Linear Multistep F10.7 Forecasting Based on Task Correlation and Heteroscedasticity

Zhen Wang1, Qinghua Hu1, Qiu Zhen Zhong2,3, and Yun Wang1

1School of Computer Science and Technology, Tianjin University, Tianjin, China, 2National Space Science Center, Chinese Academy of Sciences, Beijing, China, 3School of Astronomy and Space Sciences, University of Chinese Academy of Sciences, Beijing, China

Abstract

Solar 10.7-cm radio flux (F10.7) is an important measure of solar radio emission activity. Accurate F10.7 forecasting plays a key role in both space weather and global environment forecasts. We discover that forecasting errors are heteroscedastic, something that is not often considered in previous models. In addition, task correlation between different forecasting steps is ignored in current multistep-ahead forecast models. In this work, we propose a linear multistep forecasting model based on the correlation between different forecasting steps and the characteristic of heteroscedasticity. Further, we introduce a variational Bayesian procedure to optimize the model. The performance of the proposed model is tested on F10.7 historical data. The results show that the proposed model improves the performance of multistep F10.7 forecasting by considering correlation and heteroscedasticity.

Plain Language Summary

The solar 10.7-cm radio flux (F10.7) is an important measure of solar radio emission activity. We introduce a model that considers the task correlation between different forecasting steps and heteroscedasticity, which are not considered before. The results on F10.7 historical data show that the proposed model has better performance.

1. Introduction

Solar radio flux at 10.7 cm (F10.7) is an important measure of solar activity, not only because it has a long history of records but also because it tracks other important emissions that form in the chromosphere and corona of the Sun (Tapping, 2013). The F10.7 index correlates with extreme ultraviolet and ultraviolet emissions. F10.7 forecasting is important for space weather and global environment forecasts. Accurate F10.7 forecasts will help to protect space satellites and electric power transmissions from the impact of solar radiation. In the National Space Science Center of the Chinese Academy of Sciences, the mid-term prediction of solar F10.7 is an important forecasting task; the prediction results can be used as input parameters for middle and upper atmosphere forecasting models.

A number of models have been designed to forecast the F10.7 flux. Liu et al. (2010) introduced a 54-order autoregressive model, which is used to forecast 27-day-ahead F10.7. Based on this autoregressive model, Wen (2010) introduced the observation and general evolution of active solar regions to improve the accuracy of the autoregressive model. Warren et al. (2017) introduced a linear regression (LR) model for the F10.7 multistep-ahead forecasting task.

Henney et al. (2012) forecasted F10.7 by using the earthside solar magnetic field distribution. They computed magnetic synoptic maps and then used a new method that provides a realistic estimation of the earthside solar magnetic field distribution to forecast future F10.7. American SOLAR2000 is an empirical solar irradiance model that includes an F10.7 linear midterm prediction model (see Tobiska, 2003; Tobiska, 2003).

In recent years, some researchers have turned to machine-learning techniques. For instance, Chatterjee (2001) introduced a multilayer feedforward neural network to forecast F10.7. Zhao and Han (2008) found the relationship between the maximum and the LR slope and proposed a method to predict the maximum F10.7 by means of the slope-maximum relationship. Huang et al. (2009) applied support vector regression, which maps the sample space from a nonlinearly separable space to a high- or infinite-dimensional linearly separable space (Hilbert space), making the separation easier in that space. For more details see Smola and Schölkopf (2004) to forecast F10.7, resulting in a good forecasting performance. Xiao et al. (2017) introduced the back
propagation (BP) neural network to forecast the F10.7. The BP neural network is a multilayer feedforward neural network, which is trained by error BP algorithm. It has the ability to handle large amount of data and the complex nonlinear relationships in F10.7 data. So it has good performance in F10.7 short-term forecasting.

The conclusion in Liu et al. (2010) implies that F10.7 data are heteroscedastic (for the regression model, the variance of random error is no longer constant for different sample points, so it is heteroscedastic). However, they did not verify the heteroscedasticity of F10.7 data, and their autoregressive model did not consider the property. Moreover, there is a strong task correlation (the task correlation is the correlation coefficient between different forecasting steps) between different forecasting tasks. Warren et al. (2017) only applied the LR model to the F10.7 forecasting task, but they did not consider the heteroscedasticity and the correlation mentioned above.

In this paper, we propose a linear multistep model based on the characteristic of heteroscedasticity and the task correlation between different steps to solve the above shortcomings in the LR model. First, we assume that the task correlation between different steps can be expressed by the covariance of the forecasting tasks. The task correlation is shown in a regression coefficient matrix. Second, we consider heteroscedasticity. A multivariate Gaussian prior with the step covariance of different magnifications was placed on different training samples to simulate heteroscedasticity. Finally, we constructed a fully conjugate Bayesian formulation and derive a variational Bayesian (VB) optimization procedure to solve the proposed model.

The remainder of this paper is organized as follows. Section 2 analyzes the characteristics of F10.7 historical data. Section 3 presents the proposed model and corresponding optimization procedure. Section 4 describes the data set and criteria used to measure the performance of the proposed model against other models and discusses the results of this evaluation. Finally, the conclusions and future work are given in section 5.

### 2. Data Analysis

Records for F10.7 go back to 1947 (Tapping, 2013). Researchers found that F10.7 records show obvious periodic oscillations approximately every 11 years. These oscillations are related to nuclear reactions in the interior of the Sun. It was also discovered that F10.7 historical data have the characteristic of short quasi-27-day activity period oscillations due to the rotation of the Sun. Figure 1 shows the corresponding periods of F10.7 (Ma, 2007). The period is short (approximately 25 days) near the Earth’s equator, whereas the period is long (approximately 35 days) in the polar regions, and the average period is around 27 days (see Bouwer, 1990; Kane, 2003).

To forecast F10.7 based on these oscillations, Liu et al. (2010) found that forecasting errors are relatively larger during certain periods, whereas in other periods they will be smaller—similar to the characteristic of heteroscedasticity. Figure 2 shows the distribution of a yearly mean relative error of F10.7 medium-term forecast in solar cycle 23 in Liu et al. (2010). We can see that the yearly mean relative error is larger with more solar activity (around 2001) than with less solar activity (around 1996).

In addition, Engle (1982) introduced an autoregressive conditional heteroscedastic (ARCH) model. The ARCH test proceeds as follows: First, we use ordinary least squares regression to compute the residual series $r_t$:

$$ y_t = \hat{y}_t $$

where $y_t$ is the observed value and $\hat{y}_t$ is the fitted value. Second, we build the ARCH(L) model based on $r_t$:

$$ r_t^2 = a_0 + a_1 r_{t-1}^2 + \ldots + a_L r_{t-L}^2 + u_t $$

where $u_t$ is white noise (white noise is a random signal having equal intensity at different frequencies, giving it a constant power spectral density). Third, we assume the null hypothesis on regression coefficient

---

**Figure 1.** Corresponding periods of F10.7 (Ma, 2007).

**Figure 2.** Distribution of yearly MRE of F10.7 midterm forecast in solar cycle 23 (Liu et al., 2010). MRE = mean relative error.
Table 1
Regression Coefficient in ARCH (10)

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>$a_0$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$a_6$</th>
<th>$a_7$</th>
<th>$a_8$</th>
<th>$a_9$</th>
<th>$a_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>−16.572</td>
<td>0.417</td>
<td>−0.077</td>
<td>0.086</td>
<td>−0.023</td>
<td>0.069</td>
<td>−0.021</td>
<td>0.009</td>
<td>0.008</td>
<td>−0.002</td>
<td>0.045</td>
</tr>
</tbody>
</table>

Note. ARCH = autoregressive conditional heteroscedastic.

$a_0 = a_2 = ...a_L = 0$. If $a_j (j = 0, ..., L)$ is under the null hypothesis; there are no ARCH effects on residuals. If there is at least one $a_j \neq 0 (j = 0, ..., L)$, the residuals are heteroscedastic.

We set $L = 10$ in this test and computed the regression coefficient $a_j (j = 0, ..., L)$ with equation (2). The regression coefficients for the ARCH (10) test on F10.7 are shown in Table 1. All regression coefficients $a_j (j = 0, ..., 10)$ are nonzero. Therefore, the F10.7 data are heteroscedastic. In this work, then, we consider the characteristic of heteroscedasticity.

At this stage, to the best of our knowledge, no researcher has considered the task correlation among different steps. In order to verify whether F10.7 forecasting is correlative between different steps, we computed the correlation coefficients between different steps. The results are presented in Figure 3. The correlation coefficients between target steps with six adjacent steps are more than 0.5. The statistics thus show a strong correlation between the different steps. Consequently, we applied task correlation to our proposed model.

3. Proposed Model

We propose a linear multistep-ahead forecasting model (CH-MF) to forecast F10.7 values. The model considers the task correlation between different steps and the characteristic of heteroscedasticity.

3.1. Prior Distributions of the Proposed Model

Usually, matrix variate distributions are effective tools to represent random matrices, in which the matrix variate Gaussian distribution plays a pivotal role as the multivariate Gaussian distribution in the family of multivariate distributions (Saito et al., 2014). If a random matrix $X \in \mathbb{R}^{m \times n}$ follows a matrix variate Gaussian distribution, the probability density function can be expressed as follows:

$$p(X | M, S, \Omega) = ((2\pi)^{-\frac{mn}{2}} | S |^{-\frac{1}{2}} | \Omega |^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(S^{-1}(X - M)\Omega^{-1}(X - M)^T) \right\}$$  \hspace{1cm} (3)

where $M \in \mathbb{R}^{m \times n}$ is the mean of the random matrix $X$, $S \in \mathbb{R}^{n \times n}$ is the rows covariance matrix, which describes the correlations of rows, $\Omega \in \mathbb{R}^{m \times m}$ is the columns covariance matrix, which describes the correlations of columns, and $\text{tr}(\cdot)$ denotes the trace of a matrix. To simplify the description, we use the following notation:

$$X \sim \mathcal{M}\mathcal{N}(X | M, S, \Omega)$$  \hspace{1cm} (4)

In our F10.7 multistep forecasting regression model, given the input $X = (X^T_1, X^T_2, ..., X^T_H) \in \mathbb{R}^{H \times D}$, the number of forecasting steps is $H$, the target matrix is represented by $Y = (Y_1, Y_2, ..., Y_H) \in \mathbb{R}^{H \times 1}$, and $E = (e_1, e_2, ..., e_H)^T \in \mathbb{R}^{H \times 1}$ is the corresponding error matrix. Thus, the multistep regression model can be formulated as

$$Y = W^T X + E$$  \hspace{1cm} (5)

where $W \in \mathbb{R}^{D \times H}$ represents the coefficient matrix.

According to equation (5), we suppose that the regression coefficient matrix $W$ and the training error matrix $E$ are random matrix variate distributions. Thus, we can embed the correlation of multistep-ahead forecasting tasks into the forecasting model. For the random matrix $E$, we set a matrix variate Gaussian prior parameterized with a mean of $0_{N \times H}$, rows covariance $\Xi \in \mathbb{R}^{N \times N}$, and columns covariance $\Lambda \in \mathbb{R}^{N \times N}$. This can be expressed as

$$E \sim \mathcal{M}\mathcal{N}(E | 0_{N \times H}, \Xi, \Lambda)$$  \hspace{1cm} (6)

Figure 3. Correlation coefficients between different steps.
Usually, we assume that samples in the F10.7 training data are independent. As such, the corresponding forecasting error in each sample is also independent. Thus, we have $\Xi = I_N$. According to the property of a matrix variate Gaussian distribution, the matrix variate Gaussian prior of $E$ degenerates into a multivariate Gaussian distribution parameterized with mean $0_{1 \times H}$ and covariance $\Lambda$. That is,

$$e_i \sim \mathcal{N}(e_i | 0_{1 \times H}, \Lambda) \quad (7)$$

From the above equation, we can see that the forecasting error is an independent and identically distributed random variable that obeys a homoscedastic multivariate Gaussian distribution. In order to deal with the heteroscedastic characteristic in F10.7 forecasting, we assume that the forecasting error of the $i$th sample obeys a Gaussian distribution with mean $0_{1 \times H} \in \mathbb{R}$ and covariance $\sigma_i \Lambda$. Then, the final prior on the error term can be expressed as

$$e_i \sim \mathcal{N}(e_i | 0_{1 \times H}, \sigma_i \Lambda) \quad (8)$$

As for $\sigma$, we assume it obeys an inverse Gamma distribution with shape parameter $\alpha_0$ and scale parameter $\beta_0$:

$$\sigma \sim \mathcal{IG}(\sigma | \alpha_0, \beta_0) \quad (9)$$

Given the target matrix $Y$ and the input matrix $X = [X_1^T, X_2^T, \ldots, X_n^T]$, the likelihood function is expressed as

$$p(Y | X, W, \sigma, \Lambda) = \prod_{i=1}^N N(Y_i | WX_i, \sigma_i \Lambda) \quad (10)$$

where $\Lambda$ is the task correlation matrix of different steps, and $\sigma$ is the heteroscedasticity of errors.

In order to emphasize the task correlation among different forecasting steps, we also assume that the regression coefficient matrix $W$ follows the matrix variate Gaussian prior with mean $V \Lambda$, rows covariance $\gamma \Lambda$, and columns covariance $\tau \Omega$:

$$W \sim \mathcal{MN}(W | V \Lambda, \gamma \Lambda, \tau \Omega) \quad (11)$$

where $V \in \mathbb{R}^{N+1 \times M}$ obey matrix variate Gaussian prior with mean $Z_0$, rows covariance $\tau \Omega$, and columns covariance $I_M$:

$$Z \sim \mathcal{MN}(Z | Z_0, \tau \Omega, I_M) \quad (12)$$

And $Z \in \mathbb{R}^{M \times H}$ also obey matrix variate Gaussian prior with mean $V_0$, rows covariance $I_M$, and columns covariance $\gamma \Lambda$:

$$V \sim \mathcal{MN}(V | V_0, I_M, \gamma \Lambda) \quad (13)$$
Table 2

<table>
<thead>
<tr>
<th>Description</th>
<th>All samples</th>
<th>Training samples</th>
<th>Test samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers</td>
<td>19,569</td>
<td>12,251</td>
<td>7,318</td>
</tr>
</tbody>
</table>

When $M$ is a positive integer, we set $M$ such that it is smaller than $H$. We also assume that the matrix variate distributions $\gamma$ and $\tau$ obey a Gamma distribution with shape parameter $\alpha_1$ ($\alpha_2$) and scale parameter $\phi_1$ ($\phi_2$):

$$
\gamma \sim \mathcal{G}(\gamma | \alpha_1, \phi_1) \\
\tau \sim \mathcal{G}(\tau | \alpha_2, \phi_2)
$$

Finally, we assume the matrix variate $\Omega \in \mathbb{R}^{(N+1) \times (N+1)}$ obeys an inverse Wishart distribution:

$$
\Omega \sim \mathcal{W}(\Omega | \Psi_0, v_0)
$$

where $\Psi_0 \in \mathbb{R}^{(N+1) \times (N+1)}$ and $v_0$ are constants.

Figure 4 shows the graphic models for the Bayesian formulation of the proposed model. We used the proposed model to infer the posterior of all parameters in the following joint probability density function:

$$
p(Y, W, V, Z, \Omega, \gamma, \tau, \sigma) = \prod_{i=1}^{N} p(Y) p(W) p(V) p(Z) p(\Omega) p(\gamma) p(\tau) p(\sigma)
$$

### 3.2. Optimization

Given the observed data $Q$ and the parameter set $\Theta$, which contains $n$ independent model parameters, $\Theta_1, \Theta_2, \ldots, \Theta_n$, the posterior of $\Theta$ (Castellaro et al., 2017) can be expressed as

$$
p(\Theta | Q) = \frac{p(Q | \Theta) p(\Theta)}{p(Q)}
$$

Based on the theory of Wan and Liang (2016), we applied a VB procedure to solve the above problem by introducing a new distribution, $q(\Theta)$, to approximate the posterior $p(\Theta | Q)$. To simplify the computation based on the framework of VB, we applied mean field theory, which assumes that all parameters in $\Theta$ are independent. Therefore, the approximated posterior $q(\Theta)$ can be factorized with some disjoint groups:

$$
q(\Theta) = \prod_{i=1}^{n} q(\Theta_i)
$$

Thus, all the posteriors of all involved variables expressed in equation (17) can be inferred by VB. The approximation for all the posterior distributions is processed in factorized form as follows:

$$
q(W, V, Z, \Omega, \gamma, \tau, \sigma) = q(W) q(V) q(Z) q(\Omega) q(\gamma) q(\tau) q(\sigma)
$$

where $q(x)$ denotes the posterior of variable $x$. According to equation (7), the posterior distribution of each variable can be obtained.

### Table 3

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Calculation</th>
<th>Evaluation criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>$\frac{1}{n} \sum_{i=1}^{n}</td>
<td>y_i - \hat{y}_i</td>
</tr>
<tr>
<td>RMSE</td>
<td>$\sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} $</td>
<td>The smaller the better</td>
</tr>
<tr>
<td>$R$</td>
<td>$\frac{\sum_{i=1}^{n} y_i \hat{y}<em>i}{\sqrt{\sum</em>{i=1}^{n} y_i^2 \sqrt{\sum_{i=1}^{n} \hat{y}_i^2}}}$</td>
<td>The larger the better</td>
</tr>
</tbody>
</table>

Note. MAE = mean absolute error; RMSE = root-mean-square error.
The posterior of $W$ obeys a matrix variate Gaussian distribution with mean of $E_W$, rows covariance $S_w$, and column covariance $\Omega_w$:

$$q(W) = MN_{(N+1)\times H}(W|E_w, S_w, \Omega_w)$$

(21)

From the above equation, the posterior of $V$ obeys a matrix variate Gaussian distribution with mean $E_v$, rows covariance $S_v$, and column covariance $\Omega_v$:

$$q(V) = MN_{(N+1)\times M}(V|E_v, S_v, \Omega_v)$$

(22)

The posterior of $Z$ obeys a matrix variate Gaussian distribution with mean $E_z$, rows covariance $S_z$, and column covariance $\Omega_z$:

$$q(Z) = MN_{M\times H}(Z|E_z, S_z, \Omega_z)$$

(23)

The posterior of $\Omega$ obeys an inverse Wishert distribution parameterized by the scale matrix $\Psi_{\text{new}}$ and the degree of freedom $v_{\text{new}}$:

$$q(\Omega) = IW(\Omega|\Psi_{\text{new}}, v_{\text{new}})$$

(24)

Therefore, based on Figure 4, the posterior of $\gamma$ is a Gamma distribution with shape parameter $\omega_{\text{new}}$ and scale parameter $\phi_{\text{new}}$:

$$q(\gamma) = G(\gamma|\omega_{\text{new}}, \phi_{\text{new}})$$

(25)
Figure 8. Comparison of errors between LR model and H-MF model by MAE (a) and RMSE (b). LR = linear regression; MAE = mean absolute error; RMSE = root-mean-square error.

We assume that the posterior distribution of $\tau$ obeys a Gamma distribution with shape parameter $\omega_{2}^{\text{new}}$ and scale parameter $\phi_{2}^{\text{new}}$:

$$q(\tau) = \mathcal{G}\left(\tau | \omega_{2}^{\text{new}}, \phi_{2}^{\text{new}}\right)$$

(26)

We also assume that $\sigma_{i}$ and $\sigma_{i}^{j}$ are independent and that the posterior of $\sigma$ is also an inverse Gamma distribution with shape parameter $\alpha_{i}^{\text{new}}$ and scale parameter $\beta_{i}^{\text{new}}$:

$$q(\sigma_{i}) = \mathcal{I}\mathcal{G}\left(\sigma_{i} | \alpha_{i}^{\text{new}}, \beta_{i}^{\text{new}}\right)$$

(27)

3.3. Correlation Estimation of Forecasting Task

For simplicity, $V_{0}$ and $Z_{0}$ are set as null matrices. According to the prior on $W$ and $Z$, the marginal distribution $p(W^{T}|Z^{T}, \Omega, \Lambda, \gamma, \tau)$ can be derived as

$$p(W^{T}|Z^{T}, \Omega, \Lambda, \gamma, \tau) = \mathcal{MN}_{N(n+1)}(W^{T}|0_{(n+1)\times n}, \tau^{-1} \Lambda + Z^{T} Z, \Omega_{\gamma}^{-1})$$

(28)

From the above equation, the task correlation $\Delta$ can be estimated by equation (13):

$$\Delta \approx \left(\tau^{-1}\right) \Lambda + \left< Z^{T} Z \right> = \left(\tau^{-1}\right) \Lambda + \Omega_{2} \text{tr}(S_{2}) + E_{Z}^{T} E_{Z}$$

(29)

In this paper, insofar as we show the strong correlation between F10.7 forecasting steps, the task correlation should be considered when developing a new multistep-ahead forecasting model. Thus, we next set $\Lambda = \text{cov}(Y)$ for our proposed model.

Figure 9. Comparison of errors between LR model, BP net, and CH-MF model by MAE (a) and RMSE (b). LR = linear regression; BP = back propagation; MAE = mean absolute error; RMSE = root-mean-square error.
4. Experimental Analysis

4.1. Data Sets and Evaluation Metrics

In this experiment, we selected F10.7 historical data from 1963 to 2016 gathered by the National Oceanic and Atmospheric Administration (NOAA; ftp://ftp.ngdc.noaa.gov/STP/GEOMAGNETIC_DATA/INDICES/KP_AP/), comprising five complete solar periods. And they are adjusted values (adjusted to 1 AU). The samples in the first three periods were used to train the models, and the rest were used to test them. The plots of the F10.7 data are shown in Figure 5, and the description of the dataset is shown in Table 2. The data are all recorded in 1-day span with 19,569 samples, in which 12,251 samples (from 3 January 1963 to 18 July 1996) are used to train models and 7,318 samples (from 19 July 1996 to 31 July 2016) are employed to test the forecasting performance of all models.

Three regression evaluation metrics were used to compare different regression models: the mean absolute error (MAE), the root-mean-square error (RMSE), and the correlation coefficient $R$. Table 3 provides definitions of the three evaluation metrics. MAE and RMSE were used to measure the deviation between the predicted and observed values, and $R$ was used to reveal the correlation between the predicted and observed values.

4.2. Experiments and Results

In experiments, all models use the previous 54 days of observed value to produce daily forecast from 1 to 27 days. In our experiment, we exclusively focused on multi-output models for F10.7 medium-term
Thus, we selected different multi-output regression models to compare their performance to that of the proposed model, including the multi-output linear regression model (LR model in Warren et al., 2017) and the BP neural network.

Since our CH-MF model considers both heteroscedasticity and task correlation, there are two variants in our proposed model. In the first variant, we assume that F10.7 forecasting errors are homoscedastic, and we just consider task correlations, denoted by C-MF. In the second variant, we assume that each step is independent of the others, and we just consider heteroscedasticity, denoted by H-MF. The goal of the two variants is to uncover the effect of the two factors.

4.2.1. Results of the Proposed CH-MF Model
Figure 6 shows the distribution of yearly MAEs for F10.7 medium-term forecasting during solar cycles 23 and 24. We can see that the MAE changes with the solar activity. The MAE is highest during the period of high solar activity (i.e., in 2001), where the yearly MAE is approximately 26. The MAE is lowest during the period of low solar activity (i.e., in 2008), where the yearly MAE is approximately 2. The above analysis further confirms the characteristic of heteroscedasticity in F10.7 data. Thus, we can conclude that F10.7 data are heteroscedastic and that the forecasting accuracy is related to solar activity.

4.2.2. Advantage of Considering Correlation and Heteroscedasticity
Because we considered both the characteristic of heteroscedasticity and task correlation in the proposed model, in this section, we discuss the effect of the two properties for F10.7 medium-term forecasting, respectively. A detailed comparison of LR, C-MF, and H-MF is provided as follows:

First, we compared the C-MF model with LR model. The forecasting errors of both models are shown in Figure 7. On the whole, the C-MF model performed better than the LR model. In terms of the MAE, the reduction of errors in the C-MF model compared to the LR model varied from 0.0238 to 0.7118. In terms of RMSE, the C-MF model outperformed the LR model by a ratio ranging from 0.0185 to 0.3622. These results confirm that considering task correlation improves the accuracy of F10.7 forecasting.

Second, we compare the H-MF model with LR model. Figure 8 shows the forecasting errors between H-MF model and the LR model. On the whole, the C-MF model performed better than the LR model, except the first step. In terms of MAE, the increase of H-MF over the LR model is located in the range [0.0602, 0.6442]. In terms of RMSE, the increase of H-MF is in the intervals [0.0353, 0.3299]. The above analysis shows that heteroscedasticity plays a pivotal role in F10.7 forecasting.

From the above comparison, we can conclude that it is advantageous to consider heteroscedasticity and task correlation for F10.7 forecasting. In the next section, we compare the proposed model with the LR model and BP neural network. Further, we discuss the performance of the proposed model when considering both the heteroscedasticity and task correlation.

4.2.3. Comparative Analysis of Forecasting Errors
In this section, we compare the performance of our proposed CH-MF with the LR model and BP neural network model. All forecasting errors are shown in Figure 9, and the correlation coefficient $R$ is presented in Table 4. Specifically, in terms of the MAE, the increase in accuracy is from 0.0248 to 0.7069 relative to the LR model and from 0.0574 to 0.9805 relative to the BP neural network.

In terms of the RMSE, the increase in accuracy of the CH-MF relative to the LR model is located in [0.0187, 0.3655]. The increase in accuracy of the CH-MF relative to the BP neural network varied from 0.0925 to 0.4615. In terms of correlation coefficient $R$, the level of their performance is comparatively high. Nevertheless, the
proposed CH-MF model obtained better correlation between predicted values and observed values than the LR model and BP neural network.

Figure 9 shows a comparison of the LR model, BP neural network, and CH-MF model. We found that the CH-MF model incurred fewer errors than LR model and BP neural network. From the first step to the ninth step, forecasting errors grew linearly with the increase in the number of forecasting steps. From the 10th to 27th steps, by contrast, forecasting errors tended to stabilize. Thus, it becomes increasingly difficult to forecast future F10.7 values with increased forecasting steps.

We compared the forecasting values with the observed values during the period of high solar activity in 2001 and the period of low solar activity in 2008 with different forecasting steps to further analyze the effective-
ness of our CH-MF model for F10.7 forecasting task, as shown in Figures 10 and 11. In general, there were more errors during high solar activity than that during low solar activity in the same forecast step, and the proposed model could better fit the observed curve during minimum solar activity. With an increase in forecasting steps, the forecasting values usually lagged longer. There were more forecasting errors in such cases, and both models could not fit the curve well with abrupt changes to F10.7.

During maximum solar activity (i.e., in 2001), there was considerable variation to the amplitude of F10.7 due to solar activity, and the F10.7 values ranged from approximately 120 to 280. Nevertheless, the MAE of the CH-MF model was smaller than that of the LR model. From the first to the ninth steps, the CH-MF model fit the observed curve well. From the ninth step onward, the CH-MF model generally fit the observed curve. During minimum solar activity (i.e., in 2008), all F10.7 values were lower, and the variation in F10.7 amplitude was small. Thus, both the CH-MF model and LR model effectively fit the observed curve.

### 4.2.4. Comparative Analysis of Predictive Accuracy

In this section, we further conduct paired t-test to test the difference between CH-MF models and the compared models at the 95% confidence intervals. With respect to the t-test, the null hypothesis states zero median error differentials, that is \((e_i^c - e_i^f)^2 = 0\), where \(e_i^c\) and \(e_i^f\) are the observed value, and \(y_i^c\) and \(y_i^f\) are the forecasting value of different models. The p-value of t-test is shown in Table 5. The t-test result between CH-MF and BP neural network demonstrates acceptance of the null hypothesis, which indicates that the forecasting performances of these two models can be considered to be the same, while the testing results between the CH-MF, LR, C-MF, and H-MF generally exhibit rejection of the null hypothesis, which means a large difference.

<table>
<thead>
<tr>
<th>Day</th>
<th>CH-MF and LR</th>
<th>CH-MF and BP</th>
<th>CH-MF and C-MF</th>
<th>CH-MF and H-MF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0237 (Reject)</td>
<td>0.0349 (Reject)</td>
<td>0.0638 (Accept)</td>
<td>0.0003 (Reject)</td>
</tr>
<tr>
<td>2</td>
<td>0.0853 (Accept)</td>
<td>0.4486 (Accept)</td>
<td>0.0135 (Reject)</td>
<td>0.9386 (Accept)</td>
</tr>
<tr>
<td>3</td>
<td>0.0781 (Accept)</td>
<td>0.8907 (Accept)</td>
<td>0.0541 (Accept)</td>
<td>0.0447 (Reject)</td>
</tr>
<tr>
<td>4</td>
<td>0.0949 (Accept)</td>
<td>0.1165 (Accept)</td>
<td>0.0241 (Reject)</td>
<td>0.1182 (Accept)</td>
</tr>
<tr>
<td>5</td>
<td>0.0715 (Accept)</td>
<td>0.3693 (Accept)</td>
<td>0.0407 (Reject)</td>
<td>0.5777 (Accept)</td>
</tr>
<tr>
<td>6</td>
<td>0.0399 (Reject)</td>
<td>0.0560 (Accept)</td>
<td>0.0110 (Reject)</td>
<td>0.3408 (Accept)</td>
</tr>
<tr>
<td>7</td>
<td>0.0205 (Reject)</td>
<td>0.0881 (Accept)</td>
<td>0.0202 (Reject)</td>
<td>0.0004 (Reject)</td>
</tr>
<tr>
<td>8</td>
<td>0.0095 (Reject)</td>
<td>0.1956 (Accept)</td>
<td>0.0129 (Reject)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>9</td>
<td>0.0043 (Reject)</td>
<td>0.0145 (Reject)</td>
<td>0.0878 (Accept)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>10</td>
<td>0.0023 (Reject)</td>
<td>0.6443 (Accept)</td>
<td>0.5397 (Accept)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>11</td>
<td>0.0011 (Reject)</td>
<td>0.3428 (Accept)</td>
<td>0.6522 (Accept)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>12</td>
<td>0.0006 (Reject)</td>
<td>0.0125 (Reject)</td>
<td>0.0766 (Accept)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>13</td>
<td>0.0004 (Reject)</td>
<td>0.8108 (Accept)</td>
<td>0.0058 (Reject)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>14</td>
<td>0.0003 (Reject)</td>
<td>0.0919 (Accept)</td>
<td>0.0029 (Reject)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>15</td>
<td>0.0002 (Reject)</td>
<td>0.3966 (Accept)</td>
<td>0.0044 (Reject)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>16</td>
<td>0.0001 (Reject)</td>
<td>0.8969 (Accept)</td>
<td>0.0208 (Reject)</td>
<td>0.0000 (Reject)</td>
</tr>
<tr>
<td>17</td>
<td>0.0001 (Reject)</td>
<td>0.8397 (Accept)</td>
<td>0.4074 (Accept)</td>
<td>0.0008 (Reject)</td>
</tr>
<tr>
<td>18</td>
<td>0.0003 (Reject)</td>
<td>0.0235 (Reject)</td>
<td>0.4693 (Accept)</td>
<td>0.0553 (Accept)</td>
</tr>
<tr>
<td>19</td>
<td>0.0005 (Reject)</td>
<td>0.0007 (Reject)</td>
<td>0.0243 (Reject)</td>
<td>0.7029 (Accept)</td>
</tr>
<tr>
<td>20</td>
<td>0.0014 (Reject)</td>
<td>0.0008 (Reject)</td>
<td>0.0003 (Reject)</td>
<td>0.2095 (Accept)</td>
</tr>
<tr>
<td>21</td>
<td>0.0035 (Reject)</td>
<td>0.0038 (Reject)</td>
<td>0.0000 (Reject)</td>
<td>0.0078 (Reject)</td>
</tr>
<tr>
<td>22</td>
<td>0.0094 (Reject)</td>
<td>0.1560 (Accept)</td>
<td>0.0000 (Reject)</td>
<td>0.0033 (Reject)</td>
</tr>
<tr>
<td>23</td>
<td>0.0195 (Reject)</td>
<td>0.0009 (Reject)</td>
<td>0.0000 (Reject)</td>
<td>0.0008 (Reject)</td>
</tr>
<tr>
<td>24</td>
<td>0.0427 (Reject)</td>
<td>0.1779 (Accept)</td>
<td>0.0000 (Reject)</td>
<td>0.0080 (Reject)</td>
</tr>
<tr>
<td>25</td>
<td>0.0761 (Accept)</td>
<td>0.3754 (Accept)</td>
<td>0.0000 (Reject)</td>
<td>0.0759 (Accept)</td>
</tr>
<tr>
<td>26</td>
<td>0.1314 (Accept)</td>
<td>0.1602 (Accept)</td>
<td>0.0000 (Reject)</td>
<td>0.3580 (Accept)</td>
</tr>
<tr>
<td>27</td>
<td>0.1653 (Accept)</td>
<td>0.7592 (Accept)</td>
<td>0.0001 (Reject)</td>
<td>0.8610 (Accept)</td>
</tr>
</tbody>
</table>
between CH-MF model and the compared models. Combined with the results of the model comparison, we can draw the conclusion that the CH-MF is an effective and reasonable model.

In summary, insofar as it considers heteroscedasticity and task correlation, our model outperformed other compared models, such as LR model and BP neural network. That is to say that it is reasonable to consider the characteristic of heteroscedasticity and task correlation among different steps in F10.7 medium-term forecasting task.

5. Conclusion and Future Work

For medium-term F10.7 forecasting, researchers seldom consider the task correlation among different steps. Moreover, we know that F10.7 forecasting errors are heteroscedastic (Liu et al., 2010). Inspired by these two points, we proposed a novel model that considers task correlation and heteroscedasticity in a linear multistep regression model (CH-MF). The results of our evaluation show that CH-MF model is more effective and reliable than LR model and BP neural network for F10.7 multistep forecasting task. From the comparison between LR model and C-MF model, LR model and H-MF model, we can see that both task correlation and the heteroscedastic assumption are useful to F medium-term forecasting task. Thus, it is reasonable to consider task correlation and to embed heteroscedasticity in F10.7 medium-term forecasting task.

As the record of F10.7 data traces back to 1947, a variety of samples can be used for this task. We can thus take full advantage of the data set and learn the laws of F10.7. Moreover, the relationship between F10.7 values may be nonlinear, and we might improve model performance by introducing deep learning to F10.7 forecasting. Thus, in future research, we will develop a deep learning model for F10.7 forecasting.

References


Tobiska, W. K. (2003). Forecasting of space environment parameters for satellite and ground system operations.


