Forecasting solar flare activity using an inferred solar stress index

Aaron Springford* David L Riegert† David J Thomson‡

Abstract

Solar activity can have appreciable damaging effects on Earth-based power transmission systems. Solar storms in particular cause large variations in Earth's magnetic field, inducing current in transmission lines and causing damage to sensitive power transformers and other equipment. Power companies are interested in predicting when such storms will occur in order to mitigate damages, but current forecast horizons are only on the order of days. We present a forecasting approach based on the hypothesis that subsurface stresses initiate solar flare events, which can result in coronal mass ejections and solar storms at Earth. Our approach uses multitaper spectral estimates of solar gravity modes to forecast a cumulative solar stress index using a very simple state-space model. In retrospective analysis, this stress index was found to be correlated with variance in log X-ray data measured near Earth by GOES spacecraft. Using this index, we generated a six month qualitative forecast of the likelihood of large flare events beginning November 1, 2013. We discuss the performance of our forecast and its utility for management of power transmission systems over an extended time horizon.

Key Words: multitaper; spectrum estimation; dynamic model; forecasting; solar

Key points

- Solar gravity mode estimates were obtained from ACE spacecraft measurements.
- Weekly solar flare activity was forecast using gravity mode estimates.
- The forecast horizon is limited to about six months, perhaps less.

1. Introduction

It is well known that solar activity, and in particular solar storms, can have an effect on the operations of many man-made systems such as electric power transmission, telecommunications, and spacecraft. The authorities that manage electric power transmission systems understandably have a keen interest in knowing when solar storms will occur to manage risks. Current forecasting approaches rely on studies of magnetic field patterns in and around active solar regions, (e.g. Georgoulis 2013; Gao et al. 2014; Reinard et al. 2010; González Hernández 2013) and focus on the timing and magnitude of solar flares emitted from a given active region. Because of the approach, such forecasts are limited to the lifetime of the solar active region, generally on the order of days to weeks.

This paper describes our attempt at predicting solar flare activity, as a precursor to damaging solar storms, over a time horizon of approximately six months. The longer time horizon requires a different approach to prediction of events. Our hypothesis is that solar gravity modes, which are believed to persist for at least thousands of years (Christensen-Dalsgaard et al. 1974), are indirect drivers of periods of high solar flare activity. The recent

*Queen's University at Kingston, Department of Mathematics and Statistics, Jeffery Hall, University Avenue, Kingston, Ontario, Canada, K7L 3N6. Email: aaron.springford@queensu.ca

†Queen's University at Kingston, Department of Mathematics and Statistics, Jeffery Hall, University Avenue, Kingston, Ontario, Canada, K7L 3N6. Email: driegert@mast.queensu.ca

‡Queen's University at Kingston, Department of Mathematics and Statistics, Jeffery Hall, University Avenue, Kingston, Ontario, Canada, K7L 3N6. Email: djt@mast.queensu.ca

finding of Ruzmaikin et al. (2011) that fast coronal mass ejections, which are often associated with solar flare activity, are clustered in time supports the idea that flare events are forced periodically. In addition, DJ Thomson observed that solar gravity mode reconstructions using Ulysses spacecraft data were associated with the Bastille Day event of July 14, 2000 (unpublished data).

Solar flares are caused by magnetic reconnection in active regions. Active regions are subject to forcing by subsurface mixing processes (Komm et al. 2013), which in turn are modulated by solar oscillations, in particular gravity modes. Solar gravity modes persist for much longer than our forecast horizon, potentially providing a reliable forecasting mechanism.

An outline of our approach is:

1. Estimate solar gravity modes (section 2.1)

2. Forecast a solar gravity mode time series (section 2.1)

3. Using the solar gravity mode series as input, forecast solar flare activity (section 2.2)

Solar gravity mode estimation (1.) is in itself a challenging proposition. Gravity mode detections have been made previously using spacecraft data (Thomson et al. 1995), but this approach is not without controversy (Appourchaux et al. 2010). We find that using magnetic field data from the ACE spacecraft and carefully considered spectrum estimation techniques, we are able to resolve a large amount of structure in the solar gravity mode band. However, factors such as frequency drift, the orbital geometry of the spacecraft, and non-stationary intermediate processes present a challenge and limit the forecast horizon. Thankfully, generating a forecast time series from gravity mode estimates (2.) is a simple operation.

Once the gravity mode time series has been forecast, we hypothesize that it is the total power (variance) in the series that is of importance for solar flare activity. Using solar gravity mode variance as our measure of solar stress, we consider two possible mechanisms for the relationship between stress and flare activity (3.): a direct (linear) relationship; and a simple cumulative stress/release mechanism. We find that the cumulative stress/release mechanism has statistically significant predictive power for near-Earth X-ray variation in retrospective analysis (section 2.3). Based on this finding, we generated a qualitative flare activity forecast for the period from 1 November 2013 to 10 May 2014 which predicted flare activity, albeit with declining accuracy over the forecast period.

2. Methods

2.1 Estimation and forecasting of solar gravity modes

We estimated the frequencies, phases, and amplitudes of solar gravity modes (g-modes) using magnetic field data collected by the ACE spacecraft. Because g-modes are spherical harmonics, they should appear as sinusoidal in the absence of interfering processes. We found that there was a large amount of structure in the g-mode band, consistent with previous studies (Thomson et al. 1995, 2001). We assume that the observed structure is indicative of the underlying solar g-mode process – in particular, we assume that signals in the g-mode band with larger estimated amplitudes contribute more to solar stress and subsequent flare events than signals with smaller estimated amplitudes. We also assume that signals which appear periodic are the result of g-modes. Thus, when determining the g-mode *signal*, we chose those frequencies in the estimated spectrum that were both periodic and had large associated amplitudes.

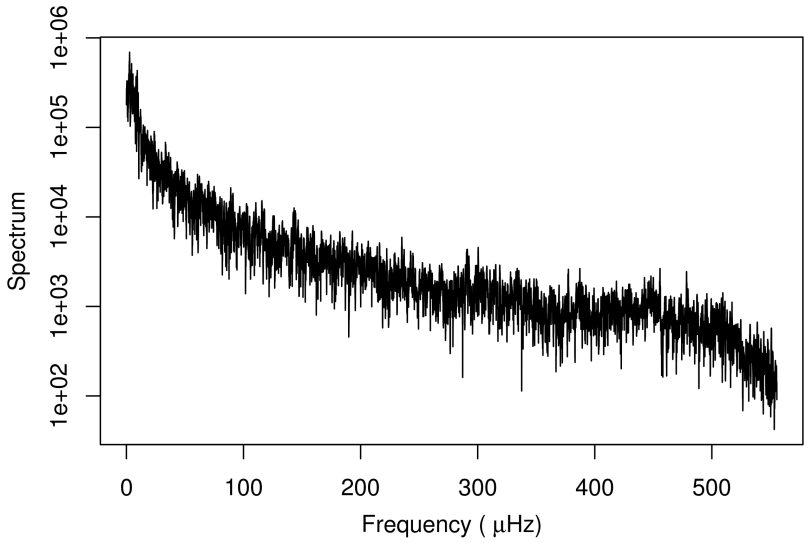

Figure 1: Spectrum of ACE magnetic field data (normal component) calculated using the multitaper method with time bandwidth parameter $NW = 4$ and $K = 7$ tapers over the time interval from January 1, 2013 to July 9, 2013.

We began by estimating the spectrum of the ACE magnetic field (normal component). Exploratory analysis suggested that an estimation period of longer than approximately six months resulted in degraded gravity mode forecasts in retrospective analysis. Thus, we chose an estimation period of 189 days (7 solar rotation cycles) when estimating spectra. Spectra were estimated using the multitaper spectrum estimator (Thomson 1982).

Returning to our original hypothesis that g-modes are drivers of solar flare events, we focus on those spectrum frequencies that are estimated to be both sinusoidal and to have significant amplitude. To establish a given signal as sinusoidal, we used the multitaper F-test (Thomson 1982). To establish the significance of a given signal amplitude, we used χ^2 quantiles of the estimated variance of the spectrum. Imagining the spectrum as a mixture of a white noise process and line components (periodic signals at a single frequency), a line component would manifest as an unusually large estimated variance.

In order to satisfy the mixture assumption, the ACE magnetic field spectra required some preliminary adjustment due to an obvious red noise background (Figure 1). Assuming that a given ACE magnetic field spectrum is a mixture of a noise process and a large number of signal processes, our goal is to adjust the noise process so that it is approximately white while leaving the signal processes intact. This is sometimes called a post-whitening approach (Thomson 2000).

Given a multitaper spectrum estimate $\hat{S}(f)$ using K tapers, we partition the set of frequencies at which the spectrum estimates were obtained into bands $[f_b, f_{b+1})$ where

$$f_{b+1} - f_b = 10\,\mu\text{Hz}$$

for all b in the theoretical g-mode band. Let $f^* \in [f_b, f_{b+1})$ be a frequency in band b. We fit a quadratic trend to the log spectrum in the band and subtract this trend, obtaining

$$\hat{S}_{\text{res}}(f^*) = \hat{S}(f^*) - \exp\{\hat{\beta}_{0,b} + \hat{\beta}_{1,b} f^* + \hat{\beta}_{2,b} f^{*2}\},$$

where $\hat{\beta}_{0,b}$, $\hat{\beta}_{1,b}$, and $\hat{\beta}_{2,b}$ are obtained by a least-squares fit over all $f^* \in [f_b, f_{b+1})$. We now assume that the resulting reshaped within-band spectrum $\hat{S}_{\text{res}}$ is composed of a mixture

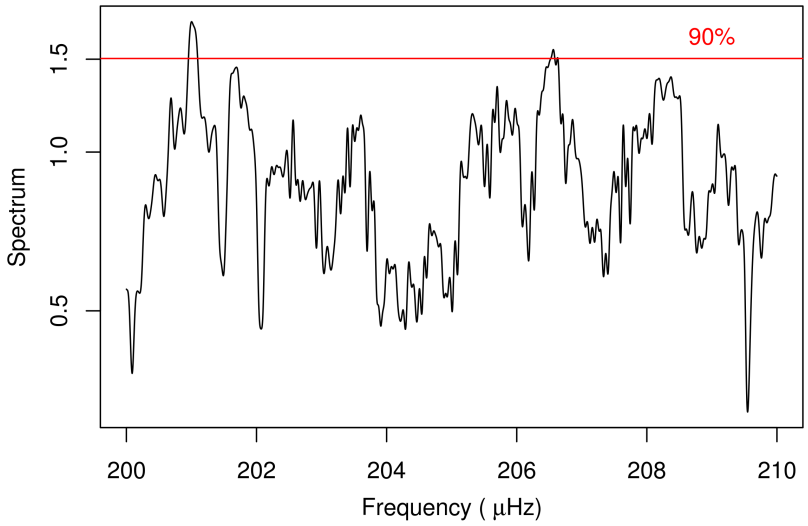

Figure 2: Example standardized spectrum on the interval 200μHz and 210μHz. The 90^{th} percentile of a χ^2_{14} is shown.

of approximately white noise and signal. However, the scale of the white noise process is not known and must be estimated. Multitaper spectrum estimates are χ^2_{2K} distributed for a white noise process with variance equal to one. To estimate the scale of the white noise process, we matched the 10^{th} percentile of the estimated reshaped spectrum $\hat{S}_{\text{res}}$ to the 10^{th} percentile of a χ^2_{2K} distribution. Within band b,

$$\hat{S}_w(f^*) = \frac{\hat{S}_{\text{res}}(f^*)}{\hat{S}_{0.10}/\chi^2_{2K,0.10}} \tag{1}$$

where $\hat{S}_{0.10}$ is the 10^{th} percentile of $\hat{S}_{\text{res}}$. We chose the 10^{th} percentile for matching because it is unlikely that the high amplitude signal processes we are interested in will affect the value of the 10^{th} percentile $\hat{S}_{\text{res}}$ within any $10\,\mu$Hz band. Having scaled the spectrum appropriately, signal components can be identified by comparison to quantiles of the χ^2_{2K} distribution. Figure 2 depicts a critical value at the 90% level.

To categorize frequencies in the band that are gravity mode signals, we select those f^* where:

1. $\hat{S}_w(f^*) > \chi^2_{2K,1-\alpha_1}$, and
2. The multitaper F-test is significant at level α_2.

We used $\alpha_1 = \alpha_2 = 0.1$ to define signals in the g-mode band, and denote their frequencies as $\{f'_0, f'_1, \ldots, f'_n\}$, their phases as $\{\phi_0, \phi_1, \ldots, \phi_n\}$, and their amplitudes as $\{A_0, A_1, \ldots, A_n\}$, where n is the total number of g-mode signals identified across all $10\,\mu$Hz bands. Using the frequencies, phases, and amplitudes we reconstruct the value of the g-mode time series $g(t)$ at time t using

$$g(t) = \sum_{i=0}^{n} A_i \sin(2\pi f'_i t + \phi_i). \tag{2}$$

2.2 Estimating the relationship between solar gravity mode time series and X-ray flux series

Solar flare activity is monitored by the GOES series of spacecraft, and flares are classified based on their peak flux in the 1-8 A band. The GOES spacecraft monitor X-ray flux in the 1-8 A band at a high time resolution; we used 1 minute resolution data for the forecasting analysis. X-ray flux was obtained from GOES-10, GOES-14, and GOES-15 satellites via SPIDR (http://spidr.ngdc.noaa.gov/spidr/ - GOES-10) and RESTful Web API (http://ngdc.noaa.gov/goes/sem - GOES-14 and GOES-15). GOES-10 data was preferred when available (from July 1, 1998 to Dec 1, 2009), followed by GOES-15 data (from September 1, 2010 to October 31, 2013), and GOES-14 data when neither of the other satellites were available.

For each estimation period, we obtained a gravity mode (g-mode) series $g(t)$ using the method described in the previous section and an X-ray flux series $h(t)$ from the GOES satellite data. Our hypothesis is that activity in $g(t)$ maps in some way to activity in $h(t)$. More specifically, we suppose that the variance (or power) of $g(t)$ and the variance (or power) of $\log(h(t))$ are related, where we log-transform the X-ray flux as a variance stabilization procedure due to the occasional extreme value. We considered aggregation of the variances at daily, weekly, and monthly resolutions. We define

$$x_j = \frac{1}{|T_j| - 1} \left[\sum_{t \in T_j} g(t)^2 - \frac{\left(\sum_{t \in T_j} g(t) \right)^2}{|T_j|} \right] \tag{3}$$

to be the variance g-mode series, where $t \in T_j$ means times t in the aggregation period T_j (daily, weekly, or monthly), and $|T_j|$ is the number of times t in the aggregation period T_j (i.e. the cardinality of T_j). Similarly, we define

$$y_j = \frac{1}{|T_j| - 1} \left[\sum_{t \in T_j} \log(h(t))^2 - \frac{\left(\sum_{t \in T_j} \log(h(t)) \right)^2}{|T_j|} \right] \tag{4}$$

to be the variance of the log X-ray series.

We examined two types of relationships between the $\mathbf{x}$ and $\mathbf{y}$ series for their ability to forecast future X-ray activity. The goal was to estimate some function $y_j = f(x_j)$ in the estimation period that is a good predictor of $\mathbf{y}$ in the forecast period, using forecast $\mathbf{x}$ as input. The first relationship we examined was a simple linear regression between $\mathbf{x}$ and $\mathbf{y}$ for daily, weekly, and monthly aggregation blocks. The second was a simple cumulative model using g-mode activity $\mathbf{x}$ as input for daily and weekly aggregation blocks (estimation periods were too short for monthly aggregation blocks using the cumulative model).

Linear forecast The linear regression forecasts involved estimating

$$y_j = \hat{\beta}_0 + \hat{\beta}_1 x_j + \epsilon_j, \tag{5}$$

where $\hat{\beta}_0, \hat{\beta}_1$ are parameters estimated using standard least-squares regression by minimizing $\sum_j \epsilon_j^2$. These $\hat{\beta}_0$ and $\hat{\beta}_1$ were estimated for each estimation period and aggregation block, and the forecasting performance of each $f(\mathbf{x}) = \hat{\beta}_0 + \hat{\beta}_1 \mathbf{x}$ was evaluated by examining the correlation between $f(\mathbf{x})$ and $\mathbf{y}$ in the forecast period. We examined the linear regression forecast for daily, weekly, and monthly aggregated $\mathbf{x}$ and $\mathbf{y}$ time series. This linear regression forecast model is consistent with the hypothesis that g-mode activity and solar flare activity are directly related on the aggregation timescale (daily, weekly, monthly).

Cumulative model forecast The cumulative model forecasts used a state-space model of the form

$$\begin{aligned} z_j^* &= z_{j-1} + \alpha x_j - \alpha\beta \\ z_j &= z_j^* - \gamma z_j^* I\left(z_j^* > \delta\right) \end{aligned} \tag{6}$$

where z is a cumulative process of x, $I()$ is the indicator function (equal to 1 if $z_j^* > \delta$, and zero otherwise), and $\alpha, \beta, \gamma, \delta, z_0$ are parameters. $\alpha > 0$ is a scaling parameter. β represents a constant (each timestep) dissipation parameter. A key feature of this cumulative model is the occasional dissipation of a fraction γ of the process $(0 < \gamma < 1)$ whenever the threshold δ is exceeded. Although an obvious simplification because γ and δ do not vary, this feature is consistent with the hypothesis that solar flares are releases of accumulated stresses due (in part) to g-mode energy. The cumulative model was fit in each estimation period by minimizing the squared difference between $\mathbf{z}$ and $\mathbf{y}$: $\sum_j (z_j - y_j)^2$. We examined the cumulative model forecast for daily and weekly aggregated $\mathbf{x}$ and $\mathbf{y}$ time series.

2.3 Retrospective analysis of forecast performance

To assess the forecast performance of the linear and cumulative model forecasts, we conducted a retrospective analysis. We estimated g-modes over 189 day periods and forecast over 189 day periods. Estimation periods were based on the availability of ACE magnetic field data, and began on January 1 1999 at 23:50:00 UTM, proceeding in twenty-eight 189 day blocks ending on June 28 2013 at 23:50:00 UTM.

Our measure of forecast performance was the correlation between forecast variance in GOES log X-ray flux and the actual variance in GOES log X-ray flux over the 189 day forecast. A null model is expected to result in a correlation of zero between forecast and actual because we expect half of the correlations to fall above zero and half to fall below zero. Under this null model, we compute the probability of observing at least the number of positive correlations that we observed if our forecast model was no better than random chance (p-value). Due to missing X-ray flux data in some time periods, we were not able to evaluate the performance of each forecasting technique for all twenty-eight 189 day time periods.

Finally, as a proof-of-concept for Bonneville Power Authority (BPA) Technology Innovation Project #290 (TIP-290), we generated a forecast for the period from November 1 2013 to May 10 2014 using ACE magnetic field data and GOES-15 log X-ray data from April 25 2013 at 23:50:00 GMT to October 31 2013 at 23:50:00 GMT. The forecast was generated using the weekly cumulative forecasting model (Figure 6). We expect to see increased flare activity when the forecast series is high. BPA was interested in how a forecast model might be integrated into operations. To make the forecast more user-friendly for management and mitigation of possible impacts at BPA, we provided a colour-coded table based on the raw forecast that characterized periods of high (red), medium (orange), and low (green) forecast flare activity. The delineations were made based on percentiles of the forecast series; weeks in the top 30% of forecast weeks were classified high, weeks in the next 30% were classified medium, and weeks in the lowest 40% were classified low. These delineations are qualitative in the sense that they don't correspond to the probability of solar events, but we expect a higher amount of flare activity as we move from green to orange to red. In the results, we compare the colour-coded forecast to actual solar flares.

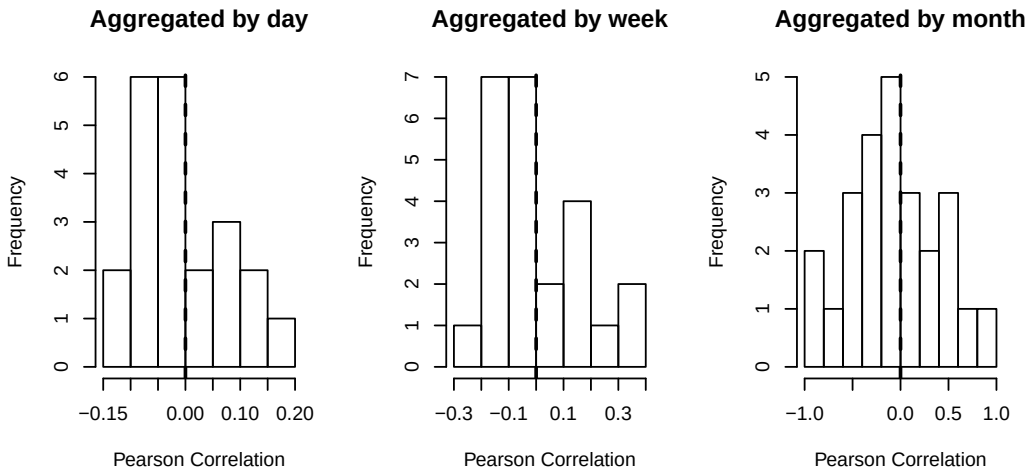

Figure 3: Distributions of forecast correlations for the linear model fits using variance of g-mode time series and variance of log X-ray time series aggregated by day, week, and month. None of the linear models display performance better than chance.

3. Results

3.1 Forecast performance

One of our hypotheses was that variation in the g-mode series was linearly related to variation in the X-ray series on a daily, weekly, or monthly timescale. Under this hypothesis, we expect there to be a reliable linear relationship between the variance of the g-mode series and the variance of the X-ray series. The distribution of linear forecast correlations are shown in Figure 3. None of the linear model forecasting techniques show performance that is better than chance – about half of the forecast correlations were postive – and thus have no consistent predictive power over the time period studied.

Although the linear model forecasts do not do consistently well in predicting variance in X-ray flux, we considered the possibility that their performance is a function of the time period studied. For example, the performance of the linear predictor may vary over the course of a solar cycle. However, an examination of the linear forecast correlations over the course of the retrospective analysis revealed that this was not the case. It appears that the linear model is simply not able to capture the relationship between the g-mode series and the X-ray flux series in a reliable way over an extended period. The linear model coefficients showed a similarly random behaviour over time.

It seems a simple linear relationship is not enough to predict variance in the log X-ray flux. Our second hypothesis was that accumulated stresses due in part to g-mode energy may be responsible for increased flare activity. We specified a simple cumulative model meant to capture the accumulation and release of energy consistent with our second hypothesis. The distribution of cumulative model forecast correlations are shown in Figure 4. The weekly aggregated cumulative model appears to perform better than chance; 18/25 forecasts have positive correlation performance. The p-value for the null hypothesis of no correlation on average is $p = 0.02$. The daily aggregation method showed a similar pattern, but not to the same level of significance. Based on an examination of the cumulative model fits in the training period, we suspect that the cumulative model we used may be too sensitive to daily variability to capture the phenomena of interest in a reliable way.

Although many of the positive correlations in Figure 4 appear to be relatively small,

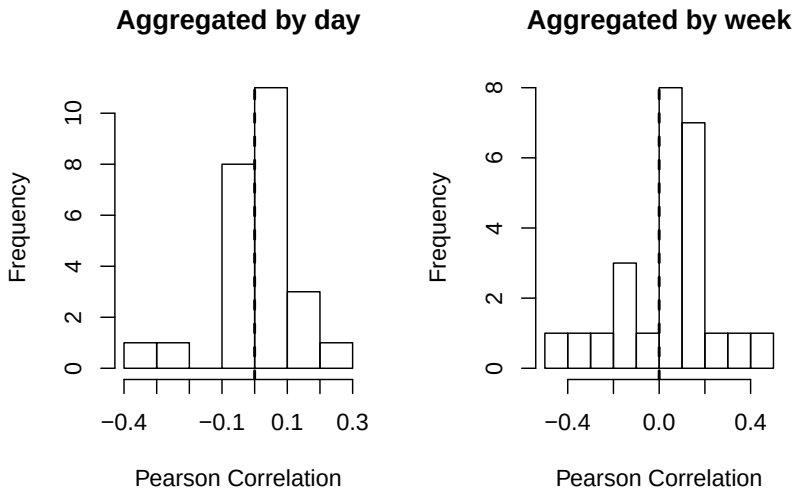

Figure 4: Distributions of forecast correlations for the cumulative model fits using variance of g-mode time series and variance of log X-ray time series aggregated by day and week. The weekly aggregated cumulative model appears to perform better than chance; 18/25 forecasts have positive correlation performance. The p-value for the null hypothesis of no correlation on average is $p = 0.02$.

even small correlations could provide useful forecasts. When the forecast correlation is large, there is obvious agreement between the shape of the forecast series and the shape of the actual historical data. However, even when the forecast correlation is small, there can be general agreement between the forecast and historical data (Figure 5). This means that although the magnitude of the forecast may be off due in part to the limitations of the simple cumulative model used, times when the forecast is *up* tend to align with times of increased variance in log X-ray flux. Thus, the forecast may be of some use in predicting the timing of events, even in cases where the magnitude of events is not well predicted.

3.2 Forecast for the period November 1 2013 to May 10 2014

For BPA TIP-290, we generated a forecast for the period from November 1 2013 to May 10 2014 (Figure 6). The outputs of the forecast are the weekly predicted values of the solar stress index z. Because this is a measure that is difficult to translate into management actions, we also provided BPA with a colour-coded weekly forecast of the propensity for large flare events (Figure 7). With a couple of M-class exceptions, the colour-coded forecast was able to predict weeks with and without large flare events for the first three-quarters of the prediction period. However, prediction performance was not good in the final quarter. We note that the purpose of the colour-coded prediction was not to demonstrate the predictive capabilities of our method (as was the purpose of the retrospective analysis). Rather, the purpose of the colour-coded prediction calendar was to provide an example of how the stress index forecast might be translated into management or operational decision-making.

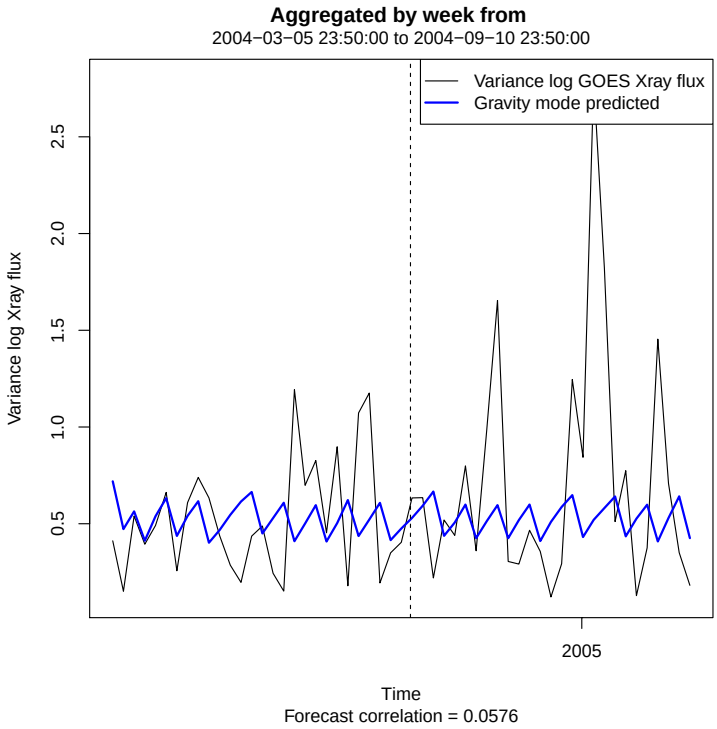

Figure 5: Example weekly cumulative model forecast for small positive correlation. The estimation period is to the left of the vertical line, and the forecast period is to the right. The amplitude of the forecast does not match the historical variance in log X-ray flux, but the timing may do well enough for prediction of events (with some misses).

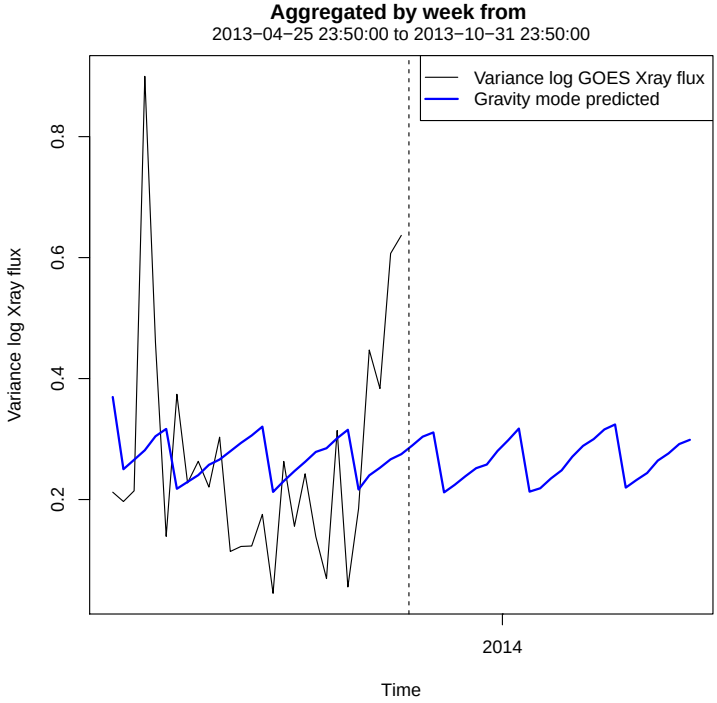

Figure 6: Cumulative model forecast for November 1 2013 to May 10 2014. The estimation period is to the left of the vertical line, and the forecast period is to the right.

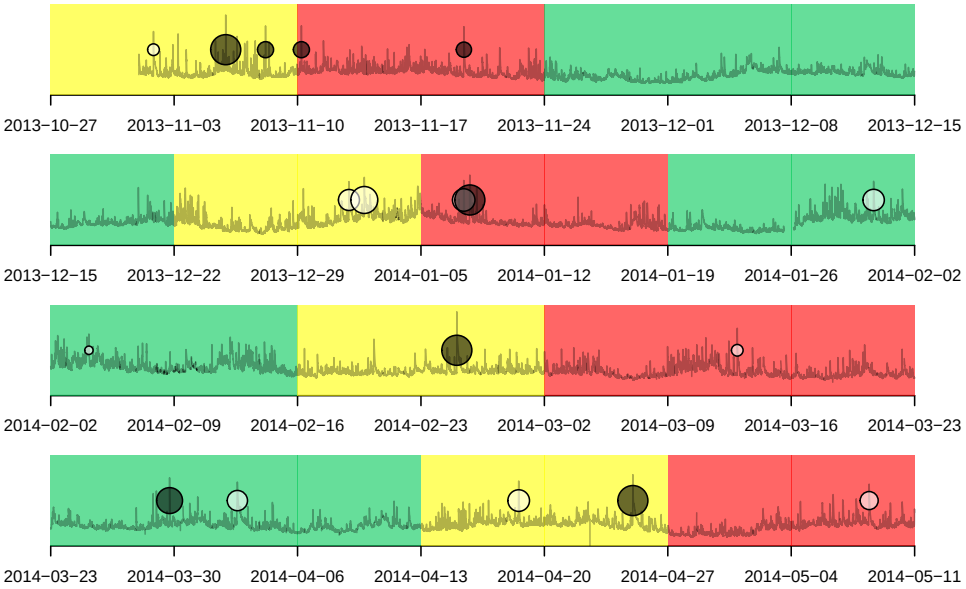

Figure 7: Cumulative model forecast for November 1 2013 to May 10 2014. Weeks in the top 30% of forecast weeks were classified high (red), weeks in the next 30% were classified medium (yellow), and weeks in the lowest 40% were classified low (green). The time series of GOES X-ray flux is also plotted. Flare events larger than class M5 are plotted as white circles, except X-class flares which are plotted as black. Relative circle sizes correspond to flare magnitudes.

4. Discussion

The approach that we have outlined here shows promise for forecasting over a timeframe that is longer than current approaches, which are based on tracking the development of active solar regions. The hypothesis that solar gravity modes provide a forcing mechanism for solar flares is based on a plausible physical process: gravity modes operate in the core of the Sun, and affect the convective layer above. In turn, the convective layer affects the development of active regions and ultimately solar flares and mass ejections. Because gravity modes are persistent, they can theoretically be predicted with a high degree of accuracy. However, the challenge is in implementation.

Obtaining estimates of solar gravity modes is far from straightforward. Although we are able to resolve many periodic components in the gravity mode band using ACE spacecraft magnetic field data, sorting out complications such as line splittings and peak broadening was beyond the scope of our current analysis. Gravity mode lines can be split for many reasons: orbital geometry of the ACE spacecraft, rotation of the solar core, differential rotation of the Sun, among others. Peak broadening implies frequency modulation of the signal, possibly due to various intermediate processes occurring in the Sun.

Without resolving these g-mode estimation issues, it may be that even a six month forecast horizon is too optimistic. Our anecdotal example forecast showed good performance for 4+ months, but seemed to struggle for the final portion of the forecast. A targeted approach to gravity mode estimation, taking account of as many complicating processes as possible, might improve forecast performance compared to our current approach, which is to ignore intermediate processes and simply run estimated periodic components in the g-mode band forward in time.

Alternatives to our simple cumulative stress model may also improve prediction performance. For example, avalanche models have recently been shown to be able to reproduce the statistical behaviour of solar flares in active regions (Strugarek and Charbonneau 2014), and may also provide an approach to modelling g-mode forcing of flares. Adjustment of the cumulative stress model to include the effects of the 11 year solar cycle would also be a good starting point for improvement. Finally, the choice of performance measure might be improved. If the goal is to predict the timing of events, the use of a classification measure such as a confusion matrix might be more appropriate than correlation between observed and predicted X-ray flux. This could affect the choice of model or even the method by which model parameters are fit to data in the training period. It could also provide management-oriented measures of performance directly e.g. probability of a flare event in a given time period.

5. Conclusion

Beginning from our hypothesis that solar flare activity is triggered by solar gravity modes, and with the goal of forecasting periods of high flare activity, we obtained solar gravity mode estimates from ACE spacecraft measurements. Using a weekly timestep, we found that a simple state-space model that accumulates solar stress and uses forecast solar gravity modes as input was able to predict variability in GOES X-ray flux where a simple linear relationship failed. Using the inferred solar stress index, we generated a qualitative forecast of solar flare activity that has promising performance. We suspect that the forecast horizon is limited to about six months, which is a substantial improvement compared to current techniques. We hope that further refinements to our technique will improve prediction performance.

6. Acknowledgments

This work was supported in part by Bonneville Power Authority Technology Innovation Project #290, and the National Science and Engineering Research Council of Canada.

References

Appourchaux, T., Belkacem, K., Broomhall, A.-M., Chaplin, W., Gough, D., Houdek, G., Provost, J., Baudin, F., Boumier, P., Elsworth, Y., et al. (2010), "The quest for the solar g modes," *The Astronomy and Astrophysics Review*, 18, 197–277.

Christensen-Dalsgaard, J., Dilke, F. W. W., and Gough, D. O. (1974), "The Stability of a Solar Model to Non-Radial Oscillations," *Monthly Notices of the Royal Astronomical Society*, 169, 429–445.

Gao, Y., Zhao, J., and Zhang, H. (2014), "A study of connections between solar flares and subsurface flow fields of active regions," *Solar Physics*, 289, 493–502.

Georgoulis, M. K. (2013), "Toward an Efficient Prediction of Solar Flares: Which Parameters, and How?" *Entropy*, 15, 5022–5052.

González Hernández, I. (2013), "Helioseismology in Space Weather," in *Astronomical Society of the Pacific Conference Series*, eds. Jain, K., Tripathy, S. C., Hill, F., Leibacher, J. W., and Pevtsov, A. A., vol. 478 of *Astronomical Society of the Pacific Conference Series*, p. 309.

Komm, R., De Moortel, I., Fan, Y., Ilonidis, S., and Steiner, O. (2013), "Sub-photosphere to solar atmosphere connection," *Space Science Reviews*, 1–33.

Reinard, A., Henthorn, J., Komm, R., and Hill, F. (2010), "Evidence that temporal changes in solar subsurface helicity precede active region flaring," *The Astrophysical Journal Letters*, 710, L121.

Ruzmaikin, A., Feynman, J., and Stoev, S. (2011), "Distribution and clustering of fast coronal mass ejections," *Journal of Geophysical Research*, 116, A04220.

Strugarek, A. and Charbonneau, P. (2014), "Predictive capabilities of avalanche models for solar flares," *Solar Physics*, 1–14.

Thomson, D., Lanzerotti, L., and Maclennan, C. (2001), "Interplanetary magnetic field: Statistical properties and discrete modes," *Journal of Geophysical Research: Space Physics (1978–2012)*, 106, 15941–15962.

Thomson, D. J. (1982), "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, 70, 1055–1096.

— (2000), "Baseline and Distribution Estimates of Complicated Spectra," in *Proc. Tenth IEEE Signal Processing Workshop*, Pocono Manor, PA: IEEE Press, pp. 414–418.

Thomson, D. J., Maclennan, C. G., and Lanzerotti, L. J. (1995), "Propagation of solar oscillations through the interplanetary medium," *Nature*, 376, 139–144.