AEROSPACE REPORT NO.
TOR-2013-00873

A Surrogate Time Series Model for the F10.7 Solar Activity Index and Sunspot Number

October 10, 2013

T. Paul O'Brien
Space Science Applications Laboratory
Physical Sciences Laboratories

Prepared for:

Air Force Research Laboratory
Kirtland AFB NM 87117-5776

Contract No. FA8802-14-C-0001

Authorized by: Engineering and Technology Group

PHYSICAL SCIENCES LABORATORIES

The Aerospace Corporation functions as an "architect-engineer" for national security programs, specializing in advanced military space systems. The Corporation's Physical Sciences Laboratories support the effective and timely development and operation of national security systems through scientific research and the application of advanced technology. Vital to the success of the Corporation is the technical staff's wide-ranging expertise and its ability to stay abreast of new technological developments and program support issues associated with rapidly evolving space systems. Contributing capabilities are provided by these individual organizations:

Electronics and Photonics Laboratory: Microelectronics, VLSI reliability, failure analysis, solid-state device physics, compound semiconductors, radiation effects, infrared and CCD detector devices, data storage and display technologies; lasers and electro-optics, solid-state laser design, micro-optics, optical communications, and fiber-optic sensors; atomic frequency standards, applied laser spectroscopy, laser chemistry, atmospheric propagation and beam control, LIDAR/LADAR remote sensing; solar cell and array testing and evaluation, battery electrochemistry, battery testing and evaluation.

Space Materials Laboratory: Evaluation and characterizations of new materials and processing techniques: metals, alloys, ceramics, polymers, thin films, and composites; development of advanced deposition processes; nondestructive evaluation, component failure analysis and reliability; structural mechanics, fracture mechanics, and stress corrosion; analysis and evaluation of materials at cryogenic and elevated temperatures; launch vehicle fluid mechanics, heat transfer and flight dynamics; aerothermodynamics; chemical and electric propulsion; environmental chemistry; combustion processes; space environment effects on materials, hardening and vulnerability assessment; contamination, thermal and structural control; lubrication and surface phenomena. Microelectromechanical systems (MEMS) for space applications; laser micromachining; laser-surface physical and chemical interactions; micropropulsion; micro- and nanosatellite mission analysis; intelligent microinstruments for monitoring space and launch system environments.

Space Science Applications Laboratory: Magnetospheric, auroral and cosmic-ray physics, wave-particle interactions, magnetospheric plasma waves; atmospheric and ionospheric physics, density and composition of the upper atmosphere, remote sensing using atmospheric radiation; solar physics, infrared astronomy, infrared signature analysis; infrared surveillance, imaging and remote sensing; multispectral and hyperspectral sensor development; data analysis and algorithm development; applications of multispectral and hyperspectral imagery to defense, civil space, commercial, and environmental missions; effects of solar activity, magnetic storms and nuclear explosions on the Earth's atmosphere, ionosphere and magnetosphere; effects of electromagnetic and particulate radiations on space systems; space instrumentation, design, fabrication and test; environmental chemistry, trace detection; atmospheric chemical reactions, atmospheric optics, light scattering, state-specific chemical reactions, and radiative signatures of missile plumes.

AEROSPACE REPORT NO.
TOR-2013-00873

A Surrogate Time Series Model for the F10.7 Solar Activity Index and Sunspot Number

October 10, 2013

T. Paul O'Brien
Space Science Applications Laboratory
Physical Sciences Laboratories

Prepared for:

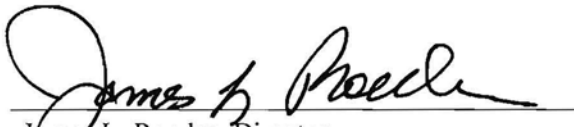
Air Force Research Laboratory
Kirtland AFB NM 87117-5776

Contract No. FA8802-14-C-0001

Authorized by: Engineering and Technology Group

A Surrogate Time Series Model for the F10.7 Solar Activity Index and Sunspot Number

Approved by:



James L. Roeder, Director
Space Sciences Department
Space Science Applications Laboratory
Physical Sciences Laboratories

Abstract

A new surrogate time series model for the monthly solar F10.7 radio flux index and the sunspot number is presented. The model incorporates varying solar cycle amplitude and length. The sunspot model is driven by the F10.7 model, with a transform and a noise term, thus preserving the observed strong correlation between F10.7 and sunspot number in any given surrogate time series. The model can be used to generate realistic scenarios of solar cycle variation over years to decades into the future. We also present a daily model of F10.7 that adds in residual random variation observed on sub-monthly timescales. Such surrogate solar activity models can be used in long-term satellite mission planning, e.g., to drive geophysical models such as atmospheric drag, solar particle intensity, and low-altitude radiation belt proton interactions with the thermosphere.

Acknowledgments

The authors acknowledge useful discussions with our colleagues at The Aerospace Corporation, especially Tim Guild, with Michael Xapsos at NASA GSFC, and Stu Huston at Atmospheric and Environmental Research, Inc.

Contents

Figures

1. Introduction

Many geophysical models are parameterized by solar activity. These models capture, in one way or another, solar cycle variations in solar emissions and eruptions and the consequences in interplanetary space and at Earth. Space system design depends on understanding these consequences as they impact satellites. Such impacts include solar cycle variation in the upper atmosphere and solar cycle dependence of the likelihood of solar energetic particle events.

With the recent development of the AE9/AP9 models of the trapped radiation at Earth [Ginet et al., 2013], it will soon become necessary to tie the Monte Carlo dynamic scenarios produced by AP9 to variations in solar activity both through modulation of the proton flux at low altitude (in and near the thermosphere) and through generation of dynamic scenarios of solar proton flux to couple to the trapped scenarios from AP9. Additionally, a solar cycle scenario-based approach may prove appropriate for bounding satellite re-entry forecasts.

The ability to generate an arbitrary number of realistic scenarios of solar activity far into the future (decades) will enable transition of more of our knowledge about how the Sun-Earth system behaves into our satellite mission design and planning. We have selected the F10.7 cm solar radio flux index [Tapping, 2013, and references therein] and the international sunspot number [Vaquero, 2007] as the two parameters most widely used to drive solar and geophysical models that satellite designers might employ. We have developed a statistical parameterization of the random fluctuations in solar activity as the solar cycle length and intensity vary from cycle to cycle, including the short-term, month-to-month variations, as well as the strong, but not perfect, correlation between F10.7 and sunspot number. To do this, we extend the F10.7 model of Xapsos et al. [2002] to incorporate stochastic variations and to produce a sunspot number. We have also produced an adjunct daily F10.7 model based on the monthly model.

2. Formulation of the Model

We begin with the Xapsos et al. [2002] model (hereafter the Xapsos model), which provides the F10.7 index as:

$$f(t) = Ag(t) + B \tag{1}$$

$$g(t) = \frac{\nu(\nu t)^{k-1}}{\Gamma(k)} e^{-\nu t} \tag{2}$$

$$F_{max} = a \times s + u \tag{3}$$

$$s = -\ln(-\ln(C)) \tag{4}$$

$$a = 33.22 \textit{ (deprecated)} \tag{5}$$

$$u = 167.44 \textit{ (deprecated)} \tag{6}$$

$$A = 5.214F_{max} - 400 \tag{7}$$

$$B = 92.3 - 0.158F_{max} \textit{ (deprecated)} \tag{8}$$

$$k = 7.562 - 0.006825F_{max} \tag{9}$$

$$\nu = 1.25 \tag{10}$$

Here, *t* is years since solar minimum, *Γ(k)* is the complete gamma function, and *C* is the desired confidence value. This set of equations provides a smoothly varying time series of F10.7 for a given confidence value. These details in Eqs. (3)–(10) were provided by M. Xapsos in a private communication. Some values, as we will see, have been deprecated, as they will be replaced by updated numbers.

We reformulate the model to introduce inter-cycle and intra-cycle variation beyond the basic pattern determined from *C*.

$$f_{t;z_t} = \left[A(C_t)g\left(\mathrm{mod}(t, T_t)\frac{\bar{T}}{T_t}; C_t\right) + B\right] \exp z_t^{(1)}. \tag{11}$$

Here we have replaced constant C with a varying C_t to allow F10.7 to vary on short timescales around the whole-cycle confidence value. Because our simulations run for multiple cycles, we have introduced a modulus to reset t at the end of each cycle. We have also allowed the cycle length T_t to vary over the simulation. The average value of T_t is given by $\bar{T}$. Finally, we include a short-term multiplicative noise factor controlled by a Gaussian random variable $z_t^{(1)}$.

The time varying parameter C_t must remain on the domain [0,1]. Thus, we will control C_t with a Gaussian random variable $z_t^{(2)}$ that we transform onto [0,1] using cumulative standard normal distribution Φ:

$$C_t = \Phi\left(z_t^{(2)}\right). \tag{12}$$

Next, we introduce a variable solar cycle length T_t, which has a Gaussian distribution with mean $\bar{T}$ and standard deviation σ_T:

$$T_t = \bar{T} + z_t^{(3)}\sigma_T. \tag{13}$$

The random variable $z_t^{(3)}$ is also Gaussian.

The sunspot number S_t is computed from F10.7 using a nonlinear transform and another noise factor:

$$S_t = \left[\max\left(0, f_{t;z_t}\exp\left(z_t^{(4)}\right) - b\right)/m\right]^{1/\gamma}. \tag{14}$$

As before, $z_t^{(4)}$ is a Gaussian random variable. Note that max(a,b) yields the larger of a or b, and is used here to ensure that the sunspot number is never negative.

The four random variables $\vec{z}$ are all Gaussian with zero mean, so they can evolve in time according to an autoregressive process. Specifically:

$$\vec{z}_{t+1} = \underline{\underline{W}}\vec{z}_t + \underline{\underline{Z}}\vec{\xi}_t, \tag{15}$$

where $\underline{\underline{W}}$ controls persistence, and $\underline{\underline{Z}}$ controls innovation, and $\vec{\xi}_t$ is a 4-element vector of Gaussian white noise with zero mean and unit variance and no correlation among the vector elements. By formulating the model this way, we allow for correlation among the random variables (z's) and (1st-order) temporal correlation among them as well.

A daily F10.7 model is useful in some cases because, for example, daily F10.7 is used as an input to MSIS [Picone et al., 2002]. Our daily F10.7 model is simply an add-on to the monthly F10.7 model. Namely:

$$f_{t,daily} = B_{daily} + \left(f_{t,monthly} - B_{daily}\right)\exp\left(z_t^{(5)}\right). \tag{16}$$

The random daily variation is captured in $z_t^{(5)}$, which varies according to another first-order auto-regressive equation:

$$z_{t+1}^{(5)} = \widehat{W}_5 z_t^{(5)} + \hat{Z}_5 \xi_{t+1}^{(5)}. \tag{17}$$

Here the time step is one day, and $\xi_t^{(5)}$ is Gaussian white noise with zero mean and unit variance. The monthly F10.7 model values are linearly interpolated onto the daily time tags assuming the monthly value occurs on the 15th of the month. Note, this model does not guarantee precise agreement between the monthly average of the daily model and the monthly model itself.

Because the form of the model was derived in part simultaneously with fitting its parameters, the next section will include some discussion of how we arrived at Eqs. (11)–(17), as well as the fitted results.

3. Derivation of the Model Parameters

First, we must determine the mean $\bar{T}$ and standard deviation σ_T of the sunspot cycle length, and show why a Gaussian is an appropriate model. Using the full record of the sunspot number from 1749 to present, we determine the length (min to min) of each of 23 solar cycles. The time of minimum is found automatically from 12-month smooth monthly sunspot number, and the maximum is taken from the original monthly value once the minima are identified. Figure 1 (left) shows the sunspot number history, the smoothing, and the identified minima. The average time between minima is 10.92 years, and the standard deviation is 1.27 years. Figure 1 (right) shows a quantile-quantile plot comparing the distribution of sunspot cycle length to quantiles of a Gaussian. The fact that the blue crosses fall near the red line indicates that the distribution of T_t is well-approximated by a Gaussian.

Similarly, we can compute the [2:3,2:3] submtarices of $\underline{\underline{W}}$ and $\underline{\underline{Z}}$ (these are the persistence and innovation parts of the sunspot length model). We denote these submatrices $\underline{\underline{\tilde{W}}}$ and $\underline{\underline{\tilde{Z}}}$. We convert the time series of solar cycle length and amplitude to Gaussian variables by replacing each entry in either time series with $\Phi^{-1}[i/(N+1)]$, where i is the entry's rank order (1 for smallest value, N for largest), N is the number of points in the time series. From these transformed variables, we compute the 2x2 covariance $\underline{\underline{\Sigma}}$ and lag covariance $\underline{\underline{R}}$. Then we have:

$$\underline{\underline{\tilde{W}}} = \left\{\underline{\underline{R}}\underline{\underline{\Sigma}}^{-1}\right\}^{1/(12\bar{T})}. \tag{18}$$

The matrix power converts from solar cycle timescale to monthly evolution.

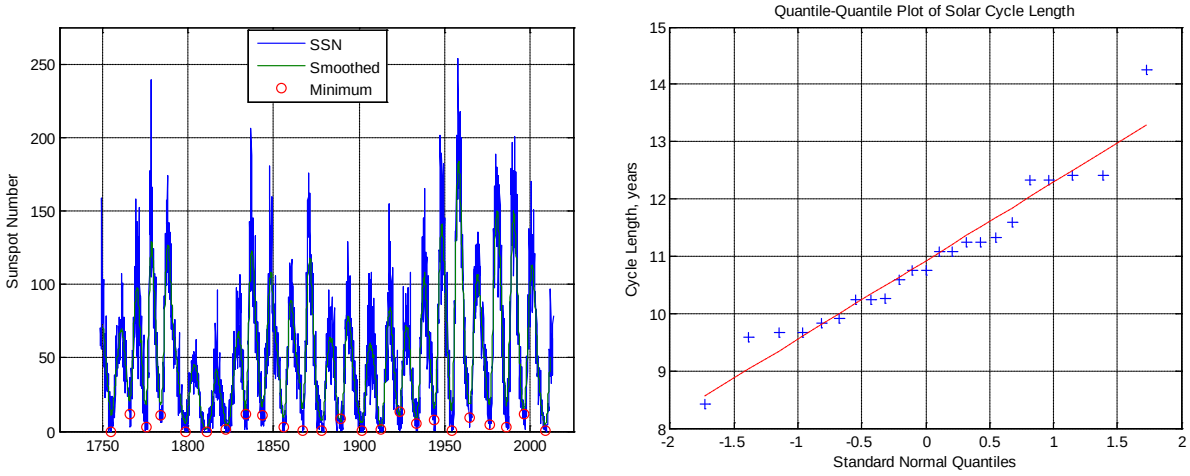


Figure 1. (Left) History of Sunspot number and automatically-identified minima. (Right) Quantiles of cycle length versus standard normal distribution (a straight line indicates a Gaussian distribution).

$$\underline{\underline{\tilde{Z}\tilde{Z}^T}} = \underline{\Sigma} - \underline{\underline{\tilde{W}\Sigma\tilde{W}^T}}. \quad (19)$$

$\underline{\tilde{Z}}$ is computed by Cholesky factorization of $\underline{\underline{\tilde{Z}\tilde{Z}^T}}$. We also store the last sunspot cycle length (12.42 years) for later use.

Next, we move on to F10.7, which is the core of the model. We compute B by taking the minimum F10.7 observed, which is 66.5. Then we repeat the solar min finding and max amplitudes for the monthly F10.7 time series (5 solar cycles). We fit each cycle to the form of (1), allowing F_{max} and B to vary, and scaling time to match $\bar{T}$. The history of F10.7 and the fit are given in Figure 2. This gives us five estimates of F_{max}, from which we can derive new a (29.3165) and u (186.1612). We also obtain the last value of F_{max} (189.59) and the date of the last minimum (1 Dec 2008, a month before the official date of the last minimum). These "last" values will be used to initialize the Monte Carlo scenarios.

Next, we look at the log residuals after the fitting of F10.7. A Q-Q plot shows they are Gaussian (Figure 3). They have a significant lag correlation, leading to W_{11} of 0.7137, and the rest of the residual is carried by a Z_{11} of 0.0905. The total log residual represented by z_1 is 0.1292, or a 12.92% relative residual, with a correlation coefficient of 0.7137 at one-month lag.

We use the F10.7 model as the basis of the sunspot number model. We had to decide whether to add a monthly noise term before or after the transform from F10.7 to sunspot number. We determined that the relative residuals of sunspot number to a rescaled fitted F10.7 were both Gaussian and strongly

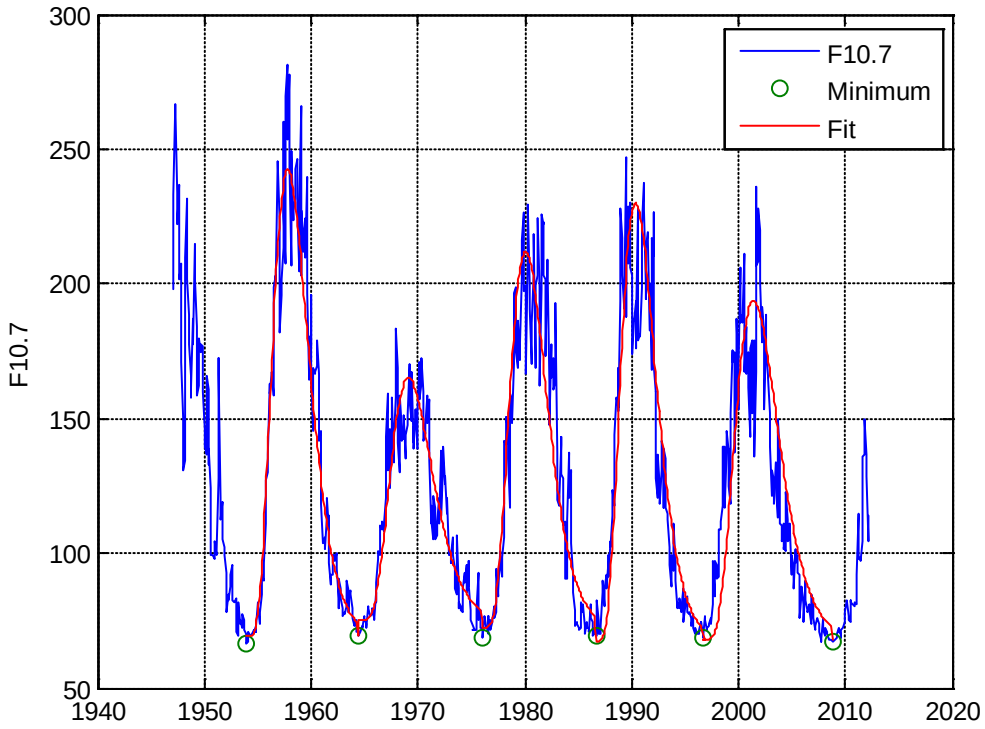


Figure 2. History of F10.7 with automatically identified minima and fit to parametric form (11).

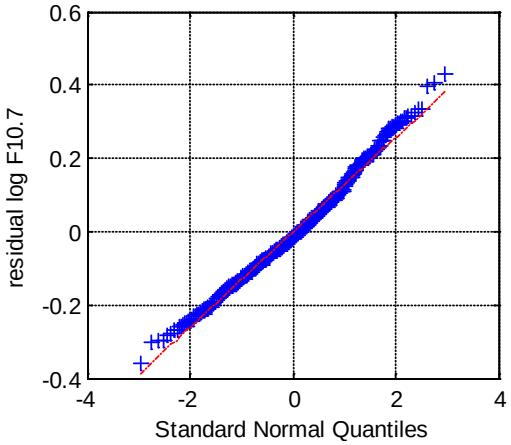


Figure 3. A Q-Q plot of F10.7 log residuals for the monthly model shows they are Gaussian.

correlated with the F10.7 residuals relative the same fitted F10.7 (Figure 4). Thus, we add the sunspot number residual after transforming the simulated F10.7, as shown in Eq. (14).

We also have to generate W_{44} and Z_{44}, which govern the residuals in the sunspot number model. First, however, we have to obtain the parameters of the transform: m, b, and γ, used in (14). The exponent γ is determined by visual fitting (i.e., trial and error) to be 1.1. Then the other parameters are determined by least-squares fitting to be m=0.5336, and b=66.71. Z_{44} and W_{44} are respectively 0.1099 and 0.6505, for a total relative residual of 14.47% in the sunspot number model and a one-month lag correlation coefficient in the log residual of 0.6505.

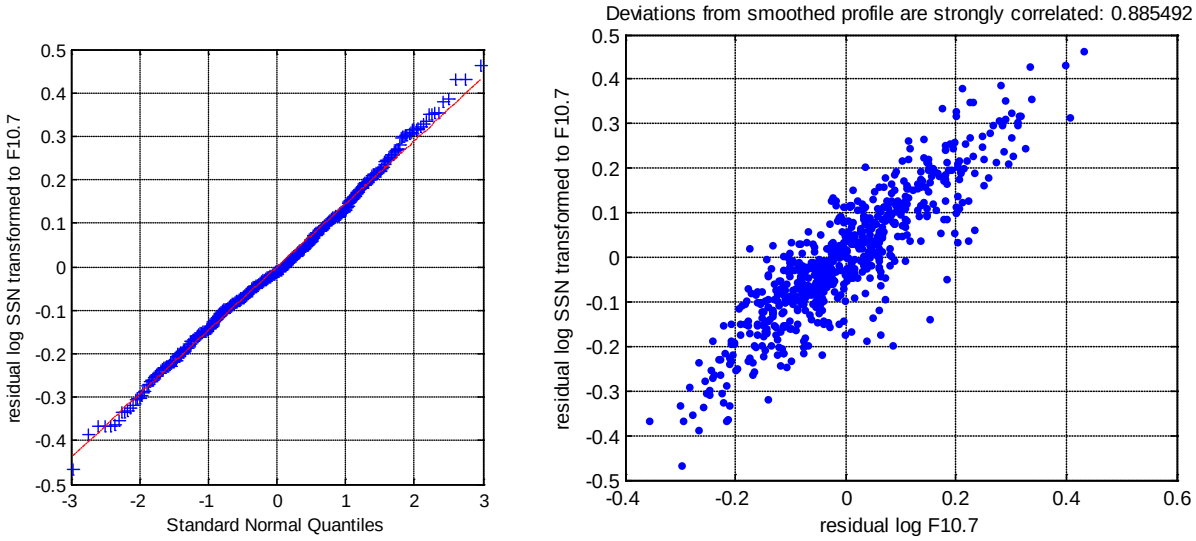


Figure 4. (Left) Q-Q plot of sunspot number residuals transformed to F10.7 domain. (Right) Deviations of sunspot number from smooth model are strongly correlated with corresponding devaitions in F10.7.

This completes the calculation of the $\underline{\underline{W}}$ and $\underline{\underline{Z}}$ matrices:

$$\underline{\underline{W}} = \begin{bmatrix} W_{11} & 0 & 0 & 0 \\ 0 & \widetilde{W}_{11} & \widetilde{W}_{12} & 0 \\ 0 & \widetilde{W}_{21} & \widetilde{W}_{22} & 0 \\ 0 & 0 & 0 & W_{44} \end{bmatrix} = \begin{bmatrix} 0.7137 & 0 & 0 & 0 \\ 0 & 0.9923 & -0.0123 & 0 \\ 0 & 0.0089 & 0.9957 & 0 \\ 0 & 0 & 0 & 0.6505 \end{bmatrix} \tag{20}$$

$$\underline{\underline{Z}} = \begin{bmatrix} Z_{11} & 0 & 0 & 0 \\ 0 & \tilde{Z}_{11} & \tilde{Z}_{12} & 0 \\ 0 & \tilde{Z}_{21} & \tilde{Z}_{22} & 0 \\ 0 & 0 & 0 & Z_{44} \end{bmatrix} = \begin{bmatrix} 0.0905 & 0 & 0 & 0 \\ 0 & 0.0749 & 0 & 0 \\ 0 & -0.0183 & 0.1230 & 0 \\ 0 & 0 & 0 & 0.1099 \end{bmatrix} \tag{21}$$

For the daily model, we obtained B_{daily} of 50 by visual fitting to obtain a Gaussian distribution of the log residuals. Figure 5 shows that the resulting residuals are indeed Gaussian. The parameters of the autoregressive model are $\widehat{W}_5$=0.9215 and $\hat{Z}_5$=0.0789, for a total residual relative error of 20.32%, with a daily lag correlation of 0.9215 for log residuals.

The model coefficients derived from monthly sunspot numbers from January 1749 to May 2013 and Penticton adjusted F10.7 from February 1947 to May 2013.

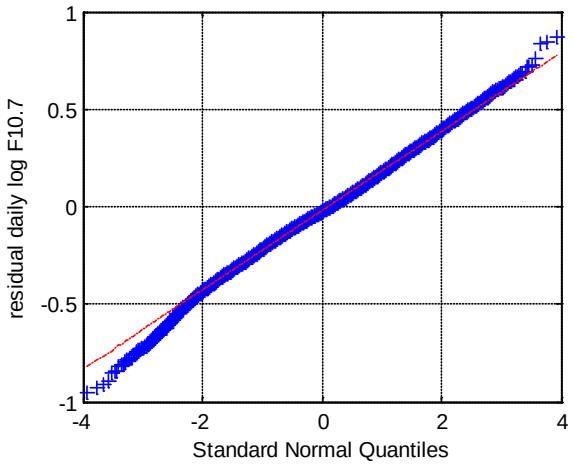


Figure 5. A Q-Q plot of F10.7 log residuals for the daily model shows that they are Gaussian.

4. Demonstration

Now that we have obtained all the needed parameters for our model, we conduct two different demonstrations: First, we will run it for many scenarios of a single solar cycle with random starting points (i.e., not stitched on to the most recent F10.7 and Sunspot data). Then we will run many scenarios starting from the most recent conditions out several cycles into the future.

Figure 6 shows many scenarios of the F10.7 and Sunspot number model. Note that due to variation in the cycle length, a couple of the cases start a new cycle in year 9 or 10, and that those are the same scenario (color) for F10.7 and sunspot number. Also note that the new model adheres to the bounds provided by the Xapsos model, with updated coefficients. Also, there is a mathematical discontinuity between cycles at solar minimum, but it is negligible compared to the random residuals controlled by z_1 and z_4.

Figure 7 shows an application of the model to the current situation: we must project forward from the current conditions several cycles into the future. We see that starting around 2050 the different scenarios start to drift out of phase, and by 2100, the phases are nearly random.

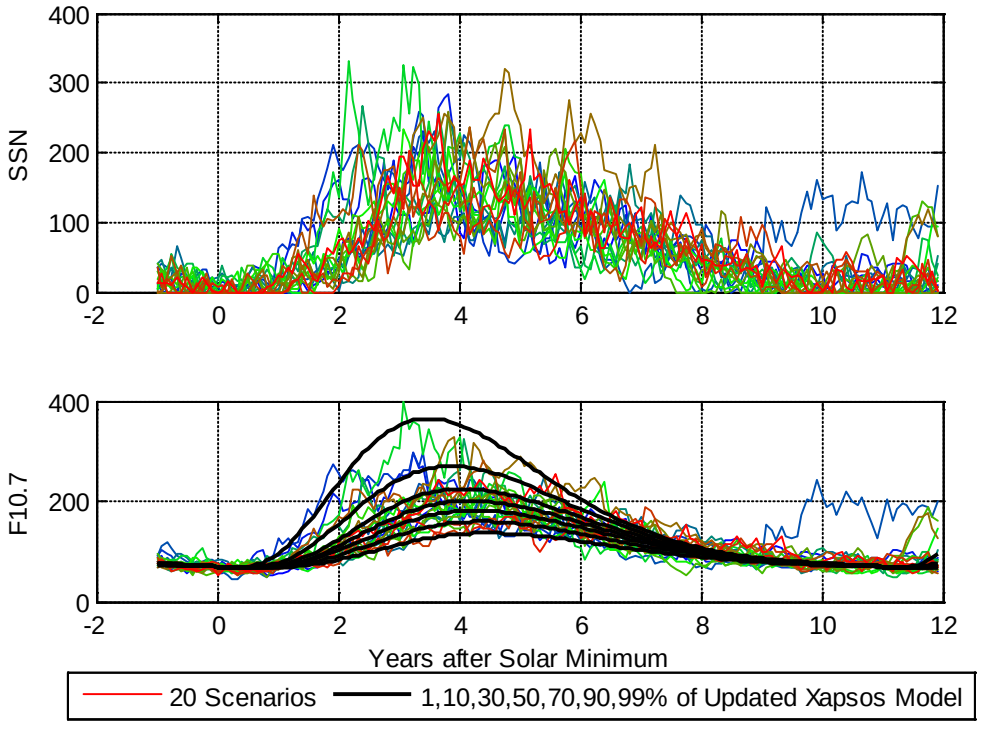


Figure 6. Many scenarios of an arbitrary solar cycle. (Top) Sunspot Number. (Bottom) F10.7. Dark curves provide percentiles of the updated Xapsos model.

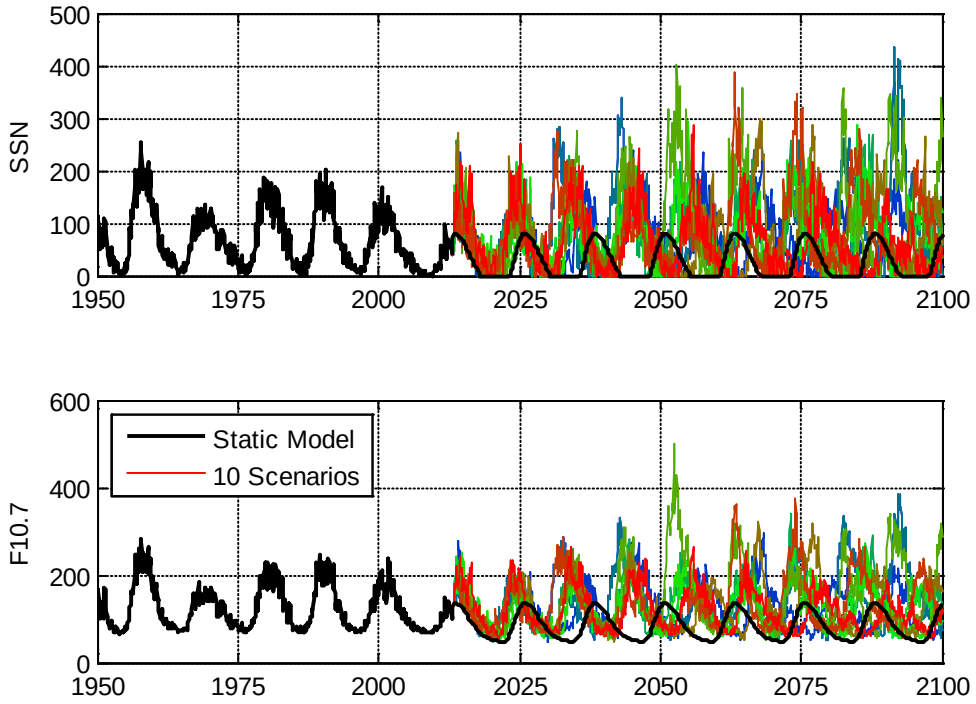


Figure 7. Ten scenarios of the solar cycle model, showing sunspot number (top) and F10.7 (bottom). All scenarios are built to match the most recent observations and continue randomly thereafter.

To make each Monte Carlo series match current conditions, we determined the set of $\vec{z}_0$ parameters that match the latest F10.7 and SSN observations. We used the date of last minimum, and the last maximum F10.7 determined in Section 3. We first make a preliminary estimate of $\vec{z}$. Our preliminary $z_0^{(1)}$ is 0; our preliminary $z_0^{(2)}$ is −0.2255 based on the last F_{max} being 189.59; our preliminary $z_0^{(3)}$ is 1.1775 based on a latest sunspot cycle length of 12.42 years determined in Section 3; finally, our preliminary estimate of $z_0^{(4)}$ is 0. We then update our estimates of $z_0^{(1)}$ to −0.3488 to match the last F10.7 of 134.2 observed for May 2013. At this point, in order to address the possibility that $z_0^{(1)}$ is very large, we update $z_0^{(1)}$ and $z_0^{(2)}$ via a least-squares fitting that minimizes the following penalty function:

$$\left(z_0^{(1)}\right)^2 + (z_0^2)^2 + \frac{\left(\log 134.2 - \log f_{t;z_0}\right)^2}{0.0167}. \tag{22}$$

The denominator in the final term is the variance of the log residual in the F10.7 model. This fitting procedure yields $z_0^{(1)}$ of −0.3730 and $z_0^{(2)}$ of −0.0781. Finally, we set $z_0^{(4)}$ to −0.0189 to match the last sunspot number of 78.7. With these initial values of $\vec{z}_0$ (−0.3730, −0.0781, 1.1775, and −0.0189), we can apply the Monte Carlo model to advance arbitrarily far into the future past May 2013. Each scenario into the future differs only in the random number seed used to generate the time series of innovations $\vec{\xi}$ in (15).

For the daily model, we similarly have to set the initial value of $z_0^{(5)}$, which we do based on the last observation in our dataset: 16 May 2013, for which F10.7 is 148.0, giving a z_5 of 0.1518. We can then simulate an arbitrary number of scenarios after 16 May 2013. We do this for 10 scenarios in Figure 8.

In both the daily and the monthly simulations, the recent slow rise to maximum drags the “static” model down, but the Monte Carlo models show more natural variability.

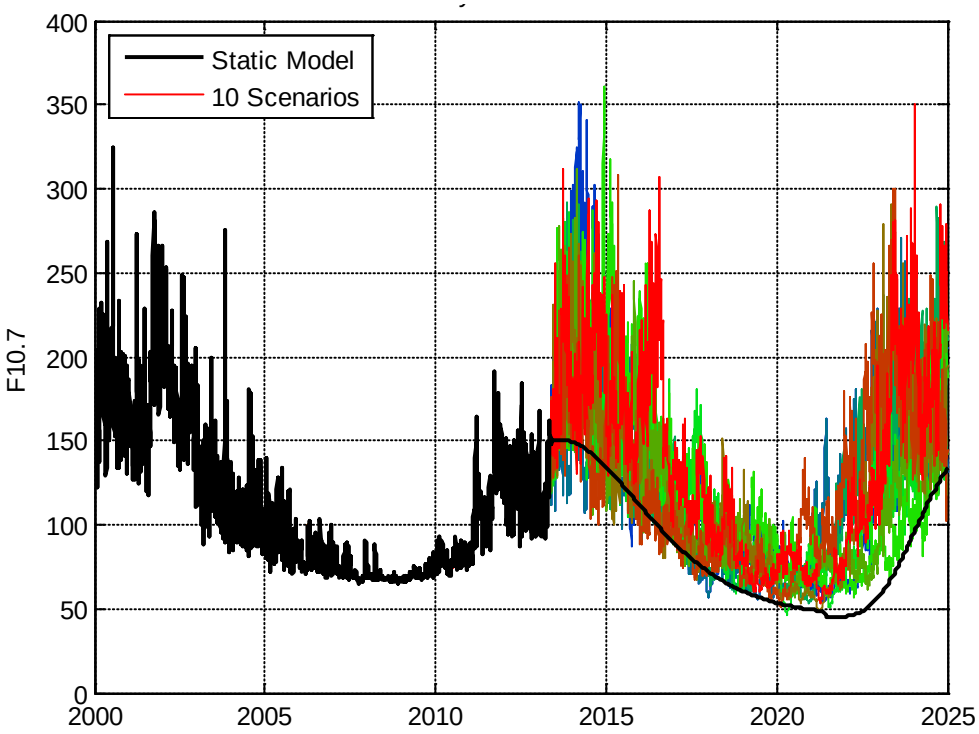


Figure 8. Ten scenarios from the daily F10.7 model. The static model does not update the daily residual or the monthly model parameters.

5. Validation

Aside from visually validating the profiles shown in Figures 6−8, we should also confirm that the marginal distribution of sunspot number and F10.7 is preserved by our simulation. Figure 9 shows three more Q-Q plots. For the bulk of the distribution, the plots lie along the expected diagonal. Some divergence at the largest sunspot numbers or F10.7 values indicates that the largest values of F10.7 or sunspot number are more (daily) or less (monthly) rare in the model than in the recent data. At this time, it is not known whether that represents an anomalous solar max or a part of the distribution that should be included in a future model. The fact that the daily and monthly models disagree in the direction of their deviation from the straight line suggests it is a benign sample size effect.

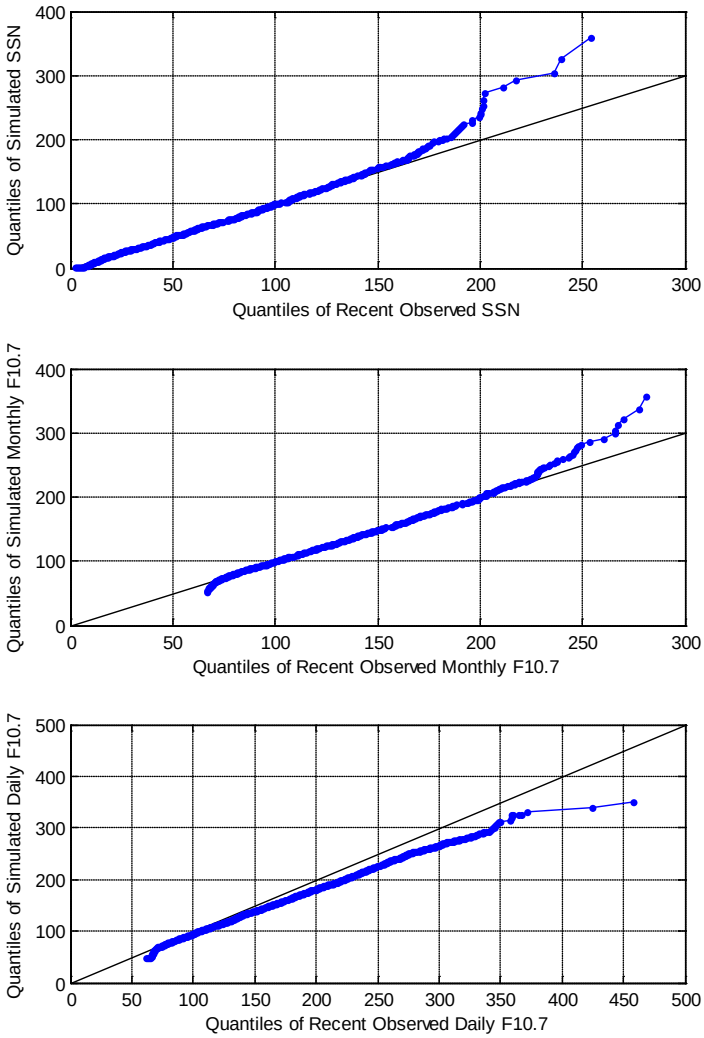


Figure 9. Q-Q plots for simulated versus observed sunspot number (top), monthly F10.7 (middle), and daily F10.7 (bottom).

6. Summary

We have developed a coupled bivariate time series model of the monthly solar activity indices F10.7 and sunspot number. We have also developed a daily F10.7 model that builds on the monthly model These models can be used to generate hypothetical scenarios of future variation in the solar cycle. Such a capability can be used, for example, for modeling long-term thermospheric variability, which is important for satellite mission planning: namely, as it applies to drag and variation of proton radiation in the South Atlantic Anomaly. These models can also be used to drive solar particle event models, which, in some cases, control the frequency and intensity of the solar particle events through solar activity indices.

References

Ginet, G.P., et al., "AE9, AP9 and SPM: New models for specifying the trapped energetic particle and space plasma environment," *Space Sci. Rev.*, doi:10.1007/s11214-013-9964-y, 2013.

Picone, J. M., A. E. Hedin, D. P. Drob, and A. C. Aikin, "NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues," *J. Geophys. Res.*, **107**(A12), 1468, doi:10.1029/2002JA009430, 2002.

Tapping, K. F., "The 10.7 cm solar radio flux (F10.7)," *Space Weather*, **11**, 394–406, doi:10.1002/swe.20064 (2013).

Vaquero, J. M., "Historical sunspot observations: A review," *Adv. Space Res.*, **40**, 929−941, 2007. Sunspot data were obtained from NOAA/NGDC: http://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-indices/sunspot-numbers/international/

Xapsos, M. A., S. L. Huston, and J. L. Barth, "Probabilistic Model for Low-Altitude Trapped-Proton Fluxes," *IEEE Trans. Nucl. Sci.*, **49**(6), December, 2002.