amcs

ADAPTIVE PREDICTIONS OF THE EURO/ZŁOTY CURRENCY EXCHANGE RATE USING STATE SPACE WAVELET NETWORKS AND FORECAST COMBINATIONS

Mietek A. Brdyś [a,b], Marcin T. Brdyś [a,*], Sebastian M. Maciejewski [c]

[a]Department of Control Systems Engineering
Gdańsk University of Technology, ul. Narutowicza 11/12, 80-952 Gdańsk, Poland
e-mail: mtbrdys@outlook.com

[b]Department of Electronic, Electrical and Systems Engineering
University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

[c]PGE Polish Energy Group, ul. Mysia 2, 00-496 Warsaw, Poland
e-mail: sebastian.michal.maciejewski@gmail.com

The paper considers the forecasting of the euro/Polish złoty (EUR/PLN) spot exchange rate by applying state space wavelet network and econometric forecast combination models. Both prediction methods are applied to produce one-trading-day-ahead forecasts of the EUR/PLN exchange rate. The paper presents the general state space wavelet network and forecast combination models as well as their underlying principles. The state space wavelet network model is, in contrast to econometric forecast combinations, a non-parametric prediction technique which does not make any distributional assumptions regarding the underlying input variables. Both methods can be used as forecasting tools in portfolio investment management, asset valuation, IT security and integrated business risk intelligence in volatile market conditions.

Keywords: currency exchange rate, artificial intelligence, state space wavelet network, Metropolis Monte Carlo, forecast combinations, data generating process.

1. Introduction

Many statistical models developed before the start of the global financial crisis of 2008 that aimed at forecasting financial and macroeconomic time series failed to act as good forecasting models after the crisis outbreak, in the market environment characterized by increased volatility. The volatility increase was driven, among other things, by financial problems of banks and other financial institutions primarily in the US and the Eurozone, the crisis on international government debt markets and deteriorating macroeconomic conditions in the Eurozone economies, as well as changing investment decisions of international investors driven by global risk-aversion.

The new, rapidly changing and highly volatile financial market environment called for more flexible forecast models that are able to react to changing global circumstances as well as to handle many variables instantaneously.

This paper presents two forecasting methods: the state space wavelet network (SSWN) and forecast combinations (FCs) models, and demonstrates how these methods can be effectively used in a changing and highly volatile market environment, facilitating investment decisions.

The first approach, i.e., the state space wavelet network, structures a model using sets of unknown parameters and lets the optimization routine seek the best fitting parameters to obtain the desired results based on historical correlations. Importantly, the SSWN model does not impose any statistical constraints or assumptions in generating predictions and therefore is suited for modelling financial time series in volatile market conditions.

The second model is the forecast combinations

---

*Corresponding author

method based on linear econometric regressions. This approach is used in applied econometrics with the aim of approximating the unknown and highly complex true market model that generates the time series of interest. More specifically, the econometric forecast combinations model combines forecasts from different single regressions producing more accurate and stable forecasts than any of the single regression models treated separately. This happens because the complex true model that generates the time series of interest is approximated by a set of single regressions, and not by one regression only. This feature also makes forecast combinations more suitable for modelling changing market conditions as compared, e.g., with single regressions.

The paper is organised as follows. In Section 2 the dynamics and evolution of the foreign exchange market as well as the EUR/PLN exchange rate pair are introduced. Section 3 determines input variables. The SSWN and FC prediction methods are presented in Sections 4 and 5, respectively. The validation results obtained by both the methods based on real data records and comparison of method performance are presented in Section 6. Conclusions in Section 7 complete the paper.

2. Dynamics and evolution of the foreign exchange market today and the EUR/PLN exchange rate

The global foreign exchange market is largely made up of banks, institutional investors, hedge funds, corporations, governments as well as currency speculators. It is an over-the-counter (OTC), decentralized market connected electronically.

The size of the global foreign exchange market has grown exponentially in the last decade. According to BIS (2013), foreign-exchange trading increased to an average of \$5.3 trillion (thousand billion) a day in April 2013. This is up from \$4.0 trillion in April 2010 and \$3.3 trillion in April 2007.[1]

As the most traded currency, the US dollar makes up 85% of the forex trading volume. At nearly 40% of the trading volume, the euro is ahead of the third place Japanese yen, which takes almost 20%. Foreign exchange swaps were the most actively traded instruments in April 2013, at \$2.2 trillion per day, followed by spot trading at \$2.0 trillion (BIS, 2013).

The Polish złoty is the currency of Poland, with a free floating exchange rate regime. According to NBP (2013), roughly 72% of all transactions concluded in April 2013 on the Polish foreign exchange market involved the Polish

złoty. The most popular Polish złoty exchange rate is the EUR to PLN rate (EUR/PLN), due to strong economic ties between Poland and the European Monetary Union (EMU) countries, which is reflected, among others, by the trade volume amounting to 50% of Poland's total foreign trade volume in 2013 (CSO, 2014).

According to NBP (2013), foreign exchange trading on the Polish foreign exchange market amounted to \$7.6 billion a day in April 2013. This was down from \$7.9 billion in April 2013. Foreign exchange swaps were the most actively traded instruments in April 2013 on the Polish foreign exchange market, at \$4.6 billion per day, followed by spot transactions at \$2.3 billion. The EUR/PLN pair constituted 55% of the net turnover of spot transactions (the second pair being EUR/USD, with a 17% share) and 61 % of FX swaps (the second pair being USD/PLN, with a 15% share) in April 2013.

Overall, the EUR/PLN pair was by far the most heavily traded currency on the Polish foreign exchange market in April 2013 in terms of all foreign exchange transactions' net turnover. Given the relative importance of the EUR/PLN exchange rate, two forecasting methods—state space wavelet network (SSWN) and forecast combinations (FCs) models—are applied to forecasting of the EUR/PLN spot exchange rate. The aim of the proposed models is to facilitate the investment decision-making of investors trading actively on the spot market and investing in instruments of shortest maturity, including FX swaps and outright forwards.

3. Selection of input variables for EUR/PLN exchange rate forecasting

The SSWN and FC models capture such features of the EUR/PLN rate dynamics as financial and macroeconomic factors volatility (e.g., government debt and financial markets current levels), and their correlation with EUR/PLN, as well as auto-regression and volatility clustering in the EUR/PLN series. In addition, the SSWN model offers effective mechanisms for handling nonlinearities, uncertainty in the inputs structure and different time scales in the EUR/PLN rate trading dynamics. However, significant structural changes in global forex flows and shifts in economic cycles require an on-line adaptation of the model internal structure (Qi and Brdyś, 2008; 2009).

Correct selection of essential inputs to the EUR/PLN rate system is a mile stone in designing the SSWN and FC models. As relations between the economic indicators on foreign exchange markets are extremely complex and almost impossible to be measured or estimated, it is impossible to choose all the factors that influence the exchange rate level considered. Therefore, it is attempted to choose only the most important factors that influence the predicted exchange rate level. Unknown, complex

---

[1] The Bank for International Settlements collected data from around 1,300 banks and other financial institutions from 53 countries on transactions (i.e., spot transactions, outright forwards, FX swaps, currency swaps and currency options) concluded in April 2013 on the foreign exchange market.

Table 1. Results of correlation analysis for the period of 11/08/2011 to 14/04/2014.

Indicator symbol	Correlation coefficient	Description of indicator
EURPLN	1.00	EUR/PLN close at the end of trading session
WIG20	−0.54	Value of the WIG20 equity index at the end of a trading session
PL106670	−0.55	Price of the 10 year maturity benchmark bond at the end of a trading session
VIX	0.61	Value of the volatility index at the end of a trading session
DAX	−0.44	Value of the DAX equity index at the end of a trading session
FTSE	−0.40	Value of the FTSE equity index at the end of a trading session
STOXX50	−0.41	Value of the EURO STOXX 50 equity index at the end of a trading session
SPX	−0.38	Value of the S&P equity index at the end of a trading session
EURUSD	0.07	EUR/USD close at the end of a trading session

and nonlinear relations between inputs and outputs of the SSWN an FC models are estimated during the process of learning, which is discussed in the following sections. A standard approach to selecting the input variables is to construct, based on qualitative knowledge, a list of potential measurable inputs and to apply a standard data correlation analysis to calculate the correlation coefficients between the input and the output considered. The final input selection is then based on the correlation coefficient values. The larger the coefficient, the higher selection priority assigned to the corresponding input. The correlation analysis based on preselected 20 input variables and future values of the EUR/PLN rate has supported the final selection of 9 variables as shown in Table 1. Other pre-selected variables were the FRA, OIS and LIBOR rates, bond yields, CDS spreads and commodity futures.

All the above-listed variables are available for the period of 11/08/2011 to 14/04/2014, and all subjected time series are raw (seasonally unadjusted) daily data. The inputs to the SSWN and FC predictors will be designed based on these variables.

It has to be emphasised that the correlation analysis is strictly valid only for linear relationships between the predicted rate and the economic indicators that are the inputs. In reality, these relationships are typically heavily nonlinear. Hence, the analysis should be seen as qualitative. The final selection of the inputs needs to be done within an iterative process, where different inputs are substituted; the predictor is validated and, based on the validation results, new inputs are produced. The process stops when the required prediction accuracy is reached.

Based on the correlation analysis, 9 variables (indicators) are selected as the base for designing inputs to the FC predictor in the following sections: EURPLN, WIG20, PL106670, VIX, DAX, FTSE, STOXX50, SPX and EURUSD. As described in the following section, the SSWN has a dedicated mechanism designed in order to achieve robustness with respect to an uncertainty in the structure of variables having an impact on the output. In order to quantify by simulation this robustness, only 5 of the indicators in Table 1 are selected to design 7 inputs

to the SSWN predictor in Section 4.2: EURPLN, WIG20, PL106670, DAX and VIX.

4. State space wavelet network predictor

Artificial intelligence models based on neural networks and/or fuzzy systems are of an increasing interest in financial engineering applications for prediction/forecasting purposes (Zhang *et al.*, 1998; Kuo *et al.*, 2001; Tsang *et al.*, 2007). In this paper, a recently developed artificial dynamic neural network with wavelet processing nodes and internal states called the state space wavelet network (SSWN) is applied. The SSWN was initially proposed for modelling nonlinear and non-stationary processes with multiple time scales in internal dynamics and non-measurable states under uncertainty in the inputs and dynamic models. It was successfully applied to input-output modelling in a state-space form of a wastewater treatment plant (Borowa *et al.*, 2007) for model predictive control purposes and on-line prediction of a future WIG20 index level as a key financial indicator of the Polish equities listed on the Warsaw Stock Exchange (WSE) (Brdyś *et al.*, 2009).

4.1. SSWN mathematical model. A general structure of the SSWN is illustrated in Fig. 1, where y_i, $i = 1, \ldots, N$, x_i, $i = 1, \ldots, M$, and u_i, $i = 1, \ldots, K$, denote network outputs, internal states and inputs, respectively. All the variables are discrete time, and the time variable is denoted by k.

Network internal states do not have to be related to states of the modelled system. In the case of unknown or unmeasurable system states, this is a great advantage of the network model. Identifying state variables of a complex system is in most cases impossible. However, artificial neural model states can still correctly describe the impact of system state variable dynamics on the system output (Zammarreno and Pastora, 1998; Kulawski and Brdyś, 2000). This vastly improves the ability of the model to approximate unknown system input-output dynamics. The EUR/PLN exchange rate reflects both the complexity of financial markets and the depth of the
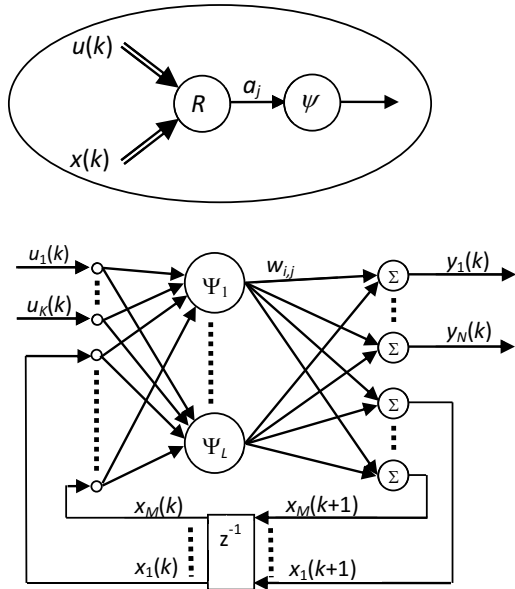
$$A(z(k), d_j, t_j) = \text{diag}(d_j)(z(k) - t_j)^T, \qquad (6)$$

$$R(z(k), d_j, t_j) = a_j(k)$$
$$= [A^T(z(k), d_j, t_j)A(z(k), d_j, t_j)]^{\frac{1}{2}}, \qquad (7)$$

where vectors $d_{i,j}$, $t_{i,j}$ and $i = 1, \ldots, M + K$ are composed of the j-th MRW parameters $d_{1,j}$, $t_{1,j}$, $j = 1, \ldots, L$, and $l = 1, \ldots, M + K$. Solving Eqns. (3)–(7) yields

$$a_j(k) = \Big(\sum_{i=1}^{M} \big(d_{i,j}(x_i(k) - t_{i,j}) \big)^2$$
$$+ \sum_{i=M+1}^{M+K} \big(d_{i,j}(u_i(k) - t_{i,j}) \big)^2 \Big)^{\frac{1}{2}}. \qquad (8)$$

It can now be seen from (8) that the components of the parameter vector d_j are scaling coefficients for the network inputs, both external and internal, produced by the internal feedback loops from the internal outputs delayed by one lag. The components of the parameter vector t_j perform translations of the inputs. Both the parameter vectors help to efficiently handle the multiple time scales in the system dynamics due to supply in demand shocks occurrence in the global economy, for example. Finally, the mapping ψ in Fig. 1 constitutes a one-dimensional Morlet wavelet function (Grossmann and Morlet, 1984):

$$\psi(a_j) = \exp\Big(-\frac{1}{2}a_j^2 \Big) \cos(5a_j). \qquad (9)$$

Equations (1)–(9) define an input-output model in a state space form with the weights $w_{i,j}$, scaling and translation factors $d_{i,j}, t_{i,j}$, respectively, that are the continuously valued model parameters to be determined, as well as numbers of the internal states M and data processing wavelons L, which are the discretely valued model structure parameters. Let us denote by w, d, t the vectors of continuously valued parameters.

Fig. 1. General structure of the SSWN.

global forex market. The SSWN input–output relationship can be written as

$$x_i(k + 1) = \sum_{j=1}^{L} w_{N+i,j} \Psi_j(x(k), u(k)),$$
$$i = 1, \ldots, M, \quad (1)$$

$$y_i(k) = \sum_{j=1}^{L} w_{i,j} \Psi_j(x(k), u(k)),$$
$$i = 1, \ldots, N, \quad (2)$$

where $w_{i,j}$, $i = 1, \ldots, N + M$, $j = 1, \ldots, L$, are the network weights to be determined and $u(k), x(k)$ are the network input and state vectors at time instant k, respectively. The network nodes that process the input information at time instants k are multidimensional radial wavelons (MRWs) (Zhang, 1992). The MRW processing mappings are denoted by Ψ_j, $j = 1, \ldots, L$. Feedforward networks with wavelet-based processing nodes were introduced by Zhang and Beneveniste (1992).

The MRW input information processing mapping Ψ is structured as a composition of two mappings R and ψ as $\Psi = R\psi$, as illustrated in Fig. 1. The mapping R is defined as follows:

$$z(k) = [x(k), u(k)], \qquad (3)$$

$$d_j = [d_{1,j}, \ldots, d_{K+M,j}], \qquad (4)$$

$$t_j = [t_{1,j}, \ldots, t_{K+M,j}], \qquad (5)$$

4.2. Determining SSWN inputs and outputs. Let us consider trading sessions during the day $k - 1$ and k. Let $k - 1$ be a discrete time instant located between the session $k - 1$ closing time and the session k opening time. For the k-th session, let $EURPLN_{cl}(k)$ and $EURPLN_{op}(k)$ denote the closing and opening exchange rate values, respectively. The following 7 inputs are selected for the SSWN to produce the k-th session (daily) prediction of $EURPLN_{cl}(k|k-1)$ of $EURPLN_{cl}(k)$ performed at instant $k - 1$ that is after closing the session $k - 1$ and at the same time before opening the session k:

- $u_1(k) \equiv EURPLN_{cl}(k - 1)$: the exchange rate value at the end of the previous trading session;

- $u_2(k) \equiv \frac{EURPLN_{cl}(k-1)-EURPLN_{cl}(k-2)}{EURPLN_{cl}(k-2)}$: the daily relative change in the EUR/PLN exchange rate value over the previous day, that is, day $k-1$;

- $u_3(k) \equiv EURPLN_{cl}(k-1) - EURPLN_{op}(k-1)$: the session exchange rate change over the previous trading session, that is, session $k-1$;

- $u_4(k) \equiv \frac{DAX_{cl}(k-1)-DAX_{cl}(k-2)}{DAX_{cl}(k-2)}$: the daily relative change in the German equity index DAX value over the previous day, that is, day $k-1$;

- $u_5(k) \equiv \frac{GOVPL10_{cl}(k-1)-GOVPL10_{cl}(k-2)}{GOVPL10_{cl}(k-2)}$: the daily relative change in the Government of Poland 10 year maturity bond price value over the previous day, that is, day $k-1$;

- $u_6(k) \equiv \frac{WIG20_{cl}(k-1)-WIG20_{cl}(k-2)}{WIG20_{cl}(k-2)}$: the daily relative change in the Polish equity index WIG20 value over the previous day, that is, day $k-1$;

- $u_7(k) \equiv \frac{VIX_{cl}(k-1)-VIX_{cl}(k-2)}{VIX_{cl}(k-2)}$: the daily relative change in the volatility index value over the previous day, that is, day $k-1$.

Due to the SSWN training properties, the input variable values are scaled to the intervals $[-1, 1]$ for the variables $u_2, \ldots, u_7$, which can take both positive and negative values and $[0, 1]$ for u_1, which is positive. At the instant $k-1$, the SSWN network operates as follows: first, the new state at k is calculated from (1) as $x_i(k) = \sum_{j=1}^{L} w_{N+i,j} \Psi_j(x(k-1), u(k-1))$, $i = 1, \ldots, M$, and next the network output is calculated according to (2). Let us notice that most of the inputs are of incremental type. Therefore, the same structure of the SSWN output is assumed. Hence,

$$y(k) = \Delta EURPLN(k|k-1)$$
$$= \frac{EURPLN_{cl}(k|k-1) - EURPLN_{cl}(k-1)}{EURPLN_{cl}(k-1)}. \tag{10}$$

As the quantities $y(k)$ and $EURPLN_{cl}(k-1)$ are known at $k-1$, the forecast $EURPLN_{cl}(k|k-1)$ is calculated from (10) in a straightforward manner.

4.3. Determining SSWN parameters. Having defined the SSWN inputs and outputs, it remains to determine suitable values of the network parameters. This is done by using the historical data and searching for parameters values such that the corresponding prediction error is minimal. The procedure is called network training (Zhang *et al.*, 1998; Borowa *et al.*, 2007). The SSWN structure for one-session-ahead prediction with parameters to be optimised is illustrated in Fig. 3.

Let us denote the initial state discharge time by J. It is determined by simulation performed for representative network parameter sets. The learning data series is then composed of the historical trading sessions used to calculate the prediction errors and the initial sessions during which the initial network state discharges. Optimising the performance function $E(w, t, d)$ with respect to the parameters is performed by a simulated annealing solver. Each iteration of the solver starts from discharging the network state initial condition for the actual parameter values. This is done by running the network over the first J sessions to discharge the initial condition applied at the beginning of session 1 and determine the initial state corresponding to the actual parameters with which the prediction error over the next N session is evaluated. The parameter search can be performed by solving the following optimisation problem:

$$\min \mathbb{E}(w, t, d, L, M) = \sum_{i=1}^{N} e_i^2(w, t, d, L, M), \tag{11}$$

$$e_i = y(i) - \frac{EURPLN_{cl}(i) - EURPLN_{cl}(i-1)}{EURPLN_{cl}(i-1)}, \tag{12}$$

where i denotes the session number, e_i stands for the prediction error for session i, E signifies the total
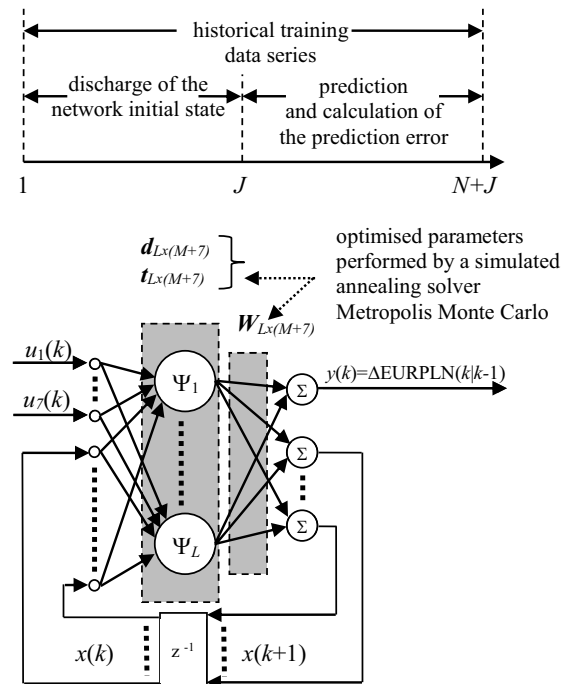

Fig. 2. Training data structure process by application of a simulated annealing solver Metropolis Monte Carlo using historical training data series to produce optimised parameters of the SSWN for one-session-ahead prediction.
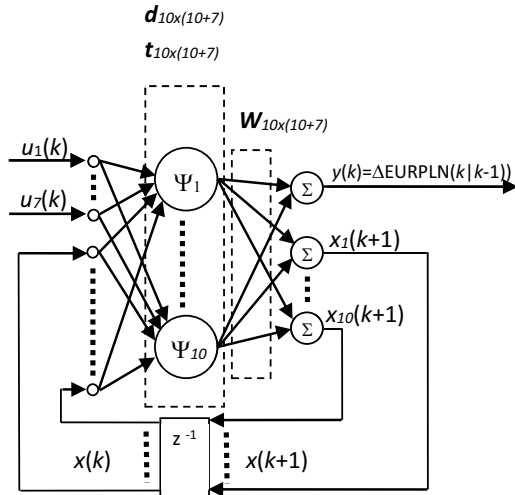
$d_{10 \times (10+7)}$

$t_{10 \times (10+7)}$

$W_{10 \times (10+7)}$

$u_1(k)$

$u_7(k)$

Ψ_1

Ψ_{10}

$y(k) = \Delta EURPLN(k|k-1))$

$x_1(k+1)$

$x_{10}(k+1)$

$x(k)$

z^{-1}

$x(k+1)$

Fig. 3. Structure of the SSWN for one session ahead prediction of the EUR/PLN exchange rate.

prediction error over N consecutive sessions. The resulting SSWN is then validated by using different data sets in order to assess its generalisation properties.

As the parameters are mixed-integer and the SSWN is described by nonlinear mapping solving, the optimisation problem is a very challenging task for any known optimisation solver. Hence, we shall separate determining the number of states M and the number of wavelons L, which are the integer valued SSWN structure parameters, from determining the continuously valued parameters of the SSWN, i.e., W, d and t. Hence, network training is structured in the form of a bi-level optimisation scheme, where at the upper level a direct intelligent search is employed to vary M and L while a dedicated powerful stochastic optimisation Metropolis Monte Carlo (MCC) algorithm (Brdyś *et al.*, 2009) supported by the simulated annealing (SA) mechanism is applied to optimise (11) and (12) under the prescribed values of M and L.

The MCC search was proposed by Metropolis *et al.* (1953). The SA component (cooling schedule) was added by Kirkpatrick *et al.* (1983). The alternative cooling schedules were proposed by Hajek (1988), Jacobson *et al.* (2005), Karafyllidis (1999) and Locatelli (2000) with some convergence analysis provided. In spite of the bi-level structuring of the solver and optimising parameters of the stochastic algorithm at the lower level, the overall computational effort was high. The good choice of the initial network structure parameters at the upper level was crucial for reducing this effort. The final structure of the one step ahead predictor is illustrated in Fig. 3 and Table 2.

Table 2. Selected structure parameters of the SSWN.

SSWN parameters	EURPLN forecast one-session-ahead
Number of inputs	7
Number of wavelons	10
Number of states	10

4.4. Adaptive prediction. A preliminary validation of the predictors on data different than those used for training have shown results not entirely satisfactory. This is mainly due to not including certain variables as the inputs, which have non-negligible influence on the predicted exchange rate value. Some of them are not included as they are not measurable; the others have not been identified. The predictor trained on a selected data set accommodates these uncertainties in the parameter values. If the uncertain inputs remain constant or they slowly vary, the predictor still performs well on different data sets. Otherwise, the parameters need to be updated on-line during the predictor operation. This leads to an adaptive predictor where initially the SSWN is trained off-line based on longer data sets as described earlier. The same training schema is then applied on-line to update the network parameters to actual values of the uncertain variables. However, the training performance function $E_k(w, t, d)$ at instant k is now modified by introducing the weights with which the prediction errors during the previous sessions $k, k-1, \ldots, k-N$, contribute to an overall prediction error over the last N trading sessions. Namely,

$$E_k(w, t, d, L, M) = \sum_{i=k-N}^{k} \omega(i) e_i^2(w, t, d, L, M), \quad (13)$$

where $\omega(i) = 2i/N$ and the i-th session e_i is defined as in (12). The weights $\omega(i)$ are linearly growing in time reaching the highest value for the last prediction error. Hence, the actual uncertainty input values are best accommodated into the resulting network parameter values $w(k), t(k), d(k)$ obtained at the instant k. The optimisation solver starts from the parameters $w(k-1), t(k-1), d(k-1)$ determined at the last time instant $k-1$.

5. Forecast combinations

The general rationale behind the use of the forecast combinations methodology in forecasting is that we do not know the true model that generates the time series of our interest. This true model is described in the econometric and statistical nomenclature as the data generating process (DGP), which is assumed to be highly complex and non-linear in its structure. The structure of the DGP is almost always not known to econometricians. Its

dynamics are often difficult to approximate by any single regression. For these reasons, single regression forecasts are very likely to be unstable over time and yield relatively poor forecasts, even if a regression is re-estimated on a timely basis. Combination of forecasts from a set of single regressions may be an attractive alternative to any single regression forecasts since it usually turns out to produce more accurate and stable forecasts over time than single regressions separately. This often happens because we approximate the complex DGP by a set of single regressions, and not by one regression only.

Similarly, the same holds for the dataset underlying the true model. Even if the structure of the DGP were known by an econometrician, the analysis would fail to achieve its goal due to data unavailability. Specifically, the data set used in regression analysis is restricted to data that are observed, can be easily and precisely quantified, and are regularly collected by some data provider. Because we do not know the structure of the DGP and we are dealing with very limited information included in the available dataset (not necessarily even being an input to the true DGP), we can only approximate the behaviour of this complex unknown system via regression analysis. Although regressions used in forecast combinations have an erroneous structure by assumption, the more regressions we use to approximate the complex non-linear DGP, the more likely it is that we approximate the true DGP with greater accuracy over a relatively longer period of time.

Particularly, below we approximate the complex DGP of the EUR/PLN series via combinations of linear regressions using 8 available explanatory variables. These regressions are estimated on the available data set. Those with the best prediction accuracy over the past trading days are used to generate final forecasts of the combinations model.

5.1. Data description and the model structure. The underlying dataset consists of 9 time series of seasonally unadjusted daily data from the period of 11/08/2011 to 14/04/2014, which is the same as in the case of the SSWN model. Before entering the regression analysis, all time series had been log-1st-pre-processed. For example, in the case of EUR/PLN, the log-1st-difference of the *EURPLN* series, denoted by *DEURPLN*, is defined as

$$DEURPLN(t) = \ln\left(\frac{EURPLN(t)}{EURPLN(t-1)}\right). \quad (14)$$

The same transformation is applied to the remaining 8 time series, *DAX*, *SPX*, *EURUSD*, *FTSE*, *PL106670*, *STOXX50*, *VIX* and *WIG20*, yielding the transformed *DDAX*, *DSPX*, *DEURUSD*, *DFTSE*, *DPL106670*, *DSTOXX50*, *DVIX* and *DWIG20*. The *DEURPLN(t)* series is a dependent variable in all regressions. This

time series is stationary for the subjected period and is characterized by the existence of volatility clusters and outliers. The 8 time series *DDAX*, *DDSPX*, *DEURUSD*, *DFTSE*, *DPL106670*, *DSTOXX50*, *DVIX* and *DWIG20* are used as regressors in regression analysis.

In total 3 different combination models using different versions of the linear autoregressive distributed lag (ARDL) model are tested for their forecast accuracy. The three ARDL regressions applied in combination models aim at capturing volatility clustering and existence of outliers in the *DEURPLN* series using a different modelling approach. ARDL regression is most suitable in the context of this analysis (over different versions of, e.g., factor models) since the underlying data set consists of a relatively small number of explanatory variables (Stock and Watson, 2004).

Each of the 3 combination models is built on one but different ARDL regression, which is modified within the given combination model by adding the 8 explanatory variables available to it. Specifically, the numerical procedure applied to any of the 3 combination models tests regressions with all possible combinations of 8 regressors being added to the base regression (initial ARDL regression with autoregressive term *DEURPLN(t − 1)* only). In this way, $\sum_{k=0}^{8} C_{k=0}^{8}$ different regressions are tested in each combination model. The base regression (regression with the smallest number of explanatory variables) of the first combination model is as follows:

$$\begin{aligned} DEURPLN(t) \\ = \rho_0 + \rho_1 DEURPLN(t-1) + \epsilon(t), \quad (15) \end{aligned}$$

where $\epsilon(t)|\vartheta(t-1) \sim N(0, \sigma^2)$ by standard assumption ($\vartheta(t-1)$ denotes the information set of all information through time $t - 1$). This assumption is used in the maximum likelihood estimation procedure for this and subsequent regressions. It will not be further examined in the present and further cases whether this particular assumption behind the regression's error term is admissible. This assumption should not discredit further forecast combination analysis for was already mentioned in the preface of Section 5 that linear regressions already have an erroneous structure because of the nonlinear complex structure of the true DGP of the *EURPLN* series.

Similarly, regression with a maximum number of regressors in the first combination model is as follows:

$$\begin{aligned} DEURPLN(t) = \rho_0 + \rho_1 DEURPLN(t-1) \\ + \sum_{i=1}^{8} \rho_{i+1} REG_i(t-1) + \epsilon(t), \quad (16) \end{aligned}$$

where $\epsilon(t)|\vartheta(t-1) \sim N(0, \sigma^2)$ by standard assumption, and $\sum_{i=1}^{8} \rho_i REG_i(t-1)$ denotes the sum of all 8 different regressors lagged by 1 trading day, multiplied by their

coefficients values: *REG1 = DDAX, REG2 = DSPX, REG3 = DEURUSD, REG4 = DFTSE, REG5 = DPL106670, REG6 = DSTOXX50, REG7 = DVIX* and *REG8 = DWIG20.*

The second combination model has a slightly changed structure in comparison to the first model, i.e., two additional dummy variables are added to the base regression in order to account for the existence of outliers in the *DEURPLN* series. By this arbitrary choice of the threshold value of 0.01 and −0.01, special treatment is given to approximately 1% of extreme observations (specifically, 31 out of 2348) in the estimation procedure. More over, this value of the threshold makes it possible to estimate all regression coefficients on samples consisting of 160 and more observations, allowing comparison between the 3 models:

$$
\begin{aligned}
DEURPLN(t) \\
= \rho_0 + \rho_1 DEURPLN(t-1) \\
+ \rho_2 DUMMY_1(t) + \rho_3 DUMMY_2(t) + \epsilon(t), \quad (17)
\end{aligned}
$$

where $\epsilon(t)|\vartheta(t-1) \sim N(0, \sigma^2)$ by assumption, and

$$
DUMMY_1 = \begin{cases} 1, & DEURPLN(t) \geq 0.01, \\ 0, & \text{otherwise,} \end{cases}
$$

$$
DUMMY_2 = \begin{cases} 1, & DEURPLN(t) \leq -0.01, \\ 0, & \text{otherwise.} \end{cases}
$$

Importantly, when forecasts for date t are generated from this type of regression further on in this paper, all dummy variables at time t are set equal to 0. The numerical procedure which tests all combinations of different sets of 8 regressors added to the above-described base regression is also applied here, in the same way as in the first combination model.

As to the third model, its regression base also consists of one regressor, as in the case of the first model; however, here the generalized autoregressive conditional heteroskedasticity, GARCH(1,1), structure is applied to the regression error term:

$$
\begin{aligned}
DEURPLN(t) \\
= \rho_0 + \rho_1 DEURPLN(t-1) + \epsilon(t), \quad (18)
\end{aligned}
$$

$$
\sigma^2(t) = \alpha_0 + \alpha_1 \sigma^2(t-1) + \alpha_2 \epsilon^2(t-1),
$$

where $\epsilon(t)|\vartheta(t-1) \sim N(0, \sigma^2(t))$ by assumption. The conditional variance $\sigma^2(t)$ (conditional on the information available at time $t-1$) is not directly observed and is assumed to change over time in the GARCH type of model, as Eqn. (18) indicates. It is assumed to depend on $\sigma^2(t-1)$ and the squared forecast error $\epsilon(t-1)$ from the previous period. All α and ρ parameters

of the model are estimated by means of maximum likelihood estimation (MLE). Specifically, the estimation incorporates an iterative procedure where the conditional variance $\sigma^2(t)$ is computed for each observation day given a set of parameters, and is inserted into the main log-likelihood function. A detailed description of the estimation procedure is given by Bollerslev (1986).

Here again, the numerical procedure testing $\sum_{k=0}^{8} C_8^k$ different regressions is applied in this model, in the same manner as in the previous two combination models.

Final forecasts obtained from the above-described three combination models depend on the set of three input parameters: r, the number of single regression forecasts pooled into the final forecast; p, the number of observations prior to forecast date that enter the estimation sample; and m, the number of one-step ahead forecasts prior to the forecast date which are taken to assess the historical forecast accuracy of a regression. Specifically, the implemented forecast combination procedure tests $\sum_{k=0}^{8} C_8^k$ different regressions for their historical forecast accuracy. Then it chooses r single regressions with the best historical forecast accuracy, estimated on the sample length of p observations and assessed on the period of m trading days prior to forecast, and pools them into one final forecast of the given combination model. The assessment of historical forecast accuracy prior to forecast date t is performed by the mean absolute error (MAE) statistics measuring the one-step-ahead forecast error over the last m trading days prior to date t:

$$
\begin{aligned}
MAE \\
= \frac{1}{m} \sum_{s=t-m-1}^{t-1} |\widehat{DEURPLN}(s; p, m, r)|\vartheta(s-1) \\
- DEURPLN(s)|, \quad (19)
\end{aligned}
$$

where $\widehat{DEURPLN}(s; p, m, r)|\vartheta(s-1)$ is the forecast value at time s given all information available at $s-1$ and parameter set values (p, m, r).

Next, r single regressions with lowest MAE out of all $\sum_{k=0}^{8} C_8^k$ different regressions are chosen to yield the final forecast of the combination model $\widehat{DEURPLN}(s; p, m, r)^{FC}|\vartheta(t-1)$ for the trading day t, given parameter values (p, m). This final forecast of the combination model is computed as the simple arithmetic average of forecasts from r regressions.

Finally, the overall forecast performance of each combination model is evaluated for the period of 01/01/2013 to 14/04/2014 also by means of the MAE

Table 3. Validation results of the accuracy of the one-session-ahead predictor.

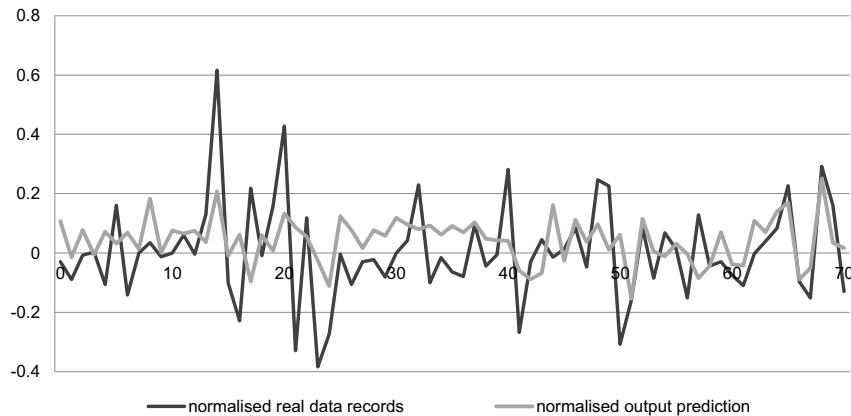| Prediction accuracy measure	Units
Minimum prediction error	0.80129%
Average prediction error scaled to range $+1/-1$	0.63561
Annualised standard deviation of the EUR/PLN exchange rate	35.34%

Fig. 4. Results of validation of the one-session-ahead predictor using normalised real data records of the EUR/PLN exchange rate and normalised forecasted output prediction of the EUR/PLN exchange rate for 73 trading sessions ahead.

statistic, which is defined as

$$MAE = \sum_{t=01/01/2013}^{14/04/2014} |\widehat{DEURPLN}(t;r)^{FC}|\vartheta(t-1) - DEURPLN(t)|. \quad (20)$$

In each of the combination models, m is set to be $7, 9, 11$ while p is set to be equal to $160, 180, 200, 220, 240$. Results of the overall forecast performance of combination models are compared for values of r equal to $1, 3, 5, 7, 9$. By pooling r forecasts from single regressions into one final forecast, combination models are expected to take advantage of forecast averaging. Some significant findings in the domain of applied econometrics on forecast combinations show that forecast averaging improves forecast accuracy of regressions; in particular,

1. combining forecasts from different single regressions is likely to improve forecasts as compared with any of the single regressions treated separately (Timmermann, 2006; Guidolin and Timmermann, 2009; Hyndman *et al.*, 2011; Kawakami, 2013);

2. combining forecasts produced by single regressions, which are estimated on the same time series but on samples of different length, is likely to improve forecasts as compared with any of the single regressions treated separately (McCracken and Clark, 2009; Pesaran and Pick, 2011);

3. combining forecasts produced by single regressions by means of simple weights often improves forecasts as compared with any of the single regressions treated separately (Rapach *et al.*, 2010; Tian and Anderson, 2014).

6. Forecast results

6.1. SSWN model. The one-session-ahead predictor was validated based on the data records composed of 73 subsequent sessions over the period of 01/01/2013 to 14/04/2014. The SSWN-based forecasts were validated on the basis on real data series recorded by the European Central Bank. The summary of the results is displayed in Table 3, validated on the accuracy and volatility of one-session-ahead predictor. Further results obtained are displayed in Fig. 4 for the one-session-ahead predictor using normalised real data records and normalised output prediction for a 73-session frame.

The validation results show that the predictor can be effectively used to perform an on-line one-session-ahead forecasts of the EUR/PLN exchange rate. Due to a high volatility exhibited in the forecast range of the EUR/PLN exchange rate, volatility jumps have decreased on-line accuracy. Indeed, the SSWN is unable to predict volatility jumps but rather adjusts accordingly on-line to efficiently accommodate predictions and learn intelligently based on current variable inputs. The historical input data used in the SSWN model were taken as actual real time series daily data from 15/08/2011 to 31/12/2013 for *EURPLN*,

PL106670, *WIG20*, *DAX* and *VIX*, comprising 673 daily time series data. This input range is rather too long for the SSWN and results suggest a prediction error higher than average. In fact, a much better data range is 72 sessions prior to the forecasted period of daily 73 sessions. This is due to the phenomenon of overloading the SSWN, which decreases the wavelons processing power during off-line learning period.

Overall, the structure of the SSW over the 73 sessions frame is constant. On the other hand, forecast combinations require a larger pool of historical data to average out volatility jumps and to optimize the unknown model based on the selected time series range. In fact, they complement the SSWN predictions although do not increase the insight of a risk manager as to what the risk drivers are. The non-linearity component is efficiently modelled by using rates of change as inputs in the SSWN structure as well as incorporating a EUR/PLN $t - 1, \ldots,$ $t - 2$ loop with increasing efficiency and accuracy. On the other hand, forecast combinations use logarithmic returns as input variables due to the statistical stability of the modelled returns.

6.2. Forecast combinations model. Forecast results generated by combination models 1–3 are depicted in Fig. 5. Results are presented for the period of 01/01/2013–14/04/2014 in terms of MAE$\times 10^2$ of the one-step-ahead forecast of the *DEURPLN* series. The MAE refers to the mean absolute error of daily EUR/PLN log-1st-difference forecast. It is calculated from (20). The horizontal axis depicts r parameter values denoting the number of best regressions entering the final forecast of the combination model. The remaining two parameters, m (the number of one-step-ahead forecasts prior to the forecast day that are taken to assess the historical forecast accuracy of a regression) and p (the number of observations prior to forecast date that enter the estimation sample), are automatically chosen by the forecast routine
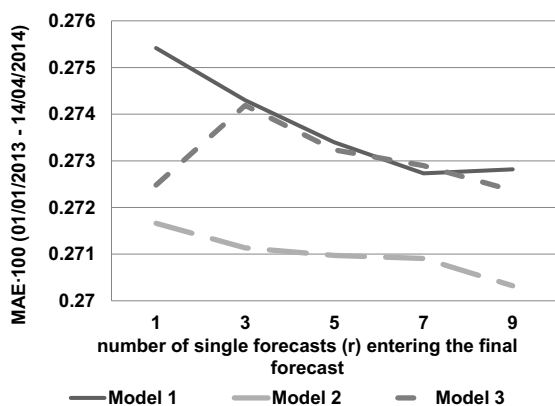

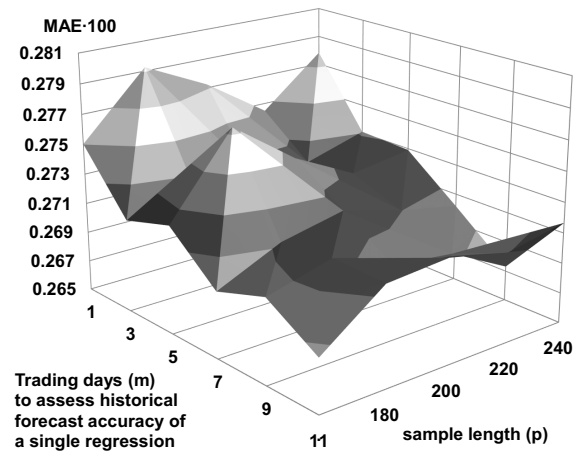Fig. 5. Forecast results of combinations models 1–3.

Fig. 6. Forecast results for combinations model 2, $r = 1$, m (number of trading days) and p (sample length) varying.

in each of the 3 combination models.

Figure 5 clearly indicates that model 2 outperforms models 1 and 3 in terms of forecast accuracy, producing MAE of forecast ranging from 27.17×10^{-4} for $r = 1$ to 27.03×10^{-4} for $r = 9$. The differences between the models' predictive power are relatively small; however, model 2 produces consistently better results for all r parameter values for this data set. Model 2 is found to produce the best forecasts in this dataset most likely due to high variability and existence of many outliers in the *DEURPLN* series.

Another interesting observation derived from Fig. 5 is that the MAE statistics of the final forecast results generally decrease with increasing values of r. This pattern in the final forecast results means that the pooling of the best forecasts from different single regressions or the same regressions estimated over different samples improves forecast accuracy of combination models.

Partial results from combination models presented in Fig. 6 allow closer analysis of forecast results with respect to the m and p parameters. Here, parameter r is set equal to 1, indicating that only one best regression is considered for the final forecast. As a general rule, higher forecast accuracy is obtained for relatively larger estimation samples (parameter p) and longer samples on which the predictive power of each regression is assessed (parameter m). Partial forecast results for the $r = 1$ range from 27.93×10^{-4} for $m = 1$, $p = 180$ to 26.73×10^{-4} for $m = 9$, $p = 24$. Due to the relatively high variability of partial model results over parameter pairs (m, p), forecast averaging is advantageous.

7. Conclusions

The results obtained so far are conclusive in adaptive nature under high volatility conditions of a multi-input

forecasted financial instrument EUR/PLN exchange rate. More simulation results and comparisons with other prediction methods are subject to further research. A novel research work is under way on application of model predictive control to on-line risk-based optimised portfolio management. The on-line predictions of financial instruments are the inputs to the model-based optimisation problem of MPC. Their accuracy is crucial for the quality of the decision produced by solving this problem in spite of the powerful receding horizon mechanism of MPC, which helps to reject the uncertainty. Currently, the SSWN and FC methods are integrated in order to optimally utilise their advantages. However, this still requires a better understanding of the advantages and disadvantages of these two approaches when applied to this application-driven problem. For the time being, we shall conclude as follows.

7.1. SSWN model. The paper has considered the prediction of the EUR/PLN currency exchange rate by using an adaptive dynamic state space wavelet network model of the exchange rate. The approach can be applied to developing tools for predicting changes of other financial indicators. The EUR/PLN exchange rate has been effectively forecasted during increased volatility periods. The main advantage of the SSWN is that it does not compose any statistical constraints or assumptions in generating credible predictions and therefore is ideal for modelling financial time series instruments in volatile market conditions. The SSWN methodology accommodates for on-line learning by adaptive learning process and searches for the best an solution of w, t, d parameters adding to the accuracy and robustness of forecasting credibility. Correlation changes between output and input variables are adjusted daily by optimising the function $E(w, t, d)$ with respect to the parameters outlined in this paper. Indeed, the non-linearity phenomenon is a more realistic reflection of financial markets returns and price movements derived from globalisation growth and integration.

Clearly, on-line update of the SSWN structure would improve the forecast performance. This would be done based on the moving measurement data window on the basis of a suitably defined indicator of the current structure potential. As the SSWN is a universal approximator, this update would be done in the time scale much slower than the SSWN parameter update fit scale, as opposed to the FC method. This is subject to future research work. The SSWN model non-parametric assumptions incorporate this characteristic with elegance and sophistication. Currencies today are subjected to more than ever changing variables, often fluctuating daily in magnitude, weights and degree of correlation. The SSWN model can deal with this scenario and act as an early warning system anticipating changes as they happen,

leading to an ex ante risk management approach.

7.2. Forecast combinations. In order to improve the results of the presented combination models, one could either alter the regression model underlying our combination models or change the set of regressors in the analysis or apply another selection criterion for the choice of the set of regressions entering the final forecast of the combination model. Introducing changes to the construction of the selection criterion is our favoured direction towards improving the predictive power of combination models, since it is considerably less dependent on the underlying dataset as compared to the structure of the regression model. As to the selection criterion defined in this paper, instead of choosing the model with the lowest MAE over the last m trading days, one could also take into account, e.g., regressions that yield higher but constant or switching forecast errors over the period of length m and/or pick up regressions which best point to the sign of the *DEURPLN* series in the period ahead. As well as that, specific changes to the algorithm choosing the best single regressions for the final forecast can also be made. This algorithm can be extended by the use of a more general autoregressive criterion of choice, where in-sample forecast errors are weighted differently with respect to their proximity to the final forecast date.

References

BIS (2013). *Foreign Exchange Turnover in April 2013: Preliminary Global Results*, Triennial Central Bank Survey, Bank of International Settlements, Basel.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**(3): 307–327.

Borowa, A., Brdyś, M.A. and Mazur, K. (2007). Modelling of wastewater treatment plant for monitoring and control purposes by state-space wavelet networks, *International Journal of Computers, Communications & Control* **2**(2): 121–131.

Brdyś, M.A., Borowa, A., Idźkowiak, P. and Brdyś, M.T. (2009). Adaptive predictions of stock exchange indices state space wavelet networks, *International Journal of Applied Mathematics and Computer Science* **19**(2): 337–348, DOI: 10.2478/v10006-009-0029-z.

CSO (2014). *Foreign Trade in January–December 2013*, Statistical Information and Elaborations, Central Statistical Office, Warsaw.

Grossmann, A. and Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape, *SIAM Journal on Mathematical Analysis* **15**(4): 723–736.

Guidolin, M. and Timmermann, A. (2009). Forecasts of US short-term interest rates: A flexible forecast combination approach, *Journal of Econometrics* **150**(2): 297–311.

Hajek, B. (1988). Cooling schedules for optimal annealing, *Mathematics of Operations Research* **13**(2): 311–329.

Hyndman, R.J., Ahmed, R.A., Athanasopoulos, G. and Shang, H.L. (2011). Optimal combination forecasts for hierarchical time series, *Computational Statistics and Data Analysis* **55**(9): 2579–2589.

Jacobson, S.H., Hall, S.N., Mclay, L.A. and Orosz, J.E. (2005). Performance analysis of cyclical simulated annealing algorithms, *Methodology and Computing in Applied Probability* **7**(2): 183–201.

Karafyllidis, I. (1999). A simulator for single-electron tunnel devices and circuits based on simulated annealing, *Superlattices and Microstructures* **25**(4): 567–572.

Kawakami, K. (2013). Conditional forecast selection from many forecasts: An application to the yen/dollar exchange rate, *Journal of the Japanese and International Economies* **28**(C): 1–18.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. (1983). Optimization by simulated annealing, *Science* **220**(4598): 671–680.

Kulawski, G.J. and Brdyś, M.A. (2000). Stable adaptive control with recurrent networks, *Automatica* **36**(1): 5–22.

Kuo, R.J., Chen, C.H. and Hwang, Y.C. (2001). An intelligent stock trading support system through integration of genetic algorithm based fuzzy neural network and artificial neural network, *Fuzzy Sets and Systems* **118**(1): 21–45.

Locatelli, M. (2000). Convergence and first hitting time of simulated annealing algorithms for continuous global optimization, *Mathematical Methods of Operations Research* **54**(2): 171–199.

McCracken, M.W. and Clark, T.E. (2009). Improving forecast accuracy by combining recursive and rolling forecasts, *International Economic Review* **50**(2): 363–395.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equations of state calculations by fast computing machines, *Journal of Chemical Physics* **21**: 1087–1092.

NBP (2013). *Turnover in the Domestic Foreign Exchange and OTC Derivatives Markets in April 2013*, National Bank of Poland, Warsaw.

Pesaran, M.H. and Pick, A. (2011). Forecast combination across estimation windows, *Journal of Business Economics and Statistics* **29**(2): 307–318.

Qi, R. and Brdyś, M.A. (2008). Stable indirect adaptive control based on discrete-time T–S fuzzy model, *Fuzzy Sets and Systems* **159**(8): 900–925.

Qi, R. and Brdyś, M.A. (2009). Indirect adaptive control based on a self-structuring fuzzy system for nonlinear modelling and control, *International Journal of Applied Mathematics and Computer Science* **19**(4): 619–630, DOI: 10.2478/v10006-009-0049-8.

Rapach, D.E., Strauss, J.K. and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *The Review of Financial Studies* **23**(2): 821–862.

Stock, J.H. and Watson, M.W. (2004). Combination forecasts of output growth in a seven-country data set, *Journal of Forecasting* **23**(6): 405–430.

Tian, J. and Anderson, H.M. (2014). Forecast combinations under structural break uncertainty, *International Journal of Forecasting* **30**(1): 161–175.

Timmermann, A. (2006). Forecast combinations, *in* G. Elliott *et al.* (Eds.), *Handbook of Economic Forecasting*, Elsevier, Amsterdam.

Tsang, P.M., Kwok, P., Choy, S.O., Kwan, R., Ng, S.C., Mak, J., Tsang, J., Koong, K. and Wong, T. (2007). Design and implementation of NN5 for Hong Kong stock price forecasting, *Engineering Applications of Artificial Intelligence* **20**(4): 453–461.

Zammarreno, J.M. and Pastora, V. (1998). State space neural network: Properties and applications, *Neural Networks* **11**(6): 1099–1112.

Zhang, G., Patuwo, B.E. and Hu, M.Y. (1998). Forecasting with artificial neural networks: The state of the art, *International Journal of Forecasting* **14**(1): 35–62.

Zhang, Q. (1992). Wavelet network: The radial structure and an efficient initialization procedure, *Technical Report LiTH-ISY-I-1423*, Linkoping University, Linkoping.

Zhang, Q. and Beneveniste, A. (1992). Wavelet networks, *IEEE Transactions on Neural Networks* **3**(6): 889–898.

Mietek A. Brdyś received the M.Sc. degree in electronic engineering as well as the Ph.D. and D.Sc. degrees in control systems from the Institute of Automatic Control at the Warsaw University of Technology in 1970, 1974 and 1980, respectively. From 1974 to 1983, he held the posts of an assistant professor and an associate professor at the Warsaw University of Technology. In 1992 he became a full professor of control systems. Between 1978 and 1995, he held various visiting faculty positions at the University of Minnesota, City University, De Montfort University and the University Polytechnic of Catalunya. From 1989, he held the post of a senior lecturer in the School of Electronic, Electrical and Computer Engineering at the University of Birmingham. From 2001 he held the post of a full professor of control systems in the Department of Automatic Control at the Gdańsk University of Technology. In 2008 he founded there the Department of Control Systems Engineering and became its head. He served as a consultant for the Honeywell Systems and Research Centre in Minneapolis, GEC Marconi and Water Authorities in the UK, France, Spain, Germany and Poland. He was the head of the Interdisciplinary Research Network on Decision Support and Control Systems at the University of Birmingham and the head of the Intelligent Decision Support and Control System Group at the Gdańsk University of Technology. His research was supported by the UK and Polish research councils, industry and the European Commission. He was the author and co-author of about 220 refereed papers and six books. His last research included intelligent decision support and control of complex uncertain systems, robust monitoring and control, and softly switched robustly feasible model predictive control. The applications included environmental systems, technological processes, autonomous intelligent vehicles and defence systems. He was a chartered engineer, a member of the IEE, a senior member of the IEEE, a fellow of the IMA and the chair of the IFAC Technical Committee on Large Scale Complex Systems. He was the IPC chair of the 11th IFAC Symposium on *Large Scale Complex Systems* (Gdańsk, 2007). Professor Brdyś passed away in July 2015.

Marcin T. Brdyś received a B.Sc. degree in economics and an M.B.A. from the University of Wales Cardiff Business School in 1997 and 1999, respectively. Since 2000 he has gained extensive experience working in European financial hubs, notably the City of London, Frankfurt and Warsaw. That included work for some of the industries' most revived financial institutions, such as Schroder Salomon Smith Barney, European Central Bank, Fimat Societe Generale International Banque SA and Pekao SA/Unicredit. His specialisation is in risk management and pricing currency, interest rates and equity derivatives. In 2008 he attained the internationally recognized *Financial Risk Manager Certified* qualification.

Sebastian M. Maciejewski received an M.Sc. degree in finance from the Vrije University of Amsterdam, a Dipl.-Vw. degree in economics from the University of Mannheim, and an M.A. degree from the Warsaw School of Economics. He is a *Certificate in Quantitative Finance* (CQF) designation holder. He has worked as an analyst at BRE Bank, Bank Peako, PwC and PKO Bank Polski in Warsaw, developing operational (LDA), market (VaR) and credit (CVaR) risk models, pricing exotic derivatives and preparing short-term macro-econometric forecasts.