francisco, A. (2005). Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy*, *79*, 65-77.
Tuller, S. E., & Brett, A. C. (1984). The Characteristics of Wind Velocity that Favor the Fitting of a Weibull Distribution in Wind Speed Analysis. *Journal of Climate and Applied Meteorology*, *23*(1), 124-134.
WRMA (2010). Weather derivatives volume plummets. www.wrma.org/pdf/weatherderivativesvolumeplummets.pdf. Accessed January 2010.
Xu, W., Odening, M., & Musshof, O. (2008). Indifference Pricing of Weather Derivatives. *American Journal of Agricultural Economics*, *90*(4), 979-993.
Yamada, Y. Simultaneous optimization for wind derivatives based on prediction errors. In *Americal Control Conference, Washington, USA, June 11-13, 2008* (pp. 350-355)
Zapranis, A., & Alexandridis, A. (2008). Modelling Temperature Time Dependent Speed of Mean Reversion in the Context of Weather Derivetive Pricing. *Applied Mathematical Finance, 15*(4), 355 - 386.
Zapranis, A., & Alexandridis, A. (2009). Weather Derivatives Pricing: Modelling the Seasonal Residuals Variance of an Ornstein-Uhlenbeck Temperature Process With Neural Networks. *Neurocomputing*, *73*, 37-48.
Zapranis, A., & Alexandridis, A. (2010). Modeling and forecasting cumulative average temperature and heating degree day indices for weather derivative pricing. *Neural Computing & Applications*, 1-15, doi:10.1007/s00521-010-0494-1.
Zapranis, A., & Refenes, A. P. (1999). *Principles of Neural Model Indentification, Selection and Adequacy: With Applications to Financial Econometrics*: Springer-Verlag.

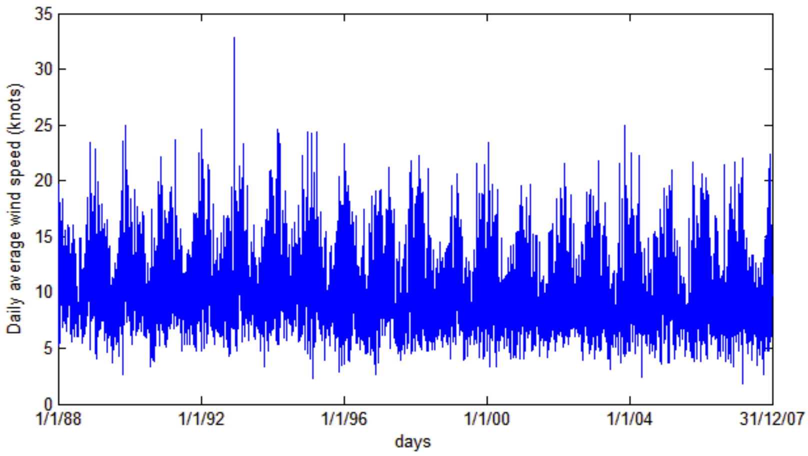
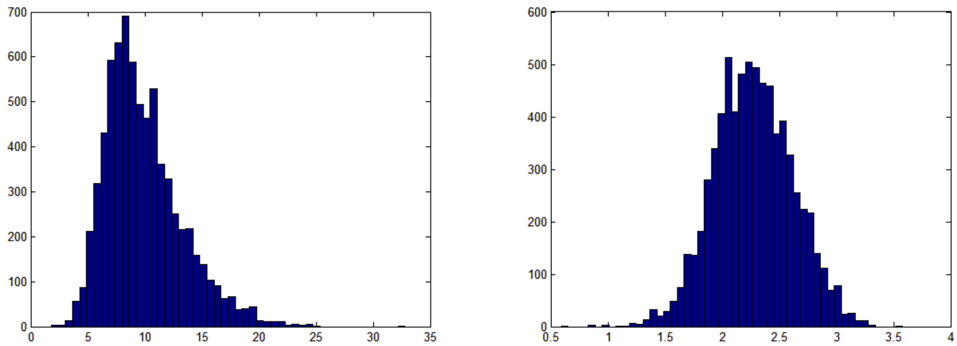
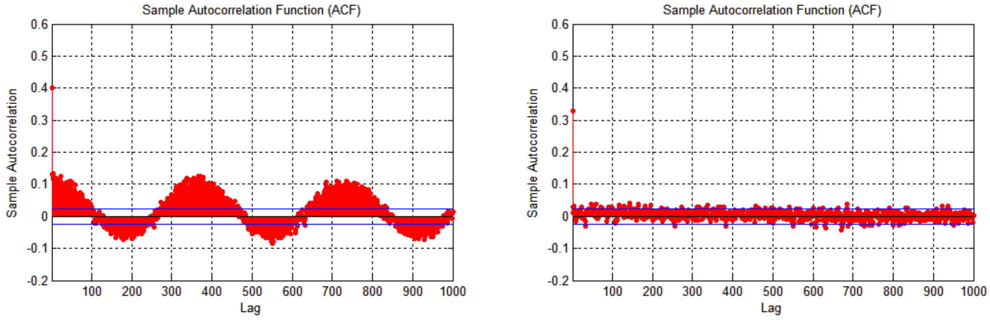


Fig. 1 Daily Average Wind Speed for New York



(a) (b)

Fig. 2 Histogram of the (a) original and (b) Box-Cox transformed data



(a) (b)

Fig. 3 The Autocorrelation Function of the transformed DAWSs in New York (a) before and (b) after removing the seasonal mean

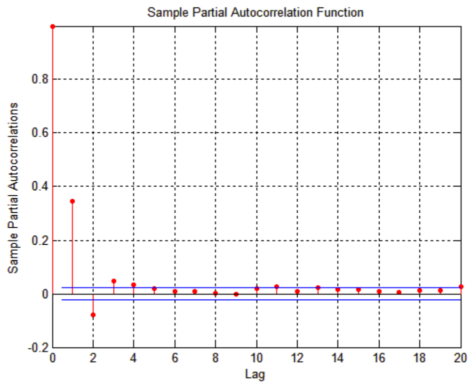
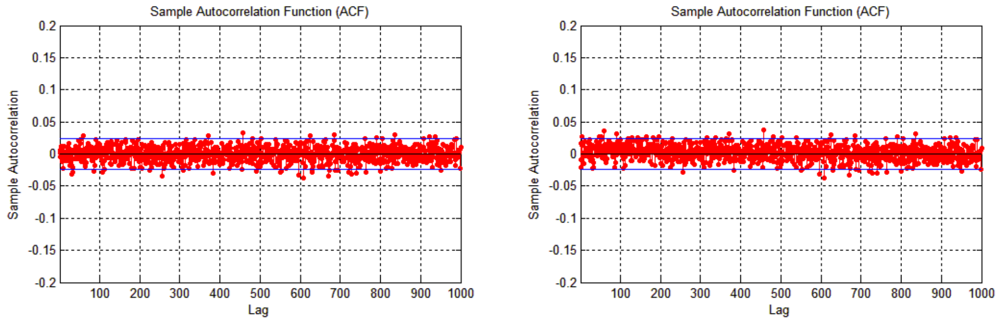
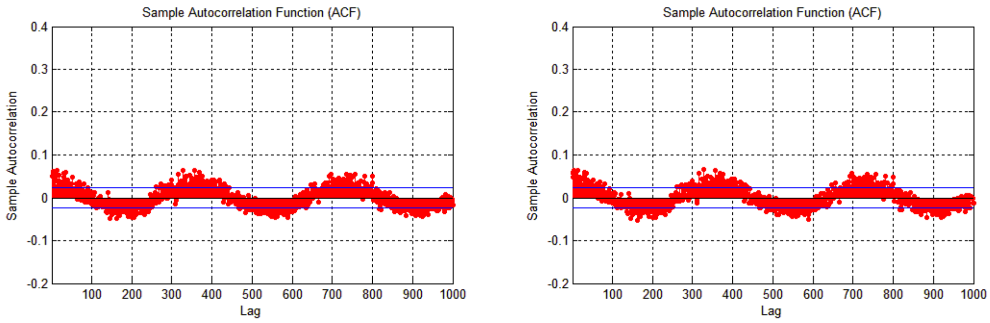


Fig. 4 The Partial Autocorrelation Function of the de-trended and de-seasonalized DAWS in New York



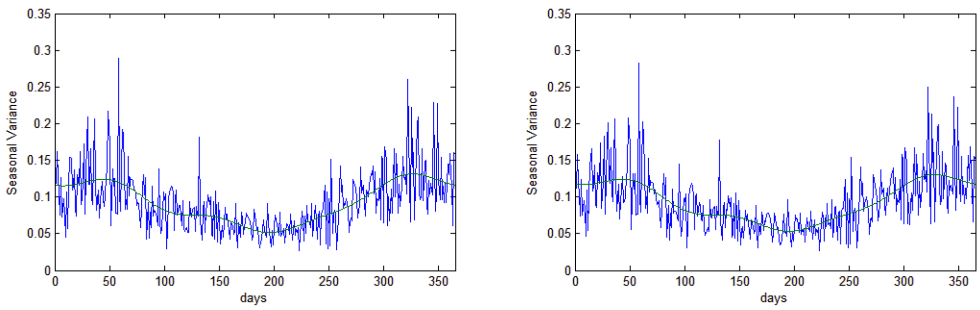
(a) (b)

Fig. 5 Autocorrelation function of the residuals of (a) the linear model and (b) the WN



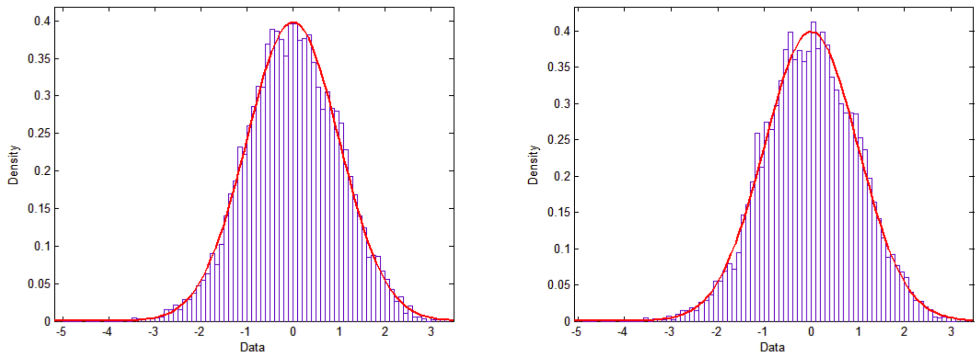
(a) (b)

Fig. 6 Autocorrelation function of the squared residuals of (a) the linear model and (b) the WN



(a) (b)

Fig. 7 Empirical and fitted seasonal variance of (a) the linear model and (b) the WN



(a) (b)

Fig. 8 Empirical and fitted Normal distribution of the final residuals of the WN

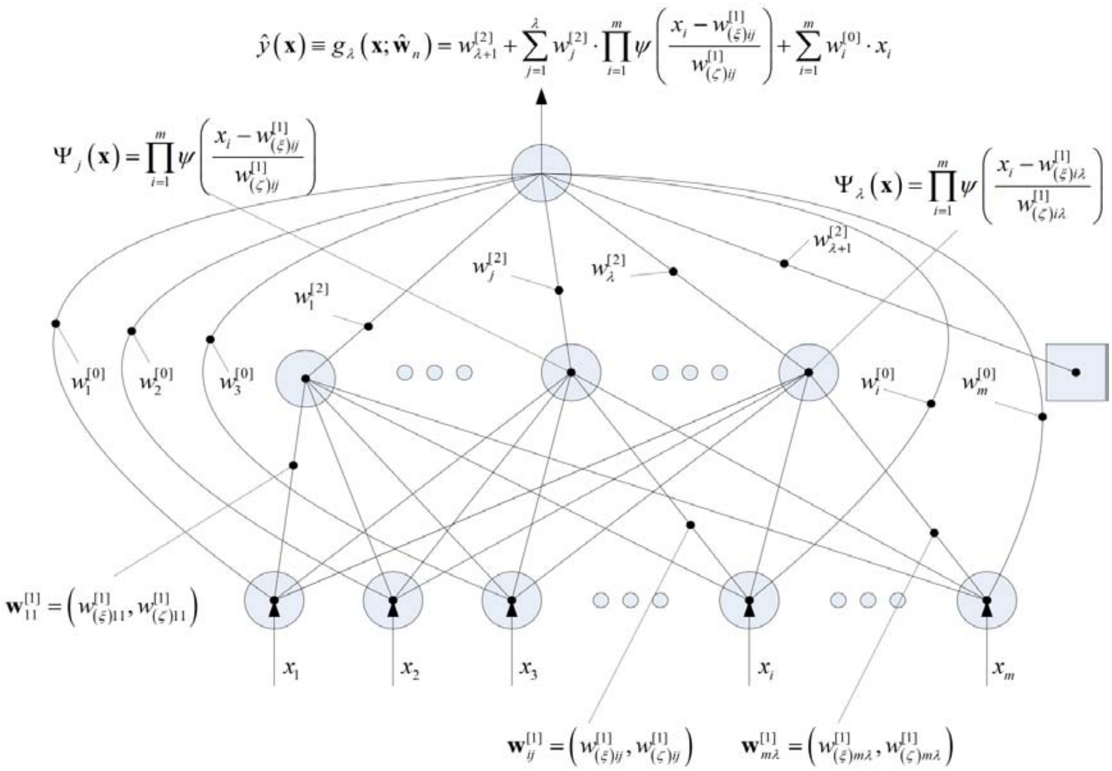


Fig. 9 Structure of a Wavelet Network

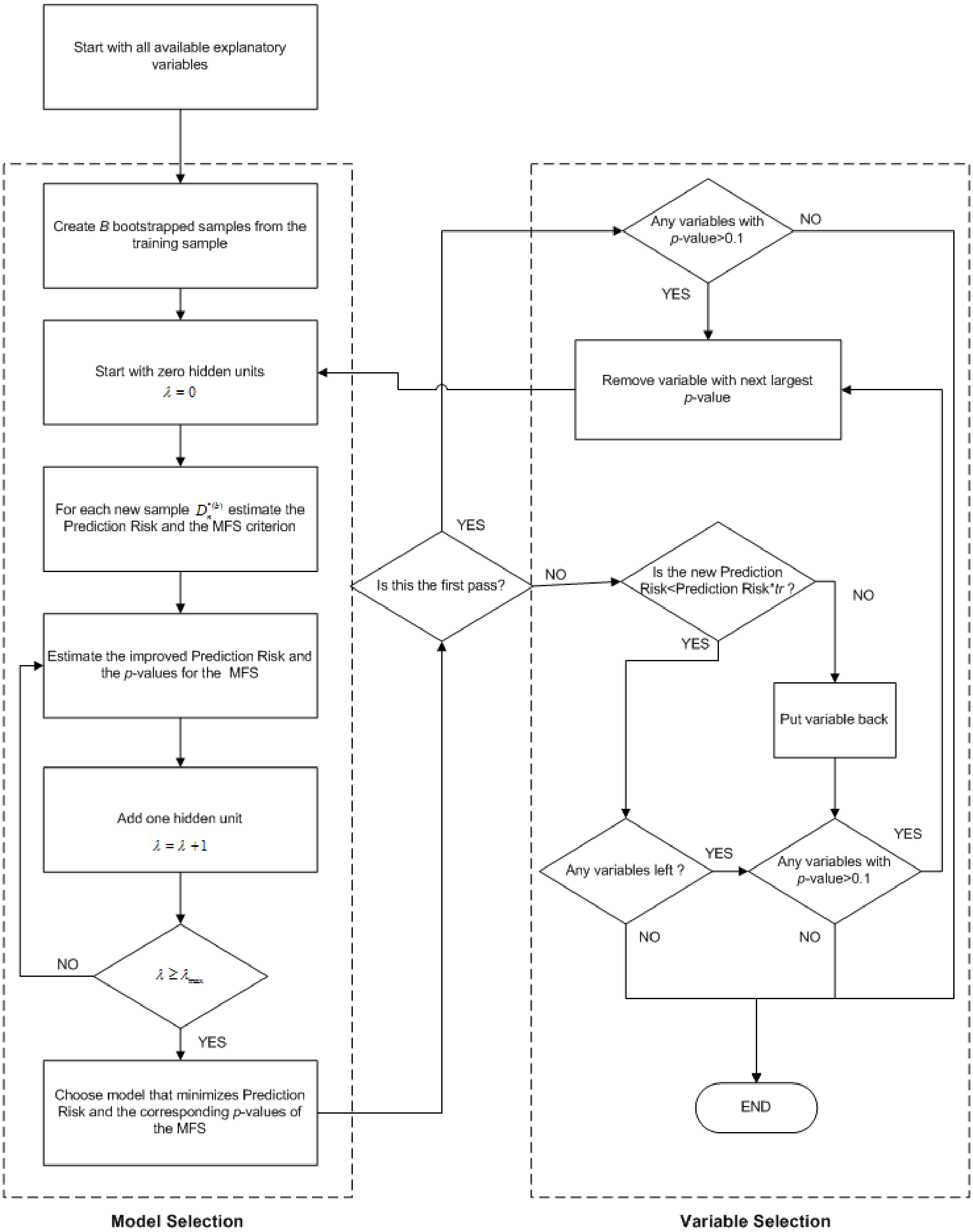


Fig. 10 Model and Variable Selection Algorith

Table 1. Descriptive statistics of the wind speed in New York.

	Mean	Median	Max	Min	StdDev.	Skew	Kurt	J-B	*p*-value
Original	9.91	9.3	32.8	1.8	3.38	0.96	4.24	1595.41	0
Transformed	2.28	2.3	3.6	0.6	0.34	0.00	3.04	0.51	1

J-B=Jarque-Bera statistic
p-value=*p*-values of the J-B statistic

Table 2. Estimated parameters of the seasonal component.

a_0	b_0	a_1	f_1	b_1	g_1
2.3632	-0.000024	0.0144	827.81	0.1537	28.9350

Table 3. Estimated parameters of the linear AR(4) model.

Parameter	AR(1)	AR(2)	AR(3)	AR(4)
Value	0.3617	-0.0999	0.0274	0.0216
p-value	0.0000	0.0000	0.0279	0.0657

Table 4. Estimated parameters of the seasonal variance in the case of the linear model.

c_0	c_1	c_2	c_3	c_4	d_1	d_2	d_3	d_4
0.0932	0.000032	-0.0041	0.0015	-0.0028	0.0358	-0.0025	-0.0048	-0.0054

Table 5. Descriptive statistics of the residuals for the linear AR(4) model.

Var	Mean	St.Dev	Max	Min	Skew	Kur.	JB	*p*-value	KS	*p*-value	LBQ	*p*-value
noise	0	1	3.32	-5.03	-0.09	3.03	10.097	0.007	1.033	0.2349	8.383	0.989

St.Dev.=Standard Deviation
JB=Jarque-Bera statistic
KS=Komogorov-Smirnov statistic
LBQ= Ljung-Box Q-statistic

Table 6. Variable selection with backward elimination in New York.

Step	Variable to remove (lag)	Variable to enter (lag)	Variables in model	Hidden Units (Parameters)	*n/p* Ratio	Empirical Loss	Prediction Risk
-			7	1 (23)	317.4	0.0467	0.0938
1	7	-	6	1 (20)	365.0	0.0467	0.0940
2	5	-	5	1 (17)	429.4	0.0467	0.0932
3	6	-	4	2 (23)	317.4	0.0467	0.0938
4	4	-	3	2 (18)	405.6	0.0468	0.0937

The algorithm concluded in 4 steps. In each step the following are presented: which variable is removed, the number of hidden units for the particular set of input variables and the parameters used in the wavelet network, the ratio between the parameters and the training patterns, the empirical loss and the prediction risk.

Table 7. Estimated parameters of the seasonal variance in the case of the WN.

c_0	c_1	c_2	c_3	c_4	d_1	d_2	d_3	d_4
0.0935	-0.000020	-0.0034	0.0014	-0.0026	0.0353	-0.0016	-.0042	-0.0052

Table 8. Descriptive statistics of the residuals for the WN model.

Var	Mean	St.Dev	Max	Min	Skew	Kur.	JB	*p*-value	KS	*p*-value	LBQ	*p*-value
noise	0	1	3.32	-4.91	-0.08	3.04	8.84	0.0043	0.927	0.3544	13.437	0.858

St.Dev.=Standard Deviation
JB=Jarque-Bera statistic
KS=Komogorov-Smirnov statistic
LBQ= Ljung-Box Q-statistic

Table 9. Out-of-sample comparison. 1 month.

	PERSISTENT	AR(4)	WN
Md.AE	2.3000	2.2147	2.1081
MAE	3.3000	2.5547	2.5026
Max AE	8.2000	7.9217	7.7590
SSE	507.9300	328.9947	320.2573
RMSE	4.0478	3.2577	3.2142
NMSE	1.5981	1.0351	1.0076
MSE	16.3848	10.6127	10.3309
MAPE	0.3456	0.2744	0.2680
SMAPE	0.3233	0.2570	0.2518
POCID	47%	80%	80%
IPOCID	33%	33%	37%
POS	100%	100%	100%

Md. AE=Median Absolute Error
MAE=Mean Absolute Error
Max AE=Maximum Absolute Error
SSE=Sum of Squared Errors
RMSE=Root Mean Square Error
NMSE=Normalized Mean Square Error
MSE= Mean Square Error
MAPE=Mean Absolute Percentage Error
SMAPE=Symmetric MAPE
POCID=Position of change in direction
IPOCID=Independent POCID
POS=Position of Sign

Table 10. Out-of-sample comparison. 2 months.

	PERSISTENT	AR(4)	WN
Md.AE	2.4000	2.7981	2.6589
MAE	3.3678	2.8126	2.7976
Max AE	11.2000	7.9345	8.0194
SSE	1054.3500	706.1806	702.4437
RMSE	4.2273	3.4596	3.4505
NMSE	1.4110	0.9450	0.9400
MSE	17.8703	11.9692	11.9058
MAPE	0.3611	0.3014	0.3001
SMAPE	0.3289	0.2798	0.2782
POCID	45%	71%	69%
IPOCID	36%	38%	38%
POS	100%	100%	100%

Md. AE=Median Absolute Error
MAE=Mean Absolute Error
Max AE=Maximum Absolute Error
SSE=Sum of Squared Errors
RMSE=Root Mean Square Error
NMSE=Normalized Mean Square Error
MSE= Mean Square Error
MAPE=Mean Absolute Percentage Error
SMAPE=Symmetric MAPE
POCID=Position of change in direction
IPOCID=Independent POCID
POS=Position of Sign