Prediction Regions for Interval-Valued Time Series[1]

Gloria GONZÁLEZ-RIVERA[2]

and

Yun LUO

Department of Economics
University of California, Riverside, CA, USA

Esther RUIZ

Department of Statistics
Universidad Carlos III de Madrid, Spain

September 21, 2019

[1]We are grateful to participants in the International Symposium on Forecasting in Cairns 2017, in the 2nd International Symposium on Interval Data Modeling, Xiamen University and in the Workshop in Financial Econometrics, in Zaragoza (Spain) 2019 for useful comments. Gloria González-Rivera acknowledges financial support from the 2015/2016 Chair of Excellence UC3M/Banco de Santander and the UC-Riverside Academic Senate grants. Esther Ruiz and Gloria González-Rivera are grateful to the Spanish Government contract grant ECO2015-70331-C2-2-R (MINECO/FEDER).

[2]Corresponding author: gloria.gonzalez@ucr.edu, (951) 827-1590

Abstract

We approximate probabilistic forecasts for interval-valued time series by offering alternative approaches. After fitting a possibly non-Gaussian bivariate VAR model to the center/log-range system, we transform prediction regions (analytical and bootstrap) for this system into regions for center/range and upper/lower bounds systems. Monte Carlo simulations show that bootstrap methods are preferred according to several new metrics. For daily S&P500 low/high returns, we build joint conditional prediction regions of the return level and volatility. We illustrate the usefulness of obtaining bootstrap forecasts regions for low/high returns by developing a trading strategy and showing its profitability when compared to using point forecasts.

Key Words: Bootstrap, Constrainted Regression, Coverage Rates, Logarithmic Transformation, QML estimation.

JEL Classification: C01, C22, C53

1 Introduction

Data sets in interval format are common in economics and finance. First, in some instances, intervals could be the only data available to the researcher. For example, when data is sensitive to privacy concerns, the records must be aggregated and measured as intervals. Wealth and income are measured as intervals in the Health and Retirement Study (HRS) and Current Population Survey (CPS), respectively; see Manski and Tamer (2002). Also, in bond markets, traders report bid/ask intervals; see Pascual and Veredas (2010) and Bien *et al.* (2011). Interval data also appears in contingent valuation surveys that aim to elicit the willingness to pay of respondents for some good not currently marketed; see, for instance, Fernández et al. (2004). In energy markets, the US Energy Information Administration provides min/max retail prices of electricity at the state level; see García-Ascanio and Maté (2010) for an application to electric power demand. The US Department of Agriculture also provides daily low and high prices on agricultural commodities and livestock; see Lin and González-Rivera (2016) and Xiong et al. (2015) for applications. Second, interval data can also appear when dealing with big data sets to reduce dimensionality. For example, in stock markets, when intra-day prices are available, one could analyze daily intervals of low/high asset prices; see, for example, Rodrigues and Salish (2015). One advantage of analyzing low/high stock prices intervals instead of all available intra-day returns is that problems related to irregular temporal spacing, strong diurnal patterns, microestructure noise or complex dependences, are avoided; see Engle and Russell (2010) for the problems associated to high frequency data. Additionally, interval prices obviously have more information than one-point prices. High and low prices have jointly information about both level and volatility. Thus, we expect more efficient estimation by using interval-valued data; see, among others, Xiong *et al.* (2017) for the advantages of high/low prices when compared to one-point prices.

In this paper, we focus on interval-valued times series defined as a collection of interval realizations ordered over time, i.e., $\{(y_{l,t}, y_{u,t})\}$ for $t = 1, ...T$, where $y_{l,t}$ is the lower bound and $y_{u,t}$ is the upper

bound of the interval at time t, such that $y_{l,t} \leq y_{u,t}$ for all t. An equivalent representation is given by $\{(C_t, R_t)\}$ where $C_t = (y_{l,t} + y_{u,t})/2$ and $R_t = y_{u,t} - y_{l,t} \geq 0$, are the center and range of the interval, respectively. Most of the econometric analysis in this area has focused on model estimation and inference and, though it is possible to construct point forecasts based on a given model or algorithm, the question of constructing probabilistic forecasts for interval data has not been addressed yet. This is the main question that we aim to analyze in this paper. There are several routes to construct probabilistic forecasts for interval valued time series, which involve some trade-offs between estimation and prediction decisions.

When dealing with lower/upper bounds systems, one needs to incorporate the constraint $y_{l,t} \leq y_{u,t}$ into the estimation.[1,2] Alternatively, dealing with the center/range system, one needs to incorporate the constraint $R_t \geq 0$.[3] In this paper, we follow Tu and Wang (2016) who overcome this restriction by log-transforming the range, and estimating the center/log-range system without imposing any distributional assumptions. However, forecasting the center/range or lower/upper bounds will be more complicated. First, for point forecasts, one needs the inverse transformation, i.e. $R_t = \exp[\log R_t]$, which itself introduces non-trivial econometric issues. Second, for a density forecast, a joint distributional assumption for the center and range or for the upper and lower bounds is required. Consequently, we propose constructing prediction regions for the center/log-range system using not only the bivariate normality assumption but also bootstrap procedures that do not require any specific assumption on the forecast error distribution; see Fresoli *et al.* (2015) for the bootstrap procedure. Based on the prediction regions of the center/log-range system, we construct prediction regions for the center/range system and for the upper/lower bounds system by implementing both analytical and numerical approaches. An important advantage of our approach is that, by focusing

[1] González-Rivera and Lin (2013) propose two-step estimators based on assuming a truncated bivariate normal density of the errors of the lower/upper bounds. The estimation of the system is complex but, if the distributional assumption is adequate, it is possible to construct a direct bivariate density forecast for the upper/lower bounds.

[2] Han, Hong and Wang (2016) propose the conditional interval (ACI) model that is based on the concept of "extended" interval for which the left bound needs not to be smaller than the right bound.

[3] Lima Neto and De Carvalho (2010) impose non-negative constraints on the parameters of the range equation, which are unnecessarily too restrictive and complicate the estimation of the system.

on prediction regions rather than on point forecasts, we avoid the biases associated with the exp-transformation of point forecasts of log-transformed variables. Note that we estimate only one unconstrained system, the center/log-range system, but our objective is the prediction regions of the center/range or upper/lower bounds system. However, if the researcher is interested just on the dynamics of these constrained systems, she needs to take into account the constraint in the estimation of the systems.

We assess the finite sample out-of-sample performance of the prediction regions considered and compare their performance according to several metrics. There is a rather extense literature on the evaluation of point multivariate forecasts; see, for example, Komunjer and Owyang (2012). Also, recently, several authors have proposed different evaluation criteria for multivariate forecast densities; see Diebold *et al.* (1998), Clements and Smith (2002), Gneiting *et al.* (2008), González-Rivera and Yoldas (2011) and González-Rivera and Sun (2015). However, the literature on evaluating multivariate prediction regions is rather thin.[4] The most basic required property is coverage so that regions are reliable when the empirical and nominal coverages are close. To our knowledge, there is one additional metric that brings the volume of the region to interact with its coverage (Golestaneh *et al.*, 2017). In this paper, we also contribute to this literature by introducing several new measures that account for (i) the location of out-of-the-region points with respect to a central point of the region, (ii) the tightness of the intervals that result from projecting the two-dimensional region into one-dimensional intervals, and (iii) the distance of the also projected out-of-the-region points to the projected one-dimensional interval. These new measures bring a notion of risk associated with the prediction region. In addition, we also provide a description of the distribution of the out-of-the-region points around the region to measure whether the region is probability-centered. We show that, even for Gaussian systems, bootstrap methods deliver the best performance, mainly when the estimation sample is small and estimation uncertainty is most relevant. For non-Gaussian systems, bootstrap regions are preferred regardless of the sample size.

[4]See Rodrigues and Salish (2015) for accuracy measures of interval-valued point forecasts.

Finally, we construct forecast regions for a time series of daily low/high S&P500 returns. In addition to illustrating the procedures described in this paper, this empirical application has interest on its own. As mentioned above, daily low/high return intervals are more informative than just daily one-point measurements (end-of-day return) as they have simultaneous information about the level and volatility and, at the same time, they avoid some of the problems often associated with high frequency intra-daily observations. Several authors show that range-based measures of volatility may be prefered to other alternative measures based either on one-point returns or high-frequency intra-day returns; see Parkinson (1980), Rogers and Satchell (1991), Rogers *et al.* (1994), Alizadeh *et al.* (2002), Brand and Diebold (2006) and Suh and Zhang (2006), among others. Furthermore, it is also important to measure the uncertainty of the volatility itself; see Vorbrink (2014), Blasques *et al.* (2016) and Ji and Shi (2018) for the relevance of measuring the uncertainty of volatility in the context of financial models. In addtition, when computing forecast intervals for future returns and volatilities, one should take into account that both quantities are usually correlated. For example, Cheung *et al.* (2009) point out that a proper specification of the range using only its own history may be inferior to a model that jointly describes the behaviour of high and low prices. In the interval approach described in this paper, we forecast jointly future low/high return intervals and construct prediction regions of the center and range of the interval at any desired horizon that do not require parametric distributional assumptions. The forecast regions incorporate the dependence between returns and volatilities when building joint forecast regions for the center/log-range or low/high systems. Overall, the main advantage of the proposed approach is that it allows for the modeling of the joint conditional density of the level and volatility of returns, which in our sample are contemporaneous and negatively correlated, and consequently allows for the construction of bivariate density forecasts. We also carry out a trading strategy that illustrates the economic advantages of taking into account the probabilistic forecasts of the interval returns instead of point forecasts.

The organization of the paper is as follows. Section 2 establishes notation by describing the

VAR model for the center/log-range system and the construction of forecast regions both with and without the normality assumption. Section 3 shows how to construct forecast regions for the center/range and upper/lower bounds systems, starting from the forecast regions of the center/log-range system. Section 4 proposes several new metrics to evaluate the performance of prediction regions while Section 5 reports Monte Carlo simulations carried out to compare the performance of the proposed procedures to construct forecast regions. Section 6 constructs prediction regions for S&P500 low/high return intervals as the basis for implementing profitable trading strategies. Section 7 concludes.

2 Forecasting the center/log-Range system

2.1 Point forecasts

Let us call $y_{c,t} \equiv C_t$, $y_{r,t} \equiv \log R_t$ and $Y_t \equiv (y_{c,t}, y_{r,t})'$. We start by fitting the following VAR(p) model for the center/log-range system

$$Y_t = A + \sum_{i=1}^{p} B_i Y_{t-i} + \varepsilon_t \tag{2.1}$$

where A and B_i are parameter matrices restricted to satisfy the usual stationarity conditions[5] and ε_t is a bivariate white noise process with covariance matrix Ω. The estimation of the parameters proceeds by LS, which is consistent and asymptotically normal under standard assumptions.[6]

Given $Y_1, ..., Y_T$, if the loss function is quadratic, optimal h-step-ahead point forecasts of Y_{T+h} are given by

$$Y_{T+h|T} = A + \sum_{i=1}^{p} B_i Y_{T+h-i|T} \tag{2.2}$$

[5]Note that the VAR(p) model is not subject to further restrictions as we are log-transforming the range.

[6]Tu and Wang (2016) used the estimator of Yao and Zhao (2013) that relies on kernel estimates of the likelihood. This estimator is computationally more demanding than LS and depends on the choice of tuning parameters. Their empirical results suggest that both estimators are very similar and, consequently, we focus on the LS estimator.

where $Y_{T+h-i|T} = Y_{T+h-i}$ for $i \geq h$. The forecast error covariance matrix is given by $W_h = \Omega + \sum_{i=1}^{h-1} \Psi_i \Omega \Psi_i'$ where the matrices Ψ_i come from the MA(∞) representation of Y_t. In practice, consistent LS estimates are plugged in $Y_{T+h|T}$ and W_h to obtain the estimated h-step-ahead point forecasts and their estimated covariance matrices, denoted by $\hat{Y}_{T+h|T}$ and $\hat{W}_h$, respectively.

2.2 Forecast regions under normality

If the center/log-range system is bivariate normal, pointwise bivariate density forecasts can be obtained as follows

$$Y_{T+h} \rightarrow N(\hat{Y}_{T+h|T}, \hat{W}_h), \tag{2.3}$$

and h-step-ahead forecast ellipses for Y_{T+h} with coverage $100 \times (1-\alpha)\%$ are given by

$$NE_{T+h} = \big[Y_{T+h} \quad \text{such that} \quad (Y_{T+h} - \hat{Y}_{T+h|T})' \hat{W}_h^{-1} (Y_{T+h} - \hat{Y}_{T+h|T}) \leq q_{1-\alpha}\big], \tag{2.4}$$

where $q_{1-\alpha}$ is the $(1-\alpha)$ quantile of the chi-square distribution with 2 degrees of freedom.

As proposed by Lutkephol (1991), we could also construct forecast regions by using Bonferroni rectangles (BR), which are simpler and rather popular among practitioners. A BR with (at least) $100 \times (1-\alpha)\%$ coverage has the following sides

$$\big[b_{c,\alpha/4}, b_{c,1-\alpha/4}\big] \equiv \big[\hat{y}_{c,T+h|T} - z_{\alpha/4}\sqrt{\hat{W}_{h,11}}, \quad \hat{y}_{c,T+h|T} + z_{\alpha/4}\sqrt{\hat{W}_{h,11}}\big] \tag{2.5}$$

$$\big[b_{r,\alpha/4}, b_{r,1-\alpha/4}\big] \equiv \big[\hat{y}_{r,T+h|T} - z_{\alpha/4}\sqrt{\hat{W}_{h,22}}, \quad \hat{y}_{r,T+h|T} + z_{\alpha/4}\sqrt{\hat{W}_{h,22}}\big], \tag{2.6}$$

where $z_{\alpha/4}$ is the $\alpha/4$-quantile of the standard normal distribution. To include the contemporaneous linear correlation between the center and log-range, the BR can be modified as in Fresoli *et al.*

(2015). The corners of the modified rectangle are as follows

$$[b_{c,\alpha/4}, \quad b_{r,\alpha/4} + p_h b_{c,\alpha/4}], \qquad [b_{c,\alpha/4}, \quad b_{r,1-\alpha/4} + p_h b_{c,\alpha/4}], \tag{2.7}$$
$$[b_{c,1-\alpha/4}, \quad b_{r,\alpha/4} + p_h b_{c,1-\alpha/4}], \quad [b_{c,1-\alpha/4}, \quad b_{r,1-\alpha/4} + p_h b_{c,1-\alpha/4}]$$

where $p_h = \hat{W}_{h,21}/\hat{W}_{h,11}$. Although the area of the modified Bonferroni rectangle (MBR) is the same as that of the BR, their coverage may be slightly different depending on the quantiles associated with the modified terms, e.g., $b_{r,\alpha/4} + p_h b_{c,\alpha/4}$, which in turn depend on the magnitude and sign of p_h. Simulation results will provide some information on the coverage rate of the MBR.

To illustrate the shapes of the three forecast regions described above, Figure 1 (top panel) plots the one-step-ahead 95% NE, BR and MBR together with 1000 realizations of Y_{T+1} generated by a VAR(4) model with Gaussian errors and a contemporaneous correlation of -0.24. The forecast regions have been obtained after estimating the parameters using T=1000 observations. We observe that both the NE and MBR regions are able to capture the negative correlation between the center and the log-range while the BR cannot inform about this correlation. Note that the BR has large empty areas without any realization of Y_{T+1}.

2.3 Forecast regions under non-normality

Often the normal distribution is not a good approximation for the distribution of the center/log-range system. In this case, we can obtain bootstrap point-wise forecast densities that incorporate parameter uncertainty without relying on any specific forecast error distribution. Note that even in the Gaussian case, if the estimation sample is not very large, the effect of parameter estimation on the forecast uncertainty may not vanish, and so the use of bootstrap forecast densities may be desired. In this paper, we implement the asymptotically valid bootstrap algorithm of Fresoli *et al.* (2015) to obtain bootstrap replicates of $Y^{*(b)}_{T+h|T}$, for $b = 1, ..., B$.

The bootstrap replicates can be used to obtain the following point-wise bootstrap ellipse with $100 \times (1-\alpha)\%$ coverage

$$BE_{T+h} = \left[Y_{T+h} | [Y_{T+h} - \bar{Y}^*_{T+h|T}]' S_{Y^*}(h)^{-1} [Y_{T+h} - \bar{Y}^*_{T+h|T}] \leq q^*_{1-\alpha}\right], \tag{2.8}$$

where $\bar{Y}^*_{T+h|T}$ is the sample mean of the B bootstrap replicates $Y^{*(b)}_{T+h|T}$, $S_{Y^*}(h)$ is the corresponding sample covariance matrix and $q^*_{1-\alpha}$ is the $(1-\alpha)$ quantile of the empirical distribution of the quadratic form $[Y^{*(b)}_{T+h|T} - \bar{Y}^*_{T+h|T}]' S_{Y^*}(h)^{-1} [Y^{*(b)}_{T+h|T} - \bar{Y}^*_{T+h|T}]$.

Point-wise bootstrap prediction regions for the center/log-range system can also be constructed as BR with at least $100 \times (1-\alpha)\%$ coverage with corners obtained from the marginal bootstrap distributions of the center and the log-range and denoted as BBR. These BBR regions can be correct for the contemporaneous correlation between the center and the log-range as in (2.7).

Note that, when the joint distribution of the center/log-range system is not symmetric, neither the BE nor the BBR need to be probability-centered; see, for example, Beran (1993) for the desirable properties of multivariate forecast regions. In this case, the BE and BBR regions will only be approximations to the true shape of the bootstrap forecasts. Alternatively, probability-centered forecast regions can be constructed using the convex hull peeling method of Tukey (1975) that consists of constructing a series of convex prediction polygons. Given a data cloud, the first layer of the Tukey convex hull is the convex polygon formed by the boundary of the data. It continues by peeling the first layer off and finding the second layer for the remaining data. This process is repeated until no convex polygon can be constructed any more. Given the two-dimensional bootstrap data cloud, $Y^{*(b)}_{T+h|T}$, we construct the Tukey nonparametric region by choosing the polygon that provides the closest coverage to the desired nominal coverage rate.

As an illustration of the bootstrap regions, Figure 1 (top panel) plots the 95% BE, BBR and BMBR regions for Y_{T+1} generated as described in the previous section. These regions are based on

B=4000 bootstrap replicates. Recall that the estimation sample size is $T = 1000$ and, consequently, the uncertainty due to parameter estimation is negligible. Therefore, the normal and bootstrap ellipses have identical shapes. The Tukey hull follows very closely these ellipses. The BBR are also very similar to their normal counterparts.

3 Regions for center/range and lower/upper systems

When forecasting interval-valued time series, often the interest is not the center/log-range system but either the center/range or the lower/upper systems. In this section, we describe analytical and numerical methods to construct prediction regions for these systems.

First, under bivariate normality of center/log-range, the bivariate density of the center/range system can be estimated as follows

$$f(y_{c,T+h}, R_{T+h}) = \frac{1}{2\pi\sqrt{|\hat{W}_h|}} \frac{1}{R_{T+h}} \exp[-\frac{1}{2}(Y_{T+h} - \hat{Y}_{T+h|T})'\hat{W}_h^{-1}(Y_{T+h} - \hat{Y}_{T+h|T})]. \tag{3.1}$$

Since the center and the range are linear combinations of the upper and lower bounds, it is easy to see that that the conditional bivariate density of the lower/upper bounds is also given by (3.1). Analytical contours for the center/range (or lower/upper bounds) system can be constructed by horizontally cutting the bivariate density in (3.1) at a value determined by the nominal coverage and obtained by numerical simulation. Figure 1 (bottom panel) plots 1000 realizations of (C_{T+1}, R_{T+1}) based on the same system described above, together with the 95% forecast region obtained using the analytical density in (3.1). We observe that, as expected, the region is not an ellipse.

Second, note that any of the prediction regions constructed for the center/log-range system can be directly transformed into a prediction region for the center/range system. For instance, consider the NE region with $100 \times (1 - \alpha)\%$ probability coverage. Its boundary is the $(1 - \alpha)$ bivariate

quantile. The boundary points (center, log-range) can be transformed into another boundary of points (center, exp(log-range)) of a prediction region for the center/range system, that is,

$$\text{T-NE}_{T+h} = \left\{ \left[(y_{c,T+h}, \exp[\log R_{T+h}])' \right] \quad \text{such that} \quad (Y_{T+h} - \hat{Y}_{T+h|T})' \hat{W}_h^{-1} (Y_{T+h} - \hat{Y}_{T+h|T}) = q_{1-\alpha} \right\}. \tag{3.2}$$

The new region will not preserve the shape of an ellipse but will have the same coverage because the exponential function is a monotonic transformation. Furthermore, the transformation has the advantage of delivering strictly positive values for the range.[7]

In Figure 1 (bottom panel), we illustrate the shape of the T-NE regions using the same simulated example considered before. The transformed shape is similar to the analytical although they are not identical.

Similarly, the BR and MBR regions can be transformed by taking the exponential transformation of the log-range intervals in (2.6) and (2.7), respectively. Figure 1 (bottom panel) illustrates the shapes of the transformed BR and MBR regions. As expected, while the transformed MBR region shows the correlation between center and range, the transformed BR region does not and some portions of its area are empty.

It is important to note that the transformation of the forecast regions constructed for the center/log-range system is not feasible when constructing prediction regions for the lower/upper bounds system because there is not a monotonic transformation from the boundary points of the center/log-range region to the boundary points of the lower/upper bounds region.

Finally, if one avoids the normality assumption of the center/log-range system and wishes to construct bootstrap forecast regions for the center/range system, the regions can also be transformed. For example, by transforming the points (center, log-range) sitting on the boundary of (2.8) to

[7]Note that, by focusing on prediction regions rather than on point forecasts, we avoid the biases associated with the exp-transformation of point forecasts of log-transformed variables, for which corrections are necessary; see, for example, Granger and Newbold (1976) and Guerrero (1993).

points (center, exp(log-range)), we obtain the following transformed bootstrap ellipse

$$\text{T-BE}_{T+h} = \left\{ \left[(y_{c,T+h}, \exp[\log R_{T+h}])' \right] \quad \text{such that} \quad (Y_{T+h} - \bar{Y}^*_{T+h|T})' S_{Y^*}(h)^{-1} (Y_{T+h} - \bar{Y}^*_{T+h|T}) = q^*_{1-\alpha} \right\} \tag{3.3}$$

Similarly, we obtain the transformed bootstrap Bonferroni and modified Bonferroni rectangles for the center/range system. Finally, we also construct the Tukey nonparametric region for the data cloud of bootstrap realizations of center and range $(\hat{y}^{*(b)}_{c,T+h|T}, \exp(\hat{y}^{*(b)}_{r,T+h|T}))'$. Figure 1 (bottom panel) plots the transformed bootstrap forecast regions of the center/range system for the same simulated system considered above.

Third, recall that the direct transformation of the boundaries of the regions for the center/log-range system cannot be implemented when constructing forecast regions for the upper/lower bounds system. This is why, for this system, we construct the following bootstrap ellipse

$$BE^{UL}_{T+h} \quad = \quad \left[Y^{UL}_{T+h} | \;\; [Y^{UL}_{T+h} - \bar{Y}^{UL*}_{T+h|T}]' S^{UL}_{Y^*}(h)^{-1} [Y^{UL}_{T+h} - \bar{Y}^{UL*}_{T+h|T}] \leq q^{UL*}_{1-\alpha} \right] \tag{3.4}$$

where $Y^{UL}_{T+h} = (y_{u,T+h}, y_{l,T+h})'$ and $\bar{Y}^{UL*}_{T+h|T}$ and $S^{UL}_{Y^*}(h)$ represents the mean vector and variance covariance matrix, respectively, of the bootstrap upper/lower bound realizations given by $y^{*(b)}_{u,T+h} = \hat{y}^{*(b)}_{c,T+h|T} + \frac{1}{2}\exp(\hat{y}^{*(b)}_{r,T+h|T})$ and $y^{*(b)}_{l,T+h} = \hat{y}^{*(b)}_{c,T+h|T} - \frac{1}{2}\exp(\hat{y}^{*(b)}_{r,T+h|T})$, respectively.

Finally, a Tukey nonparametric region can be constructed for the data cloud of bootstrap realizations of upper and lower bounds $(y^{*(b)}_{u,T+h}, y^{*(b)}_{l,T+h}))'$. Note that for this system, we do not construct Bonferroni rectangles because they may contain unfeasible subregions of points where the lower bound is greater than the upper bound.

4 Out-of-sample evaluation of prediction regions

Suppose that we construct h-step ahead prediction regions with nominal coverage $100 \times (1-\alpha)\%$, from $t = 1, ..., N$ and want to evaluate them. As in the case of loss functions, it is only the objective of the forecaster that will define which criterium is the most appropriate. At the most basic level, the forecaster will aim for reliability, that is, those prediction regions that provide the closest coverage to the nominal. The **average coverage rate** is defined as follows

$$C_{(1-\alpha)} = \frac{1}{N} \sum_{t=1}^{N} I_t^{(1-\alpha)} \tag{4.1}$$

where $I_t^{(1-\alpha)}$ is an indicator variable that is equal to 1 if the observed outcome falls within the prediction region and 0 otherwise. Furthermore, following Golestaneh *et al.* (2017), we combine reliability with sharpness, a preference for regions with small volume. The forecaster would prefer a lower **average coverage-volume score** given by

$$CV_{(1-\alpha)} = \Big| \frac{1}{N} \sum_{t=1}^{N} \big[I_t^{(1-\alpha)} - (1-\alpha) \big] \times \big[V_t^{(1-\alpha)} \big]^{\frac{1}{2}} \Big| \tag{4.2}$$

where $V_t^{(1-\alpha)}$ is the volume of the prediction region at time t.

Another aspect to the evaluation of forecast regions is to consider the observations outside of the region and to assess how far they are from a central point within the prediction region. We propose the following **average outlier distance**

$$O_{(1-\alpha)} = \frac{1}{G} \sum_{t=1}^{N} \big[1 - I_t^{(1-\alpha)} \big] \times D(y_t, M_t) \tag{4.3}$$

where G is the number of observations outside the region, D is a distance measure (e.g. Euclidean distance) of each outside-the-region outcome, y_t, from M_t, which is the median of the realizations generated within the region at each time t. To obtain M_t, we implement the definition of median

in a multi-dimensional setting introduced by Zuo (2003), known as 'projection depth median'. The outside-the-region observations can be considered 'risk' that the forecaster has to bear and, in this sense, he would like to minimize $O_{(1-\alpha)}$. For two regions with similar coverage, the forecaster will choose that with a lower average outlier dispersion.

We also evaluate the prediction region by the sharpness or tightness of the intervals that result from projecting the two-dimensional region into one-dimensional intervals. We draw a large number of directions, which are given by the lines drawn from the zero origin of the unit circle to any point in its boundary. For each direction, we find the two bounding tangent lines to the prediction region that are perpendicular to that direction. We calculate the length of the projected interval bounded by the tangent lines; see Figure 2 (top panel) for a graphical representation. Denote $d_i \in \Upsilon$ as the i^{th} direction in Υ, where Υ is the set of all directions, and let D be the number of directions. The **average length of the projected intervals** associated with the prediction region is

$$P_{(1-\alpha)} = \frac{1}{N}\sum_{t=1}^{N} P_t \tag{4.4}$$

where P_t is the average projection length over all directions at time t given by $P_t = \frac{1}{D}\sum_{i=1}^{D}(u_{d_i} - l_{d_i})$ with u_{d_i} and l_{d_i} being the upper and lower bounds of the projected interval in the i^{th} direction. The forecaster would prefer prediction regions that deliver tight projected intervals. We also consider the realized data points over the prediction period in conjunction with the projected intervals. The **average distance of the projected outliers** associated with the prediction region is

$$OP_{(1-\alpha)} = \frac{1}{N}\sum_{t=1}^{N} OP_t \tag{4.5}$$

where OP_t is the average distance of the projected outliers to the projected interval given by $OP_t = \frac{1}{D}\sum_{i=1}^{D}[(l_{d_i} - x_{d_i})I(x_{d_i} < l_{d_i}) + (x_{d_i} - u_{d_i})I(x_{d_i} > u_{d_i})]$ with x_{d_i} being the coordinate of the data point projected on the i^{th} direction and $I(\bullet)$ being an indicator function. The forecaster prefers

prediction regions with projected outliers close to the projected intervals. Finally, if the length of the projected interval is large, we expect the distance of the projected outliers to the interval to be smaller. To take into account this trade-off, we propose a combined criteria $POP_t = P_t \times OP_t$ so that, over the prediction sample, the average trade-off associated with the prediction region is

$$POP_{(1-\alpha)} = \frac{1}{N}\sum_{t=1}^{N} POP_t \tag{4.6}$$

A smaller $POP_{(1-\alpha)}$ would be preferred by the forecaster.

Our final assessment of prediction regions is whether they are probability-centered. We check whether the points outside of the prediction region are evenly distributed around the region R. In this case, we expect the following statistic to be close to zero

$$S_{(1-\alpha)}(M_t) = \frac{1}{D}\sum_{i=1}^{D} |C_u(d_i, M_t)) - C_l(d_i, M_t)| \tag{4.7}$$

where D is the number of directions that pass through M_t, $C_u(d_i, M_t)) = \{\#x|(x \in R^c) \cap (x \in H_u(d_i, M_t))\}$ and $C_l(d_i, M_t) = \{\#x|(x \in R^c) \cap (x \in H_l(d_i, M_t))\}$ with $H_u(d_i, M_t)$ being the half-plane above the line generated by the direction d_i, $H_l(d_i, M_t)$ being the half-plane below the same line and R^c the complementary region to R; see Figure 2 (bottom panel) for a graphical representation. Note that $S_{(1-\alpha)}(M_t)$ will not be feasible with real data (we will have only one realized observation at time t that could be in or out of the prediction region). However, in a simulated environment, we will be able to assess whether each prediction region is probability-centered.

5 Monte Carlo Simulations

In this section, we perform Monte Carlo simulations to evaluate the prediction regions for interval valued time series according to the criteria described in Section 4. We report just one case for

small samples here and extensive simulation cases in the supplementary material.

We generate $R = 500$ replicates of the center/log-range system from a VAR(4) with $T = 50$ observations and its parameters chosen to replicate the dynamics of S&P500 interval returns. The center and log-range errors are distributed according to a Skewed-Student-t with 5 degrees of freedom and asymmetry parameter -0.5 and a normal distribution, respectively.[8] Note that the distributional assumptions are on the marginal densities of the errors of each equation and we do not know the exact bivariate densities. For the system to have the desired marginal density functions and the stated correlation structure, we have generated bivariate errors from a Gaussian copula and re-transform the PITs of the corresponding univariate normal variates according to the Student-t, to obtain the new error variates, which need to be adjusted to have the desired mean and variance. We construct one-step-ahead prediction regions with 95% nominal coverage and simulate 1000 future values of the required vector at time $T + 1$, i.e. center/log-range, center/range, and upper/lower bounds. The empirical coverage is calculated as the proportion of these values that falls within the constructed prediction regions. The number of bootstrap samples is $B = 2000$, and the number of directions to calculate the average length of the projected intervals and outliers is $D = 100$.[9]

Table 1 reports the evaluation criteria for the Monte Carlo experiments when the prediction regions are constructed by all procedures described in Section 2 for the three systems (center/log-range, center/range and upper/lower bounds). First, note that regardless of the particular system being forecast, the regions based on the normality assumption for the center/log-range system have coverages below the 95% nominal and their coverage-volume scores, CV, are nearly double than those of the regions based on bootstrap procedures. Therefore, it seems that in cases of small samples and/or non-Gaussian interval data, bootstrap procedures should be used instead of relaying on normality. Comparing the alternative bootstrap regions considered, we can observe that bootstrap

[8]The Skewed-Student distribution is defined as proposed by Hansen (1994).

[9]The parameters of the VAR(4) model are reported in the Supplementary material available online. Results for other VAR models, error distributions and sample sizes are also reported in the Supplementary material.

ellipsoid have minimum coverage-volume scores, although the differences among them are very small. However, the average outlier distance, O, is lower for the Tukey regions than for the other bootstrap regions. When constructing the regions for the upper/lower bounds, the Tukey convex hull has also smaller trade-off between sharpness and risk, POP and centeredness. Therefore, it seems that in the presence of asymmetric distributions as that considered in this paper, the Tukey convex hull provide adequate regions for interval data.

6 Prediction regions for S&P500 low/high returns

In this section, we construct joint prediction regions for the center (level) and log-range (volatility) of daily S&P500 return intervals that take into account their interaction without making particular assumptions on their joint distribution.

6.1 Joint modeling of returns and volatilities

Several authors propose modeling the interaction between volatility and asset returns fitting bivariate models to returns and realized variances. Among them, Takahasi *et al.* (2009) propose an Stochastic Volatility model in which both returns and realized variances depend on a latent volatility. The presence of this latent volatility force the authors to estimate the models and extract the volatility using computationally intensive Bayesian procedures. Shephard and Sheppard (2010) also propose modeling returns and realized volatilities by fitting a VAR model called High-frequency-based-volatility (HEAVY) and obtain the entire predictive distribution of returns based on drawing with replacement pairs from the joint distribution of standardized returns and realized volatilities. More recently, Catania and Proietti (2019) propose bivariate score driven models for returns and realized volatilities.

The methodology proposed in this paper has several advantages with respect to bivariate models for returns and realized volatilities. First, we avoid issues related to microstructure noise associated

with the construction of realized volatilities. As mentioned in the Introduction, several authors advocate using the range as a measure of volatility. Second, using bootstrap prediction regions, we avoid potential unrealistic assumptions about the joint distribution of returns and volatilities; see, for example, Catania and Proietti (2019) who assume that returns and log-volatilities have a joint Student-t distribution with ν degrees of freedom implying that the marginal univariate distributions of returns and volatilities are both Student-t with ν degrees of freedom. Third, the proposed methodology is a natural framework to incorporate other stylized facts as, for example, the effect of past volatilities on returns. Last but not least, the proposed methodology is computationally less intensive than alternative approaches and thus, it allows for a straightforward implementation and can be easily extended to other models for the center/log-range system as VARMA or cointegrated systems.

6.2 Model estimation

We collect intervals of low/high S&P500 prices, (P_t^L, P_t^U) observed daily from January 2, 2009 to April 20, 2018 for a total of 2341 observations. Since prices are non-stationary, we construct daily stationary intervals of low/high returns, (y_t^L, y_t^U) by calculating the daily minimum and maximum return with respect to the closing price of the previous day, i.e. $y_t^L = P_t^L - P_{t-1}^C$ and $y_t^U = P_t^U - P_{t-1}^C$, where P_t^C is the closing price at day t. Figure 3 (top panel), which plots a kernel estimate of the bivariate unconditional density of the center and the range, shows that the center exhibits fat tails and it is slightly skewed to the left. Furthermore, according to the sample descriptive statistics reported in the Supplementary material, we can observe that the log-range is only slightly skewed to the right and has a coefficient of kurtosis of about 3. The correlation between the center and log-range is about -0.10. With respect to temporal dependence, as expected, the Q-statistics for the center indicate no autocorrelation while those for the range and log-range indicate high autocorrelation mimicking the autocorrelations often observed in the end-of-the day returns and in their squared returns, respectively.

We split the total sample into an estimation sample from January 2, 2009 to December 31, 2016 (2014 observations) and a prediction/evaluation sample from January 1, 2017 to April 20, 2018 (327 observations). Using the estimation sample, we obtain the following LS estimates of the (significant) parameters of a VAR(6) model as selected by the SIC criteria (robust standard errors in parenthesis)[10]

$$\hat{C}_t = 0$$

$$\log\hat{R}_t = \underset{(0.013)}{-0.17} C_{t-1} - \underset{(0.014)}{0.09}\, C_{t-2} - \underset{(0.014)}{0.06}\, C_{t-3} - \underset{(0.014)}{0.04}\, C_{t-4} +$$

$$+ \underset{(0.023)}{0.17}\, logR_{t-1} + \underset{(0.023)}{0.22}\, logR_{t-2} + \underset{(0.023)}{0.16}\, logR_{t-3} + \underset{(0.023)}{0.09}\, logR_{t-4} + \underset{(0.023)}{0.1}\, logR_{t-5} + \underset{(0.022)}{0.11}\, logR_{t-6}$$

As expected, all regressors (lagged center and lagged log-range) in the equation for the center are not statistically significant. On the contrary, the equation for the log-range presents interesting dynamics. The center Granger-causes the log-range so that the lagged centers are negatively correlated with the current log-range, i.e. positive and large changes in the center return today predict a narrower range tomorrow. This is similar to the leverage effect in a conditional variance equation. Another relevant aspect, in agreement with the ACF/PACF profiles, is the strong and statistically significant autoregressive nature of the log-range. The estimated VAR model captures the main stylised facts often observed in financial returns: i) persistence of volatility; ii) heavy tails of returns; and iii) negative dependence between volatility and past returns. The goodness of fit for the log-range equation is high with an adjusted $R^2 = 0.52$. In the Supplementary Material, we report the results of the residual diagnosis. First, we observe that the residuals are all clear of any autocorrelation. Furthermore, the center residuals and log-range residuals are contemporaneous negatively correlated with a correlation coefficient of -0.17. Finally, we formally test for conditional bivariate normality by implementing the Generalized AutoContouR (G-ACR) (in-sample) tests based on the Probability Integral Transformations (PIT) of the joint density under the null hypothesis of bivariate normality (González-Rivera and Sun, 2015). We also report the results of the t-statistics ($t_{k,\alpha}$) that canvas the density from the 1% to the 99% PIT autocontours

[10] Only significant values are reported.

for lags $k = 1, 2, ...5$. The null hypothesis is strongly rejected at the 5% significance level for mostly all but the 10%, 90% and 95% autocontours. The portmanteau test C_k also reinforces the strong rejection of bivariate normality. Figure 3 (bottom panel) plots the autocontours of the contemporaneous PITs $(\text{center}_t, \text{log-range}_t)$. Under the correct null hypothesis, the distribution of the PITs should be uniformly distributed within these autocontour squares. It is obvious that this is not the case.

6.3 Out-of-sample forecast regions

We evaluate the out-of-sample performance of the one-step-ahead 95% prediction regions from January 1, 2017 to April 20, 2018 (327 observations). The results are reported in Table 2. For the system center/log-range, the bootstrap MBR regions offer the best coverage C with empirical rates of mostly 95% and they are the most reliable with the lowest average coverage-volume scores CV. They also provide the tightest projected one-dimensional regions measured by POP. However, the Tukey convex hull regions provide the lowest average outlier distance O. For the system center/range, we find that the transformed modified bootstrap Bonferroni rectangle is the best performer according to most metrics C, CV and POP. For the system upper/lower bounds, the Tukey convex hull offers the best coverage and the lowest scores for O and POP. As expected, the analytic methods are not reliable as they tend to undercover. On the contrary, the bootstrap ellipse and its transformed region tend to overcover. All these results are very consistent with the Monte Carlo findings of the previous section.

In Figures 4 and 5, we plot the one-step-ahead 95% prediction regions for the center/log-range and center/range systems respectively. We choose six random dates over the prediction sample (March 15, May 11, August 30, December 8, 2017 and February 22, April 6, 2018). In all six dates, the one-step-ahead realized values of the (center, log-range) and (center, range) fall within the regions; only the realized values on December 8, 2017 and April 6, 2018 are slightly more extreme and they

fall towards the boundaries of the prediction regions. For the center/log-range system, the normal ellipse and the bootstrap ellipse are very similar but in the center/range system, the bootstrap ellipse tends to be wider adapting to the kurtosis of the center and the asymmetry of the range. The differences among the Bonferroni rectangles are more obvious in the center/range system. In the center/log-range system, the Tukey convex hull has a cone shape over all the six dates though the shape becomes more irregular in the center/range system.

It is important to point out that when using point data, we can obtain a single forecast density for future returns. However, by using interval data, we are able to provide a wider characterization of returns by constructing the conditional forecast density of returns conditioning on volatility or the joint bivariate density of returns and volatilities.

6.4 Trading strategy

Using the bootstrap prediction regions for the daily S&P500 high/low returns obtained in the previous section, we develop a trading strategy for interval data that extends that proposed by He *et al.* (2010) for point forecasts. The proposed trading strategy exploits the probability distribution of forecasts. Consider the following ratio $s_t = \frac{\left|O_t - \hat{y}_{l,t+h}\right|}{\left|\hat{y}_{u,t+h} - O_t\right|}$, where O_t is the opening return at day t, calculated using the opening price at day t with respect to the closing price at day $t-1$ and $\hat{y}_{l,t+h}$ and $\hat{y}_{u,t+h}$ are the low and high return forecasts, respectively. If $s_t < 1$, then the return is more likely to go up than down in the next h days. If this is observed for several days, it is reasonable to believe that the market is forming an upward trend and a "buy alert signal" should be generated. A similar argument can be applied to the "sell alert signal". While He *et al.* (2010) compute s_t using point forecasts of $\hat{y}_{u,t+h}$ and $\hat{y}_{l,t+h}$, we compare the probability of $s_t > 1$ (sell signal) and $s_t < 1$ (buy signal). Figure 6 illustrates the proposed trading strategy. Notice that s_t is the absolute value of the slope of any line that connects point $A \equiv (O_t, O_t)$ and any other point below the 45 degree line. The ellipse represents the h-step-ahead prediction region of the high/low

returns. The slope of line AB is equal to (minus) one and it is perpendicular to the 45 degree line. Hence, the area under the 45 degree line is divided into two areas by the line AB into two areas: $s_t > 1$ to the left of line AB, and $s_t < 1$ to the right of line AB. Therefore, counting the bootstrap realizations in the two subareas of the prediction region, we can compare the estimated probability of $s_t < 1$ with that of $s_t > 1$ for a given $100 \times (1-\alpha)\%$ confidence region. Then, the trading strategy consists of the following steps:

- At day t, plot Figure 6 based on O_t and the h-step-ahead prediction region of high and low returns. Within the prediction region, if the number of bootstrap realizations (obtained as in equation (3.4)) on the right hand side of the line AB is larger than that on the left hand side of the line AB, a "buy alert signal" is generated.

- If $Prob(s_t < 1) > Prob(s_t > 1)$ is observed for m consecutive days beginning with day t, buy the asset on day $t + m - 1$ using the closing price on that day.

- After buying the asset, on any other day d, watch for the"sell alert signal", that is, the number of bootstrap realizations on the left hand side of the line AB should be larger than that on the right hand side of the line AB within the prediction region. If the "sell alert signal" is observed for m consecutive days from day d, sell the asset on day $d + m - 1$ using the closing price on that day. Otherwise, hold the asset.

We evaluate this trading strategy over the out-of-sample period (Jan. 1, 2017 to Apr. 20, 2018) using the bootstrap ellipse and Tukey Convex Hull prediction regions with a 95% nominal coverage. For the implementation, the choice of m should not be too small because it will introduce substantial noise in trading but it should not be too large either because we could miss profitable trades. We consider $m = 1, 2, 3, 4$ and $h = 1, 2, 3$. We apply a transaction cost of 0.1%, and we annualize the profit/loss for each trade because each trade will have a different holding period. The annualized return is calculated as $AR_t = (\frac{y_{c,t+j} - y_{c,t}}{y_{c,t}} - 0.001)(\frac{365}{j})$ where $y_{c,t}$ and $y_{c,t+j}$, $(j > 0)$, are the closing prices for the buying and selling days, respectively. The investor can buy the asset

again before the previous bought asset is sold. At the end of the evaluation period, if there are still assets that have not been sold, these assets will not be considered when calculating the profits.

Table 3 reports the average and the max and min of AR_t together with the percentage of trades with positive returns for all cases. HKW is the trading strategy in He et. al. (2010) based on point forecasts of the high/low returns. For all cases but two, the averaged annualized returns are positive. The choice of m is very relevant because the gap between the max and min annualized returns narrows as m increases for all h. A large m means that the investor is looking for a stronger signal and, though she may miss some trades with extreme positive returns, she will also avoid those extreme negative returns that can be catastrophic. There is also a monotonic positive relation between m and the percentage of trades with positive returns. For average annualized returns, the performance of BE is better than that of TH in most cases, and the performance of BE or TH is better than HKW in particular when $m = 4$.

7 Conclusions

Often time series interval measurements offer a more complete description of a data set because each observation has joint information on the level and the dispersion of the process under study. However, statistical analysis of interval-valued data requires that the natural order of the interval is preserved. Though there are several works that consider the problem of estimation with constraints, we are not aware of any work that considers the construction of probabilistic forecasts for interval-valued data satisfying the natural constraint in each period of time. Our contribution lies on approximating a probabilistic forecast of an interval-valued time series by offering alternative approaches to construct bivariate prediction regions of the center and the range, or the lower and upper bounds, of the interval.

To overcome the positive constraint of the range, we estimate a Gaussian bivariate system for the center/log-range system, which delivers QML properties for our estimators. However, the

interest of the researcher is not the prediction of the center/log-range but the center/range or upper/lower bounds of the interval. By implementing either analytical or bootstrap methods we directly transform the prediction regions for the center/log-range system into those for the center/range and upper/lower bounds systems. It is important to remark that we do not focus on point forecast purposely. By focusing on prediction regions rather than on point forecasts, we avoid the biases associated with the exp-transformation of the point forecasts of log-transformed variable. A prediction region for the center/log-range does not need any bias correction when transformed into a prediction region of the center/range system because the quantile is preserved under a monotonic transformation like the exp-transformation. These transformed prediction regions can have very irregular shapes even in the most straightforward scenario of bivariate normality of the center/log-range system. If a central point forecast is of interest, the researcher can always calculate the centroid of the region.

Beyond the standard coverage rate, we propose several new metrics to evaluate the performance of prediction regions. We introduce a notion of risk to the evaluation of the regions by considering the location of the out-of-the-region outcomes with respect to some central point in the region. The researcher would like to minimize risk once the empirical coverage of the region is close to the nominal coverage. We have considered Gaussian and non-Gaussian systems and our recommendation leans towards bootstrap methods, even for Gaussian systems. Bootstrap ellipses and their transformed are best when the joint distribution of the center/log-range system is symmetric. If it is not, then bootstrap Tukey hull regions will be preferred. In summary, we find that simulation-based methods are the most reliable.

We analyze a time series of the daily low/high return intervals of the S&P500 index. We model and predict the joint conditional density of the return level and volatility. We show that the bootstrap regions have best properties. Furthermore, we carry out a trading strategy to illustrate the economic advantages of taking into account the probabilistic forecast of the low/high returns

instead of point forecasts.

References

[1] Alizadeh, S., M. W. Brandt, & F.X. Diebold (2002). Range-based Estimation of Stochastic Volatility Models. *Journal of Finance*, 57(3), pp. 1047-1091.

[2] Beran, R. (1993). Probability-Centered Prediction Regions. *Annals of Statistics*, 21(4), pp. 1967-1981.

[3] Bien, K., I. Nolte & W. Pohlmeier (2011). An inflated multivariate integer count hurdle model: An application to bid and ask quote dynamics. *Journal of Applied Econometrics*, 26, pp. 669-707.

[4] Blasques, F., S.J. Koopman, K. Lasak & A. Lucas (2016). In-sample confidence bands and out-of-sample forecast bands for time-varying parameters in observation-driven models. *International Journal of Forecasting*, 32, pp. 875-887.

[5] Brandt, M.W. & F.X. Diebold (2006). A no-arbitrage approach to range-based estimation of return covariance and correlations. *Journal of Business*, 79, pp. 61-74.

[6] Catania, L., & T. Proietti (2019). Forecasting volatility with time-varying leverage and volatility effects. *CEIS Tor Vergata Research Paper Series*, Vol. 17, issue 1, no.450.

[7] Clements, M.P. & J. Smith (2002). Evaluating multivariate forecast densities: a comparison of two approaches. *International Journal of Forecasting*, 18(3), pp. 397-407.

[8] Cheung, Y.L., Y.W. Cheung, & A.T. Wan (2009). A high-low model of daily stock price ranges. *Journal of Forecasting*, 28(2), pp. 103-119.

[9] Diebold, F.X., A. Han & K. Tay (1999). Multivariate density forecast evaluation and calibration in financial risk management: High-frequency returns on foreign exchange. *Review of Economics and Statistics*, 81(4), pp. 661-673.

[10] Fernández, C., León, C.J., Steel, M.F.J., & Vázquez-Polo, F.J. (2004). Bayesian analysis of interval data contingent valuation models and pricing policies. *Journal of Business & Economic Statistics*, 22(4), pp. 431-442.

[11] Fresoli, D., E. Ruiz, & L. Pascual (2015). Bootstrap Multi-step Forecasts for Non-Gaussian VAR Models. *International Journal of Forecasting*, 31, pp. 834-848.

[12] García-Ascanio, C. & C. Maté (2010). Electric power demand forecasting using interval time series: A comparison between VAR and iMLP. *Energy Policy*, 38(2), pp. 715-725.

[13] Gneiting, T., L.I. Stanberry, E.P. Grimit, L. Held & N.A. Johnson (2008). Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble prediction of surface winds. *TEST*, 17(2), pp. 211-235.

[14] Golestaneh, F., P. Pinson, R. Azizipanah-Abarghooee, & H.B. Gooi (2018). Ellipsoidal Prediction Regions for Multivariate Uncertainty Characterization. *IEEE Transactions on Power Systems*, 33(4), pp. 4519-4530.

[15] González-Rivera, G., & W. Lin (2013). Constrained Regression for Interval-valued Data. *Journal of Business and Economic Statistics*, 31(4), pp. 473–490.

[16] González-Rivera, G., & Y. Sun (2015). Generalized Autocontours: Evaluation of Multivariate Density Models. *International Journal of Forecasting*, 31(3), pp. 799-814.

[17] González-Rivera, G., & E. Yoldas (2011). Autocontour-based evaluation of multivariate predictive densities. *International Journal of Forecasting*, 28(2), 328-342.

[18] Granger, C.W.J. & P. Newbold (1976). Forecasting Transformed Series. *Journal of the Royal Statistical Society, B*, 38, pp. 189-203.

[19] Guerrero, V.M. (1993). Time Series Analysis Supported by Power Transformation. *Journal of Forecasting*, 12, pp. 37-48.

[20] Han, A., Y. Hong & S. Wang (2016). Autoregressive conditional models for interval-valued time series data. In R. Hill, G. González-Rivera & T. Lee (eds.), *Advances in Econometrics*, vol. 36, Emerald Group Publishing Limited, pp. 417-460.

[21] Hansen, B.E. (1994). Autoregressive conditional density estimation. *International Economic Review*, 35(3), pp. 705-730.

[22] He, A.W.W., J.T.K. Kwok & A.T.K. Wan (2010). An empirical model of daily highs and lows of West Texas intermediate crude oil prices.*Energy Economics*, 32, pp. 1499-1506.

[23] Ji, S. & X. Shi (2018). Reaching goals under umbiguity: Continuous-time optimal portfolio selection.*Statistics and Probability Letters*, 137, pp. 63-69.

[24] Komunjer, I. & M. Owyang (2012). Multivariate forecast evaluation and rationality testing.*Review of Economics and Statistics*, 94(4), pp. 1066-1080.

[25] Lima Neto, E. & F. de Carvalho (2010). Constrained Linear Regression Models for Symbolic Interval-Valued Variables. *Computational Statistics and Data Analysis*, 54, pp. 333-347.

[26] Lin, W. & G. González-Rivera (2016). Interval-valued time series models: estimation based on order statistics exploring the Agriculture Marketing Service data. *Computational Statistics & Data Analysis*, 100, pp. 694-711.

[27] Lutkephol, H. (1991). *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

[28] Manski, C.F, & E. Tammer (2002). Interence on regressions with interval data on a regressor or outcome. *Econometrica*, 70(2), pp. 519-546.

[29] Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business*, 53(1), pp. 61-65.

[30] Pascual, L., J. Romo & E. Ruiz (2005). Bootstrap Prediction Intervals for Power Transformed Time Series. *International Journal of Forecasting*, 21, pp. 219-235.

[31] Pascual, R. & D. Veredas (2010). Does the open limit order book matter in explining informational volatility?. *Journal of Financial Econometrics*, 8(1), pp. 57-87.

[32] Rodrigues, P. & N. Salish (2015). Modelling and forecasting interval time series with threshold models. *Advances in Data Analysis and Classification*, 9, pp. 41-57.

[33] Rogers, L.C.G. & S.E. Satchell (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability*, 1(4), pp. 504-512.

[34] Rogers, L.C.G., S.E. Satchell & Y. Yoon (1994). Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Applied Financial Economics*, 4, pp. 241-247.

[35] Shephard, N. & K. Sheppard (2010). Realising the future: Forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 25, pp. 197-231.

[36] Takahashi, M., Y. Omori and T. Watanabe (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics and Data Analysis*, 53, pp. 2404-2426.

[37] Tu, Y. & Y. Wang (2016). Center and Log Range Models for Interval-valued Data with An Application to Forecast Stock Returns. Working paper.

[38] Tukey, J. (1975). Mathematics and Picturing Data, in *Proceedings of the 1975 International Congress of Mathematics*, 2, pp. 523-531.

[39] Vorbrink, J. (2014). Financial markets with volatility uncertainty. *Journal of Mathematical Economics*, 53, pp. 64-78.

[40] Xiong, T., Y. Bao, Z. Hu & L. Zhang (2015). A combination method for interval forecasting of agricultural commodity futures prices. *Knowledge-Based Systems*, 77, pp. 92-102.

[41] Xiong, T., C. Li & Y. Bao (2017). Interval-valued time series forecasting using a novel hybrid Holt and MSVR model. *Economic Modelling*, 60, pp. 11-23.

[42] Yao, W. & Z. Zhao (2013). Kernel Density-based Linear Regression Estimate. *Communications in Statistics. Theory and Methods*, 42(24), pp. 4499-4512.

[43] Yeh, A.B:, & Singh, K. (1997). Balanced confidence regions based on Tukey's depth and the bootstrap. *Journal of the Royal Statistical Society, Series B (Methodological)*, 59(3), 639-652.

[44] Zuo, Y. (2003). Projection-based Depth Functions and Associated Medians. *Annals of Statistics*, 31(5), pp. 1460-1490.

	EVALUATION CRITERIA							
CENTER/log-RANGE	C	$V^{1/2}$	CV	O	P	OP	POP	S
Normal ellipse	0.851	1.999	0.190	1.529	2.219	0.018	0.037	0.081
Bonferroni rectangle	0.861	2.098	0.177	1.541	2.641	0.010	0.026	0.075
Modified Bonferroni rectangle	0.862	2.098	0.176	1.539	2.674	0.010	0.025	0.075
Bootstrap ellipse	0.935	2.543	0.083	1.986	2.824	0.008	0.021	0.036
Bootstrap Bonferroni rectangle	0.937	2.705	0.086	1.844	3.420	0.004	0.012	0.037
Modified Bootstrap Bonferroni rectangle	0.937	2.705	0.086	1.839	3.466	0.004	0.011	0.038
Tukey convex hull	0.914	2.412	0.105	1.732	2.834	0.008	0.020	0.050

CENTER/RANGE system	C	$V^{1/2}$	CV	O	P	OP	POP	S
Analytical method	0.851	2.081	0.194	1.656	2.446	0.022	0.049	0.084
T-Normal ellipse	0.851	2.136	0.202	1.635	2.518	0.020	0.047	0.081
T-Bonferroni rectangle	0.861	2.256	0.190	1.649	2.899	0.013	0.034	0.075
T-Modified Bonferroni rectangle	0.862	2.280	0.191	1.643	3.049	0.012	0.032	0.074
T-Bootstrap ellipse	0.935	2.806	0.091	2.122	3.385	0.009	0.026	0.036
T-Bootstrap Bonferroni rectangle	0.937	3.010	0.096	1.981	3.885	0.005	0.016	0.037
T-Modified Bootstrap Bonferroni rectangle	0.936	3.153	0.104	1.959	4.363	0.004	0.015	0.038
Tukey convex hull	0.914	2.711	0.117	1.851	3.385	0.009	0.026	0.050

UPPER/LOWER system	C	$V^{1/2}$	CV	O	P	OP	POP	S
Analytical method	0.851	2.081	0.194	1.929	3.033	0.031	0.087	0.088
Bootstrap ellipse	0.936	2.947	0.099	2.675	4.101	0.015	0.055	0.054
Tukey convex hull	0.914	2.712	0.117	2.196	4.052	0.013	0.044	0.048

Table 1: Monte Carlo evaluation of one-step-ahead 95% prediction regions for a VAR(4) model with Skewed-Student-5 center and normal log-range.

	EVALUATION CRITERIA						
CENTER/log-RANGE system	C	$V^{1/2}$	CV	O	P	OP	POP
Normal ellipse	0.954	2.22	0.009	1.61	2.43	0.004	0.009
Bonferroni rectangle	0.945	2.31	0.011	1.47	2.90	0.002	0.006
Modified Bonferroni rectangle	0.948	2.31	0.004	1.45	2.91	0.002	0.005
Bootstrap ellipse	0.960	2.36	0.025	1.69	2.59	0.003	0.007
Bootstrap Bonferroni rectangle	0.948	2.47	0.004	1.38	3.10	0.001	0.003
Modified Bootstrap Bonferroni	0.951	2.47	0.003	1.36	3.12	0.001	0.002
Tukey convex hull	0.945	2.14	0.010	1.33	2.50	0.002	0.005

CENTER/RANGE system	C	$V^{1/2}$	CV	O	P	OP	POP
Analytical method	0.936	1.81	0.032	1.52	2.21	0.009	0.019
T-Normal ellipse	0.954	1.89	0.002	1.71	2.26	0.006	0.013
T-Bonferroni rectangle	0.945	1.98	0.018	1.45	2.58	0.004	0.010
T-Modified Bonferroni rectangle	0.948	1.99	0.013	1.44	2.63	0.003	0.008
T-Bootstrap ellipse	0.960	2.01	0.008	1.85	2.42	0.004	0.010
T-Bootstrap Bonferroni rectangle	0.948	2.19	0.013	1.31	2.87	0.003	0.007
T-Modified Bootstrap Bonferroni	0.951	2.21	0.002	1.24	2.95	0.001	0.004
Tukey convex hull	0.945	2.05	0.025	1.24	2.49	0.004	0.009

UPPER/LOWER system	C	$V^{1/2}$	CV	O	P	OP	POP
Analytical method	0.936	1.81	0.032	1.68	3.17	0.008	0.027
Bootstrap ellipse	0.960	2.21	0.021	1.67	3.58	0.005	0.017
Tukey convex hull	0.945	2.05	0.025	1.31	3.61	0.003	0.010

Table 2: Evaluation of one-step-ahead 95% prediction regions for S&P500 returns forecasted from January 1, 2017 to April 20, 2018.

	m=1			m=2			m=3			m=4		
	HKW	BE	TH	HKW	BE	TH	HKW	BE	TH	HKW	BE	TH
h=1												
Averaged annualized returns	26.16%	33.12%	33.14%	-5.71%	10.89%	5.94%	43.49%	42.52%	38.04%	38.76%	40.25%	40.25%
max annualized returns	853.58%	853.58%	853.58%	239.83%	239.83%	239.83%	155.24%	155.24%	155.24%	85.18%	85.18%	85.18%
min annualized returns	-3760.12%	-3760.12%	-3760.12%	-1621.35%	-1621.35%	-1621.35%	-116.43%	-116.43%	-116.43%	4.64%	4.64%	4.64%
% of trades with positive returns	61.04%	61.54%	61.54%	78.26%	81.48%	80.77%	75.00%	75.00%	75.00%	100%	100%	100%
h=2												
Averaged annualized returns	29.86%	30.39%	29.15%	2.02%	-1.53%	0.79%	45.44%	48.10%	44.43%	38.76%	38.76%	38.76%
max annualized returns	853.58%	853.58%	853.58%	239.83%	239.83%	239.83%	155.24%	155.24%	129.53%	85.18%	85.18%	85.18%
min annualized returns	-3760.12%	-3760.12%	-3760.12%	-1621.35%	-1621.35%	-1621.35%	-116.43%	-116.43%	-116.43%	4.64%	4.64%	4.64%
% of trades with positive returns	61.54%	62.82%	61.25%	76.00%	76.00%	76.00%	75.00%	71.43%	71.43%	100%	100%	100%
h=3												
Averaged annualized returns	29.92%	11.48%	3.37%	4.55%	9.53%	3.22%	42.37%	20.67%	17.26%	43.01%	63.75%	61.42%
max annualized returns	853.58%	853.58%	853.58%	239.83%	239.83%	239.83%	175.56%	175.56%	175.56%	85.18%	102.78%	102.78%
min annualized returns	-3760.12%	-3760.12%	-3760.12%	-1621.35%	-1621.35%	-1621.35%	-116.43%	-116.43%	-116.43%	20.14%	26.47%	26.47%
% of trades with positive returns	62.67%	58.97%	57.69%	80.77%	78.57%	77.78%	88.89%	71.43%	57.14%	100%	100%	100%

Table 3: Trading strategy comparison for S&P500 average annualized returns over the out-of-sample period from January 1, 2017 to April 20, 2018.

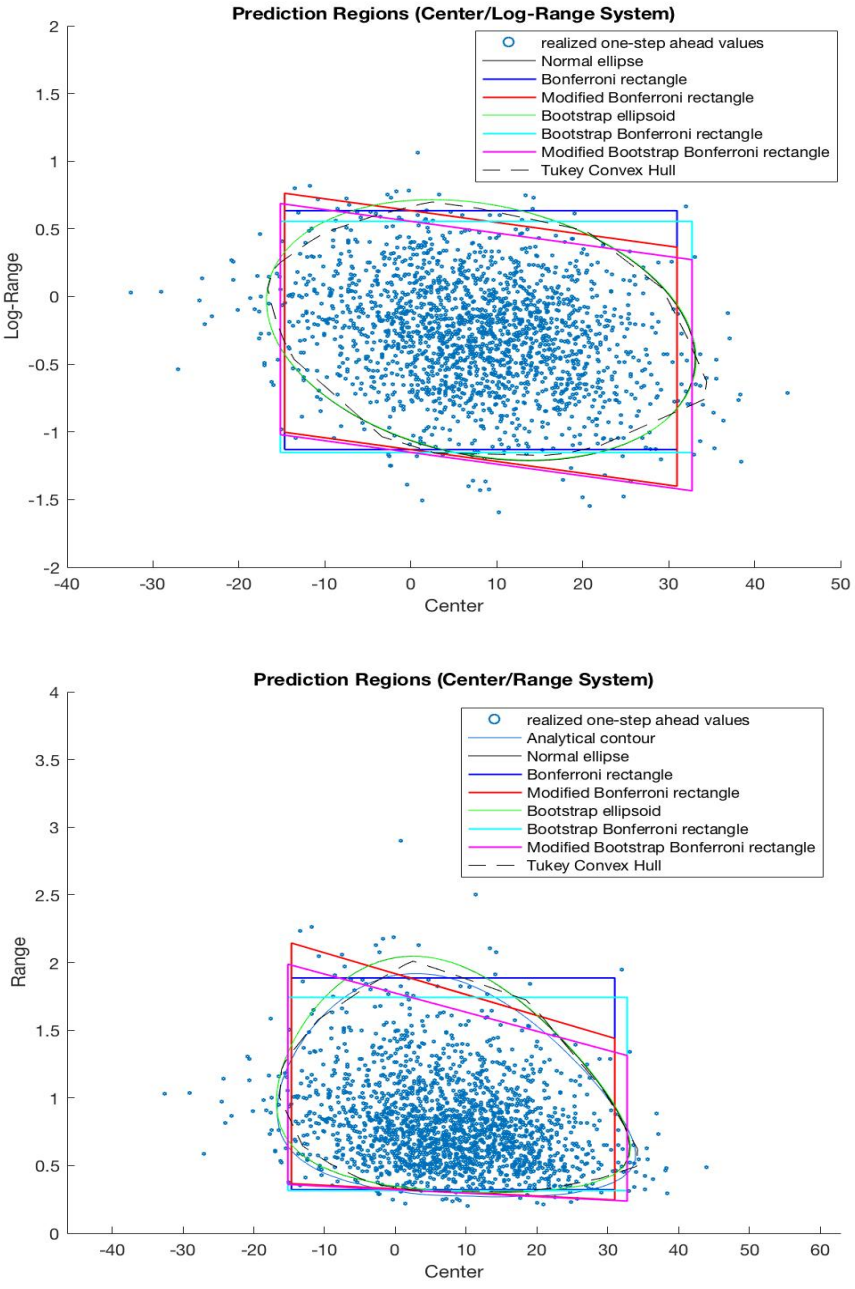

Figure 1: 95% prediction regions for the center/log-range system (top panel) and center/range system (bottom panel) obtained from a simulated VAR(4) model with parameters: $a_1 = -0.93, b_{11}^{(1)} = 0.34, b_{11}^{(2)} = -0.15, b_{11}^{(3)} = 0.03, b_{11}^{(4)} = -0.06, b_{12}^{(1)} = -0.50, b_{12}^{(2)} = 0.13, b_{12}^{(3)} = -0.16, b_{12}^{(4)} = 0.92, a_2 = 0.08, b_{21}^{(1)} = -0.01, b_{21}^{(2)} = b_{21}^{(3)} = b_{21}^{(4)} = 0, b_{22}^{(1)} = 0.09, b_{22}^{(2)} = 0.18, b_{22}^{(3)} = 0.15, b_{22}^{(4)} = 0.08$ and Gaussian errors with variances given by $\sigma_1^2 = 111.24$ and $\sigma_2^2 = 0.16$ and covariance $\sigma_{12} = -1.02$. The sample size is $T = 1000$.

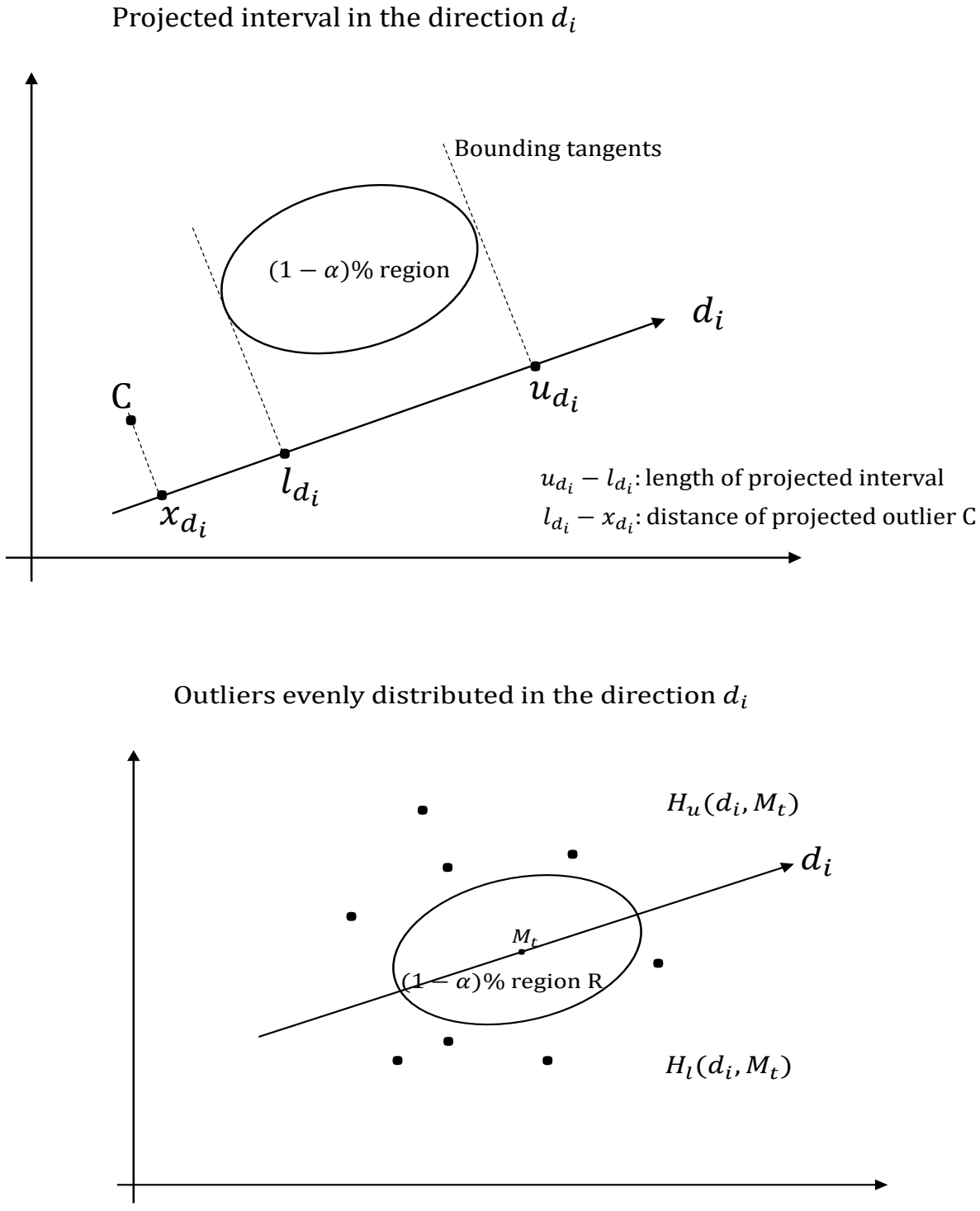

Figure 2: Projected interval and projected outliers (top panel). Outlier distribution around a region (bottom panel)

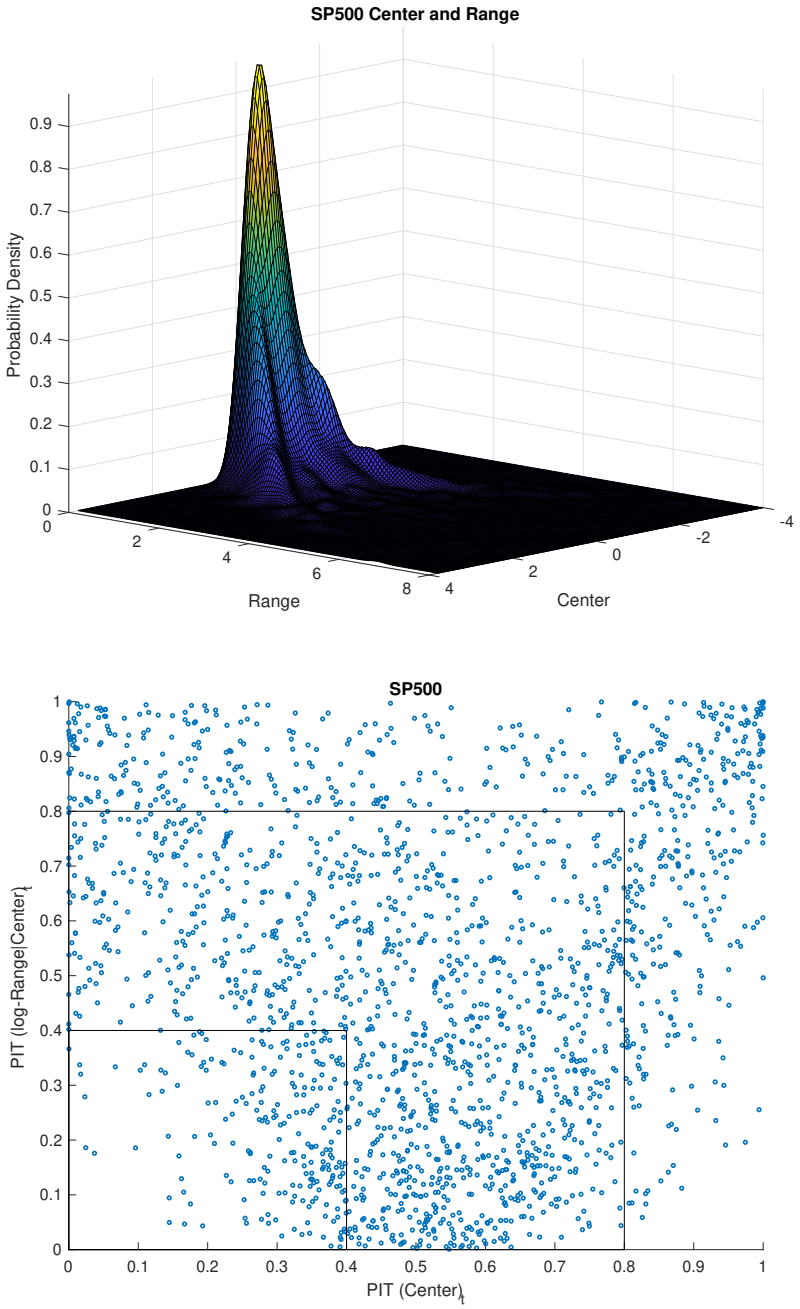

Figure 3: Unconditional bivariate density (top panel) of S&P500 low/high return intervals and G-ACR specification tests for conditional bivariate normality of residuals of the center and log-range VAR(6) model (bottom panel).

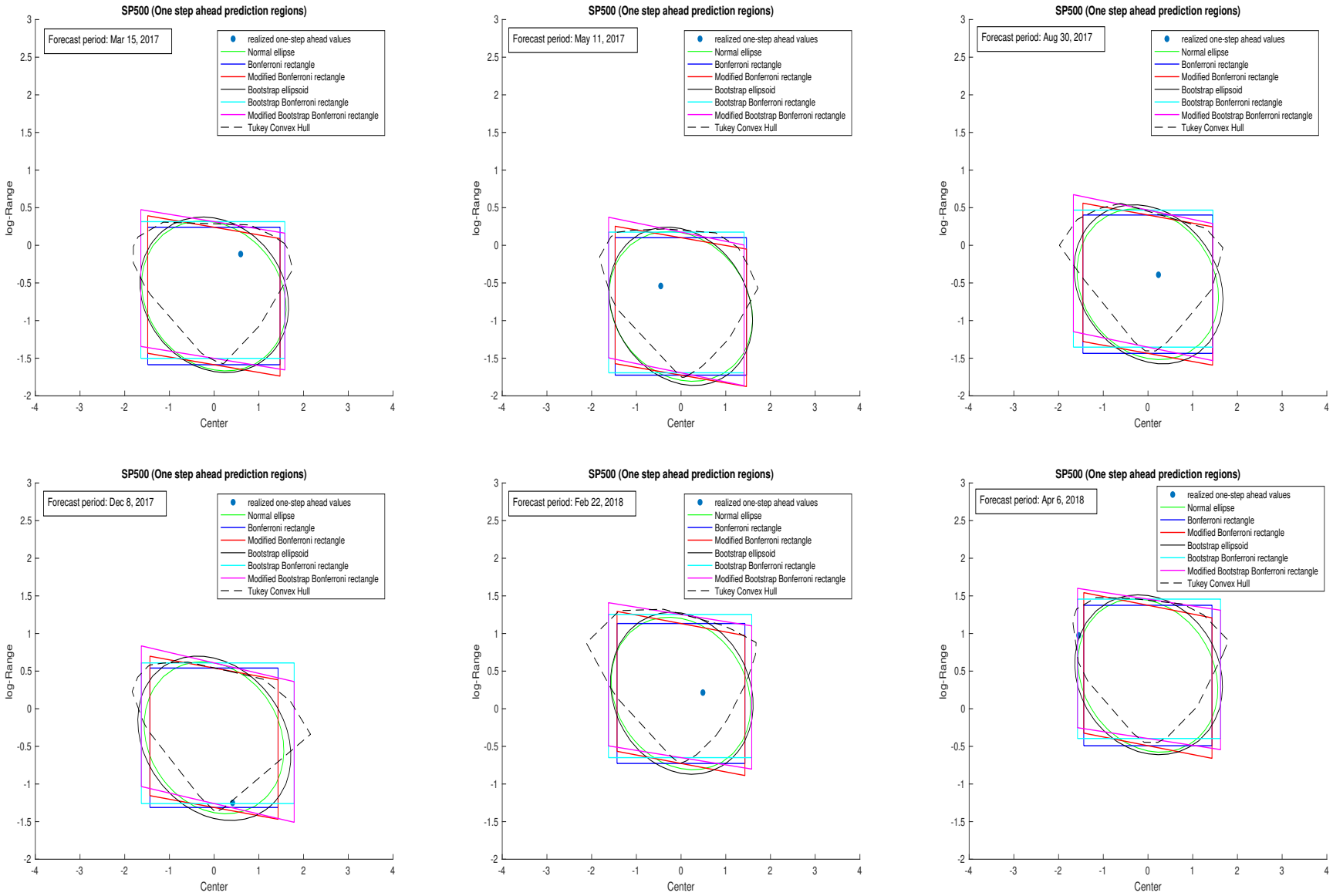

Figure 4: One-step-ahead 95% prediction regions for the center/log-range system of the SP500 return intervals corresponding to different dates of the out-of-sample period.

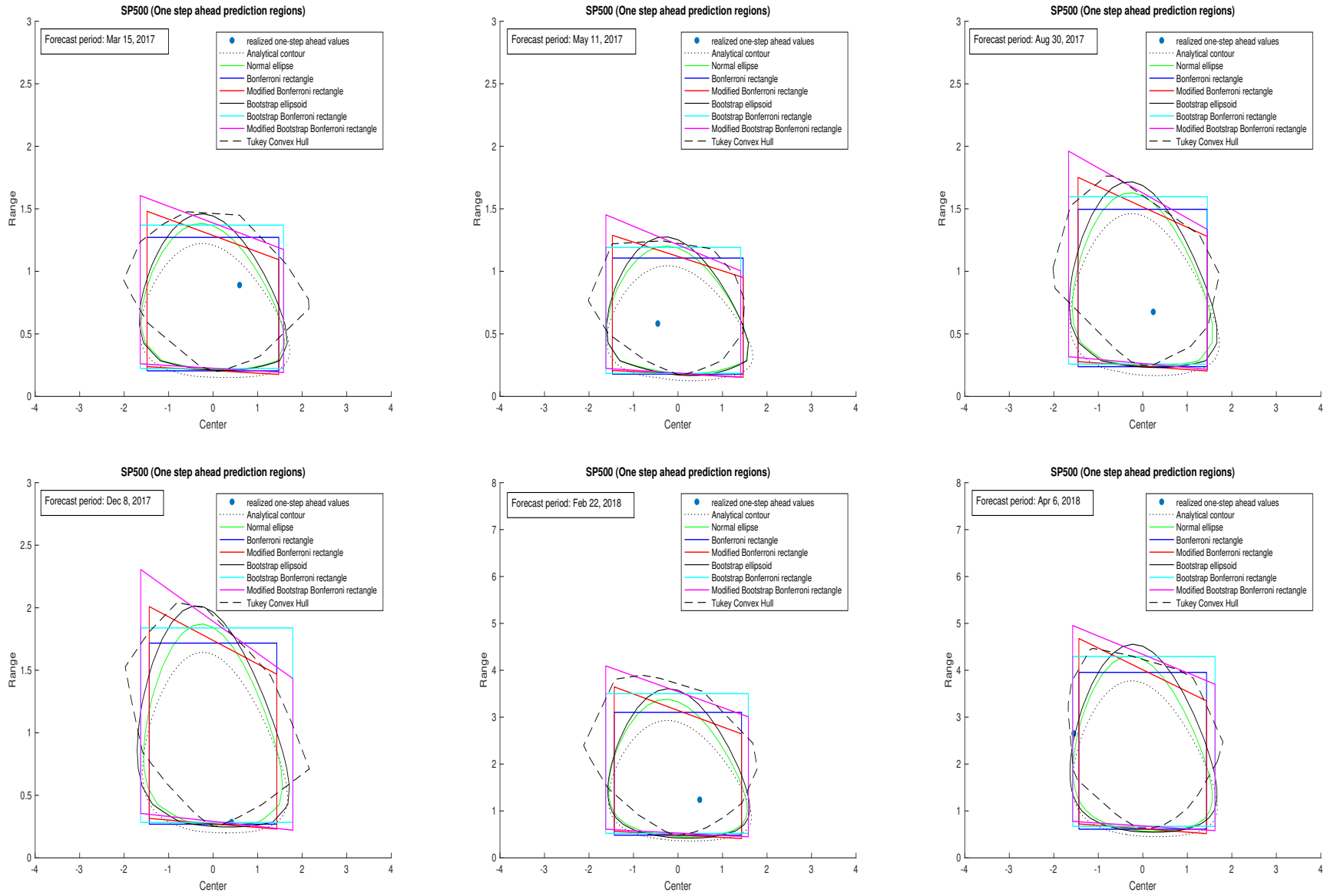

Figure 5: One-step-ahead 95% prediction regions for the center/range system of the SP500 return intervals corresponding to different dates of the out-of-sample period.

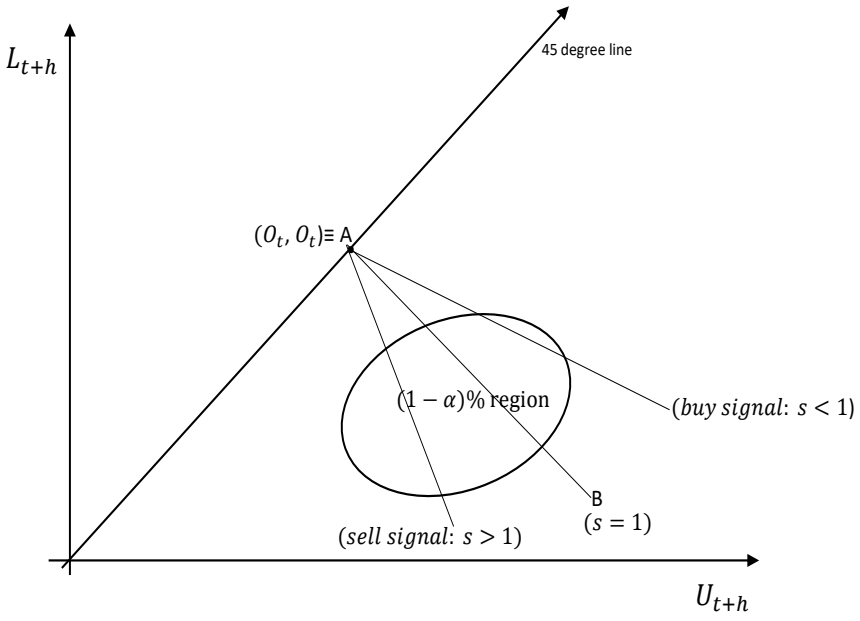

Figure 6: Buy and sell signals from trading strategy.