UNIVERSITÀ DEGLI STUDI ROMA TRE

COLLANA DEL DIPARTIMENTO DI ECONOMIA

FORECASTING WITH GARCH MODELS UNDER STRUCTURAL BREAKS: AN APPROACH BASED ON COMBINATIONS ACROSS ESTIMATION WINDOWS

Davide De Gaetano

ISSN 2279-6916 Working papers

(Dipartimento di Economia Università degli studi Roma Tre) (online)

Working Paper n° 219, 2017

I Working Papers del Dipartimento di Economia svolgono la funzione di divulgare tempestivamente, in forma definitiva o provvisoria, i risultati di ricerche scientifiche originali. La loro pubblicazione è soggetta all'approvazione del Comitato Scientifico.

Per ciascuna pubblicazione vengono soddisfatti gli obblighi previsti dall'art. l del D.L.L. 31.8.1945, n. 660 e successive modifiche.

Copie della presente pubblicazione possono essere richieste alla Redazione.

REDAZIONE:
Dipartimento di Economia
Università degli Studi Roma Tre
Via Silvio D'Amico, 77 - 00145 Roma
Tel. 0039-06-57335655 fax 0039-06-57335771
E-mail: dip_eco@uniroma3.it
http://dipeco.uniroma3.it

DIPARTIMENTO DI ECONOMIA

FORECASTING WITH GARCH MODELS UNDER STRUCTURAL BREAKS: AN APPROACH BASED ON COMBINATIONS ACROSS ESTIMATION WINDOWS

Davide De Gaetano

Forecasting with GARCH models under structural breaks: an approach based on combinations across estimation windows

Davide De Gaetano [*]

Department of Economics

University of Roma Tre (Italy)

davide.degaetano@uniroma3.it

Abstract

This paper proposes some weighting schemes to average forecasts across different estimation windows to account for structural changes in the unconditional variance of a GARCH (1,1) model. Each combination is obtained by averaging forecasts generated by recursively increasing an initial estimation window of a fixed number of observations v. Three different choices of the combination weights are proposed. In the first scheme, the forecast combination is obtained by using equal weights to average the individual forecasts; the second weighting method assigns heavier weights to forecasts that use more recent information; the third is a trimmed version of the forecast combination with equal weights where a fixed fraction of forecasts with the worst performance are discarded. Simulation results show that forecast combinations with high values of v are able to perform better than alternative schemes proposed in the literature. An application to real data confirms the simulation results

Keywords: Forecast combinations, Structural breaks, GARCH models.

JEL Code: C530, C580, G170

[*]I wish to thank Professor M. Barbieri for her useful suggestions and comments. I am also grateful to Professor G. De Luca and Professor A. Parisi for their insightful and stimulating comments and suggestions on the first draft which led to an improved final version. The usual disclaimer applies.

1 Introduction

The Generalized Autoregressive Conditional Heteroskedastic model GARCH (p,q) has been found to be a particularly useful parametrization in modelling and forecasting financial time series. Numerous applications of GARCH models have already appeared in the statistical and econometric literature, since the seminal papers by Engle and Bollerslev (1986).

Many studies (see for example Andreou and Ghysels, 2009) agree that the existence of structural breaks in the volatility can be the cause of estimation problems and forecast failures. If structural breaks are present in the data generating process but they are not considered in the specification of the model, the analysis could be biased toward a spurious persistence.

The hypotheses that structural changes may misinterpret the persistence estimation in GARCH models has been first highlighted by Diebold (1986). Lamoureux and Lastrapes (1990) have confirmed the Diebold's idea by allowing for changing states of the constant term of the conditional variance of a GARCH(1,1) model.

Mikosch and Stărică (2004) have explained how changes in the unconditional variance can be the cause of two stylized facts observed in long log-return series: the long-range dependence in volatility and the integrated GARCH (IGARCH) effect. In the first case the sample autocorrelation functions of the absolute values of the log returns series and their squares are all positive, decay relatively fast at the first lags and tend to stabilize around a positive value for larger lags. Concomitantly, the periodograms blow up at frequencies near zero. In the second case, whereas for shorter samples the estimated parameters of the GARCH(1,1) model sum to values significantly different from one, for longer samples their sum becomes close to one. This motivated the introduction of the integrated GARCH(1,1) model, IGARCH(1,1) by Engle and Bollerslev (1986) as a possible generating process for returns. As a consequence, the fitted GARCH model often appears to be very close to an IGARCH model only because structural breaks in unconditional volatility are ignored.

Hillebrand (2005) has shown that if structural changes in the conditional volatility process are not taken into account in the model specification, they could cause a substantial overestimation of the autoregressive parameters of the conditional variance.

The presence of structural breaks is even more fateful in a forecasting context and it constitutes one of the primary reasons for forecast failures in practice.

In particular, West and Cho (1995) have shown that a GARCH (1,1) that allows for structural breaks in the unconditional variance, could have better forecasting performances with respect to methods which do not take into account their presence. This result has been confirmed by Stărică *et al.* (2005) in the context of long-horizon forecasts of stock return volatility and by Rapach and Strauss (2008) for eight daily U.S. dollar exchange rate return series.

A common strategy to handle parameter instability, in the context of GARCH models, is to select a single estimation window in which only a fraction of the most recent observations is used to estimate the parameters and to generate the forecasts. However, while in the regression framework solutions have been proposed to derive an optimal estimation window size (Pesaran and Timmerman, 2007; Pesaran *et al.*, 2013; Giraitis *et al.*, 2013; Inoue *et al.*, 2017), in a GARCH context this is not an easy task. One possible approach is to identify the last structural break in the series and to use only observations over the post-break period. This procedure could be not reliable in many applications because it depends on the correct identification of the last break, which is unknown in terms of location and magnitude.

Another approach is to use an estimation window whose size is proportional to the number of observations in the sample, independently of the last break location (Rapach and Strauss, 2008). However the forecasting performance is sensitive to the choice of the observation window due to the bias-variance trade off. More precisely a relative long estimation window reduces the forecast error variance but increases its bias; on the other hand, a short estimation window produces an increase in the forecast error variance although the bias decreases.

In order to solve the problems arising with the choice of a single estimation window, it can be useful to consider forecast combinations generated by the same model but over different estimation windows. As highlighted by Pesaran and Pick (2011) in the case of random walks and for a linear regression model, this strategy could be superior to forecasts generated by a single estimation window. However, while this procedure is suitable in a regression context and has been widely used in the econometric literature, in the context of GARCH models some problems arise. They are related to the very high length of the involved series, to the more complex generation of forecasts and to the possible large estimation uncertainty.

The aim of this paper is to propose some alternative weighting schemes that, in the spirit of Pesaran and Timmerman (2007), average forecasts across estimation windows, in order to account for structural breaks in the unconditional variance of a GARCH (1,1) model. Common to all the proposed forecast combinations is that the individual forecasts are obtained by expanding the length of an initial estimation window backwards of a fixed number of observations. The length of the initial window is introduced to allow a convergent estimation of a GARCH model whereas the length of the expanding windows controls the number of individual forecasts entering into the combination. The proposed combinations are usually effective to account for structural breaks even in the common situation when their location and magnitude are not known a priori. At the same time, they ensure low computation costs, avoiding the estimation of thousands of GARCH models that could be obtained when all possible window sizes are used. The proposed schemes differ for the specification of the combination weights: in the first scheme, the forecast combination is obtained by using equal weights to average the individual forecasts; the second weighting method simply assigns heavier weights to forecasts that use more recent information; the third is a trimmed version of the forecast combination with equal weights where a fixed fraction of forecasts with the worst performances are discarded.

The paper is organized as follows. Section 2 introduces the problem of forecasting under structural breaks and briefly overviews the most used methods. Starting from the single estimation window approach, the use of forecasting combinations is addressed, highlighting their advantages. In Section 3, the proposed forecast combinations across estimation windows are illustrated and discussed. In section 4 a Monte Carlo experiment has been implemented. The aim is to evaluate how the performance of the proposed forecast combinations, with different values of the tuning parameter and different choices of the weights, are influenced by the location and the size of the break. Moreover, through the implemented experiment, the forecast combinations have been also compared to some alternative schemes proposed in the literature. The forecasting performances of all the considered forecast combinations have been compared in terms of a loss functions, through the model confidence set procedure, proposed by Hansen *et al.* (2011). In section 5 the effectiveness of the proposed forecast combinations is evaluated on the same data set used by Rapach and Strauss (2008). It consists of daily returns of the U.S. dollar against the currencies of Canada, Denmark, Germany, Japan, Norway, Switzerland and the daily returns corresponding to the U.S. trade-weighted exchange rate. Some final remarks close the paper.

2 Forecast methods

In the paper we will refer to a GARCH (1,1) model since it is a very parsimonious model and usually it is adequate to obtain good performances in terms of fitting and forecasting (Hansen and Lunde, 2005).

The canonical GARCH (1,1) model for the zero-mean series a_t can be expressed in the form:

$$a_t = h_t^{0.5}\epsilon_t \tag{1}$$

with

$$h_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 h_{t-1} \tag{2}$$

where $\{\epsilon_t\}$ is a sequence of *i.i.d.* random variables with zero mean and unit variance. Conditions on α_0, α_1 and β_1 need to be imposed for the previous equation to be well defined. In particular,

$\alpha_0 > 0$ and $\alpha_1, \beta_1 \geq 0$ are imposed to ensure that the conditional variance h_t is positive. Moreover $\alpha_1 + \beta_1 < 1$ ensures that the process is stationary. For a GARCH (1,1) process, the unconditional variance is defined as $\sigma^2 = \alpha_0/(1 - \alpha_1 - \beta_1)$.
The parameters are estimated by using the Quasi Maximum Likelihood Estimation in which the likelihood corresponding to the assumed distribution of ϵ_t, is maximized under the previous assumptions. In the case of stable GARCH(1,1) model, the one-step-ahead forecast at time $T+1$, obtained by using the observations from 1 to T is defined as

$$\widehat{h}_{T+1} = \widehat{\alpha_0} + \widehat{\alpha_1} a_T^2 + \widehat{\beta}_1 h_T \tag{3}$$

As it is well known, the presence of structural breaks can have serious consequences on the performance of the forecasts. To deal with this problem, many procedures have been suggested in the econometric literature.
One of the most popular approach, in the context of GARCH models, is to select a single estimation window to generate a single forecast. In this perspective, a possible approach is to identify the presence of potential breaks in the data, by means of the numerous tests proposed in the econometric literature, and to predict with a GARCH model estimated using only the data from the last break (Rapach and Strauss, 2008). However, this approach is not free from criticism. First of all, the detection of breaks could be imprecise especially in the context of financial time series in which the classical tests proposed in the literature could not distinguish between break points and extreme observations (Ross, 2013). Secondly, as pointed out by Pesaran and Timmerman (2007) in the regression framework, the forecasts generated by this scheme is likely to be biased and may not minimize the mean square forecast error even when the last break date is correctly detected. This suggests that the pre-break observations could be useful for forecasting even after the break. Moreover, when the last break is small and the variance parameter increases at the break point it could be advisable to use also a fraction of the pre-break observations. Finally, if the last detected break point is very close to the last observation, the sample used for the estimation of the parameters of the forecast model is relatively short causing a very large estimation uncertainty. This is a strong weakness in GARCH model in which the parameters estimations need quite a high number of observations to converge.
Alternatively it is possible to consider forecasting methods which implicitly take into account the presence of breaks. In these cases, it is not necessary to specify the number and the locations of the breaks but they are taken into account in the generation of forecasts by using a GARCH model estimated with a fixed sample size. The forecasting performance of these methods is obviously sensitive to the choice of the observation window which should be chosen in a way to balance the trade off between bias and

variability. More specifically, using data from different data generating processes, may lead to biased parameter estimates and forecasts; such bias can accumulate leading to large mean square forecast errors. This bias could be reduced using only that data relevant to the present. On the contrary, in order to reduce the heterogeneity, it could be appropriate to reduce the sample but this could increase the variance of the parameter estimates. This increase in variance maps into the forecast errors and, again, it causes an increment of the mean square forecast error. Hence, a plausible estimation window should guarantee an accurate estimate of the GARCH parameters and a not extensively use of observations from different regimes (Clark and McCracken, 2009).

However, in this approach, the identification of the optimal estimation window size is not an easy task. In the context of regression, Pesaran and Timmermann (2007) propose some methods to select the window size in the case of multiple discrete breaks when the errors of the model are serially uncorrelated and the regressors are strictly exogenous; Pesaran *et al.* (2013) derive optimal weights under continuous and discrete breaks in the case of independent errors and exogenous regressors; Giraitis *et al.* (2013) propose to select a tuning parameter to downweight older data by using a cross-validation-based method in the case of models without regressors; Inoue *et al.* (2017) suggest to choose the optimal window size that minimizes the conditional mean square forecast error (MSFE).

For the GARCH models the selection of this single estimation window remains an open issue and in many empirical studies, it is arbitrarily determined. For example, Rapach and Strauss (2008) propose different specifications for this parameter; in particular, in their analysis, the estimation window size is fixed to one-half and one-quarter of the length of the in sample period.

To overcome the problem of the selection of a single estimation window, an alternative strategy can be used. It is based on combinations of forecasts generated from the same model but over different estimation windows.

2.1 Forecast combinations

Starting from the pioneeristic paper of Bates and Granger (1969), there has been a growing interest, in the econometric literature, for the combinations of forecasts obtained by estimating a number of alternative models over the same sample period. Clemen (1989) has provided a review highlighting that forecast accuracy can be substantially improved through the combination of multiple individual forecasts. Makridakis and Hibon (2000), by analysing 3003 time series, have concluded that the accuracy of forecast combinations have better performances on average with respect to each specific method which is present in the combination. The same conclusions have been confirmed in the papers by Stock and Watson (2001,2004) and Marcellino (2004) in the context of economic and financial variables. More recently, Timmerman (2006) has theoretically analysed the factors that determine the advantages in combing forecasts focusing, in particular, on model misspecification, instability and estimation error.

In a regression framework, Pesaran and Timmerman (2007) have proposed forecast combinations formed by averaging across forecasts generated by using all possible window size subject to a minimum length requirement. Based on the same idea, more complex forecasting schemes have been proposed (see, for example, Tian and Anderson, 2014 and Pesaran *et al.*, 2013).

The idea of forecast averaging over estimation windows has been fruitfully applied also in macroeconomic forecasting, in particular in the context of vector autoregressive models with weakly exogenous regressors (Assenmacher-Wesche and Pesaran, 2008; Pesaran *et al.*, 2009), and in the context of GDP growth on the yield curve (Schrimpf and Wang, 2010).

Pesaran and Pick (2011) have discussed the theoretical advantages of using such combinations considering random walks with breaks in the drift and volatility and a linear regression model with a break in the slope parameter. They have shown that averaging forecasts over different estimation windows leads to a lower bias and root mean square forecast error than forecasts based on a single estimation window for all but the smallest breaks. Similar results are reported in Clark and McCracken (2009); they have highlighted that, in presence of structural breaks, averaging forecasts obtained by using all the observations in the sample and forecasts obtained by using a window can be useful for forecasting[1].

All the discussed approaches are feasible for linear regression models and moderate sample size. However, when dealing with GARCH models, they became not suitable because of the estimation of thousands models just to form a single combination forecast.

In the light of this result, Rapach *et al.* (2008), in the case of stock return volatility, have proposed the following forecast combinations which they have found to be especially effective for forecasting in presence of breaks.

Rapach and Strauss forecast combination.

This scheme, labelled as RS combined, is a combination which averages individual forecasts obtained with four different estimation windows. The first two have fixed sizes proportional to the number of observations in the sample; in particular they are equal to one-half and one-quarter of the size of the sample period; the third window is formed by observations from the final break point to the end of sample. In this case, the last break has been identified by using the test proposed by Sansó *et al.* (2004). The last estimation window uses all the observations in the sample.

[1] In this case forecasts from only two different windows are combined and so this procedure can be seen as a limited version of that proposed by Pesaran and Timmermann (2007)

Trimmed Rapach and Strauss forecast combination.

This scheme, denoted as Trimmed-RS combined, is the average of the individual forecasts after excluding the highest and lowest from the previous scheme.

Clark and McCracken forecast combination.

The last scheme, denoted as CM combined, is the average of two individual forecasts: the first uses all the observations in the sample and the second only the observations from 0.75T to T. This is in the spirit of Clark and McCracken (2009).

3 Forecast combinations across estimation windows in GARCH(1, 1) models

In this section we propose some alternative forecast combinations which use different estimation windows to account for structural breaks, in the context of GARCH models.

Let T be the number of observations in the sample and ω the minimum acceptable estimation window size. The forecast combination at time $T+1$, denoted with $\widehat{h}_{T+1}$, is obtained by averaging across forecasts generated by increasing recursively, of a fixed number v, the minimum estimation window ω. More precisely:

$$\widehat{h}_{T+1} = \sum_{\tau=0}^{k-1} c_\tau \widehat{h}_{T+1}^{[T-w-\tau v:T]} \tag{4}$$

where $\widehat{h}_{T+1}^{[T-w-\tau v:T]}$ is the one-step-ahead forecast obtained by using the observations from $(T-w-\tau v)$ to T; c_τ are combination weights and

$$k = \left\lceil \frac{T-\omega}{v} \right\rceil \tag{5}$$

being $\lceil x \rceil$ the smallest integer greater than or equal to x.

In the equation (4), the individual forecasts are generated by expanding the length of the estimation window backwards of v observations after reserving the most recent ω observations. Therefore, in the combination, the last ω observations are used in all the forecasts, whereas the observations at the beginning of the sample are used less.

The proposed forecast combination scheme depends on the parameters ω and v which have to be fixed and to the weights c_τ.

The estimation windows should not be smaller than a minimum length ω. This parameter should be set in such a way that allows the parameter estimation of the GARCH (1,1) model to converge. Hwang and Valls Pereira (2006) have shown that the estimates of the popular GARCH(1,1) model are significantly negatively biased in small samples and that in many cases converged estimates are not possible with Bollerslev's non-negativity conditions. Considering the size of biases and convergence errors, they have proposed at least 500 observations for GARCH(1,1) models.

The parameter v controls the number of observations which are added to the minimum estimation window ω and, as a consequence, to the number of individual forecasts which enters in the combination. The lower the value of v, the more individual forecasts enter in the combination. Moreover if the location of the last break is near to the end of the sample, the higher the value of v, the less the number of windows containing many pre-break observations are in the combination scheme. This could be advantageous since the forecasts generated by using many pre-break observations could be biased especially when the size of the breaks is high. Note that the logic behind this approach is similar to that proposed by Pesaran and Timmermann (2007) in a regression context. In this latter case, the parameter v is set equal to 1 since the number of observations is generally not so high and the linearity of the models ensures the feasibility, in term of computational costs, of the generation of the many individual forecasts involved. As previously pointed out, in the case of GARCH models, the choice of $v = 1$ is unrealistic; in this context, a value for v should guarantee the effectiveness of the forecast combination in accounting for possible structural breaks in the series and, at the same time, it should ensure not too high computational costs[2].

The selection of an optimal value of v is an open issue. However, it is possible to select it by simulations, searching, for example, the value which minimizes a fixed loss function. This is the approach which will be used in this paper.

As far as the vector of the weights $c_0, c_1, \ldots, c_{k-1}$ is concerned, as it is usual in the literature of forecast combination, we will impose that they satisfy the constrains:

$$c_\tau \geq 0 \quad \tau = 0, 1, \ldots, k-1 \qquad \text{and} \qquad \sum_{\tau=0}^{k-1} c_\tau = 1 \tag{6}$$

These conditions ensure that the equation (4) is a convex linear combination of the individual forecasts. Moreover we assume that the weights are such that

$$c_0 \geq c_1 \geq \ldots \geq c_{k-1} \tag{7}$$

[2]For example, in the case of a GARCH(1,1) model, consider a time series with $T = 4000$ observations and fix $\omega = 800$. The one-step haed forecast at time $T + 1$ obtained by using a value of $v = 50$, is generated by averaging 64 individual forecasts.

being c_τ the weight associated with the forecast obtained by using the observations from $(T-w-\tau v)$ to T. This assumption ensures that the forecasts obtained by the most recent observations have higher weights than those further away from the forecasting origin. Although all the sequences satisfying the previous assumptions can be used in the equation (4), some trivial choices have been proposed.

Mean window forecast combination with equal weights.

In this scheme, the forecast combination is obtained by using equal weights to average the individual forecasts. More precisely, the weights are defined as:

$$c_\tau = \frac{1}{k} \qquad \tau = 0, 1, \ldots, k-1 \tag{8}$$

and, as a consequence, the proposed forecast combination (4) is:

$$\widehat{h}_{T+1} = \frac{1}{k}\sum_{\tau=0}^{k-1} \widehat{h}_{T+1}^{[T-w-\tau v:T]} \tag{9}$$

Many researches (see for example Pesaran and Timmermann, 2007) have highlighted the advantages of the equally weighted forecast combination; this scheme is easy to compute and often has performances as good as more complicated schemes, also when there is uncertainty about the presence of structural breaks in the data.

Mean window forecast combination with location weights.

In this scheme, heavier weights are assigned to forecasts that use more recent information. As in Tian and Anderson (2014), we have proposed the following linear function of the location of time τ in the full sample:

$$c_\tau = \frac{k-\tau}{\sum_{\tau=0}^{k-1}(k-\tau)} = \frac{2(k-\tau)}{k(k+1)} \qquad \tau = 0, 1, \ldots, k-1 \tag{10}$$

As a consequence, the proposed forecast combination (4) is:

$$\widehat{h}_{T+1} = \frac{1}{\sum_{\tau=0}^{k-1}(k-\tau)}\sum_{\tau=0}^{k-1}(k-\tau)\widehat{h}_{T+1}^{[T-w-\tau v:T]} \tag{11}$$

The use of these weights implies that heavier weights are placed on the forecasts which are based on more recent parts of the sample.

Trimmed mean window forecast combination.

In this scheme, rather than combining the full set of forecasts, a fraction of those with the worst performance is discarded. In particular a fraction of the highest and lowest individual forecasts is excluded from the analysis. The remaining individual forecasts are then combined by using an equal weighting scheme. This approach is in the spirit of Armstrong (1989) and has been used by Rapach *et al.* (2008) in the context of forecasting stock return volatility. This weighting scheme could be useful since it is less sensitive to possible implausible forecasts.

4 Monte Carlo experiment

The aim of this simulation experiment is twofold. First our objective is to evaluate the effectiveness of the proposed procedures in accounting for structural breaks and the effects of the choice of the tuning parameter v. The results of this check will be used to identify an empirical value v^* of v such that the corresponding forecast combinations have better forecasting performances compared to other alternative ones. The second aim is to compare the proposed forecast combinations in correspondence of v^* with those proposed by Rapach and Strauss (2008) and Rapach *et al.* (2008).
As data generating process we have used a GARCH (1,1) with a single break point and, as a consequence, two different regimes are present:

$$a_t = h_t^{0.5} \epsilon_t \tag{12}$$

$$\begin{aligned} h_t = & \left[\alpha_{01} + \alpha_{11} a_{t-1}^2 + \beta_{11} h_{t-1}\right] I_{(t \le \tau)}(t) \\ & + \left[\alpha_{02} + \alpha_{12} a_{t-1}^2 + \beta_{12} h_{t-1}\right] I_{(t > \tau)}(t) \end{aligned} \tag{13}$$

where $(\alpha_{01}, \alpha_{11}, \beta_{11})$ and $(\alpha_{02}, \alpha_{12}, \beta_{12})$ are the parameters of the GARCH(1,1) model respectively in the first and in the second regime, τ is the location of the break point and $I_A(x)$ is the indicator function defined as:

$$I_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases} \tag{14}$$

Three different parameter specifications of the model (13) have been considered:

$$\text{M1}: \begin{cases} (\alpha_{11}, \beta_{11}) = (0.10, 0.80) & \alpha_{01} \text{ such that } \sigma_1^2 = 2 \\ (\alpha_{12}, \beta_{12}) = (0.05, 0.90) & \alpha_{02} \text{ such that } \sigma_2^2 = 3 \end{cases}$$

$$\text{M2}: \begin{cases} (\alpha_{11}, \beta_{11}) = (0.15, 0.70) & \alpha_{01} \text{ such that } \sigma_1^2 = 1 \\ (\alpha_{12}, \beta_{12}) = (0.10, 0.80) & \alpha_{02} \text{ such that } \sigma_2^2 = 2 \end{cases}$$

$$\text{M3}: \begin{cases} (\alpha_{11}, \beta_{11}) = (0.15, 0.70) & \alpha_{01} \text{ such that } \sigma_1^2 = 1 \\ (\alpha_{12}, \beta_{12}) = (0.05, 0.90) & \alpha_{02} \text{ such that } \sigma_2^2 = 3 \end{cases}$$

where σ_1^2 and σ_2^2 are, respectively, the unconditional variances in the first and in the second regime. Concerning the break, five different locations τ have been considered; they have been fixed as function of the time series length T. In particular:

$$\tau = \psi T \quad \text{with } \psi \in \{0.50, 0.60, 0.70, 0.80, 0.90\}$$

Therefore, the simulation experiment consists of fifteen different specifications of the model (13): three arising from the parameter specifications and five from the different locations of the break. Moreover two different series lengths $T \in \{3000, 4000\}$ have been considered. For each model specification and for both the sample sizes, 500 time series have been generated.

In order to evaluate the forecasting performances of the proposed forecast combinations for each model specification, an initial sub-sample, composed by the data from $t = 1$ to $t = R$, is used to estimate the model and the 1-step ahead out-of-sample forecast is produced. The sample is increased by one, the model is re-estimated using data from $t = 1$ to $t = R + 1$ and 1-step ahead forecast is produced. The procedure continues until the end of the available out-of-sample period. In the following, R has been fixed so that the number of out of sample observations is 300.

Common to all the considered forecast combinations is the specification of the minimum acceptable estimation window size ω. This parameter has been fixed equal to 800 in order to guarantee a convergent estimation of the GARCH(1,1) parameters.

For all the fifteen model specifications, for both sample sizes and for each of the 500 simulated series, the forecast combinations have been compared in terms of their predictive performances by using the model confidence set (Hansen *et al.*, 2011).

As pointed out in the Appendix A, this procedure is able to construct a set of combinations which exhibit the same predictive ability, at a given level of confidence, in terms of a predefined loss function. The choice of this function is arbitrary and depends on the nature of the competitors. However, as detailed by Patton (2011), when comparing conditional variance forecasts, the use of imperfect, but conditionally unbiased, volatility proxy can lead to objectionable results. In this context, where squared returns are used as proxy of the realized volatility, the most used loss function is the QLIKE function defined as:

$$\text{QLIKE}_t = \frac{\tilde{h}_t}{\widehat{h}_t} - \log\left(\frac{\tilde{h}_t}{\widehat{h}_t}\right) - 1 \tag{15}$$

where $\tilde{h}_t$ is some realized volatility measure and $\hat{h}_t$ is the punctual volatility forecast. It yields a ranking of volatility forecasts that is robust to noise in the proxy; moreover it is able to better discriminate among models and it is less affected by the most extreme observations in the sample. Once the analysis has been performed for all the 500 simulated series, the relative frequencies a given forecast combination enters in the Model Confidence Set (MCS) have been determined.

4.1 Evaluation of the proposed forecast combinations

In order to evaluate the effect of different values of the tuning parameter v on the proposed forecast combinations, the following different values of v

$$v \in \{50, 100, 200, 300, 400, 500, 600, 700, 800, 900\}$$

have been considered. This range of values allows to evaluate how the proposed forecast combinations vary with this parameter. For each value of v, the proposed forecast combinations with equal weights, defined in (9) and with location weights defined in (11) have been considered. Moreover, when it is reasonable, a trimmed version, with a fraction of trimmed forecasts equal to 0.2, has been also included in the analysis.

Moreover, to assess the effectiveness of the proposed combinations, a benchmark forecasting method has been introduced in the analysis. As a natural benchmark, a GARCH (1,1) model estimated by using an expanding window method has been considered. This method uses all the available observations and, as a consequence, it ignores the presence of structural breaks. This choice is optimal in situations with no breaks and it is appropriate for forecasting when the data is generated by a stable model.

As previously pointed out, the model confidence set has been employed and it has been determined the relative frequency a given forecast combination enters in the set of superior model. Two values of the confidence levels, $\alpha = (0.10, 0.25)$, for both the statistics used in the MCS procedure (see Appendix A for details) have been considered.

Tables 1, 2 and 3 report the relative frequencies for the parameter specification M1, M2 and M3, respectively, for $T = 3000$; the same parameter specifications for $T = 4000$ are reported in Tables 4, 5 and 6.

From Table 1, it is evident that when the size of break is not so high (specification M1) the proposed

combinations have good predictive performances for all the values of the tuning parameter v.
The relative frequencies of the proposed forecast combinations are higher than those of the expanding method, which does not take into account the presence of the break. This behaviour is more evident in correspondence of a confidence level equal to 0.25 and when the location of the break approaches the end of the sample. With reference to the three specifications of the combination weights, the combinations based on location weights seem to have better performances with respect to the other two weighting schemes, independently from the break locations. The tuning parameter v seems not so relevant when the break is far from the end of the sample, $\tau = (0.5T, 0.6T)$; in the other cases, the combinations with high values v and with the same weighting schemes, have relative frequencies significantly higher than those with small values of v. In particular the combination with location weights and $v = 800$ seems to have the best performances in all the cases.
In the case of specification M2 reported in Table 2, with respect to the previous case, the expanding method has performances even worse than the proposed combinations. Moreover, except for the case in which the break is located in the middle of the sample, the relative frequencies it enters in the set of superior model are very low and they decrease as the break location approaches the end of the sample. With respect to the three weighting schemes, it is more evident that the forecast combinations with location weights in general outperform those with the equal weights and the trimmed ones. Finally, regarding the tuning parameter v, the Table shows that, again, high values of v improve the forecasting performance of the considered combinations.
When considering the M3 specification, reported in Table 3, it is also more evident that not accounting for the break can have dramatic consequences in forecasting. Indeed, the behaviour of the expanding window method is very similar to that observed in the previous Table but the values of the relative frequencies are significantly reduced, especially when the confidence level is $\alpha = 0.25$. Moreover, it is confirmed that the forecast combinations with location weights and high values of the tuning parameter, especially $v = 800$, have better performances with respect all the other schemes.
In Tables 4, 5 and 6 the relative frequencies for the parameter specification M1, M2 and M3, respectively, for $T = 4000$ are reported. For all the model specifications, the results support the effectiveness of the proposed combinations with respect to the expanding procedure. Indeed, the expanding method enters in the set of superior model with very low relative frequencies if the comparison is made with the corresponding results obtained with $T = 3000$. The general behaviour of the proposed forecast combinations is confirmed: also in this case, for all the model specifications, the location schemes associated with a high values of the tuning parameter v seem to have better performances, especially when the break is located at the end of the sample size. The relative frequencies corresponding to these combinations are all very large especially when the confidence level is $\alpha = 0.25$.

4.2 A comparison with alternative forecast combinations

The proposed forecast combinations have been compared with some alternative combination schemes proposed in the econometric literature.

In particular the forecast combinations with equal weights and location weights associated with a tuning parameter equal to 800 (labelled, respectively Mean Wind 800 and Mean Wind 800 Loc) have been considered; this choice is motivated by the good performances highlighted by these two schemes in the simulations employed in the previous section. Moreover the forecast combinations proposed by Rapach and Strauss (2008) and Rapach *et al.* (2008), reviewed in section (2.1), have been examined; they are the Rapach and Strauss forecast combination (RS Mean), its trimmed version (RS Trim) and the Clark and McCraken combination (CM combined). Also in this analysis, the procedure based on an expanding window has been considered as benchmark. Again, the model confidence set has been used and the relative frequency a given forecast combination enters in the set of superior model has been determined.

Tables 7, 8 and 9 report the relative frequencies for the parameter specification M1, M2 and M3 for $T = 3000$. From their comparison, it is evident that all the considered forecast combinations have better performances with respect to the expanding window procedure. The RS Mean trimmed seems to have the best performance, for all the parameter specifications M1, M2, M3, when the break location is far from the end of the sample $\tau = (0.5T, 0.6T)$. The relative frequencies associated with these forecast combinations decrease when τ increases, especially for M2 and M3 specification. The RM mean seems to have a similar behaviour of the corresponding trimmed version but its performances are lower for $\tau = (0.5T, 0.6T)$ and greater for $\tau = (0.7T, 0.8T, 0.9T)$. The CM combined seems to have good performances for all the values of τ in the case of M1 specification; however, for the other parameter specifications M2 and M3, it seems to have good performances for $\tau = (0.7T, 0.8T)$, it gets worse for $\tau = (0.5T, 0.6T)$ reaching low relative frequencies for $\tau = 0.9T$.

As shown in the previous subsection, the Mean Wind 800 has, in general, worse performances than the Mean Wind 800 Loc which outperforms all the other forecast combinations, especially when the break location is near to the end of the sample, for all the parameter specifications M1, M2 and M3. Moreover, it behaves well also in all the other cases.

The same parameter specifications for $T = 4000$ are reported in Tables 10, 11 and 12. For all the model specifications, the results confirm the behaviour of the competing forecast combinations observed when $T = 3000$. As expected, when T increases, the MCS procedure is able to better discriminate among forecast combinations; in this case there is even more evidence that the Mean Wind 800 Loc outperforms all the other forecast combinations, especially when τ increases.

5 An application to real data

In order to evaluate the effectiveness of the proposed forecast combinations on real data, we have considered the same data set used by Rapach and Strauss (2008). It consists of daily returns of the U.S. dollar against the currencies of Canada, Denmark, Germany, Japan, Norway, Switzerland and the daily returns corresponding to the U.S. trade-weighted exchange rate. The series (Figure 1) cover the period from 1/2/1908 to 8/13/2005 [3].

Following West and Cho (1995) the unconditional and conditional mean of these eight series can be considered zero while Rapach and Strauss (2008) have given evidence for modelling them as GARCH(1, 1) processes. Moreover, they have applied the iterated cumulative sum of squares algorithm (Inclan and Tiao, 1994) on the test proposed by Sansó *et al.* (2004) to detect possible breaks in the unconditional variance of the series. The exact locations of the structural breaks are reported in Table 13. In particular, the algorithm selects a single structural break in the unconditional variance of returns for Germany; two structural breaks for Japan, Norway and Switzerland; three structural breaks for Canada and the U.S.; four structural breaks for the U.K. As pointed out by Rapach and Strauss, they appear to correspond to significant economic events, giving evidence that the proposed test strategy is able to identify correctly the breaks. Moreover the GARCH (1,1) parameter estimates, for the full sample and for each sub-sample identified by the structural breaks, highlight significant shifts in the intercept of the GARCH (1,1) model. These shifts cause significant changes in the unconditional variance across regimes justifying the relevance of structural breaks in the analysis of the considered series and the use of forecasting techniques which account for them.

In the out-of-sample exercise, 1-step-ahead forecasts have been considered; they are based on the sub-sample formed of the last 300 observations. As in the simulation experiment, we have compared the forecasts of volatility generated by the forecast combinations Mean Wind 800, Mean Wind 800 Loc, and those proposed by Rapach and Strauss (2008) and Rapach *et al.* (2008) defined in the previous section and labelled RS Mean, RS Trim and CM combined.

Again, the model confidence set has been used to construct the set of combinations which exhibits the same predictive ability, in terms of QLIKE loss function. Both the statistics proposed in the MCS procedure (see Appendix A for details) have been considered.

The results are reported in Table 14. With regard to Canada, the MCS reduces to a singleton containing only the RS Mean trimmed combination. Because the last surviving element is never eliminated, from Theorem 1 of Hansen *et al.* (2011), the associated probability approaches 1 as $T \to \infty$. In this case, the only information which could be deduced is that the loss associated to the RS Mean trimmed

[3]The data can be freely downloaded from `http://sites.slu.edu/rapachde/home/research`

combination is the lowest. For all the other countries, there is strong evidence that accounting for structural breaks in the unconditional variance leads to out-of-sample forecasting gain: the expanding method is always out of the MCS. The composition of MCS changes across countries; this could be due to the different numbers, locations and magnitudes of the structural breaks in the series. However the Mean Wind 800 Loc forecast combination always enters in the set of superior model for both the statistics.

6 Concluding remarks

This paper has proposed three new combination schemes, which take into account for structural breaks in the unconditional variance of a GARCH(1, 1) model. They are obtained by averaging forecasts based on different estimation windows generated by recursively increasing an initial window of a fixed number of observations v. In the first scheme, the forecast combination is obtained by using equal weights to average the individual forecasts; the second weighting method simply assigns heavier weights to forecasts that use more recent information; the third is a trimmed version of the forecast combination with equal weights where a fixed fraction of forecasts with the worst performance are discarded.

The simulation experiment, carried out to evaluate the effect of the tuning parameter v on the performances of the proposed forecast combinations, has shown that the second weighting scheme based on an high value of v ($v = 800$) outperforms all the other forecast combinations with low or moderate values of v and with different weighting schemes. This result is particularly evident when the location of the structural break is near the end of the sample.

Through the Monte Carlo experiment, it has been verified that the forecast combination with location weights and with the value of v previously identified seems to better perform also with respect to some other forecast combinations proposed in the literature.

An empirical application, based on daily returns of the U.S. dollar against the currencies of Canada, Denmark, Germany, Japan, Norway, Switzerland and the daily returns corresponding to the U.S. trade-weighted exchange rate, has confirmed these findings.

In any case, several different aspects should be further explored to get a better insight into the usage of the proposed forecasting combinations. From a computational point of view, a more extensive simulation experiment should be implemented in order to evaluate the proposed forecasting schemes when more than one break are present in the data generating process. From a statistical point of view, it should be assessed if different approaches, other then simulations, can be used to determine a plausible value for the tuning parameter.

References

[1] Armstrong J.S. 1989. Combining forecasts: The end of the beginning or the beginning of the end?. *International Journal of Forecasting* **5(4)**: 585-588.

[2] Andreou E, Ghysels E. 2009. Structural breaks in financial time series. In: Anderson TG, Davis RA, Kreiss JP, Mikosch T (Eds). *Handbook of financial time series*. Springer Berlin Heidelberg: 839-870.

[3] Assenmacher-Wesche K, Pesaran MH. 2008. Forecasting the Swiss economy using VECX models: An exercise in forecast combination across models and observation windows. *National Institute Economic Review* **203(1)**: 91-108.

[4] Bates JM, Granger CW. 1969. The combination of forecasts. *Journal of the Operational Research Society*, **20(4)**: 451-468.

[5] Clark TE, McCracken MW. 2009. Improving forecast accuracy by combining recursive and rolling forecasts. *International Economic Review*, **50(2)**: 363-395.

[6] Clemen RT. 1989. Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, **5(4)**: 559-583.

[7] Diebold FX. 1986. Modelling the Persistence of Conditional Variance: A Comment. *Econometric Reviews*, **5(1)**: 51-56.

[8] Engle, RF, Bollerslev, T. 1986. Modelling the persistence of conditional variances. *Econometric Reviews* **5(1)**: 1-50.

[9] Giraitis L, Kapetanios G, Price S. 2013. Adaptive forecasting in the presence of recent and ongoing structural change. *Journal of Econometrics* **177(2)**: 153–170.

[10] Hansen PR, Lunde A. 2005. A forecast comparison of volatility models: does anything beat a GARCH (1,1)? *Journal of Applied Econometrics* **20(7)**: 873-889.

[11] Hansen PR, Lunde A, Nason JM. 2011. The model confidence set. *Econometrica* **79(2)**: 453-497.

[12] Hillebrand E. 2005. Neglecting parameter changes in GARCH models. *Journal of Econometrics* **129(1)**: 121-138.

[13] Hwang S, Valls Pereira PL. 2006. Small sample properties of GARCH estimates and persistence. *The European Journal of Finance* **12(6-7)**: 473-494.

[14] Inclan C, Tiao GC. 1994. Use of Cumulative Sums of Squares for Retrospective Detection of Changes in Variance. *Journal of American Statistic Association* **89(427)**: 913-923.

[15] Inoue A, Jin L, Rossi B. 2017. Rolling window selection for out-of-sample forecasting with time-varying parameters. *Journal of Econometrics* **196(1)**: 55-67.

[16] Lamoureux CG, Lastrapes WD. 1990. Persistence in variance, structural change, and the GARCH model. *Journal of Business & Economic Statistics* **8(2)**: 225-234.

[17] Marcellino M. 2004. Forecast pooling for European macroeconomic variables. *Oxford Bulletin of Economic and Statistics* **66(1)**: 91-112.

[18] Makridakis S, Hibon M. 2000. The M3-Competition: results, conclusions and implications. *International Journal of Forecasting* **16(4)**: 451-476.

[19] Mikosh T, Stărică C. 2004. Nonstationarities in Financial Time Series, the Long -range Dependence and the IGARCH Effects. *Review of Economics and Statistics* **86(1)**: 378-390.

[20] Patton AJ. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* **160(1)**: 246-256.

[21] Pesaran MH, Pick A. 2011. Forecast combination across estimation windows. *Journal of Business and Economic Statistics* **29(2)**: 307-318.

[22] Pesaran MH, Pick A, Pranovich M. 2013. Optimal forecasts in the presence of structural breaks. *Journal of Econometrics* **177(2)**: 134-152.

[23] Pesaran MH, Schuermann T, Smith LV. 2009. Forecasting economic and financial variables with global VARs. *International Journal of Forecasting* **25(4)**: 642-675.

[24] Pesaran MH, Timmermann A. 2007. Selection of estimation windows in the presence of breaks. *Journal of Econometrics* **137(1)**: 134-161.

[25] Rapach DE, Strauss JK. 2008. Structural breaks and GARCH models of exchange rate volatility. *Journal of Applied Econometrics* **23(1)**: 65-90.

[26] Rapach DE, Strauss JK, Wohar, ME. 2008. Forecasting stock return volatility in the presence of structural breaks. In: Rapach DE, Wohar, ME (Eds). *Forecasting in the Presence of Structural Breaks and Model Uncertainty*. Elsevier-Emerald Series Frontiers of Economics and Globalization, Bingley, UK: 381-416.

[27] Ross GJ. 2013. Modelling financial volatility in the presence of abrupt changes. *Physica A: Statistical Mechanics and its Applications* **392(2)**: 350-360.

[28] Sansó A, Arragó V, Carrion-i-Silvestre J.L. 2004. Testing for Change in the Unconditional Variance of Financial Time Series. *Revista de Economia Financiera* **4(1)**: 32-53.

[29] Schrimpf A, Wang Q. 2010. A reappraisal of the leading indicator properties of the yield curve under structural instability. *International Journal of Forecasting* **26(4)**: 836-857.

[30] Stărică C, Herzel S, Nord T. 2005. Why does the GARCH(1,1) model fail to provide sensible longer-horizon volatility forecasts? *Manuscript, Chalmers University of Technology*.

[31] Stock JH, Watson, MW. 2001. Vector autoregressions. *The Journal of Economic Perspectives* **14(4)**: 101-115.

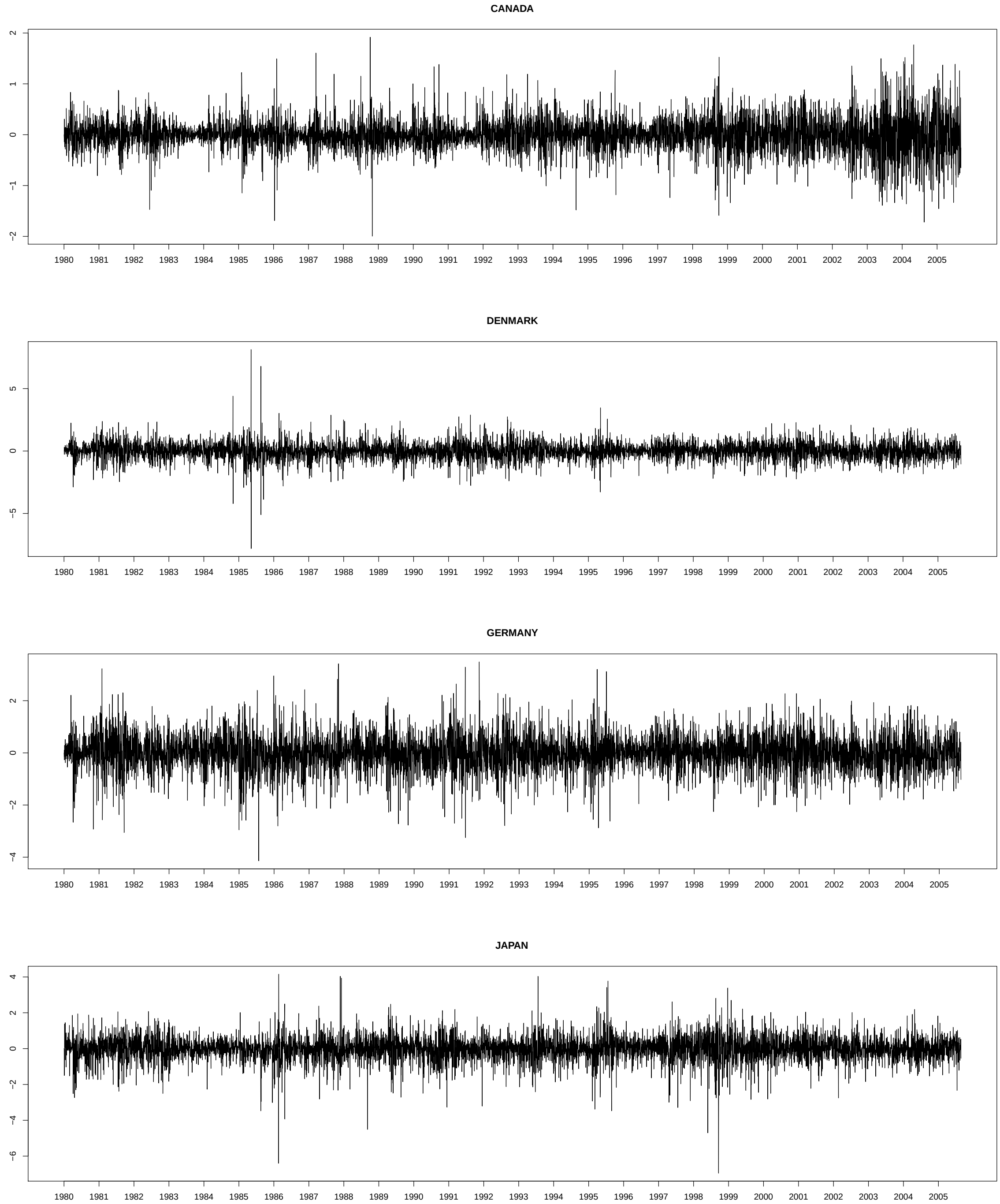
[32] Stock JH, Watson, MW. 2004. Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting* **23(6)**: 405-430.

[33] Tian J, Anderson HM. 2014. Forecast Combinations under Structural Break Uncertainty. *International Journal of Forecasting* **30(1)**: 161-175.

[34] Timmermann A. 2006. Forecast combinations. In: Elliot G, Granger CWJ, Timmermann, A (Eds). *Handbook of Economic Forecasting*. Elsevier: 135-196.

[35] West, KD, Cho, D. 1995. The predictive ability of several models of exchange rate volatility. *Journal of Econometrics*, **69(2)**: 367-391.

7 Figures and Tables



continued on next page

continued from previous page

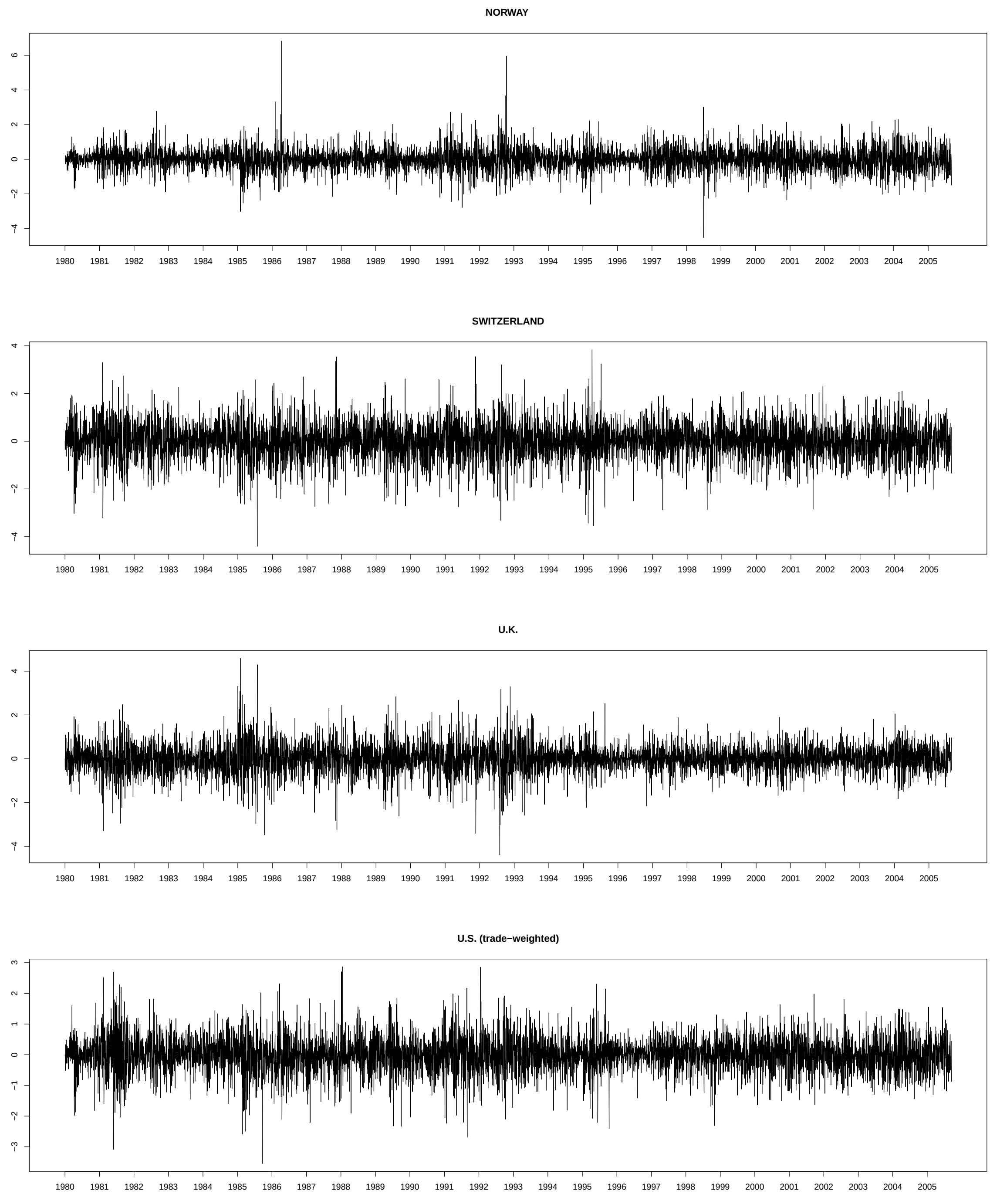


Figure 1: U.S. dollar exchange rate returns

Table 1: Evaluation of the proposed forecast combinations for parameter specification M1, T=3000.

MODEL M1	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=3000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.96	0.88	0.77	0.75	0.92	0.87	0.73	0.71	0.93	0.84	0.80	0.67	0.83	0.72	0.60	0.49	0.93	0.83	0.66	0.67
Mean wind 50-Equ	0.95	0.93	0.82	0.78	0.92	0.92	0.79	0.79	0.93	0.86	0.85	0.74	0.88	0.78	0.68	0.58	0.89	0.78	0.71	0.66
Mean wind 50-Loc	0.97	0.99	0.92	0.94	0.94	0.96	0.89	0.92	0.94	0.92	0.87	0.82	0.91	0.88	0.78	0.74	0.91	0.84	0.76	0.75
Mean wind 50-Trim	0.95	0.92	0.84	0.79	0.92	0.89	0.74	0.71	0.92	0.85	0.84	0.68	0.86	0.74	0.66	0.51	0.89	0.76	0.72	0.63
Mean wind 100-Equ	0.96	0.93	0.83	0.76	0.93	0.92	0.81	0.78	0.94	0.88	0.85	0.75	0.86	0.79	0.66	0.57	0.90	0.77	0.72	0.67
Mean wind 100-Loc	0.98	0.99	0.92	0.95	0.95	0.96	0.87	0.91	0.95	0.94	0.88	0.83	0.91	0.89	0.79	0.75	0.92	0.85	0.77	0.76
Mean wind 100-Trim	0.97	0.92	0.83	0.79	0.93	0.87	0.75	0.75	0.93	0.86	0.84	0.72	0.84	0.74	0.62	0.52	0.88	0.76	0.72	0.63
Mean wind 200-Equ	0.96	0.92	0.80	0.78	0.93	0.91	0.82	0.80	0.94	0.87	0.85	0.75	0.88	0.83	0.69	0.56	0.91	0.78	0.73	0.65
Mean wind 200-Loc	0.99	0.99	0.92	0.94	0.95	0.97	0.88	0.94	0.95	0.95	0.89	0.82	0.94	0.93	0.79	0.78	0.92	0.86	0.82	0.77
Mean wind 200-Trim	0.97	0.92	0.81	0.79	0.93	0.90	0.78	0.76	0.92	0.87	0.84	0.70	0.86	0.77	0.66	0.51	0.88	0.77	0.73	0.67
Mean wind 300-Equ	0.94	0.91	0.81	0.81	0.94	0.93	0.79	0.79	0.94	0.88	0.87	0.76	0.89	0.83	0.70	0.61	0.90	0.80	0.73	0.67
Mean wind 300-Loc	0.99	0.99	0.94	0.96	0.96	0.99	0.90	0.92	0.97	0.95	0.91	0.87	0.95	0.92	0.83	0.81	0.92	0.88	0.79	0.77
Mean wind 300-Trim	0.92	0.90	0.81	0.79	0.93	0.91	0.79	0.77	0.94	0.87	0.84	0.70	0.86	0.74	0.66	0.53	0.88	0.79	0.71	0.67
Mean wind 400-Equ	0.97	0.93	0.82	0.75	0.94	0.91	0.81	0.84	0.92	0.89	0.82	0.77	0.90	0.85	0.71	0.62	0.88	0.78	0.72	0.67
Mean wind 400-Loc	0.98	0.99	0.93	0.94	0.98	0.99	0.90	0.94	0.96	0.95	0.90	0.89	0.98	0.96	0.84	0.84	0.90	0.88	0.81	0.77
Mean wind 400-Trim	0.97	0.95	0.83	0.77	0.92	0.90	0.76	0.80	0.93	0.88	0.81	0.71	0.87	0.78	0.71	0.54	0.86	0.77	0.66	0.68
Mean wind 500-Equ	0.97	0.95	0.84	0.79	0.91	0.92	0.83	0.81	0.94	0.89	0.85	0.81	0.96	0.87	0.71	0.63	0.87	0.81	0.69	0.65
Mean wind 500-Loc	0.99	0.99	0.95	0.95	0.97	0.99	0.92	0.92	0.98	0.98	0.92	0.91	1.00	0.97	0.87	0.88	0.93	0.92	0.78	0.83
Mean wind 500-Trim	0.97	0.93	0.84	0.79	0.93	0.92	0.79	0.76	0.94	0.88	0.85	0.77	0.94	0.82	0.71	0.59	0.90	0.79	0.67	0.66
Mean wind 600-Equ	0.97	0.95	0.80	0.80	0.93	0.91	0.80	0.77	0.94	0.91	0.86	0.79	0.90	0.82	0.69	0.60	0.90	0.84	0.73	0.70
Mean wind 600-Loc	0.97	0.98	0.91	0.94	0.97	0.99	0.89	0.93	0.98	0.98	0.91	0.92	0.97	0.98	0.89	0.88	0.96	0.91	0.86	0.84
Mean wind 700-Equ	0.94	0.95	0.81	0.82	0.92	0.95	0.81	0.84	0.95	0.95	0.90	0.83	0.96	0.90	0.78	0.75	0.92	0.88	0.75	0.71
Mean wind 700-Loc	0.98	0.98	0.94	0.95	0.97	0.99	0.92	0.97	1.00	1.00	0.95	0.96	1.00	0.99	0.96	0.98	0.97	0.97	0.91	0.91
Mean wind 800-Equ	0.98	0.96	0.81	0.85	0.93	0.92	0.81	0.79	0.94	0.95	0.87	0.82	0.95	0.90	0.77	0.75	0.90	0.88	0.73	0.69
Mean wind 800-Loc	0.97	0.99	0.92	0.95	0.98	1.00	0.94	0.96	0.99	1.00	0.97	0.99	1.00	0.99	0.98	0.99	0.97	0.97	0.90	0.93
Mean wind 900-Equ	0.96	0.94	0.81	0.82	0.94	0.95	0.84	0.78	0.94	0.93	0.89	0.84	0.95	0.88	0.75	0.67	0.90	0.86	0.72	0.66
Mean wind 900-Loc	0.98	0.96	0.90	0.91	0.96	1.00	0.92	0.95	0.98	1.00	0.94	0.97	0.99	0.99	0.95	0.96	0.96	0.97	0.88	0.91

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 2: Evaluation of the proposed forecast combinations for parameter specification M2, T=3000.

MODEL M2	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=3000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.87	0.85	0.74	0.64	0.83	0.71	0.60	0.41	0.71	0.48	0.36	0.23	0.62	0.33	0.35	0.17	0.53	0.43	0.26	0.25
Mean wind 50-Equ	0.92	0.87	0.83	0.75	0.89	0.79	0.72	0.56	0.76	0.58	0.48	0.33	0.75	0.40	0.39	0.23	0.60	0.49	0.32	0.29
Mean wind 50-Loc	0.99	0.97	0.93	0.92	0.98	0.93	0.88	0.85	0.88	0.68	0.58	0.49	0.82	0.64	0.54	0.41	0.70	0.63	0.50	0.47
Mean wind 50-Trim	0.93	0.88	0.82	0.73	0.86	0.71	0.66	0.49	0.74	0.49	0.43	0.26	0.67	0.34	0.36	0.19	0.59	0.45	0.29	0.28
Mean wind 100-Equ	0.93	0.87	0.83	0.78	0.89	0.78	0.73	0.58	0.76	0.58	0.47	0.35	0.76	0.39	0.42	0.24	0.62	0.49	0.30	0.27
Mean wind 100-Loc	0.99	0.97	0.94	0.94	0.98	0.94	0.89	0.86	0.90	0.69	0.60	0.49	0.85	0.63	0.56	0.41	0.70	0.65	0.50	0.48
Mean wind 100-Trim	0.93	0.88	0.83	0.72	0.88	0.73	0.66	0.48	0.73	0.53	0.43	0.27	0.69	0.33	0.37	0.19	0.58	0.45	0.30	0.25
Mean wind 200-Equ	0.94	0.87	0.83	0.79	0.90	0.80	0.76	0.58	0.84	0.59	0.48	0.34	0.76	0.38	0.44	0.22	0.61	0.48	0.33	0.25
Mean wind 200-Loc	0.98	0.97	0.93	0.94	0.98	0.96	0.89	0.86	0.90	0.73	0.66	0.51	0.86	0.72	0.58	0.44	0.73	0.65	0.48	0.50
Mean wind 200-Trim	0.94	0.88	0.83	0.73	0.87	0.72	0.70	0.53	0.73	0.50	0.44	0.29	0.68	0.33	0.36	0.22	0.62	0.45	0.34	0.24
Mean wind 300-Equ	0.95	0.90	0.83	0.79	0.90	0.79	0.73	0.60	0.81	0.61	0.53	0.38	0.75	0.41	0.45	0.24	0.64	0.50	0.37	0.34
Mean wind 300-Loc	0.98	0.97	0.93	0.95	0.98	0.96	0.90	0.86	0.92	0.75	0.69	0.58	0.87	0.68	0.66	0.50	0.79	0.73	0.55	0.54
Mean wind 300-Trim	0.94	0.89	0.80	0.79	0.89	0.76	0.74	0.57	0.74	0.58	0.45	0.35	0.71	0.36	0.38	0.22	0.60	0.50	0.36	0.29
Mean wind 400-Equ	0.95	0.91	0.84	0.82	0.91	0.82	0.79	0.61	0.84	0.60	0.53	0.36	0.76	0.49	0.47	0.22	0.62	0.50	0.33	0.27
Mean wind 400-Loc	0.99	0.99	0.93	0.94	0.96	0.96	0.92	0.91	0.90	0.83	0.74	0.61	0.87	0.75	0.66	0.58	0.76	0.73	0.54	0.56
Mean wind 400-Trim	0.97	0.91	0.86	0.81	0.91	0.75	0.74	0.55	0.75	0.54	0.45	0.31	0.72	0.40	0.39	0.20	0.63	0.48	0.32	0.27
Mean wind 500-Equ	0.95	0.89	0.86	0.82	0.90	0.82	0.79	0.62	0.84	0.65	0.60	0.41	0.76	0.50	0.49	0.29	0.58	0.49	0.31	0.28
Mean wind 500-Loc	0.98	0.99	0.95	0.96	0.95	0.95	0.89	0.91	0.93	0.88	0.78	0.71	0.90	0.81	0.73	0.61	0.77	0.74	0.58	0.62
Mean wind 500-Trim	0.96	0.91	0.86	0.80	0.92	0.81	0.76	0.59	0.86	0.59	0.59	0.38	0.77	0.44	0.44	0.27	0.62	0.50	0.32	0.28
Mean wind 600-Equ	0.90	0.91	0.82	0.80	0.89	0.81	0.74	0.61	0.83	0.66	0.58	0.40	0.77	0.52	0.47	0.24	0.64	0.55	0.39	0.33
Mean wind 600-Loc	0.99	0.99	0.92	0.97	0.97	0.97	0.89	0.90	0.93	0.87	0.78	0.70	0.93	0.86	0.68	0.71	0.84	0.82	0.66	0.68
Mean wind 700-Equ	0.97	0.93	0.85	0.84	0.93	0.84	0.81	0.72	0.88	0.76	0.67	0.53	0.83	0.62	0.55	0.37	0.70	0.65	0.42	0.39
Mean wind 700-Loc	0.99	1.00	0.93	0.96	0.99	1.00	0.97	0.98	1.00	1.00	0.96	0.95	0.97	0.98	0.91	0.89	0.96	0.97	0.81	0.82
Mean wind 800-Equ	0.91	0.91	0.81	0.79	0.94	0.90	0.76	0.69	0.86	0.72	0.66	0.56	0.81	0.61	0.53	0.34	0.68	0.59	0.43	0.35
Mean wind 800-Loc	0.99	1.00	0.92	0.97	0.99	1.00	0.98	0.98	1.00	1.00	0.98	0.99	1.00	1.00	0.97	0.97	0.98	0.98	0.90	0.89
Mean wind 900-Equ	0.90	0.91	0.77	0.76	0.93	0.86	0.76	0.63	0.89	0.74	0.64	0.53	0.85	0.60	0.55	0.34	0.69	0.60	0.46	0.35
Mean wind 900-Loc	0.95	0.97	0.82	0.91	0.97	0.99	0.93	0.93	1.00	1.00	0.94	0.94	0.95	0.96	0.87	0.86	0.92	0.95	0.80	0.86

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 3: Evaluation of the proposed forecast combinations for parameter specification M3, T=3000.

MODEL M3	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=3000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.82	0.77	0.62	0.58	0.80	0.66	0.54	0.35	0.63	0.40	0.26	0.20	0.41	0.23	0.15	0.06	0.10	0.10	0.04	0.03
Mean wind 50-Equ	0.89	0.82	0.73	0.64	0.84	0.72	0.66	0.50	0.69	0.44	0.34	0.26	0.68	0.34	0.31	0.10	0.23	0.18	0.08	0.07
Mean wind 50-Loc	0.97	0.97	0.92	0.94	0.90	0.90	0.83	0.75	0.80	0.57	0.46	0.36	0.82	0.60	0.48	0.38	0.43	0.40	0.31	0.32
Mean wind 50-Trim	0.84	0.78	0.68	0.59	0.80	0.66	0.51	0.39	0.66	0.41	0.29	0.20	0.56	0.29	0.26	0.10	0.15	0.15	0.06	0.07
Mean wind 100-Equ	0.89	0.82	0.74	0.63	0.83	0.74	0.68	0.53	0.68	0.44	0.36	0.27	0.68	0.34	0.32	0.11	0.22	0.18	0.07	0.07
Mean wind 100-Loc	0.98	0.98	0.91	0.96	0.90	0.94	0.84	0.76	0.81	0.63	0.48	0.38	0.82	0.61	0.51	0.39	0.43	0.41	0.34	0.32
Mean wind 100-Trim	0.84	0.79	0.69	0.60	0.82	0.66	0.53	0.43	0.65	0.41	0.33	0.21	0.58	0.29	0.26	0.10	0.18	0.16	0.05	0.07
Mean wind 200-Equ	0.90	0.84	0.75	0.66	0.86	0.74	0.67	0.54	0.72	0.45	0.37	0.29	0.70	0.36	0.34	0.11	0.21	0.18	0.08	0.07
Mean wind 200-Loc	0.98	0.98	0.94	0.96	0.91	0.93	0.87	0.79	0.84	0.66	0.52	0.43	0.85	0.66	0.57	0.45	0.49	0.45	0.37	0.32
Mean wind 200-Trim	0.85	0.81	0.70	0.60	0.82	0.68	0.56	0.45	0.65	0.41	0.30	0.25	0.61	0.32	0.27	0.11	0.21	0.17	0.08	0.06
Mean wind 300-Equ	0.89	0.86	0.73	0.66	0.84	0.75	0.67	0.56	0.73	0.50	0.42	0.29	0.73	0.38	0.35	0.15	0.23	0.18	0.10	0.10
Mean wind 300-Loc	0.98	0.99	0.95	0.96	0.94	0.95	0.90	0.80	0.84	0.73	0.59	0.48	0.87	0.73	0.63	0.54	0.48	0.51	0.37	0.34
Mean wind 300-Trim	0.86	0.85	0.72	0.64	0.83	0.70	0.61	0.51	0.68	0.43	0.36	0.26	0.69	0.35	0.33	0.14	0.20	0.18	0.10	0.13
Mean wind 400-Equ	0.93	0.89	0.80	0.68	0.85	0.78	0.71	0.56	0.75	0.55	0.41	0.35	0.76	0.41	0.36	0.18	0.20	0.19	0.05	0.05
Mean wind 400-Loc	0.98	0.99	0.97	0.97	0.93	0.94	0.86	0.82	0.87	0.76	0.65	0.52	0.86	0.80	0.65	0.55	0.48	0.50	0.37	0.38
Mean wind 400-Trim	0.88	0.83	0.72	0.64	0.85	0.72	0.67	0.50	0.72	0.47	0.35	0.25	0.71	0.35	0.31	0.14	0.20	0.19	0.10	0.07
Mean wind 500-Equ	0.95	0.89	0.81	0.76	0.86	0.81	0.75	0.56	0.76	0.55	0.48	0.35	0.76	0.43	0.37	0.21	0.19	0.17	0.06	0.07
Mean wind 500-Loc	0.99	0.99	0.97	0.98	0.94	0.95	0.87	0.81	0.91	0.79	0.70	0.60	0.90	0.89	0.69	0.66	0.46	0.52	0.30	0.39
Mean wind 500-Trim	0.94	0.89	0.81	0.69	0.82	0.77	0.68	0.49	0.73	0.50	0.44	0.32	0.75	0.44	0.38	0.17	0.19	0.17	0.07	0.04
Mean wind 600-Equ	0.89	0.87	0.79	0.78	0.87	0.73	0.65	0.58	0.79	0.58	0.45	0.37	0.77	0.40	0.38	0.18	0.21	0.18	0.08	0.08
Mean wind 600-Loc	0.98	0.99	0.94	0.98	0.96	0.96	0.88	0.81	0.91	0.88	0.73	0.70	0.90	0.88	0.69	0.62	0.56	0.58	0.46	0.47
Mean wind 700-Equ	0.87	0.94	0.77	0.76	0.91	0.88	0.75	0.70	0.84	0.66	0.55	0.45	0.83	0.55	0.47	0.33	0.27	0.26	0.14	0.13
Mean wind 700-Loc	0.99	1.00	0.92	0.99	0.99	0.99	0.95	0.96	0.98	0.98	0.97	0.95	1.00	1.00	0.97	0.98	0.90	0.90	0.75	0.77
Mean wind 800-Equ	0.88	0.87	0.74	0.69	0.89	0.88	0.76	0.71	0.82	0.64	0.58	0.43	0.79	0.48	0.42	0.24	0.22	0.22	0.09	0.09
Mean wind 800-Loc	0.97	0.99	0.87	0.93	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.98	0.97	0.98	0.90	0.92
Mean wind 900-Equ	0.87	0.86	0.74	0.67	0.89	0.86	0.80	0.63	0.86	0.65	0.55	0.43	0.77	0.43	0.42	0.22	0.20	0.21	0.08	0.09
Mean wind 900-Loc	0.93	0.95	0.85	0.90	1.00	0.99	0.95	0.95	0.97	0.99	0.95	0.95	0.99	1.00	0.78	0.87	0.68	0.80	0.47	0.52

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 4: Evaluation of the proposed forecast combinations for parameter specification M1, T=4000.

MODEL M1	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=4000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.89	0.89	0.74	0.77	0.93	0.93	0.77	0.78	0.93	0.84	0.75	0.57	0.89	0.79	0.70	0.49	0.79	0.66	0.59	0.44
Mean wind 50-Equ	0.94	0.91	0.80	0.82	0.95	0.96	0.85	0.89	0.93	0.89	0.80	0.73	0.88	0.83	0.70	0.54	0.81	0.68	0.61	0.52
Mean wind 50-Loc	0.99	0.99	0.91	0.96	0.99	0.99	0.91	0.96	0.96	0.93	0.87	0.87	0.93	0.92	0.79	0.70	0.88	0.81	0.70	0.57
Mean wind 50-Trim	0.94	0.92	0.82	0.81	0.95	0.95	0.82	0.83	0.92	0.89	0.81	0.63	0.89	0.79	0.70	0.52	0.82	0.64	0.58	0.45
Mean wind 100-Equ	0.93	0.90	0.82	0.83	0.95	0.96	0.88	0.88	0.93	0.89	0.83	0.72	0.88	0.83	0.70	0.55	0.83	0.69	0.59	0.48
Mean wind 100-Loc	0.99	0.99	0.91	0.94	0.99	0.99	0.93	0.96	0.96	0.94	0.87	0.87	0.94	0.92	0.80	0.75	0.90	0.85	0.71	0.60
Mean wind 100-Trim	0.94	0.93	0.81	0.82	0.95	0.96	0.83	0.83	0.91	0.88	0.80	0.63	0.89	0.79	0.70	0.50	0.83	0.63	0.57	0.44
Mean wind 200-Equ	0.95	0.92	0.82	0.81	0.97	0.96	0.87	0.88	0.94	0.91	0.83	0.70	0.90	0.83	0.71	0.56	0.85	0.71	0.58	0.46
Mean wind 200-Loc	0.98	0.99	0.91	0.94	0.99	0.99	0.92	0.97	0.96	0.96	0.91	0.89	0.96	0.93	0.84	0.75	0.90	0.87	0.74	0.64
Mean wind 200-Trim	0.93	0.94	0.80	0.81	0.94	0.96	0.84	0.83	0.92	0.86	0.78	0.63	0.91	0.80	0.71	0.51	0.82	0.63	0.56	0.44
Mean wind 300-Equ	0.95	0.90	0.81	0.81	0.98	0.97	0.85	0.87	0.94	0.91	0.84	0.73	0.90	0.86	0.74	0.60	0.81	0.69	0.61	0.47
Mean wind 300-Loc	0.97	0.99	0.91	0.95	1.00	0.99	0.93	0.98	0.96	0.94	0.90	0.90	0.94	0.94	0.86	0.82	0.88	0.89	0.72	0.64
Mean wind 300-Trim	0.94	0.94	0.80	0.80	0.94	0.96	0.84	0.84	0.92	0.87	0.78	0.67	0.89	0.79	0.72	0.51	0.84	0.65	0.61	0.45
Mean wind 400-Equ	0.93	0.91	0.79	0.81	0.97	0.96	0.88	0.90	0.93	0.89	0.82	0.73	0.89	0.85	0.70	0.60	0.82	0.70	0.59	0.44
Mean wind 400-Loc	0.97	0.99	0.91	0.95	1.00	0.99	0.93	0.96	0.96	0.96	0.90	0.90	0.95	0.95	0.85	0.81	0.91	0.88	0.78	0.73
Mean wind 400-Trim	0.94	0.94	0.83	0.81	0.96	0.97	0.85	0.85	0.91	0.88	0.80	0.69	0.89	0.81	0.73	0.55	0.83	0.68	0.60	0.45
Mean wind 500-Equ	0.95	0.93	0.84	0.84	0.96	0.98	0.83	0.89	0.95	0.92	0.86	0.74	0.91	0.87	0.73	0.66	0.85	0.73	0.64	0.50
Mean wind 500-Loc	0.97	0.99	0.93	0.96	0.99	0.99	0.90	0.98	0.96	0.97	0.93	0.92	0.96	0.97	0.91	0.86	0.95	0.92	0.80	0.75
Mean wind 500-Trim	0.95	0.93	0.86	0.87	0.92	0.96	0.84	0.84	0.92	0.89	0.81	0.64	0.90	0.84	0.69	0.61	0.85	0.70	0.63	0.52
Mean wind 600-Equ	0.95	0.93	0.82	0.80	0.96	0.98	0.86	0.89	0.95	0.93	0.88	0.76	0.91	0.89	0.77	0.65	0.84	0.75	0.66	0.52
Mean wind 600-Loc	0.96	0.99	0.92	0.95	0.99	0.98	0.91	0.96	0.98	0.97	0.95	0.94	0.96	0.98	0.89	0.89	0.92	0.91	0.79	0.80
Mean wind 700-Equ	0.94	0.92	0.81	0.83	0.98	0.97	0.83	0.90	0.93	0.92	0.84	0.72	0.92	0.88	0.79	0.66	0.83	0.74	0.64	0.51
Mean wind 700-Loc	0.97	0.99	0.90	0.95	1.00	0.99	0.91	0.97	0.97	0.97	0.95	0.93	0.96	0.98	0.90	0.91	0.92	0.96	0.78	0.81
Mean wind 800-Equ	0.95	0.97	0.84	0.85	0.98	0.98	0.85	0.94	0.95	0.93	0.87	0.79	0.92	0.90	0.80	0.69	0.83	0.82	0.67	0.55
Mean wind 800-Loc	0.98	0.98	0.93	0.96	1.00	1.00	0.94	0.99	1.00	1.00	0.98	1.00	1.00	1.00	0.98	0.99	0.96	0.99	0.94	0.97
Mean wind 900-Equ	0.93	0.91	0.79	0.82	0.97	0.96	0.84	0.89	0.94	0.91	0.83	0.72	0.93	0.90	0.83	0.73	0.83	0.77	0.67	0.55
Mean wind 900-Loc	0.97	0.99	0.90	0.94	0.98	0.99	0.85	0.95	0.98	0.99	0.96	0.96	0.99	0.99	0.93	0.96	0.93	0.94	0.81	0.82

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 5: Evaluation of the proposed forecast combinations for parameter specification M2, T=4000.

MODEL M2	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=4000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.91	0.91	0.72	0.70	0.83	0.76	0.62	0.50	0.83	0.53	0.40	0.24	0.41	0.22	0.16	0.07	0.37	0.19	0.10	0.08
Mean wind 50-Equ	0.95	0.95	0.83	0.77	0.90	0.84	0.76	0.65	0.91	0.67	0.59	0.38	0.58	0.30	0.27	0.13	0.46	0.24	0.17	0.13
Mean wind 50-Loc	0.96	0.99	0.92	0.95	0.95	0.97	0.86	0.91	0.94	0.89	0.80	0.65	0.62	0.46	0.42	0.25	0.55	0.42	0.31	0.26
Mean wind 50-Trim	0.94	0.94	0.80	0.73	0.91	0.78	0.69	0.57	0.87	0.59	0.43	0.29	0.50	0.25	0.20	0.12	0.45	0.23	0.15	0.09
Mean wind 100-Equ	0.95	0.96	0.83	0.76	0.92	0.84	0.76	0.64	0.92	0.68	0.59	0.38	0.58	0.30	0.27	0.14	0.46	0.24	0.17	0.11
Mean wind 100-Loc	0.95	0.99	0.92	0.96	0.94	0.98	0.88	0.92	0.94	0.92	0.82	0.68	0.62	0.44	0.46	0.26	0.59	0.43	0.31	0.26
Mean wind 100-Trim	0.94	0.94	0.81	0.75	0.90	0.79	0.70	0.57	0.86	0.59	0.43	0.29	0.52	0.26	0.21	0.13	0.44	0.22	0.16	0.09
Mean wind 200-Equ	0.95	0.97	0.82	0.76	0.90	0.85	0.79	0.66	0.92	0.73	0.62	0.41	0.57	0.33	0.27	0.15	0.48	0.24	0.19	0.14
Mean wind 200-Loc	0.96	0.99	0.91	0.94	0.96	0.99	0.89	0.93	0.94	0.95	0.81	0.70	0.63	0.50	0.47	0.26	0.61	0.45	0.36	0.28
Mean wind 200-Trim	0.95	0.95	0.79	0.75	0.90	0.81	0.74	0.59	0.88	0.60	0.48	0.31	0.52	0.26	0.21	0.13	0.46	0.23	0.17	0.10
Mean wind 300-Equ	0.95	0.96	0.82	0.77	0.91	0.83	0.76	0.69	0.91	0.72	0.64	0.40	0.58	0.32	0.28	0.15	0.48	0.23	0.19	0.13
Mean wind 300-Loc	0.96	0.99	0.91	0.95	0.96	1.00	0.87	0.94	0.97	0.95	0.86	0.76	0.64	0.52	0.48	0.28	0.63	0.49	0.37	0.25
Mean wind 300-Trim	0.93	0.95	0.80	0.74	0.90	0.82	0.74	0.62	0.89	0.59	0.48	0.35	0.53	0.27	0.23	0.13	0.46	0.24	0.17	0.10
Mean wind 400-Equ	0.95	0.96	0.83	0.81	0.93	0.86	0.78	0.67	0.92	0.76	0.68	0.46	0.57	0.35	0.33	0.17	0.48	0.25	0.19	0.14
Mean wind 400-Loc	0.97	0.99	0.91	0.96	0.97	1.00	0.94	0.97	0.98	0.95	0.88	0.76	0.67	0.58	0.51	0.34	0.64	0.50	0.38	0.28
Mean wind 400-Trim	0.94	0.96	0.84	0.79	0.92	0.86	0.77	0.62	0.90	0.67	0.57	0.38	0.57	0.29	0.23	0.14	0.46	0.26	0.20	0.13
Mean wind 500-Equ	0.95	0.97	0.83	0.78	0.90	0.86	0.79	0.68	0.94	0.74	0.64	0.46	0.59	0.34	0.34	0.17	0.51	0.24	0.23	0.12
Mean wind 500-Loc	0.97	0.99	0.91	0.94	0.97	1.00	0.89	0.95	0.96	0.95	0.87	0.78	0.68	0.61	0.54	0.38	0.66	0.59	0.42	0.31
Mean wind 500-Trim	0.95	0.97	0.83	0.79	0.92	0.85	0.77	0.67	0.90	0.64	0.54	0.36	0.54	0.28	0.23	0.13	0.46	0.24	0.18	0.11
Mean wind 600-Equ	0.96	0.95	0.82	0.78	0.92	0.90	0.81	0.73	0.94	0.73	0.68	0.49	0.63	0.38	0.37	0.19	0.51	0.28	0.22	0.13
Mean wind 600-Loc	0.97	0.99	0.91	0.95	0.97	1.00	0.95	0.97	0.97	0.96	0.88	0.78	0.75	0.64	0.54	0.42	0.73	0.63	0.47	0.41
Mean wind 700-Equ	0.94	0.95	0.83	0.78	0.93	0.87	0.79	0.72	0.93	0.74	0.58	0.42	0.62	0.38	0.35	0.19	0.51	0.24	0.26	0.16
Mean wind 700-Loc	0.96	0.99	0.88	0.94	0.99	1.00	0.97	0.99	0.95	0.95	0.87	0.79	0.74	0.67	0.60	0.48	0.73	0.67	0.54	0.42
Mean wind 800-Equ	0.95	0.99	0.86	0.89	0.94	0.92	0.82	0.75	0.93	0.80	0.72	0.53	0.63	0.43	0.41	0.27	0.54	0.32	0.28	0.18
Mean wind 800-Loc	0.98	1.00	0.95	0.98	0.99	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.98
Mean wind 900-Equ	0.95	0.99	0.82	0.79	0.95	0.85	0.76	0.65	0.93	0.79	0.73	0.48	0.65	0.41	0.42	0.26	0.52	0.29	0.22	0.16
Mean wind 900-Loc	0.97	1.00	0.92	0.98	0.98	1.00	0.90	0.94	0.99	0.98	0.93	0.88	0.89	0.89	0.75	0.68	0.80	0.81	0.63	0.64

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 6: Evaluation of the proposed forecast combinations for parameter specification M3, T=4000.

MODEL M3	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=4000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.86	0.83	0.67	0.63	0.84	0.72	0.60	0.41	0.60	0.45	0.34	0.22	0.39	0.25	0.16	0.10	0.02	0.01	0.00	0.00
Mean wind 50-Equ	0.91	0.87	0.80	0.73	0.88	0.78	0.70	0.59	0.74	0.60	0.50	0.34	0.53	0.33	0.29	0.18	0.06	0.03	0.01	0.00
Mean wind 50-Loc	0.98	0.98	0.95	0.92	0.98	0.98	0.86	0.87	0.85	0.81	0.70	0.62	0.64	0.58	0.44	0.38	0.16	0.09	0.06	0.03
Mean wind 50-Trim	0.89	0.86	0.75	0.67	0.83	0.70	0.64	0.45	0.69	0.53	0.35	0.25	0.50	0.29	0.21	0.14	0.03	0.02	0.00	0.00
Mean wind 100-Equ	0.92	0.87	0.82	0.72	0.90	0.78	0.72	0.58	0.77	0.61	0.53	0.36	0.52	0.33	0.29	0.18	0.05	0.03	0.01	0.00
Mean wind 100-Loc	0.98	0.98	0.94	0.94	0.96	0.98	0.85	0.88	0.85	0.83	0.73	0.62	0.65	0.58	0.43	0.38	0.15	0.09	0.06	0.03
Mean wind 100-Trim	0.91	0.87	0.78	0.68	0.84	0.74	0.64	0.47	0.69	0.54	0.38	0.26	0.51	0.29	0.26	0.14	0.04	0.02	0.00	0.00
Mean wind 200-Equ	0.93	0.85	0.82	0.74	0.89	0.80	0.73	0.61	0.80	0.63	0.58	0.39	0.54	0.34	0.32	0.20	0.05	0.03	0.02	0.00
Mean wind 200-Loc	0.98	0.98	0.96	0.94	0.97	1.00	0.92	0.93	0.88	0.85	0.79	0.65	0.65	0.59	0.47	0.40	0.18	0.11	0.06	0.04
Mean wind 200-Trim	0.90	0.84	0.73	0.71	0.84	0.75	0.66	0.43	0.72	0.54	0.43	0.26	0.52	0.30	0.24	0.14	0.04	0.02	0.00	0.00
Mean wind 300-Equ	0.91	0.88	0.79	0.72	0.89	0.80	0.73	0.63	0.81	0.66	0.59	0.39	0.55	0.36	0.33	0.20	0.07	0.03	0.03	0.01
Mean wind 300-Loc	0.97	0.98	0.93	0.94	0.98	1.00	0.90	0.95	0.89	0.84	0.81	0.71	0.70	0.65	0.49	0.45	0.19	0.18	0.07	0.05
Mean wind 300-Trim	0.92	0.86	0.78	0.70	0.86	0.75	0.63	0.51	0.71	0.55	0.45	0.27	0.52	0.30	0.27	0.17	0.06	0.03	0.01	0.00
Mean wind 400-Equ	0.92	0.88	0.80	0.72	0.91	0.81	0.75	0.60	0.77	0.65	0.57	0.42	0.56	0.37	0.33	0.22	0.05	0.02	0.02	0.00
Mean wind 400-Loc	0.98	0.98	0.94	0.95	0.99	1.00	0.95	0.97	0.91	0.87	0.85	0.73	0.70	0.67	0.54	0.49	0.18	0.16	0.07	0.07
Mean wind 400-Trim	0.91	0.90	0.79	0.72	0.86	0.78	0.68	0.59	0.72	0.59	0.49	0.30	0.52	0.34	0.32	0.16	0.06	0.03	0.02	0.00
Mean wind 500-Equ	0.94	0.88	0.86	0.77	0.91	0.84	0.73	0.65	0.80	0.62	0.57	0.37	0.57	0.38	0.35	0.22	0.07	0.04	0.03	0.01
Mean wind 500-Loc	0.98	0.99	0.97	0.96	1.00	1.00	0.95	0.99	0.90	0.88	0.82	0.71	0.75	0.68	0.53	0.51	0.27	0.25	0.15	0.14
Mean wind 500-Trim	0.93	0.90	0.84	0.77	0.87	0.77	0.70	0.52	0.72	0.59	0.46	0.28	0.52	0.34	0.30	0.17	0.07	0.04	0.02	0.01
Mean wind 600-Equ	0.92	0.90	0.79	0.73	0.90	0.85	0.78	0.69	0.83	0.63	0.62	0.45	0.60	0.41	0.35	0.29	0.10	0.06	0.03	0.00
Mean wind 600-Loc	0.97	0.98	0.91	0.92	1.00	1.00	0.98	1.00	0.93	0.88	0.82	0.75	0.80	0.74	0.57	0.60	0.36	0.31	0.22	0.19
Mean wind 700-Equ	0.89	0.87	0.75	0.70	0.89	0.83	0.76	0.67	0.80	0.65	0.59	0.41	0.57	0.39	0.34	0.24	0.07	0.02	0.04	0.01
Mean wind 700-Loc	0.92	0.96	0.87	0.91	1.00	1.00	0.98	1.00	0.89	0.87	0.82	0.75	0.77	0.72	0.53	0.57	0.28	0.24	0.09	0.11
Mean wind 800-Equ	0.93	0.93	0.80	0.74	0.92	0.88	0.76	0.71	0.82	0.77	0.66	0.47	0.61	0.49	0.37	0.27	0.08	0.04	0.04	0.02
Mean wind 800-Loc	0.98	0.99	0.96	0.99	0.99	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99	1.00	1.00	0.99	0.99
Mean wind 900-Equ	0.93	0.93	0.83	0.80	0.89	0.81	0.69	0.58	0.85	0.73	0.67	0.44	0.62	0.43	0.37	0.28	0.08	0.03	0.03	0.02
Mean wind 900-Loc	0.99	1.00	0.97	0.97	0.94	0.98	0.83	0.93	0.98	0.97	0.96	0.91	0.87	0.87	0.73	0.77	0.45	0.44	0.25	0.26

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 7: Comparison among forecast combinations for parameter specification M1, T=3000.

MODEL M1	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=3000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.89	0.87	0.71	0.75	0.88	0.86	0.77	0.71	0.86	0.85	0.69	0.63	0.67	0.64	0.45	0.40	0.79	0.78	0.59	0.60
RS Mean	0.99	1.00	0.97	0.97	0.96	0.97	0.90	0.93	0.95	0.92	0.85	0.85	0.88	0.92	0.71	0.70	0.85	0.87	0.69	0.73
RS Mean Trim	0.98	0.99	0.92	0.99	0.97	0.99	0.92	0.96	0.94	0.93	0.84	0.82	0.77	0.77	0.62	0.57	0.84	0.81	0.72	0.69
CM combined	0.93	0.95	0.90	0.89	0.96	0.94	0.89	0.90	0.99	0.97	0.87	0.86	0.99	0.99	0.94	0.91	0.94	0.93	0.82	0.84
Mean wind 800-Equ	0.97	0.97	0.88	0.88	0.93	0.96	0.85	0.86	0.97	0.94	0.81	0.83	0.87	0.90	0.69	0.67	0.85	0.89	0.65	0.71
Mean wind 800-Loc	0.98	0.98	0.90	0.96	0.96	0.96	0.91	0.92	0.99	0.98	0.96	0.96	0.96	0.97	0.89	0.89	0.96	0.97	0.90	0.91

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 8: Comparison among forecast combinations for parameter specification M2, T=3000.

MODEL M2	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=3000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.82	0.80	0.57	0.54	0.73	0.71	0.42	0.36	0.54	0.47	0.31	0.23	0.33	0.28	0.15	0.14	0.40	0.34	0.21	0.22
RS Mean	0.97	0.97	0.86	0.85	0.90	0.91	0.82	0.83	0.98	0.95	0.81	0.77	0.63	0.69	0.40	0.40	0.62	0.60	0.41	0.46
RS Mean Trim	0.96	0.97	0.92	0.94	0.99	0.99	0.97	0.96	0.88	0.83	0.66	0.60	0.51	0.47	0.30	0.30	0.52	0.50	0.38	0.34
CM combined	0.94	0.91	0.74	0.71	0.89	0.86	0.76	0.72	0.99	0.99	0.97	0.94	0.94	0.95	0.85	0.86	0.89	0.88	0.73	0.72
Mean wind 800-Equ	0.92	0.91	0.76	0.74	0.85	0.83	0.73	0.68	0.88	0.90	0.63	0.62	0.58	0.65	0.33	0.34	0.56	0.58	0.35	0.40
Mean wind 800-Loc	0.98	0.96	0.82	0.83	0.94	0.94	0.85	0.85	1.00	1.00	0.97	0.93	0.96	0.96	0.91	0.91	0.91	0.93	0.78	0.82

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 9: Comparison among forecast combinations for parameter specification M3, T=3000.

	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
MODEL M3	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=3000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.73	0.70	0.48	0.49	0.72	0.68	0.39	0.33	0.48	0.37	0.22	0.18	0.07	0.06	0.01	0.01	0.02	0.05	0.01	0.01
RS Mean	0.93	0.92	0.82	0.79	0.98	0.98	0.85	0.84	0.94	0.93	0.68	0.69	0.39	0.42	0.13	0.11	0.30	0.29	0.14	0.14
RS Mean Trim	0.99	0.99	0.97	0.98	1.00	1.00	0.99	0.99	0.79	0.73	0.45	0.44	0.25	0.22	0.07	0.05	0.18	0.18	0.12	0.12
CM combined	0.88	0.84	0.66	0.68	0.92	0.90	0.74	0.72	0.99	0.99	0.93	0.93	0.84	0.83	0.63	0.62	0.66	0.66	0.52	0.55
Mean wind 800-Equ	0.89	0.87	0.69	0.65	0.88	0.89	0.65	0.67	0.80	0.82	0.49	0.51	0.46	0.51	0.24	0.24	0.25	0.28	0.05	0.09
Mean wind 800-Loc	0.94	0.93	0.82	0.77	0.96	0.96	0.83	0.84	0.98	0.95	0.89	0.88	0.97	0.97	0.95	0.94	0.90	0.90	0.82	0.82

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 10: Comparison among forecast combinations for parameter specification M1, T=4000.

	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
MODEL M1	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
T=4000	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.85	0.87	0.66	0.69	0.89	0.89	0.72	0.68	0.84	0.81	0.63	0.58	0.71	0.70	0.49	0.47	0.66	0.63	0.47	0.41
RS Mean	0.94	0.96	0.82	0.84	0.99	0.99	0.91	0.92	0.95	0.97	0.87	0.90	0.84	0.85	0.64	0.66	0.70	0.70	0.53	0.52
RS Mean Trim	0.97	0.98	0.92	0.97	1.00	1.00	0.96	0.97	0.96	0.98	0.94	0.93	0.79	0.82	0.60	0.55	0.71	0.67	0.58	0.54
CM combined	0.92	0.95	0.73	0.74	0.96	0.99	0.80	0.81	0.94	0.95	0.86	0.84	0.92	0.91	0.76	0.71	0.82	0.81	0.65	0.60
Mean wind 800-Equ	0.92	0.95	0.81	0.83	0.92	0.97	0.83	0.86	0.94	0.93	0.85	0.82	0.93	0.94	0.74	0.72	0.78	0.78	0.59	0.58
Mean wind 800-Loc	0.94	0.95	0.87	0.88	0.98	1.00	0.88	0.94	0.96	0.98	0.90	0.93	0.99	1.00	0.99	1.00	0.98	0.99	0.96	0.96

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 11: Comparison among forecast combinations for parameter specification M2, T=4000.

MODEL M2	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
T=4000	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.85	0.82	0.61	0.56	0.77	0.74	0.40	0.42	0.59	0.57	0.30	0.25	0.22	0.23	0.07	0.06	0.20	0.15	0.08	0.09
RS Mean	0.98	0.97	0.92	0.88	0.95	0.96	0.86	0.88	0.95	0.95	0.88	0.84	0.58	0.48	0.33	0.30	0.38	0.34	0.23	0.22
RS Mean Trim	0.99	1.00	0.97	0.98	0.99	1.00	0.95	0.98	1.00	1.00	0.97	0.99	0.44	0.41	0.25	0.27	0.37	0.28	0.19	0.21
CM combined	0.95	0.91	0.80	0.76	0.95	0.90	0.76	0.75	0.95	0.93	0.82	0.79	0.62	0.62	0.40	0.38	0.57	0.57	0.35	0.37
Mean wind 800-Equ	0.96	0.95	0.85	0.84	0.90	0.91	0.71	0.74	0.84	0.84	0.60	0.61	0.61	0.60	0.33	0.37	0.43	0.44	0.20	0.19
Mean wind 800-Loc	0.98	0.97	0.91	0.93	0.96	0.98	0.89	0.91	0.97	0.98	0.87	0.85	1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 12: Comparison among forecast combinations for parameter specification M3, T=4000.

MODEL M3	$\tau = 0.5T$				$\tau = 0.6T$				$\tau = 0.7T$				$\tau = 0.8T$				$\tau = 0.9T$			
T=4000	$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$		$\alpha = 0.10$		$\alpha = 0.25$	
	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ	R	SQ
Expanding Wind	0.81	0.81	0.57	0.56	0.73	0.69	0.41	0.40	0.51	0.53	0.22	0.24	0.14	0.17	0.03	0.06	0.01	0.01	0.00	0.00
RS Mean	0.95	0.97	0.85	0.86	0.98	0.97	0.86	0.84	0.93	0.91	0.81	0.76	0.35	0.38	0.22	0.19	0.08	0.07	0.04	0.04
RS Mean Trim	1.00	1.00	0.96	0.97	0.99	1.00	0.95	0.95	0.99	1.00	0.98	0.98	0.35	0.34	0.14	0.13	0.06	0.07	0.04	0.03
CM combined	0.88	0.88	0.73	0.73	0.94	0.91	0.78	0.78	0.90	0.86	0.72	0.69	0.48	0.47	0.32	0.32	0.19	0.19	0.08	0.09
Mean wind 800-Equ	0.90	0.91	0.76	0.78	0.89	0.90	0.69	0.69	0.74	0.75	0.55	0.51	0.48	0.47	0.28	0.29	0.07	0.08	0.03	0.03
Mean wind 800-Loc	0.96	0.98	0.85	0.88	0.97	1.00	0.92	0.92	0.91	0.94	0.79	0.76	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Note: Entries denote the relative frequencies a given combination enters in the MCS at level α for the statistics R and SQ; τ indicates the location of the break.

Table 13: Break dates identified by Rapach and Strauss (2008) for the U.S. dollar exchange rate returns series.

Canada	**Denmark**	**Germany**	**Japan**
02/01/1980	02/01/1980	02/01/1980	02/01/1980
03/08/1998	20/09/1993	26/12/1996	05/05/1997
10/03/2003	29/08/2000		30/03/2000
26/04/2004	06/04/2001		
31/08/2005	31/08/2005	31/08/2005	31/08/2005

Norway	**Switzerland**	**U.K.**	**U.S.(t.w.)**
02/01/1980	02/01/1980	02/01/1980	02/01/1980
17/12/1990	03/10/1995	05/02/1984	19/09/1995
31/08/1992	24/08/1998	25/09/1985	26/12/1996
		05/03/1991	13/04/2000
		14/10/1995	
31/08/2005	31/08/2005	31/08/2005	31/08/2005

Table 14: MCS p-values for the U.S. dollar exchange rate returns series

	Canada		**Denmark**		**Germany**		**Japan**	
	R	SQ	R	SQ	R	SQ	R	SQ
Expanding wind	0.00	0.00	0.10	0.10	0.09	0.09	0.11	0.11
RS Mean	0.00	0.00	0.22	0.22	0.22	0.22	0.02	0.02
RS Mean Trim	1.00	1.00	0.44	0.44	1.00	1.00	0.02	0.02
CM combined	0.00	0.00	0.14	0.14	0.09	0.09	0.64	0.64
Mean wind 800-Equ	0.00	0.00	0.33	0.33	0.26	0.26	0.17	0.17
Mean wind 800-Loc	0.00	0.00	1.00	1.00	0.91	0.91	1.00	1.00

	Norway		**Switzerland**		**U.K.**		**U.S. (tw)**	
	R	SQ	R	SQ	R	SQ	R	SQ
Expanding wind	0.01	0.01	0.00	0.00	0.21	0.21	0.06	0.06
RS Mean	0.33	0.33	0.01	0.01	0.28	0.28	0.71	0.71
RS Mean Trim	0.63	0.63	1.00	1.00	1.00	1.00	1.00	1.00
CM combined	0.06	0.06	0.00	0.00	0.28	0.28	0.90	0.90
Mean wind 800-Equ	0.25	0.25	0.00	0.00	0.25	0.25	0.14	0.14
Mean wind 800-Loc	1.00	1.00	0.42	0.42	0.25	0.25	1.00	1.00

Appendix A. The model confidence set

The objective of this procedure is to determine which methods, from an initial set M^0 of methods indexed by $i = 1, \dots, M^0$, exhibit the same predictive ability in term of a loss function, given a level of confidence. In this approach, the comparison is made by using a loss function which is used to rank competing methods in term of forecasting accuracy.

Let us consider $\hat{M}*$ as the collection of the best methods, M^0 the initial collection of all the methods and $L_{i,t}$ the loss function associated with the method i in period t.

Define the relative performance variables as $d_{ij,t} = L_{i,t} - L_{j,t} \quad \forall i, j \in M^0$ and assumes that $E(d_{ij,t})$ is finite and does not depend on t. The set of the best models is defined by:

$$\hat{M}* = \left\{ i \in M^0 : E(d_{ij,t}) \leq 0 \quad \forall j \in M^0 \right\} \tag{16}$$

In order to determine $\hat{M}*$, a sequence of significance tests is made and the models that result to be significantly inferior to other element of M^0 are eliminated. The hypothesis that is being tested is:

$$H_{0,M} : E(d_{ij,t}) = 0 \quad \text{for all } i, j \in M \subset M^0 \tag{17}$$

$$H_{A,M} = E(d_{ij,t}) \neq 0 \quad \text{for some } i, j \in M \subset M^0$$

Alternatively, the previous hypothesis can be also formulated as:

$$H_{0,M} : E(d_{i.,t}) = 0 \quad \text{for all } i \in M \subset M^0 \tag{18}$$

$$H_{A,M} = E(d_{i.,t}) \neq 0 \quad \text{for some } i \in M \subset M^0$$

where $d_{i.,t} = (m-1)^{-1} \sum_{j \in M} d_{ij,t}$.

The MSC is a stepwise procedure which starts by setting $M = M_0$. The test $H_{0,M}$ is then implemented at level α. If $H_{0,M}$ is not rejected $\hat{M}^*_{1-\alpha} = M$; if $H_{0,M}$ is rejected an object from M is eliminatee and the procedure is repeated until $H_{0,M}$ is not rejected. The set $\hat{M}^*_{1-\alpha}$ is defined as the ”superior set of models” and it contains the surviving method.

Despite its sequential nature, the MCS procedure does not accumulate type I error. This is due to the fact that the test stops when the first hypothesis is not rejected.

Let define the following two t-statistics that form the basis of tests of hypothesis (17) and (18):

$$t_{ij} = \frac{\overline{d}_{ij}}{\sqrt{v\hat{a}r(\overline{d}_{ij})}} \qquad t_{i.} = \frac{\overline{d}_{i.}}{\sqrt{v\hat{a}r(\overline{d}_{i.})}} \tag{19}$$

where:

$$\overline{d}_{ij} = \frac{1}{n}\sum_{t=1}^{n} d_{ij,t} \qquad \overline{d}_{i.} = \frac{1}{m}\sum_{j\in M} \overline{d}_{ij} \tag{20}$$

are, respectively, the relative sample loss between ith and jth models and the sample loss of the i-th model relative to the average across models; $v\hat{a}r(\overline{d}_{ij})$ and $v\hat{a}r(\overline{d}_{i.})$ denote estimates of the variance of $\overline{d}_{ij}$ and $\overline{d}_{i.}$ respectively.

The two hypothesis (17) and (18) respectively, map naturally into the two test statistics:

$$T_{R,M} = \max_{i,j\in M} |t_{ij}| \qquad T_{SQ,M} = \max_{i\in M} t_{i.} \tag{21}$$