Prediction intervals for heteroscedastic series by Holt-Winters methods

Paolo Chirico
DIGSPES Alessandria
University of Eastern Piedmont, Italty
paolo.chirico@uniupo.it

Abstract The paper illustrates a procedure to calculate prediction intervals in case of heteroscedasticity using Holt-Winters methods. The procedure has been applied to the Italian daily electricity prices (PUN) of the year 2014; then the prediction intervals have compared to those provided by an ARIMA-GARCH model. The intervals obtained with HW methods have been very similar to the others, but easier to calculate. Moreover, the HW procedure is more flexible in dealing with periodic volatility as proved in the case study.

Key words: Holt-Winters methods, heteroscedasticity, prediction intervals.

1 Introduction

The Holt-Winters (HW) methods are forecast methods grounding on recursive formulas for updating the structural components of time series: the local mean level, L_t; the local trend, T_t; the local seasonal index, I_t, if the series is affect by periodic effects. According to the composition model of the structural components, the HW methods can be either additive or multiplicative. In the first case, when a new observation, y_t becomes available, the forecast of y_{t+k}, denoted by $\widehat{y}_t(k)$, is:

$$\widehat{y}_t(k) = L_t + kT_t + I_{t+k-hS} \qquad h = int(k/S) + 1 \tag{1}$$

after having updated the structural components:

$$L_t = \alpha(Y_t - I_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1}) \tag{2}$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \tag{3}$$

$$I_t = \gamma(Y_t - L_t) + (1-\gamma)I_{t-S} \tag{4}$$

where $0 \leq \alpha, \beta, \gamma \leq 1$ are smoothing coefficients, and S is the length of the seasonal cycle. The additive HW provides optimal forecast if the series derives from a SARIMA process of order $(0, 1, S+1)(0, 1, 0)_S$ (Yar and Chatfield, 1990) [1].

Post print copy

[1] Instead, the multiplicative HW is not optimal for any linear process. In case of multiplicative composition of the structural components, it is worthwhile to apply the additive HW to the log-serie.

Originally, the HW methods were not thought for prediction intervals (PI*s*), but over time, some authors have proposed ways to get PI*s*. The proposal of Yar and Chatfield (1990) for the additive method allows to get the same PI*s* of the equivalent SARIMA model. They define the *k*-steps-ahead prediction error as:

$$e_t(k) = \sum_{i=0}^{k-1} v_i e_{t+k-i} \tag{5}$$

where $e_t = y_t - \widehat{y}_{t-1}(1)$ is the one-step-ahead (OSA) prediction error, and:

$$v_i = \begin{cases} 1 & i = 0 \\ \alpha + i\alpha\beta & i \neq 0, S, 2S, ... \\ \alpha + i\alpha\beta + (1-\alpha)\gamma & i = S, 2S, ... \end{cases} \tag{6}$$

If the OSA prediction errors are uncorrelated and Gaussian $N(0, \sigma^2)$, the *k*-steps-ahead prediction intervals, at $1 - \theta$ confidence level, is:

$$\widehat{y}_t(k) \pm z_{(\theta/2)} \sqrt{PMSE_t(k)} \tag{7}$$

$$PMSE_t(k) = \sigma^2 \sum_{i=0}^{k-1} v_i^2 \tag{8}$$

Nevertheless this proposal assumes the OSA prediction errors are homoscedastic, that is not realistic in some cases like electricity prices (see Section 3).

2 An HW method for heteroscedastic series

Let's suppose the OSA prediction errors are heteroscedastic, and the volatility level of them, denoted by H_t, is only locally approximative constant. Following the criterion of the Exponential Smoothing, the update of the volatility level may be formulated as:

$$H_t = \lambda e_t^2 + (1-\lambda)H_{t-1} \tag{9}$$

where $0 \leq \lambda \leq 1$ is a smoothing coefficient.

According to the criterion: $\widehat{\sigma}_{t+1}^2 = H_t$, the volatility forecasts are:

$$\widehat{\sigma}_{t+1}^2 = \lambda e_t^2 + (1-\lambda)\widehat{\sigma}_t^2 \tag{10}$$

$$\widehat{\sigma}_{t+k}^2 = \widehat{\sigma}_{t+1}^2 \quad k = 2, 3, ... \tag{11}$$

The equation (10) is known in Statistical Process Control as *Exponentially Weighted Moving Variance*, EWMV, (Harris, 1993). It is formally similar to the IGARCH model, but there are conceptual differences: the IGARCH model, as well as all models of GARCH family, shapes the conditional heteroscedasticity of homoscedastic processes (Engle and Bollerslev, 1986); the method (10) treats a general heteroscedasticity where the volatility can be viewed as a stochastic process. Then the method (10) is conceptually more similar to the methods for stochastic volatility (Taylor, 1994).

Taking into account the (11), under the assumption of uncorrelated OSA prediction errors, the $PMSE_t(k)$ becomes:

$$PMSE_t(k) = \sum_{i=0}^{k-1} v_i^2 \sigma_{t+k-i}^2 \tag{12}$$

$$= \sigma_{t+1}^2 \sum_{i=0}^{k-1} v_i^2 \tag{13}$$

where the coefficients v_is have been defined in (6).

This formula is very similar to (8), but presents a fundamental difference: now the $PMSE_t(k)$ depends on the last predicted volatility; that means more accurate PIs in case of heteroscedastic e_t^2.

Periodic Volatility In case of seasonal data, heteroscedasticity can follow a seasonal pattern (*seasonal volatility*): e.g. daily electricity prices present seasonal patterns in the price levels and in the price volatilities. A way to deal with such a case in GARCH framework was proposed by Koopman et al. (2007) who suggested to scale volatility by means of seasonal factors. According to that idea, but remaining in HW framework, the method (10) should be changed in the multiplicative following HW method[2]:

$$\begin{aligned}
\widehat{\sigma}_{t+k}^2 &= H_t \cdot X_{t+k-hS} \qquad h = int(k/S) + 1 \\
H_t &= \lambda(e_t^2/X_{t-S}) + (1-\lambda)H_{t-1} \\
X_t &= \lambda_x(e_t^2/H_t) + (1-\lambda_x)X_{t-S}
\end{aligned} \tag{14}$$

where X_t represents the seasonal effect on volatility, and $0 \leq \lambda_x \leq 1$ is the corresponding smoothing coefficient.

Note that now the formula (12) cannot be reduced into the formula (13) because $\widehat{\sigma}_{t+k}^2$ is only periodically constant, then $PMSE_t(k)$ presents cyclical fluctuations.

The smoothing coefficients The value $(1-\lambda)$ can be viewed as the persistence degree of volatility: when $(1-\lambda) = 1$, volatility (or its level) is constant (max persistence); when $(1-\lambda) = 0$, each variance/level does not depend on any previous variances/levels (null persistence). Similarly, $(1-\lambda_x) = 1$ means that the seasonal pattern on volatility is constant, and $(1-\lambda_x) = 0$ means that there is no seasonal pattern on volatility.

The smoothing coefficients, α, β, γ, λ and λ_x, could be chosen minimising the following sum:

$$\sum[\ln \sigma_t^2 + u_t^2] \tag{15}$$

where $u_t = e_t/\sigma_t$ are the standardised OSA prediction errors.

If $u_t \sim N(0,1)$, the criterion (15) is equivalent to the maximum likelihood, but should not be viewed as an estimation method. Indeed, according to the theory of HW methods, the smoothing coefficients are not process parameters, but instruments to perform time series forecasting. Nevertheless, in case of financial prices, the prediction standardised errors are generally leptokurtic/heavy tailed like standardised Student's t-distributions: $t_v\sqrt{(v-2)/v}$. In that case, it is more appropriate to minimise the following sum:

$$\sum[\ln(\sigma_t^2) + \ln(1 + \frac{u_t^2}{v-2})^{v+1}] \tag{16}$$

where v are the degrees of freedom of the Student's distribution[3].

---

[2] Here the multiplicative HW should be preferred to the additive one because it avoids negative values for volatility forecasts.

[3] Minimizing the sum (16) is nearly equivalent to the maximum likelihood in case of Student't distribution (Bollerslev, 1987).

3 Prediction Intervals for the Italian PUN

We have considered the series of the daily electricity price (PUN), cleared by the Italian wholesale electricity market in 2014. As all daily electricity market prices, Italian PUN is characterized by strong seasonality affecting the price level and price volatility too (Figure 1). On the series we have calculated one step ahead prediction intervals using the method described in the previous sections, and using a suitable ARIMA-GARCH model. Then we have compared the results obtained with both approaches.
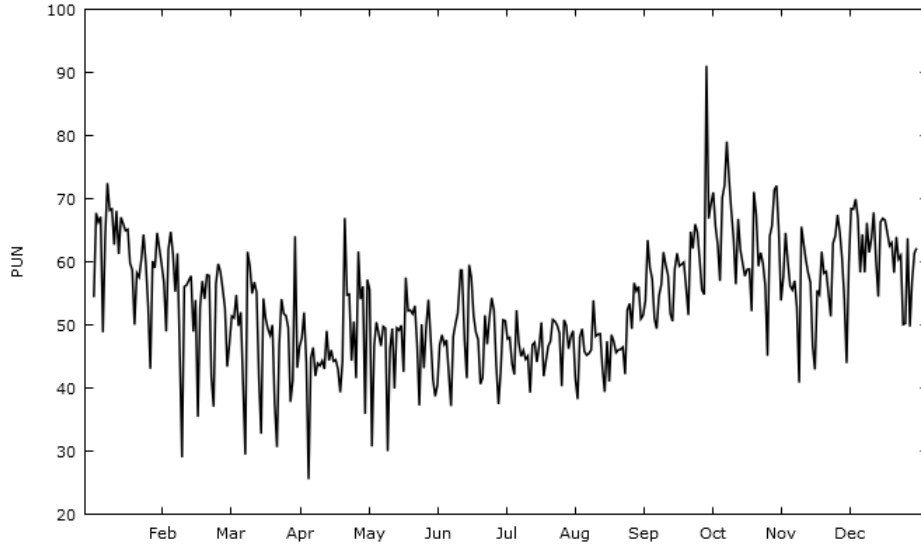
Fig. 1 The Italian PUN in 2014

3.1 The HW approach

Since the presence of spikes and jumps makes the standardised prediction errors of electricity prices strongly leptokurtic, the smoothing coefficients of the HW method have been chosen minimising the sum (16). Table 1 reports the optimal values of the smoothing coefficients, the degree of freedom and the standard deviation of the OSA prediction errors (RMSE)[4].

We can note that the smoothing coefficients present some interesting features: $\beta = 0$ means that the price level does not present any locally persistent drift; $\gamma = 0$ means that the seasonal effects on the price level are periodically constant (deterministic seasonality); $\lambda_x = 0$ means that the seasonality on the volatility is deterministic too[5]. Finally, the low value of the degree of freedom (7.4) confirms the assumption of leptokurtic standardised prediction errors.

We have compared the standardized prediction errors, u_t, to the standardised Student' t-distribution with 7.4 degrees of freedom by means of the quantile-quantile plot (Figure 2).

---

[4] The optimal values are reported without inferential information (standard errors, p-values, etc.) because the coefficients are not interpreted here as process parameters.

[5] The daily effects on PUN have become quite constant since 2013, but it was not in the previous years (Chirico, 2016)

Table 1 Smoothing coefficients

α	β	γ	λ_1	λ_x	v	$RMSE$
0.272	0.000	0.000	0.037	0.000	7.40	5.066
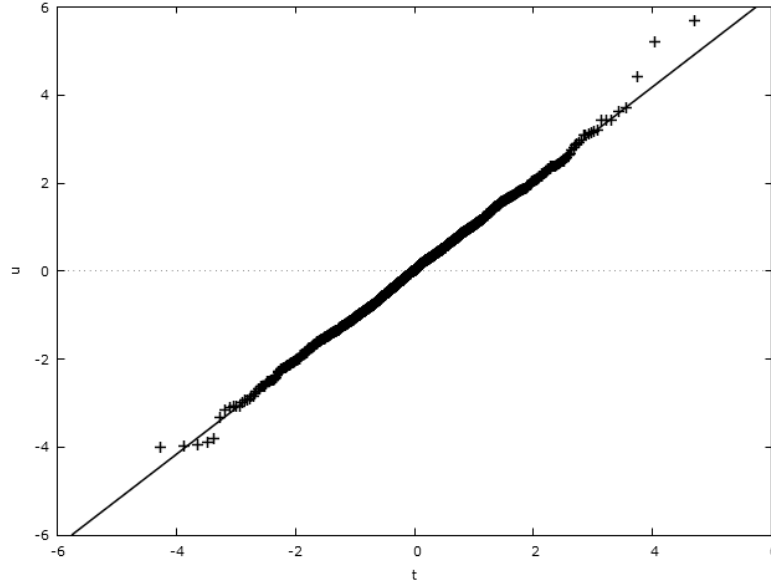
Fig. 2 Q-Q plot of the standardised errors vs standardised Student's t

We can note that the u_ts correspond almost well to the theoretical distribution except some errors at the distribution extremities. That is due to the presence of spikes and jumps which cannot be consistent with the other data.

Therefore the one-step ahead prediction intervals (at $1 - \theta$ confidence level) have been calculated with the following formulas:

$$LB: \ P\hat{U}N_t(1) - t^*_{\theta/2}\sqrt{PMSE_t(1)} \tag{17}$$

$$UB: \ P\hat{U}N_t(1) + t^*_{\theta/2}\sqrt{PMSE_t(1)} \tag{18}$$

where $t^*_{\theta/2}$ is the $\theta/2$-th percentile of the standardised Student's t distribution with 7.4 degrees of freedom.

### 3.2 The ARIMA-GARCH approach

We have repeated the study in 3.1 using an ARIMA-GARCH model. Similar to the method illustrated in the previous subsection, we have fitted the PUN series with an ARIMA-EGARCH model[6] with daily regressors (d_j) and Student's standardised errors (u_t):

---

[6] We have chosen an EGARCH model to avoid the risk of negative variance because that risk is high including seasonal regressors in classic GARCH model.

$$z_t = \Delta^d PUN_t = \mu_t + \varepsilon_t$$

$$\varepsilon_t = \sigma_t u_t \qquad u_t \sim t_v \sqrt{(v-2)/v}$$

$$\mu_t = \delta + \sum_{j=1}^{p} \phi_j z_{t-j} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{j=1}^{6} s_{1,j} d_j \qquad (19)$$

$$\ln \sigma_t^2 = \omega + \sum_{j=1}^{P} \beta_j \ln \sigma_{t-j}^2 + \sum_{j=1}^{Q} (\alpha_j |\varepsilon_{t-j}| + \gamma_j \varepsilon_{t-j}) + \sum_{j=1}^{6} s_{2,j} d_j$$

More specifically awe have identified an ARIMA(6,1,0)-EGARCH(1,1) model with daily regressors, that we have estimated using the *GIG* package of *Gretl* (Lucchetti and Balietti, 2016). For space reasons, we don't report here all the estimation output of the model, but only the standard deviation of the OSA prediction errors, $RMSE = 5.062$. This statistic is practically equal to one of the HW approach, that means the HM method fits the series as good as the ARMA-EGARCH model.

Finally, we have calculated the PIs of PUN series in the period September, 2014-December, 2014 using both methods (Figure 3). Although both intervals capture the 90% of the series in the period, the intervals may be quite different some days. Actually the interval by ARMA-EGARCH model (thin line) seem more sensitive to spike and jumps than one by HW methods does (thick line); this characteristic is typical of EGARCH models because outliers (jumps and spike) produce exponential effects on volatility. On the other hand, the HW procedure provides an interval whose width is quite regular: the seasonality on volatility is periodically constant ($\lambda_x = 0$), and the non-seasonal volatility, H_t, is quite smoothed ($\lambda = 0.037$).
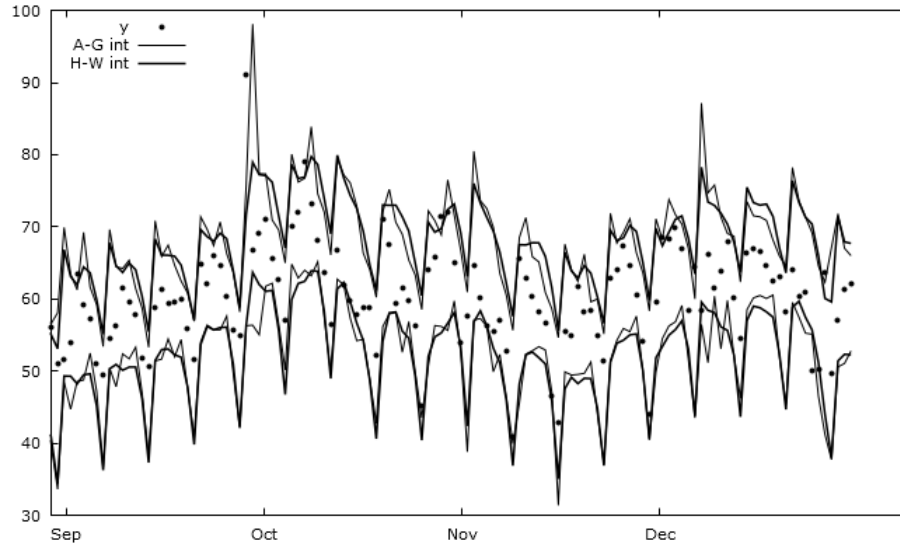
Fig. 3 Prediction intervals

4 Final Considerations

The presence of heteroscedasticity is one of the reasons for preferring ARIMA-GARCH models to HW methods. Nevertheless heteroscedasticity can be treated by HW methods, if some adjustments are taken; the paper illustrates a method to do it. The method has been applied to a case study, and the PI*s* obtained with this method have been similar to the those provided by an ARIMA-EGARCH model; some days the HW procedure seems to work better than ARIMA-EGARCH model too. The most considerable strengths of using HW methods are the ease-of-use and the ability of dealing with several type of heteroscedasticity: stochastic volatility, periodic stochastic volatility, periodically constant volatility. On the other hand, the weakness is that the HW methods are optimal only for specific ARIMA processes although works quite well with a lot of real processes; moreover the HW procedure does not deal with the issue of asymmetry. Nevertheless other versions of the method above-illustrated are possible, and it is our intention to check them.

References

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. The review of economics and statistics, **69**(3), 542–547.

Chirico, P. (2016). Which Seasonality in Italian Daily Electricity Prices? A Study with State Space Models. In Alleva, G., Giommi, A. (Eds.): Topics in Theoretical and Applied Statistics, 275–284. Springer.

Engle, R. F., Bollerslev, T. (1986). Modelling the persistence of conditional variances. Economet. Rev., **5** (1), 1–50.

Koopman, S. J., Ooms, M., Carnero, M. A. (2007). Periodic seasonal Reg- ARFIMA- GARCH models for daily electricity spot prices. J. Am. Stat. Ass., **102**(477), 16–27.

Lucchetti, R. J., Balietti, S. (2016). The gig package. Package Guide of Gretl.

MacGregor, J. F., Harris, T. J. (1993). The exponentially weighted moving variance. Journal of Quality Technology, **25**(2), 106-118.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. Mathematical finance, **4**(2), 183-204.

Yar, M., Chatfield, C. (1990). Prediction Interval for the Holt-Winters forecasting procedure. Int. J. Forecast., **6**, 127–137. North-Holland