

CentER 

Discussion Paper

No. 2010–107

ROBUST CONTROL CHARTS FOR TIME SERIES DATA

By Christophe Croux, Sarah Gelper and Koen Mahieu

October 2010

ISSN 0924-7815

Robust control charts for time series data

Christophe Croux

K.U. Leuven & Tilburg University

Sarah Gelper

Koen Mahieu

Erasmus University Rotterdam

K.U. Leuven

Abstract

This article presents a control chart for time series data, based on the one-step-ahead forecast errors of the Holt-Winters forecasting method. We use robust techniques to prevent that outliers affect the estimation of the control limits of the chart. Moreover, robustness is important to maintain the reliability of the control chart after the occurrence of alarm observations. The properties of the new control chart are examined in a simulation study and on a real data example.

Keywords: Control chart, Holt-Winters, Non-stationary time series, Outlier detection, Robustness, Statistical process control.

1 Introduction

Control charts are used to detect anomalies in processes. They are most often used to monitor production-related processes. Samples taken from such processes are assumed to be independent and identically distributed. Many business-related processes, for instance sales volumes or product prices, behave very differently. They typically contain a trend, local or global, and serial correlation. In this paper we propose a control chart aimed at detecting aberrant observations, or outliers, in such time series. The control chart needs to be robust, meaning that the presence of outliers in the series should not harm its performance. Outliers may be present in the “training period”, being the part of the series used to determine the control limits, or in the “test period”, being the part of the series where outliers should be flagged as alarm observations.

The control chart we propose belongs to the class of special-cause control charts (Alwan and Roberts, 1988), where forecast errors of a prediction method are subject to regular control chart techniques. If an observation has a large deviation from its predicted value, an unexpected event occurred, and an alarm should be given. A large difference between the observed and predicted value implies an unusually large forecasting error, and the corresponding observation will be flagged as an outlier in the control chart of the forecast errors. In this paper, we apply the Holt-Winters forecast method, which is a widely used and simple procedure to forecast time series containing trends and serial correlation. Moreover, it is known to yield good forecast performance for many types of times series encountered in business and industry (see Gardner (2006) and De Gooijer and Hyndman (2006) for recent applications). Every single forecast error is then monitored on an X -chart for individual data. Since the control chart should be resistant to outliers, a robust version of both the Holt-Winters forecasting procedure and the X -chart needs to be used.

Several types of control charts for time series data have been proposed in the literature. To deal with serial correlation in stationary processes ARMA-charts (Jiang, Tsui, and Woodall, 2000) were proposed, or data transformations (Wang, 2005). CuScore charts (Box and Ramírez, 1992) can be of use for non-stationary series; see Nembhard and Changpetch (2007) for a recent application. Vander Wiel (1996) proposed control charts for processes that wander, based on integrated moving average models. The Holt-Winters method is an important example of this approach. As the other proposals, the approach of Vander Wiel (1996) is sensitive to outliers in the training sample and test sample. This problem is remediated by following the robust approach taken in this paper.

Robust versions of the standard \bar{X} and R chart are given in Rocke (1989), Rocke (1992) and Tatum (1997), where the mean and scale of the process are estimated robustly for setting up the control limits. More recently, robust control charts for multivariate processes were proposed, see Vargas (2003), Alfaro and Ortega (2009), and Stefatos and Hamza (2009). This paper contributes to this literature by providing a robust control chart for time series typically encountered in business environments, containing stochastic trends, outliers, and strong serial correlation.

In Section 2 control charts based on the Holt-Winters method are briefly explained, while Section 3 describes the robust version. The Holt-Winters method, both the standard and the robust version, has the advantage of being easy to implement by means of a simple updating scheme. Section 4 presents a simulation study that demonstrates the advantages of a robust approach. First, the robust version

gives reliable control limits in cases where the training sample contains outliers. Secondly, if an isolated outlier occurs in the test sample, the non-robust control chart might yield a sequence of false alarms right after the occurrence of the outlier. The robust approach does not suffer from this drawback. Section 5 presents an application with real data. Finally, some conclusions are made in Section 6.

2 Control charts using the Holt-Winters method

In this section, we construct a control chart for a time series $\{y_t\}$. At every time point we make a one-step-ahead forecast by the Holt-Winters forecasting algorithm. The one-step-ahead forecast errors of this procedure are plotted on a control chart. Control limits to monitor new incoming observations are constructed from the training sample.

2.1 The forecasting algorithm

Consider a time series observed up to time point $t - 1$. The Holt-Winters method makes a prediction of the series at time t , denoted by $\hat{y}_{t|t-1}$. Once the true value y_t is observed, the one-step-ahead forecast error is given by

$$e_t = y_t - \hat{y}_{t|t-1}. \quad (1)$$

Denote the local level of the series by $\{\alpha_t\}$ and the local trend by $\{\beta_t\}$. The Holt-Winters algorithm estimates these unknown values by

$$\begin{aligned} \hat{\alpha}_t &= \lambda_1 y_t + (1 - \lambda_1) (\hat{\alpha}_{t-1} + \hat{\beta}_{t-1}) \\ \hat{\beta}_t &= \lambda_2 (\hat{\alpha}_t - \hat{\alpha}_{t-1}) + (1 - \lambda_2) \hat{\beta}_{t-1} \end{aligned} \quad (2)$$

resulting in the forecast

$$\hat{y}_{t|t-1} = \hat{\alpha}_{t-1} + \hat{\beta}_{t-1}. \quad (3)$$

The parameters λ_1 and λ_2 in (2) are smoothing parameters, with values between zero and one. The larger the smoothing parameters, the less the local level and/or local trend series are smoothed, and the more weight the current value of y_t has on the prediction of the next value.

The recursive relations (2) are started up after a short startup period of length m . A linear trend line is fitted to the data in the startup period for estimating the

local level and trend at $t = m$. The smoothing parameters λ_1 and λ_2 are selected by minimizing the sum of squared forecast errors over a training period of length n :

$$(\lambda_1^{\text{opt}}, \lambda_2^{\text{opt}}) = \underset{(\lambda_1, \lambda_2)}{\operatorname{argmin}} \sum_{t=m+1}^n (y_t - \hat{y}_{t|t-1})^2. \quad (4)$$

Note that the observations in the startup period are not used for determining the smoothing parameters.

2.2 Control chart

As motivated in Section 1, the one-step-ahead forecast errors $e_t = y_t - \hat{y}_{t|t-1}$ are monitored to detect anomalies in the outcome of the process. We proceed as for a regular X -chart, and assume normality of the forecast errors. The control limits are computed from the forecast errors e_t in the training period, with $t = m + 1, \dots, n$. Denote S^2 the mean of the squared prediction errors in the training period, excluding the startup values:

$$S^2 = \frac{\sum_{t=m+1}^n e_t^2}{n - m}.$$

Since the target value for the forecast errors is zero, the following $(1 - \alpha)$ Upper and Lower Control Limits are obtained:

$$\begin{aligned} \text{UCL} &= z_{\alpha/2} \times S \\ \text{LCL} &= -z_{\alpha/2} \times S, \end{aligned} \quad (5)$$

with $z_{\alpha/2}$ the $(1 - \alpha/2)$ -quantile of the standard normal distribution.

For every new observations y_t in the test period, where $t > n$, one updates the recursion relations (2) and computes the new prediction error e_t as in (1). The value of e_t is then plotted in the control chart. If it falls out of the control limits, then the observed value deviates substantially from the predicted value, indicating an unexpected change in the process. Note that the observations in the test period influence future prediction errors via the Holt-Winters forecast formulas in (2), but do not alter the control limits.

It is of crucial importance that the process is completely in control during the training period, otherwise the scale estimate S is inflated by the presence of outliers in the training period. If outliers are present in the training sample, it is necessary to follow a robust approach, as discussed in the next section.

3 Control charts using the Robust Holt-Winters method

The control chart presented in the previous section is sensitive to outlying observations in several ways. First, the Holt-Winters forecasts are adversely affected by outliers, since the estimated local level and trend in (2) depend on past values of the observed series, and thus also on possible outliers. Second, if outliers are present in the training sample, the selected smoothing parameters λ_1 and λ_2 might be biased. Finally, due to an inflated scale of the forecast errors, outliers may lead to wider control limits.

3.1 The forecasting algorithm

The shortcomings of the standard Holt-Winters algorithm have been addressed by Gelper *et al.* (2009). They present a robust version of Holt-Winters method, consisting of the application of the standard procedure on a cleaned version of the observed series. Furthermore, they stress the importance of a robust selection of the smoothing parameters and a robust estimation of the starting values of the algorithm. We briefly review their approach.

A cleaned version y_t^* of the time series y_t is obtained as

$$y_t^* = \psi_k \left(\frac{y_t - \hat{\alpha}_{t-1} - \hat{\beta}_{t-1}}{\hat{\sigma}_t} \right) \hat{\sigma}_t + \hat{\alpha}_{t-1} + \hat{\beta}_{t-1} \quad (6)$$

where $\psi_k(y) = \max(-k, \min(y, k))$ is the Huber ψ -function with boundary value k , and $\hat{\sigma}_t$ is a local scale estimate of the one-step-ahead forecast errors. We set $k = 2$, such that the influence of forecast errors larger than two times the local scale estimate is bounded. Note that for k tending to infinity, one gets the standard Holt-Winters method again. The local scale estimate is computed recursively as

$$\hat{\sigma}_t^2 = \lambda_\sigma \rho_k \left(\frac{y_t - \hat{\alpha}_{t-1} - \hat{\beta}_{t-1}}{\hat{\sigma}_{t-1}} \right) \hat{\sigma}_{t-1}^2 + (1 - \lambda_\sigma) \hat{\sigma}_{t-1}^2 \quad (7)$$

where $\rho_k(\cdot)$ is a bounded loss function with boundary value $k = 2$, see Gelper *et al.* (2009). Forecast errors larger than two times the local scale estimate $\hat{\sigma}_{t-1}$ will have a bounded influence on the next scale estimate. We allow for a slowly varying scale, which is translated in a choice of λ_σ rather close to zero, for example $\lambda_\sigma = 0.3$.

The recursive scheme of the robust Holt-Winters algorithm then becomes, analogous to (2),

$$\begin{aligned}\hat{\alpha}_t &= \lambda_1 y_t^* + (1 - \lambda_1)(\hat{\alpha}_{t-1} + \hat{\beta}_{t-1}) \\ \hat{\beta}_t &= \lambda_2 (\hat{\alpha}_t - \hat{\alpha}_{t-1}) + (1 - \lambda_2) \hat{\beta}_{t-1},\end{aligned}\tag{8}$$

where y_t^* follows equation (6).

There are two further issues that need to be addressed. First, the startup values $\hat{\alpha}_m$ and $\hat{\beta}_m$ need to be computed using a robust regression fit to the data in the startup period. For this robust fit the repeated median estimators is used. The median absolute deviation of the residuals with respect to this robust fit yields the starting value $\hat{\sigma}_m^2$ for the scale estimates. The second issue is the robust selection of the smoothing parameters. Instead of minimizing the sum of squares of the one-step-ahead forecast errors over the training sample, we minimize a robust performance measure

$$(\lambda_1^{\text{opt}}, \lambda_2^{\text{opt}}) = \underset{(\lambda_1, \lambda_2)}{\operatorname{argmin}} s_0^2 \sum_{t=m+1}^n \rho_k \left(\frac{y_t - \hat{y}_{t|t-1}}{s_0} \right).\tag{9}$$

This performance measure corresponds to a τ -scale estimator, see for example Maronna *et al.* (2006), which combines a high robustness with a high statistical efficiency. The auxiliary scale estimate s_0 in (9) is given by $s_0 = \operatorname{med}_{m+1 \leq t \leq n} |y_t - \hat{y}_{t|t-1}|$, the median absolute deviation of the prediction errors. The loss function used in (9) is

$$\rho_k(y) = \min(k^2, y^2) = \psi_k(y)^2,\tag{10}$$

with $k = 2$.

3.2 Control chart

Similar as in subsection 2.2, we plot the one-step-ahead forecast errors $e_t = y_t - \hat{y}_{t|t-1}$, for $t > n$, in the control chart, but now with forecasts obtained from the robust Holt-Winters algorithm. The control limits are determined from a robust scale estimate of the one-step-ahead forecast errors from the training period. We use again the τ -estimator of scale

$$\tau^2 = c_k \frac{s_0^2}{n - m} \sum_{t=m+1}^n \rho_k \left(\frac{e_t}{s_T} \right),\tag{11}$$

where $s_0 = \operatorname{med}_{m+1 \leq t \leq n} |e_t|$, and ρ_k as in (10). The consistency factor c_k in the above equation ensures that τ^2 is a consistent estimator of the population variance of a normal distribution. For $k = 2$, we have $c_k = 1.404$. We construct approximate

$(1 - \alpha)$ control limits as

$$\begin{aligned} \text{UCL} &= z_{\alpha/2} \times \tau \\ \text{LCL} &= -z_{\alpha/2} \times \tau \end{aligned} \tag{12}$$

with $z_{\alpha/2}$ as the $(1 - \alpha/2)$ -quantile of the standard normal distribution. Hence, the formula for the control limits (5) and (12) are similar, except that the robust method uses a robust τ -scale estimator instead of a standard deviation (centered at zero). The use of the τ -scale estimator guarantees the resistance to outliers in the training period.

In the next section a simulation study is carried out to verify whether the constructed control limits for the standard and the robust Holt-Winters method have comparable properties if the training sample is clean. In the presence of outliers in the training sample, we expect the robust method to perform better.

4 Simulation study

In this simulation study we take a closer look at the *size*, the *power* and the *false detection rate* of the proposed robust control chart, and compare them to the same characteristics of the non-robust version. The *size* of a control chart is the probability that an observation in the test sample falls out of the control limits when the process is actually in control, i.e. when no outliers are present in the test sample. It is supposed to be approximately 5%, when taking $\alpha = 0.05$ in (5) and (12). The *power* is the probability that an alarm observation is detected. In a simulation study, the power equals the proportion of generated outliers in the test sample that are detected. The percentage of observations in the test sample that are not generated as outliers, but that are falsely identified as alarm observations, is the *false detection rate*. The false detection rate is like the size of the control charts, but for a process that is not fully under control.

4.1 Simulation design

We generate series $\{y_t\}$ from a local linear trend model. This model generates non-stationary and correlated time series, and allows for trends. It is a reasonable model for many series encountered in the practice of business and economics. At this model the forecast errors of the Holt-Winters method form an i.i.d. sequence

from a normal distribution (e.g. Chatfield and Yar (1988)). The local linear trend model is defined as

$$Y_t = \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \quad (13)$$

where α_t , the local level, and β_t , the local trend, are given by

$$\begin{cases} \alpha_t = \alpha_{t-1} + \beta_{t-1} + \eta_t, & \eta_t \sim N(0, \sigma_\eta^2), \\ \beta_t = \beta_{t-1} + \nu_t, & \nu_t \sim N(0, \sigma_\nu^2). \end{cases} \quad (14)$$

The noise terms ε_t , η_t and ν_t are independent, and serially uncorrelated. In the simulation study, we take $\sigma_\varepsilon = 1$ and $\sigma_\eta = \sigma_\nu = 0.1$, leading to a smaller variance for the local level and trend compared to the variance in the measurement equation (13).

The first n observations from a generated series constitute the training sample, including a startup period of length $m = 10$. The next 200 generated observations constitute the test sample. From the training data the control limits are constructed as outlined in Sections 2 and 3. Outliers, hence alarm observations, are induced in the test sample by adding the value $k\sigma_\varepsilon$ to 10% of randomly selected observations in the test sample. For simulating the power, we take $k = 5$, corresponding to outliers of moderate size. To study the robustness of the size and power with respect to outliers in the training sample, the outliers are generated in the same way as for the test sample. The percentage of outliers in the training data are taken as 0% (clean), 2%, and 5%, respectively.

4.2 Simulation Results

The simulated power and size are summarized in Table 1 for the control charts based on the standard and the robust Holt-Winters algorithm. The reported numbers are averages over 1000 simulation runs. Training samples of length $n = 50$ and $n = 100$ are considered. The simulation study, although rather modest in the number of considered simulation settings, is quite time consuming, since the optimal selection of the smoothing parameters, see (4) and (9), needs to be performed for every single time series.

As we see from Table 1 there is a small size distortion in the clean case, when no outliers are present in the training sample, and this for both methods. For the larger sample size, the simulated size is already much closer to 5%. In the presence of outliers in the training sample, the size of the non-robust chart is close to zero in most cases, while the size of the robust method is kept much more stable. The

Table 1: Size and Power of the control chart, with $\alpha = 5\%$, for the standard (HW) and the robust Holt-Winters (RHW) procedures, for different sizes of the training sample (50 and 100), and for different percentages of outliers in training sample.

Size	Clean		2%		5%		
	<i>n</i>	HW	RHW	HW	RHW	HW	RHW
50		.073	.086	.044	.073	.026	.067
100		.059	.063	.027	.052	.009	.037

Power	Clean		2%		5%		
	<i>n</i>	HW	RHW	HW	RHW	HW	RHW
50		.903	.900	.843	.874	.784	.850
100		.903	.902	.842	.881	.757	.853

breakdown of the size for the non-robust method is due to two reasons. The main reason is that the scale S in (5) is inflated due to the outliers, resulting in too wide control limits. Furthermore, the outliers cause a bias in the selection of the smoothing parameters, leading to a less optimal smoothing.

Consider now the power in Table 1. If the training sample is clean, the two methods have a comparable power to detect the generated alarm observations. When 2% or 5% of outliers are present in the training sample, they do affect the detection power of the control charts, and we see that the robust Holt-Winters clearly outperforms the standard one. The difference in power is about 10% in presence of only 5% of moderate outliers in the training sample. When increasing k , the size of the outliers, and the percentage of outliers in the training sample, these differences become even much more pronounced.

Next, we compare the false detection rate of the two control charts. We follow the same simulation scheme as before, but only consider clean training data to ensure that the size of the two methods remains comparable. We vary the magnitude of the outliers by taking $k = 5, 10, 15$ and 20 . A first observation is that the false detection rate does not seem to depend much on n . Furthermore, for $k = 5$ the robust method is only slightly better than the non-robust method. For larger values of k , in contrast, the robust procedure has a much lower false detection rate. This is due to the following reason. For large values of k , an outlying observation in the test sample affects the outcome of the next forecast in the Holt-Winters procedure.

Table 2: False detection rate of the control chart, with $\alpha = 5\%$, for the standard (HW) and the robust Holt-Winters (RHW) algorithm, and for different lengths n of the training sample. The training data are clean, and outliers are created in the test data by shifting 10% of the observations upwards by k units.

shift size	HW				RHW			
	$k = 5$	$k = 10$	$k = 15$	$k = 20$	$k = 5$	$k = 10$	$k = 15$	$k = 20$
$n = 50$	0.097	0.167	0.206	0.232	0.084	0.088	0.083	0.085
$n = 100$	0.093	0.161	0.203	0.229	0.079	0.080	0.081	0.082

Hence it is likely to induce a large forecasting error for the next observation as well, even if the latter one is not an outlier. The propagation effect of outliers does not occur with the robust approach, since it uses “cleaned observations” in the update formulas (8).

5 Real data example

We illustrate the proposed method by means of a real data example. We use the Housing Data from the book of Diebold (2001), containing monthly data for housing starts and completions from January 1968 until June 1996. Figure 1 shows both series and the corresponding control charts obtained by the standard and the robust Holt-Winters procedure. The upper panel of every graph shows the observed data y_t , together with the smoothed series $\hat{y}_{t|t-1}$. The lower panel of every graph contains the control chart. As explained in section 2, the prediction errors e_t are plotted in the control chart. The vertical dashed lines in the chart indicate the end of the startup period and of the training period. Control limits are indicated for the test period. We used a startup period of length $m = 20$, a training period of length $n = 150$. The other 192 observations serve as the test sample. We immediately notice two large outliers, one near 1971 and another near 1977, in the training sample of the housing starts. Moreover, both time series may contain other smaller outliers. By applying robust exponential smoothing, we expect that the results remain stable in the presence of these outliers.

We first take a closer look at the housing starts in Figure 1(a)-(b). The smoothing parameters (λ_1, λ_2) selected from the training period are $(0.3, 0.2)$ for the standard

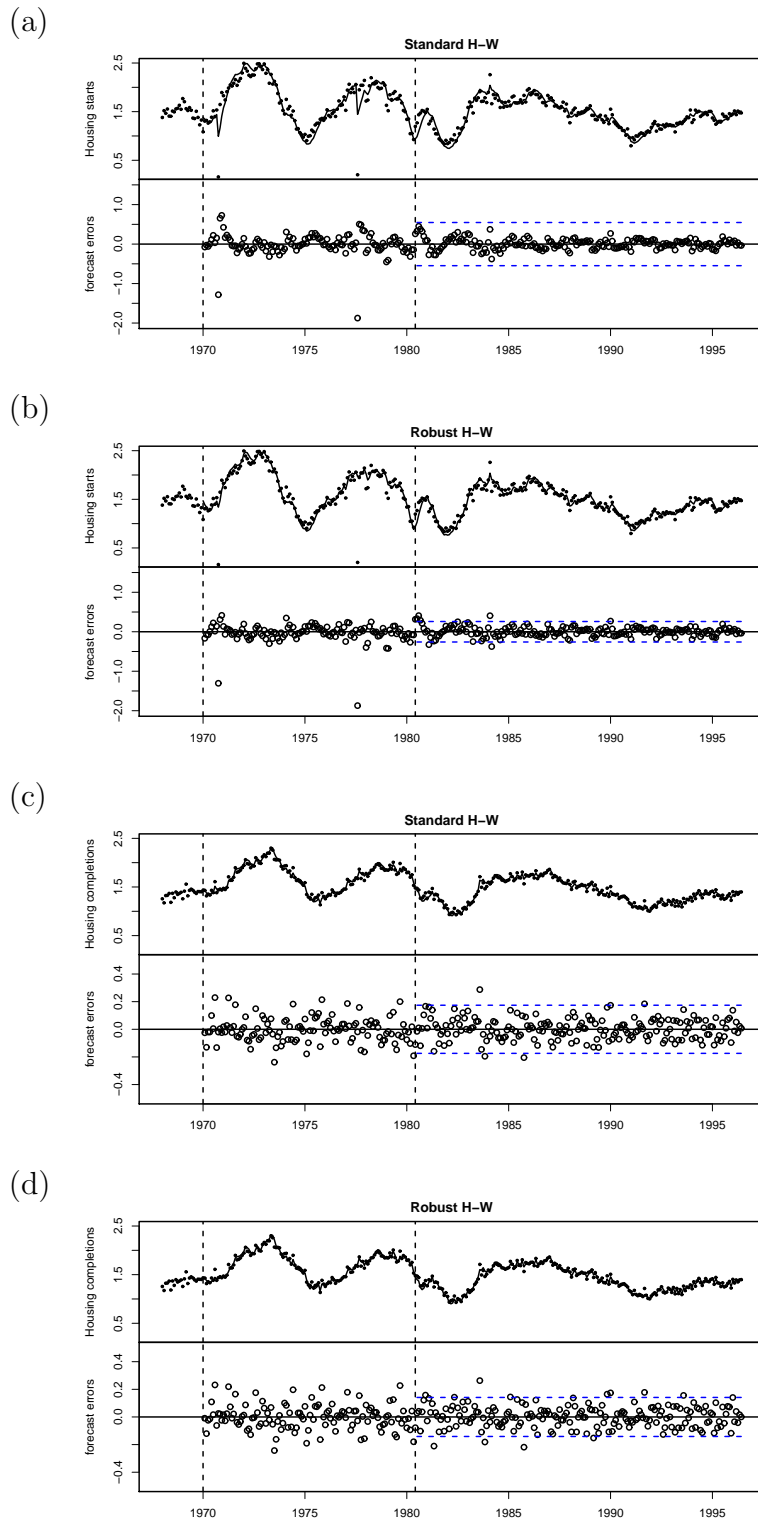


Figure 1: In each graph, the upper panel presents the observed data (dots) and the smoothed series (solid line), while the lower panel presents the control chart, with forecast errors (points) and control limits (dashed lines). The housing starts series - containing outliers in the training period - is analyzed with standard Holt-Winters in (a) and with robust Holt-Winters method in (b). Similarly for the housing completions series. This series does not contain outliers in the training period, explaining the resemblance between (c) and (d).

method and $(0.4, 0.2)$ for the robust method. These parameters are reasonable, since they are close to zero and the parameter corresponding to the trend is smaller than the one for the level. The standard method tends to select smaller parameters than the robust version, since outliers may bias the parameters towards zero. The direct effect of the large outliers in the standard Holt-Winters case can be seen at the time points subsequent to the large outliers. The forecasts for these time points are attracted by the outlier and therefore, in this case, the values are too small. The robust method clearly yields better forecasts in the outlier's subsequent time points. Also in terms of mean squared one-step-ahead forecast errors (MSFE), computed over the test period, the robust method outperforms the standard one with an MSFE of 0.0144 against 0.0175. The control chart shows that the control limits are inflated by the outliers in the non-robust case, and all observations in the test sample remain within the control bounds. This leads to a loss of power of the non-robust method to detect outliers. For example, the alarm observation in January 1984 is only detected by the robust method.

The housing completions in Figure 1(a)-(b), do not show important outliers in the training period. The selected smoothing parameters for this series are $(0.5, 0.2)$ in the standard case and $(0.6, 0.3)$ in the robust case. Similar to the housing starts, the standard parameters are somewhat smaller than the robust ones. The MSFE for both methods are comparable, with 6.01×10^{-3} in the standard case and 6.09×10^{-3} in the robust case. Since the series do not contain important outliers, both the standard and the robust control charts look similar. Nevertheless, the robust control chart has slightly tighter control limits. The percentage of observations exceeding the control limits is 2% for the standard method, which seems overly optimistic, and 6.1% for the robust method, closer to the expected size of $\alpha = 5\%$.

We conclude that it is worthwhile to use the robust method. If there are no major outliers, the performance of both methods is quite similar. However, if major outliers are present, in particular in the training period, the robust method is more reliable.

6 Conclusion

In this paper, we propose to use the one-step-ahead forecast errors of the Holt-Winters forecasting method as the input of a control chart for monitoring non-stationary processes. Holt-Winters forecasting is a frequently used method for univariate time series with a stochastic trend and a seasonality pattern. In this study,

we did not explicitly deal with seasonality, but the method can easily be adapted for that purpose. The main advantages are that the method is very easy to implement and still performs well compared to other more complicated forecasting techniques. We compared the control charts based on the standard Holt-Winters method with a robust version.

The simulation study shows that with clean training data, the standard and robust control charts perform almost equally well in terms of size and power. On the other hand, when we add a small amount of moderate outliers to the training sample the non robust method loses power, and suffers from severe size distortion. Even when the training data are clean, the false detection rate of the robust procedures is much lower if major outliers are present in the test sample. Finally, the application to real data shows that the proposed robust method is easy to put in practice. The graphical displays in Figure 1, representing both the observed series and the control chart in the same graph, are an easy-to-use visual aid for managers to monitor business processes.

There are several open questions, which we consider as areas for future research. If a structural break appears in the time series, the control chart should be readjusted and new control limits need to be computed. Another limitation of the presented approach is the normality assumption on the forecast errors (excluding outliers). Different distributional assumptions will lead to different formula for the control limits. We refer to Yang and Chen (2009) and Chen and Yeh (2009) for recent contributions to these problems, although in different settings.

References

- Alfaro, J. and Ortega, J. (2009). A comparison of robust alternatives to Hotelling's T^2 control chart. *Journal of Applied Statistics*, **36**(12), 1385–1396.
- Alwan, L. C. and Roberts, H. V. (1988). Time-series modeling for statistical process control. *Journal of Business & Economic Statistics*, **6**(1), 87–95.
- Box, G. and Ramírez, J. (1992). Cumulative score charts. *Quality and Reliability Engineering*, **8**(1), 17–27.
- Chatfield, C. and Yar, M. (1988). Holt-Winters forecasting: some practical issues. *The Statistician*, **37**(2), 129–140.
- Chen, F. L. and Yeh, C. H. (2009). Economic statistical design of non-uniform

- sampling scheme \bar{X} control charts under non-normality and Gamma shock using genetic algorithm. *Expert Systems with Applications*, **36**(5), 9488–9497.
- De Gooijer, J. and Hyndman, R. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, **22**(3), 443–473.
- Diebold, F. X. (2001). *Elements of forecasting*. South-Western, second edition.
- Gardner, E. (2006). Exponential smoothing: the state of the art - part ii. *International Journal of Forecasting*, **22**(4), 637–666.
- Gelper, S., Fried, R., and Croux, C. (2009). Robust forecasting with exponential and Holt-Winters smoothing. *Journal of Forecasting*, **29**, 285–300.
- Jiang, W., Tsui, K.-L., and Woodall, W. H. (2000). A new SPC monitoring method: the arma chart. *Technometrics*, **42**(4), 399–410.
- Maronna, R. A., Martin, R. D., and Yohai, V. J. (2006). *Robust statistics: theory and methods*. Wiley.
- Nembhard, H. B. and Changpetch, P. (2007). Directed monitoring using cuscore charts for seasonal time series. *Quality and Reliability Engineering International*, **23**, 219–232.
- Rocke, D. M. (1989). Robust control charts. *Technometrics*, **31**(2), 173–184.
- Rocke, D. M. (1992). \bar{X}_q and r_q charts: robust control charts. *The Statistician*, **41**(1), 97–104.
- Stefatos, G. and Hamza, A. B. (2009). Fault detection using robust multivariate control chart. *Expert Systems with Applications*, **36**(3), 5888–5894.
- Tatum, L. G. (1997). Robust estimation of the process standard deviation for control charts. *Technometrics*, **39**(2), 127–141.
- Vander Wiel, S. (1996). Monitoring processes that wander using integrated moving average models. *Technometrics*, **38**(2).
- Vargas, J. A. (2003). Robust estimation in multivariate control charts for individual observations. *Journal of Quality Technology*, **35**(4).
- Wang, F. (2005). A simple data transformation of auto-correlated data for SPC. *International Journal of Production Research*, **43**(5), 981–989.

Yang, S.-F. and Chen, W.-Y. (2009). Controlling over-adjusted process means and variances using VSI cause selecting control charts. *Expert Systems with Applications*, **36**(3), 7170–7182.