



Time-Causal Decomposition of Geomagnetic Time Series Into Secular Variation, Solar Quiet, and Disturbance Signals

By E. Joshua Rigler

Open-File Report 2017–1037

**U.S. Department of the Interior
U.S. Geological Survey**

U.S. Department of the Interior
RYAN K. ZINKE, Secretary

U.S. Geological Survey
William H. Werkheiser, Acting Director

U.S. Geological Survey, Reston, Virginia: 2017

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment—visit <http://www.usgs.gov/> or call 1-888-ASK-USGS (1-888-275-8747).

For an overview of USGS information products, including maps, imagery, and publications, visit <http://store.usgs.gov/>.

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this information product, for the most part, is in the public domain, it also may contain copyrighted materials as noted in the text. Permission to reproduce copyrighted items must be secured from the copyright owner.

Suggested citation:

Rigler, E.J., 2017, Time-causal decomposition of geomagnetic time series into secular variation, solar quiet, and disturbance signals: U.S. Geological Survey Open-File Report 2017–1037, 26 p., <https://doi.org/10.3133/ofr20171037>.

ISSN 2331-1258 (online)

Acknowledgments

I am grateful to Johns Hopkins Applied Physics Laboratory's SuperMAG program for making their extensive database of baseline-removed geomagnetic data easily accessible. I thank Carol A. Finn and Jeffrey J. Love from the U.S. Geological Survey, and James P. McCollough from the Air Force Research Laboratory, for reviewing this manuscript. I thank Abram Claycomb, Timothy White, Jeremy Fee, Michelle Guy, and Lynda Lastowka for their help and guidance in adapting research-grade software into a modern, robust, standards-compliant, open-source software tool that will certainly find utility beyond the U.S. Geological Survey's Geomagnetism Program.

Contents

Acknowledgments.....	iii
Abbreviations.....	iv
Abstract.....	1
Introduction.....	1
Mathematical Theory.....	2
Simple Exponential Smoothing.....	2
Holt's Linear Trend Forecast.....	3
Holt-Winters Seasonal Forecast.....	4
Prediction Intervals.....	5
Spike Detection and Adaptive Baselines.....	6
Numerical Algorithm and Practical Considerations.....	6
Verification and Validation.....	7
Is it Correct?.....	7
Is it Useful?.....	10
What Next?.....	19
Summary and Conclusions.....	20
References Cited.....	21
Glossary.....	23
Appendix 1. Pseudocode.....	24

Figures

1. Full time series of the synthetic observations used to test Holt-Winters algorithm.....	8
2. Holt-Winters decomposition and tracking of the synthetic observations.....	9
3. Holt-Winters decomposition and tracking of geomagnetic variations near Honolulu, Hawai'i.....	11
4. SuperMAG decomposition and tracking of geomagnetic variations near Honolulu, Hawai'i.....	13
5. Holt-Winters decomposition and tracking of geomagnetic variations near Boulder, Colorado.....	15
6. SuperMAG decomposition and tracking of geomagnetic variations near Boulder, Colorado.....	16
7. Holt-Winters decomposition and tracking of geomagnetic variations near Fairbanks, Alaska.....	17
8. SuperMAG decomposition and tracking of geomagnetic variations near Fairbanks, Alaska.....	18

Abbreviations

AE	auroral electrojet (a geomagnetic activity index)
BOU	Boulder (a magnetic observatory at Boulder, Colorado)
CMO	College (a magnetic observatory at University of Alaska, Fairbanks)
DIST	magnetic disturbance
Dst	disturbance storm time
HON	Honolulu (a magnetic observatory located near Honolulu, Hawai'i)
NaN	not-a-number (IEEE floating point value for undefined values)
PI	prediction interval
SQ	solar quiet
SV	secular variation
USGS	U.S. Geological Survey

Time-Causal Decomposition of Geomagnetic Time Series Into Secular Variation, Solar Quiet, and Disturbance Signals

By E. Joshua Rigler

Abstract

A theoretical basis and prototype numerical algorithm are provided that decompose regular time series of geomagnetic observations into three components: secular variation; solar quiet, and disturbance. Respectively, these three components correspond roughly to slow changes in the Earth's internal magnetic field, periodic daily variations caused by quasi-stationary (with respect to the sun) electrical current systems in the Earth's magnetosphere, and episodic perturbations to the geomagnetic baseline that are typically driven by fluctuations in a solar wind that interacts electromagnetically with the Earth's magnetosphere. In contrast to similar algorithms applied to geomagnetic data in the past, this one addresses the issue of real time data acquisition directly by applying a time-causal, exponential smoother with "seasonal corrections" to the data as soon as they become available.

Introduction

The magnetic field measured at a given location on the Earth's surface varies across a range of time scales. Earth's internal dynamo evolves slowly on timescale of decades to millennia. Earth's magnetosphere-ionosphere dynamo, itself modulated by solar variations and Earth's rotation, can drive periodic magnetic "tides" with time scales from hours to years. Finally, Earth's magnetosphere-ionosphere system responds in a nonlinear fashion to episodic variations in the solar wind to produce magnetic storms and substorms, with time scales from minutes to days. All these temporal variations in the Earth's magnetic field superimpose to produce the dynamic geomagnetic field measured at any given moment.

Geomagnetic time series are often decomposed into three constituents that correspond roughly to the variations just described. These are (1) a slow secular variation (SV; associated with Earth's internal dynamo), (2) a periodic solar quiet (SQ; associated with magnetic tides) variation, and (3) an aperiodic magnetic disturbance (DIST; associated with magnetic storms and substorms). SV is easy to isolate because its time scales are much longer than the other two constituents, and is often treated as a low-order deterministic model. SQ variations possess shorter time scales, but if they are assumed to be deterministic, they can be predicted given appropriate initial and boundary conditions. The remainder constitutes aperiodic and transient DIST.

Periodic tides can be easily described in terms of a Fourier series. The variations of highest relevance to space weather have periods of 1 day or less, but can still be substantially modulated by a superposition of lower-frequency signals. As a result, an accurate deterministic prediction requires more, usually much more, than 1 day's worth of data as input. Still, Fourier decomposition has been a

standard approach to geomagnetic time series analyses in the past (Sugiura, 1964; Love and Gannon, 2009; Love and Rigler, 2014).

In fact, Fourier decomposition has been adapted to generate real-time operational estimates of SQ (Gannon and others, 2011). However, careful consideration of the data interval during which the Fourier transform is estimated is needed. If the data window is too short, modulation of daily and higher frequencies is not well modeled. But if the data window is too long, not only do computational requirements increase, but nonstationarities in the geomagnetic time series can, and have, led to consequential artifacts in SQ predictions. The most obvious of these artifacts arise due to discontinuous baseline shifts in magnetic observatory time series caused by technological and (or) cultural changes to an observatory's local magnetic environment (McCollough and others, 2015).

This is not a newly discovered phenomenon, and there is, in fact, a very mature field of applied mathematics dedicated to adaptive, causal time series analysis to which we can look for better and more robust methods to predict periodic signals. Perhaps the simplest form of adaptive time series analysis we could employ is something commonly referred to as “exponential smoothing.” This report describes the mathematics underlying a specific form of exponential smoothing and its real-time application that decomposes U.S. Geological Survey (USGS) geomagnetic data into an adaptive baseline that tracks SV, periodic SQ variations, and a remaining DIST relative to these quasi-deterministic signals that can help identify magnetic storms and substorms in real time.

Mathematical Theory

Exponential smoothing of time series has been used in countless research, engineering, economic, sociological, political, and other applications. Although its utility has been empirically demonstrated time and again during the last half century or more, it has only been in the last couple decades that exponential smoothing has been normalized in form, withstood rigorous mathematical scrutiny, and been tied directly to well-known statistical time series models. A major contributor to the recent maturation of this subdiscipline of applied mathematics is R. J. Hyndman. The notation in his free online textbook (Hyndman and Athanasopoulos, 2013) is used to provide a very brief overview of exponential smoothing that culminates in an algorithm that can be used to decompose a time series into a trend, a repeating “seasonal” pattern, and a residual.

Simple Exponential Smoothing

Exponential smoothing is a form of causal time-series filtering that can be used to estimate the most likely observation at time $t+1$, given observations up to time t . In its simplest form, exponential smoothing is a weighted average of the most recent observation, and the previous weighted average (Brown, 1956, 1963):

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

where y_t is the observation at time t ,
 \hat{y}_t is the predicted observation at time t , and
 α is the forgetting factor between 0 and 1.

Equation 1 is a recursive formulation, which is preferred for algorithmic implementation. If equation 1 is expanded into an infinite series, it becomes clear why we refer to this as “exponential smoothing”:

$$\hat{y}_{t+1|t} = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} + \dots \quad (2)$$

While α is constant, the term $(1 - \alpha)$ changes exponentially for older and older observations, and because this term is by construction less than 1, older observations are weighted exponentially less than newer observations. As α nears unity, the algorithm's "memory" of the older observations disappears.

An alternative, but equivalent, formulation decomposes equation 1 into two steps:

$$\begin{aligned}\hat{y}_{t+1|t} &= \ell_t \\ \ell_t &= \alpha y_t + (1 - \alpha)\ell_{t-1}\end{aligned}\tag{3}$$

where ℓ_t is the level, or instantaneous baseline, at time t .

By rearranging equation 3, we obtain the so-called error-correction formulation (Hyndman and Athanasopoulos, 2013):

$$\begin{aligned}\hat{y}_{t+1|t} &= \ell_t \\ \ell_t &= \ell_{t-1} + \alpha(y_t - \ell_{t-1}) \\ \ell_t &= \ell_{t-1} + \alpha e_t\end{aligned}\tag{4}$$

where e_t is the one-step prediction error for time t , or $y_t - \hat{y}_{t|t-1}$.

The value of this error-correction formulation will become clearer when we present a numerical algorithm that includes simple exponential smoothing, as well as additional terms that model trend and seasonal variations to better match realistic observations.

Holt's Linear Trend Forecast

Equation 3 can be augmented to include a linear trend (Holt, 1957):

$$\begin{aligned}\hat{y}_{t+h|t} &= \ell_t + hb_t \\ \ell_t &= \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}\end{aligned}\tag{5}$$

where h is the forecast horizon,
 b_t is the slope at time t , and
 β^* is the slope forgetting factor between 0 and 1.

In words, a forecast is the level plus a slope multiplied by the number of discrete time steps. The level is still a weighted average of the observation at time t and the one-step prediction from time $t-1$ to t . The slope itself is the exponentially smoothed one-step difference in the baseline from time $t-1$ to time t . Similar to simple exponential smoothing, there is an error-correction formulation:

$$\begin{aligned}\hat{y}_{t+h|t} &= \ell_t + hb_t \\ \ell_t &= \ell_{t-1} + b_{t-1} + \alpha e_t \\ b_t &= b_{t-1} + \alpha\beta^* e_t\end{aligned}\tag{6}$$

Note that the β^* notation simply follows Hyndman and Athanasopoulos (2013) who reserve the term β for a more generic smoothing parameter in their innovations state-space model. We do not use this model but retain their notation for consistency.

A naïve implementation of Holt's linear trend method tends to over-forecast, especially for larger forecast horizons h . One way to address this is to dampen the trend, or in other words, force the slope toward zero for larger forecast horizons. A minor tweak to equation 5 gives:

$$\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t \\
\ell_t &= \alpha y_t + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1}) \\
b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}
\end{aligned} \tag{7}$$

where ϕ is the dampening factor between 0 and 1.

If $\phi = 1$, equation 7 is identical to the traditional Holt linear method; if $\phi = 0$, equation 7 is identical to simple exponential smoothing. As before, rearranging terms leads to an error-correction formulation:

$$\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t \\
\ell_t &= \ell_{t-1} + \phi b_{t-1} + \alpha e_t \\
b_t &= \phi b_{t-1} + \alpha \beta^* e_t
\end{aligned} \tag{8}$$

A need to forecast far into the future is not anticipated, but if such forecasting is necessary, there will likely be data gaps. Some of these data gaps may be long enough that if a constant slope is used to extrapolate across the gap, the algorithm may become unstable as the forecast increases/decreases without bound.

Holt-Winters Seasonal Forecast

Finally, it is common for time series to exhibit a repeating pattern during a fixed interval. Seasonal variations are among the most familiar form of repetition, but a “seasonal” correction can be applied at any interval, even a 24-hour day. Charles C. Holt and his student, Peter R. Winters, share credit for formalizing a powerful seasonal correction tool that has stood the test of time, and is still used frequently today (Holt, 1957; Winters, 1960). Augmenting equation 5 gives:

$$\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + hb_t + s_{t-m+h_m^+} \\
\ell_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\
b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \\
s_t &= \gamma^*(1 - \alpha)(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma^*(1 - \alpha))s_{t-m}
\end{aligned} \tag{9}$$

where m is the number of discrete time steps within a repeating interval,
 h_m^+ is the modulo of $(h-1)$ with m plus 1 (that is, $[(h-1) \bmod m] + 1$),
 s_t is the seasonal correction for time t , and
 γ^* is the seasonal correction forgetting factor between 0 and 1.

In words, the appropriate seasonal correction is added to an h -step prediction. The seasonal correction for time t is removed from the current observation before updating the level for time t , and the slope is updated using these seasonally corrected levels. Finally, the seasonal correction is updated using yet another exponential smoothing parameter γ^* . (Similar to β^* , the asterisk is used only for consistency with Hyndman and Athanasopoulos [2013].) Rearranging terms and including a trend-dampening factor leads to the error-correction formulation that will ultimately be implemented algorithmically:

$$\begin{aligned}
\hat{y}_{t+h|t} &= \ell_t + (\phi + \phi^2 + \dots + \phi^h)b_t + s_{t-m+h_m^+} \\
\ell_t &= \ell_{t-1} + \phi b_{t-1} + \alpha e_t \\
b_t &= \phi b_{t-1} + \alpha \beta^* e_t \\
s_t &= s_{t-m} + \gamma^*(1 - \alpha)e_t
\end{aligned} \tag{10}$$

A very astute and careful reader may have noticed that the one-step prediction error will not change if an arbitrary offset is added to the level and subtracted from the seasonal correction. This means that, while the final predictions will not be changed, the separation of the level from the seasonal component is not unique and may lead to undesirable equal/opposite drifts in both. These undesirable drifts can be mitigated by assuming that the sum of the seasonal corrections is zero during a repeating interval. To enforce this, a re-leveling adjustment can be calculated at each time step, added to the level, and subtracted from the seasonal correction (Archibald and Koehler, 2003; Gardner, 2006):

$$r_t = \frac{\gamma^*(1 - \alpha)}{m} e_t \quad (11)$$

Note that this adjustment can either be applied with each iteration, or simply accumulated and applied at the end. The latter may prove to be more efficient for certain interpreted programming languages that are known for slow loops but which have vector-optimized functions for relatively simple mathematical operations.

Prediction Intervals

Every forecast has an associated prediction error e_t . Assuming a real measurement is available for time $t-1$ before time t , this residual allows forecasts to be made using the error-correction formulations described previously. This formulation is convenient for causally smoothing a time series, but it also allows simulation beyond one step if e_t is treated as a statistical value with its own distribution. By performing many simulations, each using e_t drawn randomly from its distribution, so-called prediction intervals (PIs) can be constructed for any number of steps into the future. Useful percentiles can then be estimated from these PIs, providing a measure of confidence in the point forecast (that is, the forecast that assumes $e_t = 0$).

Any number of distributions for e_t may be assumed. The most accurate distribution may be obtained by simply storing every e_t and then drawing a random sample from this distribution during a simulation. Such “bootstrapping” is well established and robust (Efron and Tibshirani, 1986) but not especially convenient in real-time processing. In addition, drawing from a distribution that is not yet adequately populated can be problematic. For the error-correction formulations presented previously, however, analytic expressions exist if e_t is assumed to be normally and independently distributed with a known variance σ^2 . For the damped Holt-Winters forecast given by equation 10, the variance of the PI h steps into the future (v_h ; Hyndman and others, 2005) is:

$$v_h = \sigma^2 \left(1 + \sum_{j=1}^{h-1} c_j^2 \right) \quad (12)$$

$$c_j = \alpha(1 + \phi_{j-1}\beta^*) + \gamma^*(1 - \alpha)d_{j,m}$$

where ϕ_{j-1} is $1 + \phi + \dots + \phi^{j-1}$, and $d_{j,m}$ is 1 if $j \bmod m$ equals 0, otherwise it is 0; in words, bump c_j each time seasonal corrections are recycled after the forecast horizon h grows longer than the number of seasons.

Here, it is assumed that σ^2 is known, which may not be the case. Variance is the standard deviation squared, and the standard deviation may be thought of as a magnitude of an expected residual.

Hence, one way to estimate and track this is to use simple exponential smoothing. All that is required is to define a forgetting factor for the expected residual magnitude:

$$\sigma_t = (1 - \alpha)\sigma_{t-1} + \alpha|e_t| \quad (13)$$

where α is the expected residual forgetting factor (possibly identical to the baseline forgetting factor, as given in equation 1), and $|e_t|$ is the magnitude of residual.

In practice, σ , and therefore σ^2 , should update in equation 13 if one-step prediction errors e_t are available and equation 12 is used for any forecast beyond one step. Note that equation 13 is not an error-correction formulation because $|e_t|$ is just another type of observation and is not the one-step prediction error. It is certainly possible to rearrange equation 13 into an error-correction formulation, but the notation required to differentiate between $|e_t|$ and e_t may be confusing.

Spike Detection and Adaptive Baselines

Imagine a threshold residual is defined as a multiple of the standard deviation (that is, a Z-score). Any e_t with an absolute value exceeding this Z-score threshold is treated as a bad data point. If this bad data point is treated as a gap (that is, the forecast is made, but level, slope, and seasonal correction parameters are not updated), σ^2 will grow with the PI simply due to the dynamical nature of the prediction equations. If the threshold residual is exceeded repeatedly, eventually the standard deviation will grow large enough that e_t no longer exceeds the threshold residual, and the sequence of observations previously treated as bad data will once again be used to update the level, slope, and seasonal correction parameters. In effect, a large DC shift in observed data, while initially treated as bad data, will eventually be recognized as a new baseline, and the adaptive algorithm will converge on it according to equation 10.

Numerical Algorithm and Practical Considerations

Pseudocode describing the full Holt-Winters exponential smoothing algorithm (algorithm 1) is presented in appendix 1. The pseudocode assumes observations happen at regular discrete intervals, so time is treated as a zero-based index into a contiguous data array. The algorithm also assumes missing data are represented by a not-a-number (NaN) value. Other than these requirements, algorithm 1 can be easily implemented in almost any modern programming language. Algorithm 1's interface consists of variables that correspond to those defined in the Mathematical Theory section of this report.

Configuration parameters define algorithm 1's expected performance. The number of discrete time steps in a "seasonal" cycle (m) and the prediction horizon (h) are dictated by the application and data being processed. The forgetting factors (α , β , ϕ , and γ) need to be specified to track a quasi-stationary baseline, trend, and (or) repeatable signal that evolve according to an expected time scale. Because we assume regular discrete samples and no explicit time steps, forgetting factors need to specify the inverse of the average age desired in discrete samples, which can be calculated as $m / (\text{samples_per_cycle} \times \text{cycles})$. For example, if hourly data are being smoothed and an hourly update to a baseline with an average age of 3 days is needed, α would be specified as $1/(24 \times 3)$, or $1/72$. Alternatively, if we assume a repeating daily signal of 24 independent hourly samples, each of which evolved with a time scale of 3 days, γ would be specified as $24/(24 \times 3)$, or $1/3$. The parameter ϕ is only relevant if the prediction horizon is nonzero, or the state is propagated forward in time deterministically in the absence of valid observations. Finally, $zThresh$ provides an ad hoc means to restrict the influence of observations to only "quiet" conditions by forcing all residuals that exceed a certain threshold to be treated as invalid, which forces a deterministic propagation of the state.

Inputs are the time series to be smoothed ($yObs$), and a set of variables that initialize the algorithm to an arbitrary initial state ($yHat0$, $a0$, $b0$, $s0$, and $sigma0$). Outputs are a smoothed or predicted time series ($yHatOut$), corresponding “seasonal” corrections ($sOut$), and a running estimate of the expected residual ($sigmaOut$), plus a set of variables that can initialize the algorithm in a subsequent call with inputs that pick up where the previous call left off. Therefore, as an inherently state-aware procedure, algorithm 1 can be easily incorporated into a real-time data processing framework to process new observations as they become available.

The state variables passed into and out of algorithm 1 are not, in general, immediately useful values for geomagnetism research and operations. As noted in the Introduction, SV describes the geomagnetic trend line at a given moment, and is usually assumed to be a measure of the Earth’s internal field. SQ is used to describe mostly daily variations that result from the Earth’s rotation beneath geospace electric currents that are mostly fixed with respect to the sun. Finally, the value that is most often of interest in space weather applications is DIST, or the remainder of the geomagnetic signal if SV and SQ are removed. DIST is typically assumed to represent Earth’s magnetic response to aperiodic solar storms and substorms. Given outputs $yHatOut$ and $sOut$, these more familiar variables are easily calculated for N observed time steps in $yObs$:

1. $SV = yHatOut - sOut$
2. $SQ = sOut$
3. $DIST = yObs - yHatOut$

Finally, while DIST may not be predictable in any deterministic manner, at least using the Holt-Winters methodology, SV and SQ are predictable, and the Holt-Winters methodology may prove useful to forecast SV and SQ for short periods into the future. To this end, the most recent state needs to be passed to algorithm 1 along with synthetic inputs $yObs$ set equal to NaN for a desired number of discrete time steps into the future.

Verification and Validation

Verification is the process of confirming that an algorithm does what is expected. Validation addresses questions related to utility, stability, and added value. In addition to demonstrating the overall quality of an algorithm, verification and validation may also motivate future work toward improvements.

Is it Correct?

To verify that the Holt-Winters algorithm works as intended, the synthetic observations presented in figure 1 were generated, then processed using a Python implementation of algorithm 1. First, the synthetic observations exhibit a simple sinusoidal oscillation with an amplitude of 10, and a period of 100 time steps, which we will refer to as a synthetic “day” from this point forward. This oscillation was allowed to repeat uncorrupted for 50 days to allow ample time for the algorithm to establish a repeating daily variation. The synthetic series then undergoes a DC shift of 20, but is otherwise uncorrupted for another 50 days. Starting on day 100, this carrier signal is superposed with a 50-day sinusoid with an amplitude of 20. Days 150 to 200 consist of the carrier signal modulated in amplitude by a full sinusoid, which leads to phase reversal as the modulating signal passes through zero. Days 200 to 250 combine the superposition and amplitude modulation, and days 250 to 300 repeat this while adding white noise with a standard deviation of 5. This synthetic time series was not designed to represent any real-world phenomena, but rather to demonstrate the capabilities of the Holt-Winters algorithm. That said, this report refers to the baseline level as SV, and to the oscillations as SQ, to simplify subsequent discussion.

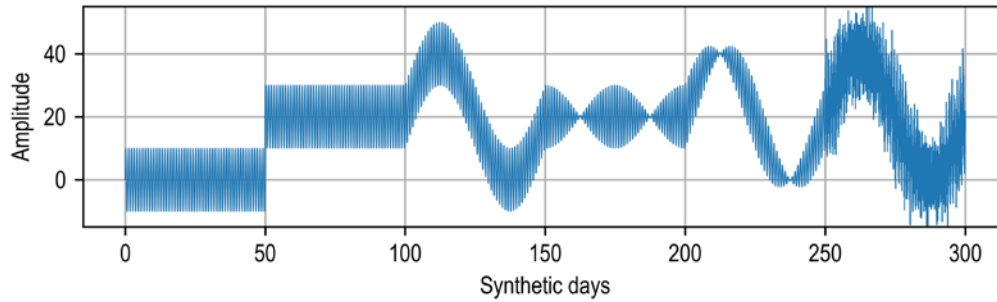


Figure 1. Full synthetic time series used to demonstrate the capabilities of the Holt-Winters algorithm.

Algorithm 1’s runtime configuration parameters were set to $m=100$, $\alpha=1/100/3$, $\beta=0$, $\gamma=100/100/3$, $\phi=1$, $h=0$, and $zThresh=6$. This memory is 300 samples, or 3 synthetic days, for SV; the memory is also 3 synthetic days for SQ, but this is only 3 samples because each element of the SQ vector is only updated just once for every cycle of length m . The forgetting factor $\beta=0$ implies an infinite memory for the slope, which prevents our default zero slope from changing. The other initial conditions also default to zero, except for σ_0 , which was set equal to the root-mean-square of the initial 50 days of synthetic data (for a sine wave, $10/\sqrt{2}$). The reason for the latter is that, if sigma is zero, any disturbance will exceed the 6-sigma z-score threshold, and be identified as a bad data point. The algorithm adapts by inflating sigma until it is large enough that such disturbances are considered valid observations, but this takes time.

Figure 2 presents the results of this Holt-Winters configuration when applied to the synthetic observations presented in figure 1. The blue trace is the synthetic data, the green trace is the slowly varying SV, and the red trace is SV plus the “daily” SQ variation. The gray-filled background represents the expected residual (+/-) between the blue and red traces multiplied by $zThresh$. If the blue trace falls outside the gray, no adjustments are made to SV or SQ, but the sigma is allowed to inflate, so eventually the blue trace always falls within the gray, and the Holt-Winters algorithm returns to updating SV and SQ. Results are stacked in 50-day panels to facilitate analysis. In panel a, the initial zero values for the SV and SQ evolve until they converge on their expected solutions. Note that SV can be nonzero, at least temporarily. Because the algorithm has no way to know whether discrepancies between predicted and measured data are the result of errors in SV or SQ, it distributes any uncertainty between them.

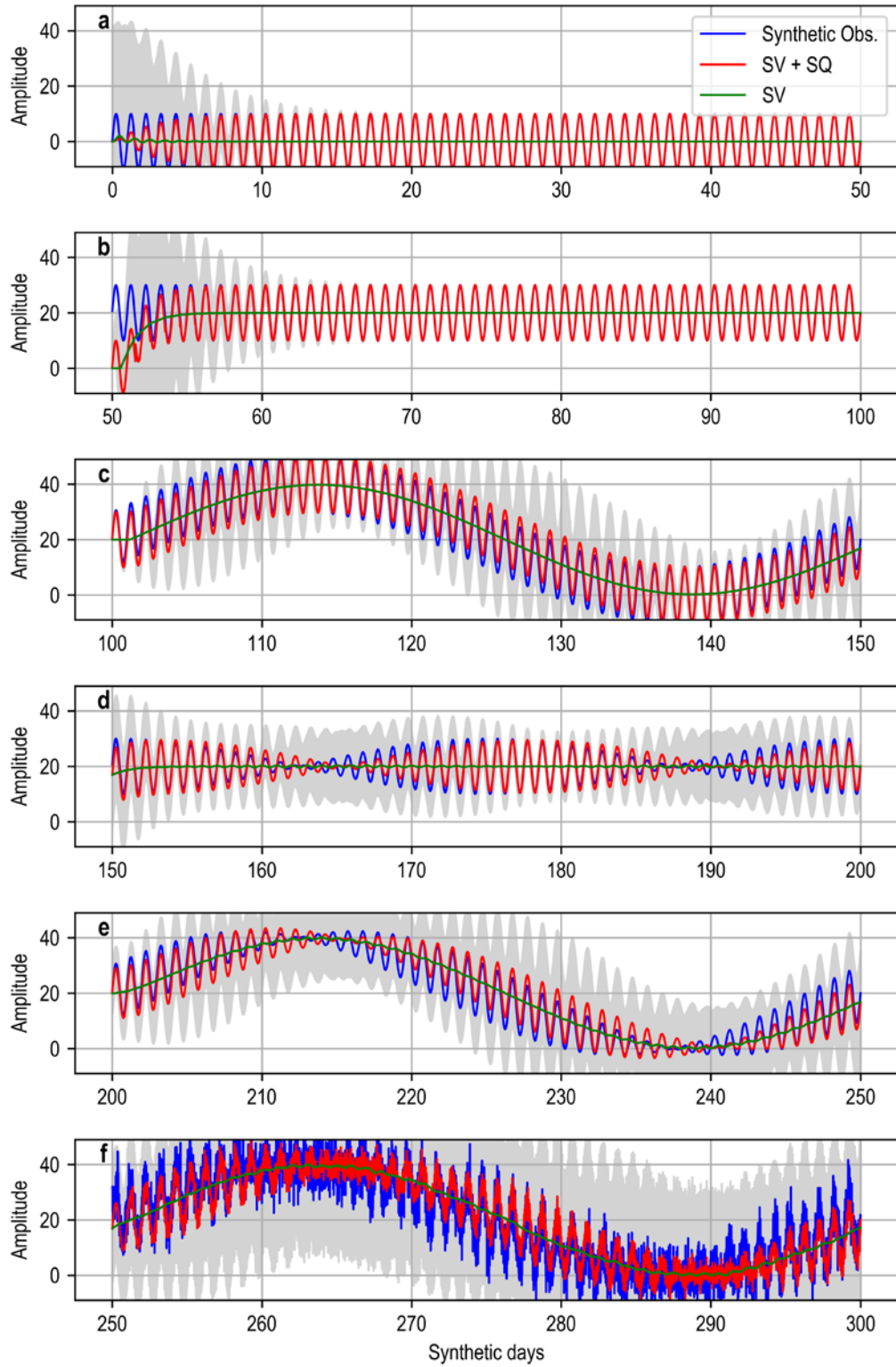


Figure 2. Holt-Winters algorithm tracking of the synthetic observations (Obs.). (Panels a–f correspond to the synthetic data described in the text, and presented in figure 1; in figure: SV, secular variation; SQ, solar quiet)

In panel b, SV adjusts to match the new baseline level, but only after an initial delay. This delay comes about because σ becomes very small while the Holt-Winters algorithm processes the uncorrupted synthetic data in panel a. Because σ is so small, the discrepancy between the old baseline and the newest data point exceeds $zThresh$, and the new data are considered invalid. As additional data are read in from the time series with the new baseline, σ grows larger until the threshold is no longer exceeded, and new data will be used to adjust SV and SQ.

Panels c–e demonstrate that the Holt-Winters algorithm can track variations in SV, variations in SQ, and variations in both together. The obvious time lags are unavoidable in a time-causal smoothing algorithm. However, nonstationarity in real geomagnetic time series often manifests with longer time scales (for example, actual seasonal, annual, or solar cycle variations), or quick-but-permanent shifts (for example, instrument reconfiguration). The former means time lags are relatively insignificant for most practical considerations; the latter means that, while there may be a brief interval after a baseline shift during which the Holt-Winters SV and SQ estimates are unreliable, they will quickly adapt to the new baseline level and daily variation.

Panel f shows what happens when random variations are superposed with the periodic carrier signal. SV is not severely impacted, and is, in fact, surprisingly similar to the SV estimated in panel e. SQ, on the other hand, is quite noisy. The reason for this apparent inconsistency is that, while SV and SQ both have approximately 3-day memories, random noise still requires many repeated observations before it averages out to a statistically significant zero mean. SV is updated with every sample, so the random noise averages out quickly. Elements of SQ are updated 100 times less often; in fact, each element is updated only 3 times within the 3-cycle time constant defined previously. This simply cannot average out to a statistically significant mean before the algorithm forgets earlier SQ values entirely. If memory is increased by reducing γ , SQ will exhibit less random noise (that is, it will be smoother), but it may not track changes such as the amplitude modulation in panels d–f as effectively. The latter is not a problem if such short-term amplitude modulation is considered disturbance.

Is it Useful?

To demonstrate that the Holt-Winters algorithm offers positive utility, we applied it to 6 months of USGS geomagnetic data for one low latitude (HON, located near Honolulu, Hawai'i), one middle latitude (BOU, located near Boulder, Colorado), and one high latitude (CMO, located at the University of Alaska, Fairbanks, referred to as “College”) magnetic observatory. Definitive versions of these data can be obtained from INTERMAGNET (<http://www.intermagnet.org>) (Love and Chuilliat, 2013). However, we compared our results with those obtained from the Applied Physics Laboratory’s SuperMAG project (Gjerloev, 2012) because, similar to the Holt-Winters algorithm, SuperMAG separates a full geomagnetic time series into a slowly varying baseline, a repeatable daily signal, and a disturbance. This is essentially where the similarities between our Holt-Winters implementation and SuperMAG end, so the algorithmic differences allow data generated by the latter to serve as an independent validation of the former. It should be noted that SuperMAG derives from definitive USGS data.

Figure 3 presents results from the HON magnetic observatory for the first 180 days of 2015, stacked in 30-day increments. The Holt-Winters runtime configuration parameters were set to $m=1440$, $\alpha=1/1440/15$, $\beta=0$, $\gamma=1440/1440/15$, $\phi=1$, $h=0$, and $zThresh=2$. The blue trace is the magnetically northward component of the three-dimensional geomagnetic vector field. As with the synthetic data, the green trace is the slowly evolving SV, the red trace is SV plus daily SQ variation, and the gray-filled background is the standard deviation of the difference between the blue and red traces multiplied by $zThresh$.

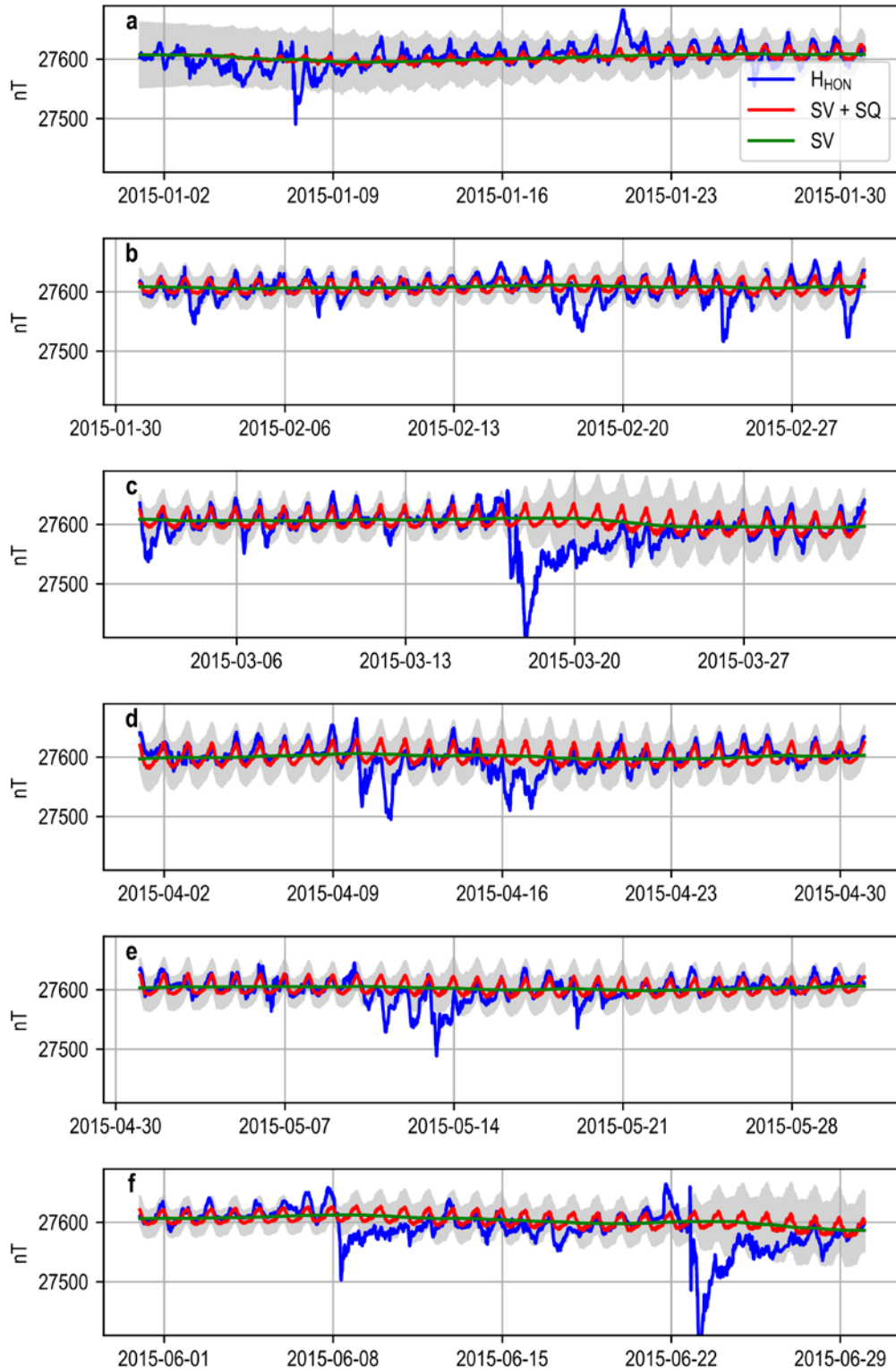


Figure 3. Holt-Winters decomposition of magnetically northward horizontal (H) geomagnetic variations at HON magnetic observatory, near Honolulu, Hawai'i. (in figure: nT, nanotesla; SV, secular variation; SQ, solar quiet)

A noteworthy feature of these data is that the red SV+SQ trace evolves from a nearly flat line to a realistic, daily variation within about 15 days, which is just as one would expect given the runtime parameter γ . This transition is not 100 percent at 15 days because 15 days is the average age of data that inform the current state. By 30 days, SQ is fully evolved, having “forgotten” its initial conditions almost entirely. The green SV baseline can also be seen to evolve in response to a slight negative shift in the observations that is a sign of very mild geomagnetic activity. As the activity subsides and the observations reflect quieter conditions, the SV trace returns to more positive values. Ideally, estimates of SV would be less sensitive to activity, but making α smaller to achieve this end means that the algorithm would be slower to adapt to the artificial changes in baseline that can arise. One can make $zThresh$ smaller, excluding more data, but this may be inadvisable during the algorithm’s spin-up period because the initial state may never evolve to capture the SQ variation.

Adjustments to the Holt-Winters algorithm sensitivity using the α parameter assume that the baseline level varies as a random walk, where α -scaled “steps” follow a Gaussian distribution. However, DIST is not generally Gaussian, and is, in fact, quite heavy tailed because of the influence of an enhanced ring current during magnetic storms. There are two magnetic storms worthy of note during this interval, the first commencing on March 17, 2015, and the second on June 22, 2015. Neither the SV baseline, nor the daily pattern of the SQ variation, change while the blue trace falls outside the gray-filled background during these storms. In addition, the gray-filled background expands during these periods of enhanced activity, accelerating a return of the blue trace back to the gray-filled background. The end effect is that the SQ variations are only adjusted to better match quieter conditions, as specified by $zThresh$. By carefully balancing the effects of α and $zThresh$, the Holt-Winters algorithm can be configured to adjust quickly to nonphysical baseline shifts, while remaining relatively insensitive to magnetic activity when estimating average quiet-time conditions. The values used here strike this balance reasonably well, but could almost certainly be refined for production runs.

Figure 4 presents the SuperMAG decomposition of the same HON data. We first note that there is no spin-up period for SuperMAG, because the SuperMAG algorithm is not a state-based algorithm. This should not be seen as a shortcoming of the Holt-Winters algorithm, but rather just a reflection of a specific use-case (time-causal smoothing) that is not addressed by SuperMAG. We also note that, compared to the Holt-Winters algorithm, the SuperMAG SV baseline is considerably less sensitive to magnetic activity. This is partly because of the acausal, time-symmetric nature of the SuperMAG smoothing algorithm, but is probably more attributable to a form of weighted averaging that deemphasizes intervals with more scatter in their data. This should certainly be seen as an advantage to the SuperMAG algorithm, and may conceivably be adapted for a time-causal implementation that would be useful in real-time scenarios. However, the method is somewhat ad hoc, and entirely empirical, meaning its general behavior cannot be predicted a priori, or extrapolated beyond the data window being processed.

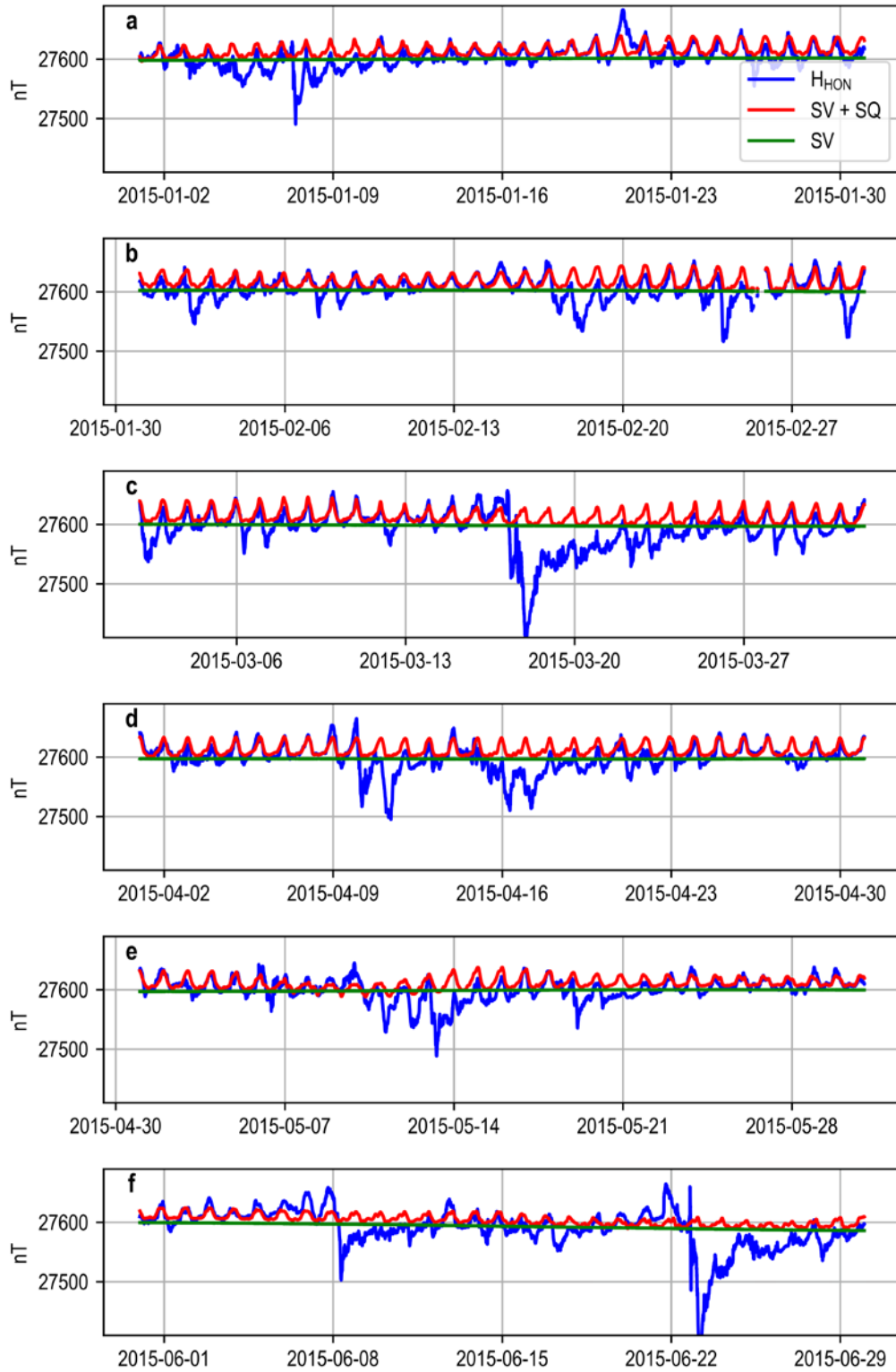


Figure 4. SuperMAG decomposition of magnetically northward horizontal (H) geomagnetic variations at HON magnetic observatory, near Honolulu, Hawai'i. (in figure: nT, nanotesla; SV, secular variation; SQ, solar quiet)

Perhaps of bigger concern with the SuperMAG algorithm than its ad hoc nature are its computational complexity and memory footprint, which are discussed in Gjerloev (2012). For example, at least 17 days of 1-minute data are required to calculate the statistics necessary to estimate each 1-minute SV baseline value. With modern computer power and ample memory, this may not be a showstopper for batch processing of existing data sets with output that will be permanently archived, but it may be too much for real-time processing. The Holt-Winters algorithm is much simpler and mathematically predictable, and the algorithm requires no more memory footprint than that needed to hold 1 day's worth of SQ corrections and a current estimate of SV.

The Holt-Winters algorithm has disadvantages compared to SuperMAG. For one, Holt-Winters SQ estimates are somewhat noisy compared to the daily variations in SuperMAG because SuperMAG fits cubic splines to hourly binned data. Holt-Winters is a more statistically rigorous approach to SQ estimation, but it is not clear how necessary this is, and we may consider adapting the technique to work with similarly binned data to achieve smoother results.

Another disadvantage is that Holt-Winters SV estimates are assumed to be the mean of the magnetic disturbance distribution, and this is somehow representative of the undisturbed background field. However, the SV estimates are not representative, because the distribution, as already noted, is very non-Gaussian. SuperMAG makes a somewhat more reasoned argument that the mode of the distribution represents quiet conditions, and indeed, the SuperMAG estimates of SV are typically several nanotesla greater than (that is, less disturbed) the Holt-Winters SV estimates. Given the parameterized statistical nature of Holt-Winters, the best we might do to emulate this is to transform the data from their natural distribution into something that more closely resembles a Gaussian distribution. In the case of geomagnetic data, it has been shown that a logarithm transform can do this quite effectively (Love and others, 2015).

Finally, the Holt-Winters algorithm imposes a constraint that SQ must be zero mean at any moment. This constraint provides a unique solution because without such a constraint, SV and SQ may drift in equal and opposite directions, but the actual prediction, and therefore DIST, would remain unchanged. There is no physical reason for SQ to be zero mean. Indeed, SuperMAG tends to produce an SQ distribution that is not, in general, zero mean, but one where the mode of the non-Gaussian SQ distribution coincides with the mode of the non-Gaussian SV distribution.

Figures 5 and 6 present similar data and results for the BOU mid-latitude magnetic observatory in Colorado, and figures 7 and 8 provide the information for the CMO high-latitude magnetic observatory in Fairbanks, Alaska. The figures demonstrate that the Holt-Winters algorithm works well across a range of magnetic latitudes, or at least that its performance suffers from a magnetic latitude dependence in a manner similar to SuperMAG. The CMO SQ estimates for both techniques are severely impacted by intense, quasi-periodic magnetic disturbances, or substorms. The SuperMAG SQ curve is certainly smoother, as should be expected based on the reasons stated previously. However, there is little reason to believe SuperMAG is any more representative of a true SQ variation than the SQ curve generated with Holt-Winters. In fact, with a carefully chosen Z-score threshold for Holt-Winters, it may be possible to exclude much of the quasi-periodic signal caused by substorms from estimates of SQ. Indeed, the Holt-Winters SQ curve appears more stable through certain periods of enhanced substorm activity studied here than the SQ curve for SuperMAG, which seems to react to enhanced activity if the activity persists for more than a few days (for example, the days following the March 17, 2015 storm). This inevitably leads to underestimates of the DIST signal during active periods, although to be fair, these discrepancies are tiny in comparison to DIST itself.

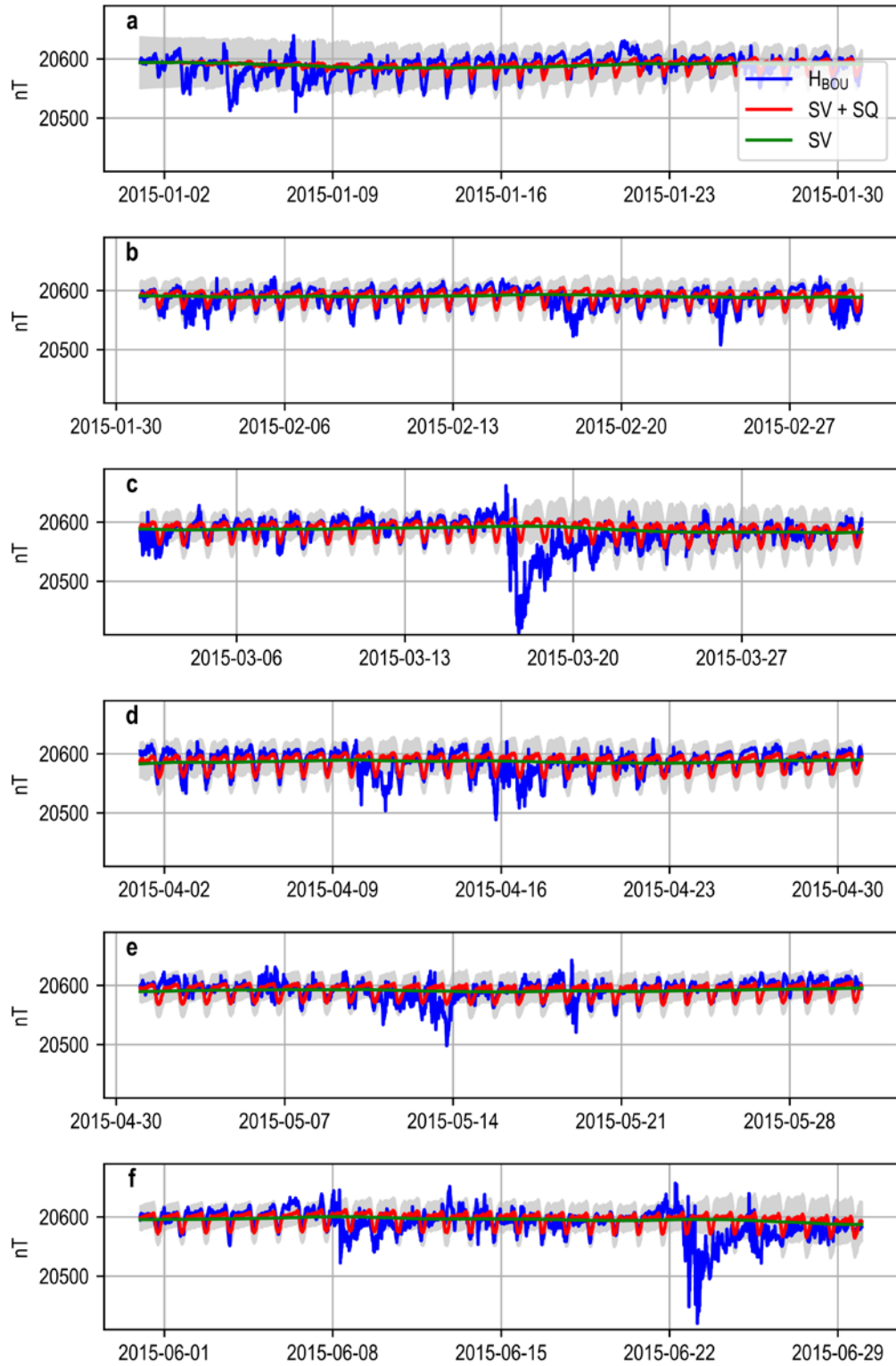


Figure 5. Holt-Winters decomposition of magnetically northward horizontal (H) geomagnetic variations at BOU magnetic observatory, near Boulder, Colorado. (in figure: nT, nanotesla; SV, secular variation; SQ, solar quiet)

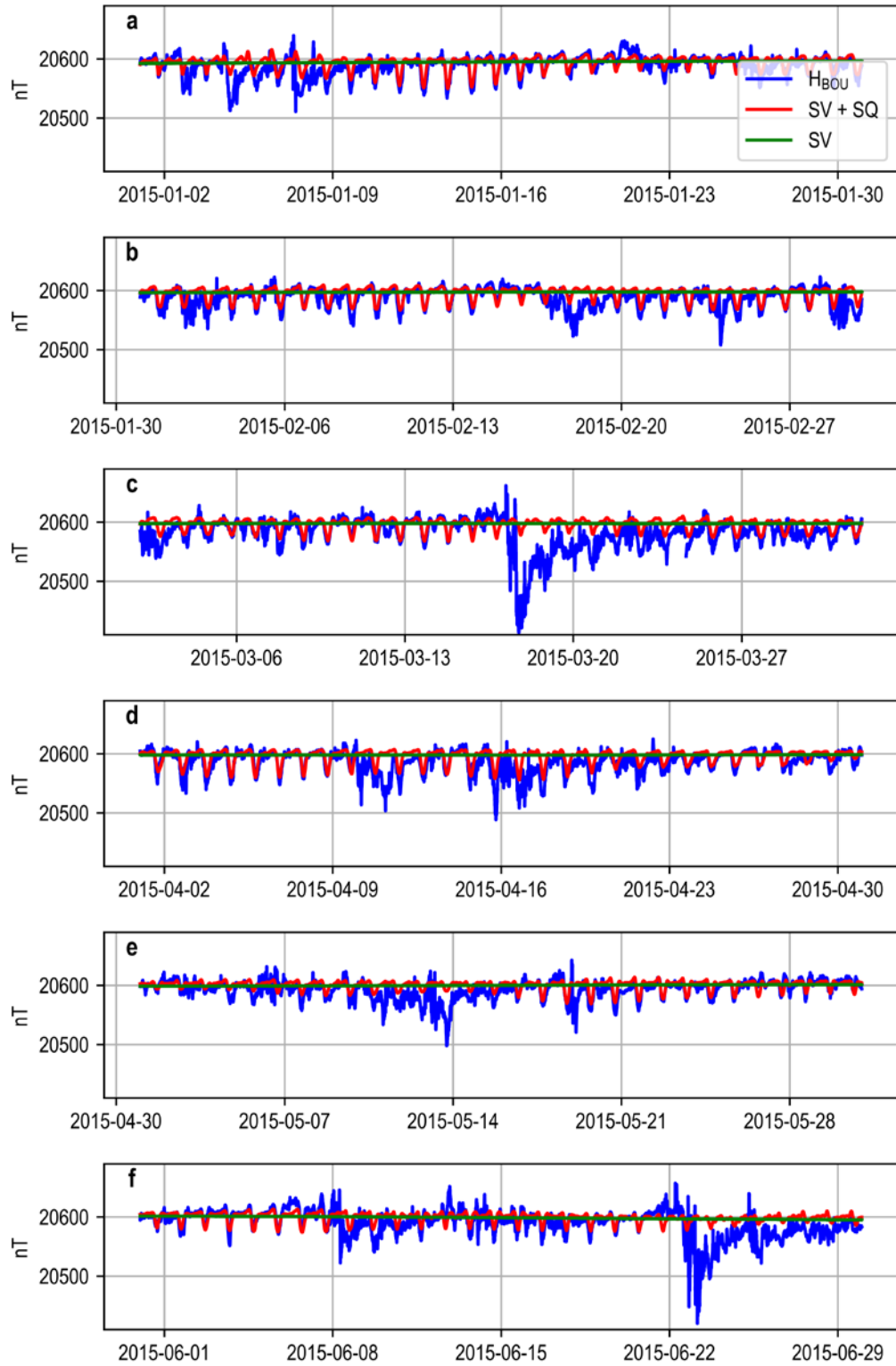


Figure 6. SuperMAG decomposition of magnetically northward horizontal (H) geomagnetic variations at BOU magnetic observatory, near Boulder, Colorado. (in figure: nT, nanotesla; SV, secular variation; SQ, solar quiet)

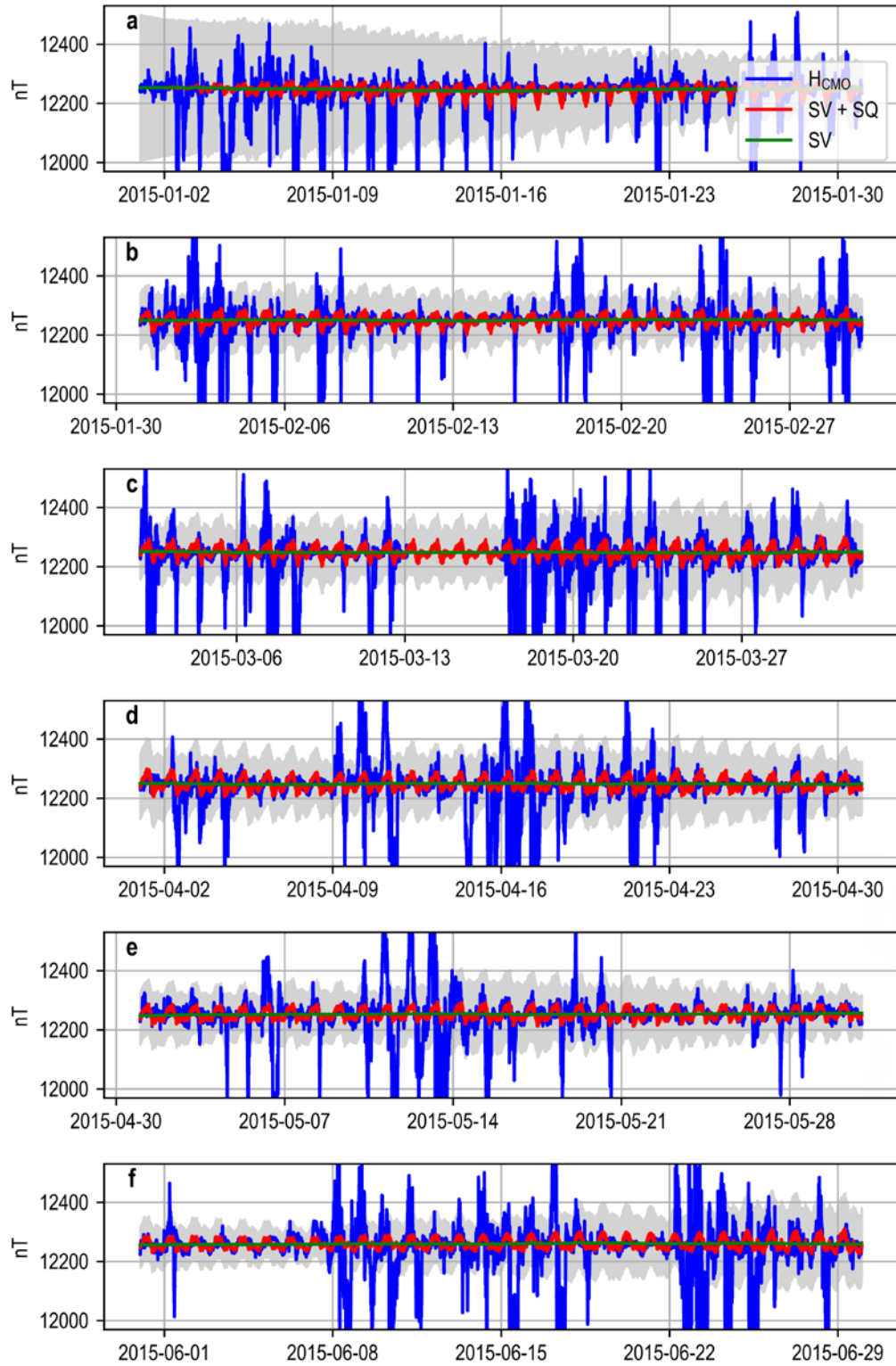


Figure 7. Holt-Winters decomposition of magnetically northward horizontal (H) geomagnetic variations at CMO magnetic observatory, near Fairbanks, Alaska. (in figure: nT, nanotesla; SV, secular variation; SQ, solar quiet)

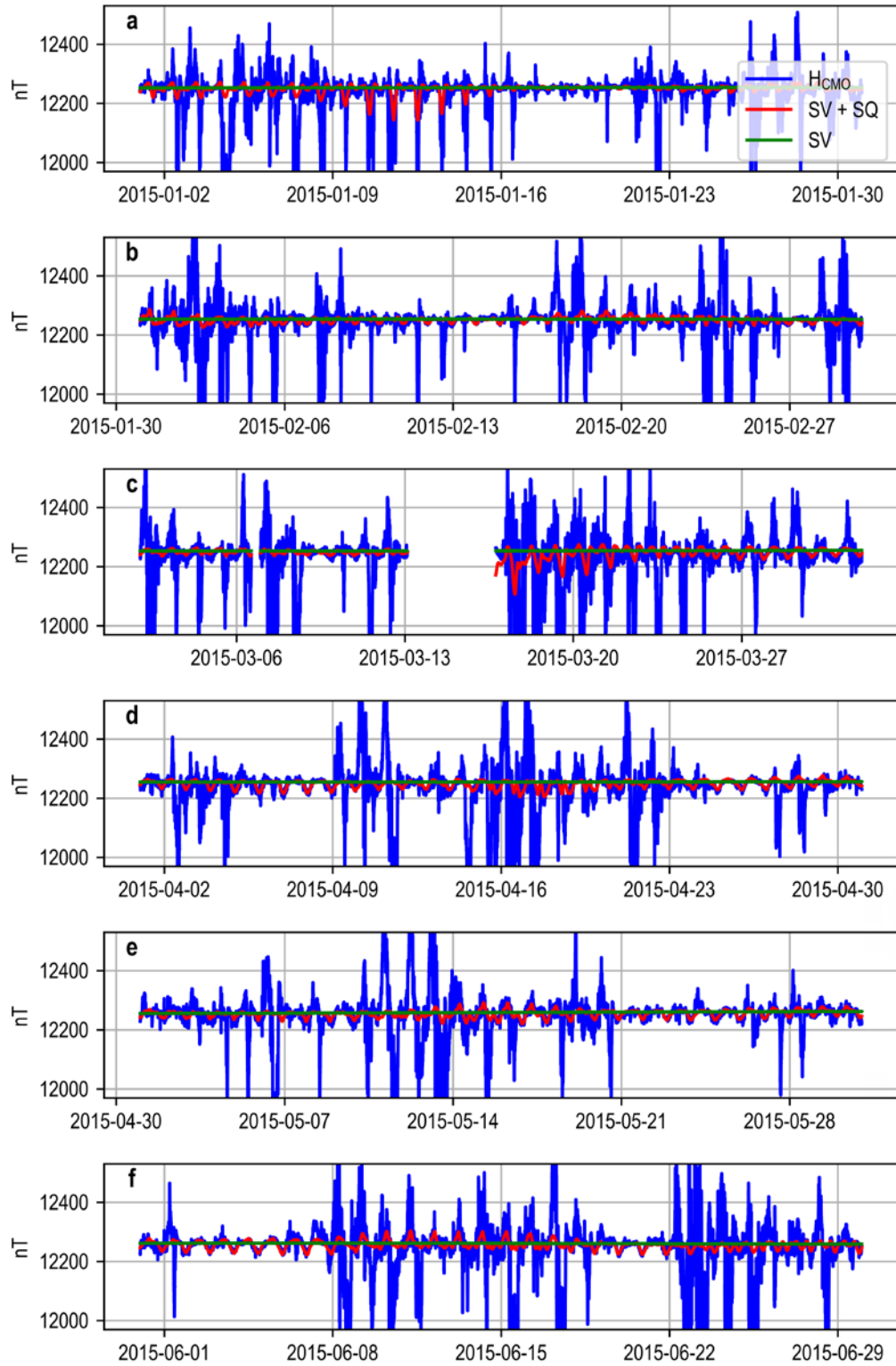


Figure 8. SuperMAG decomposition of magnetically northward horizontal (H) geomagnetic variations at CMO magnetic observatory, near Fairbanks, Alaska. (in figure: nT, nanotesla; SV, secular variation; SQ, solar quiet)

What Next?

The Holt-Winters algorithm does have certain disadvantages. Some of these disadvantages, such as the inherent lag in its time-causal response, must be weighed against the already noted advantage of being able to handle real-time data as input. Other disadvantages, such as the noisiness of the SQ signal compared to the SQ signal obtained from SuperMAG, warrant closer inspection. From a statistical point of view, each element of SQ is its own time series that corresponds to a given local time, with a sampling interval of 1 day. This constitutes a multivariate state with 1,440 variables, which in turn requires a 1440×1440 covariance matrix to describe its uncertainty. The Holt-Winters algorithm implicitly assumes this covariance matrix is diagonal (that is, the uncertainty of each state is uncorrelated with every other state), but we know from experience that geomagnetic time series exhibit autocorrelations that correspond to a coherence in local time, so the covariance matrix should contain off-diagonal elements. If we had such a matrix, we could statistically smooth—or more formally, regularize—SQ. However, obtaining this matrix is nontrivial, and inverting it is computationally expensive, possibly prohibitively so.

Fortunately, a simpler, but ad hoc, addition to our Holt-Winters algorithm can be designed to distribute the impact of each DIST across a range of SQ elements when SQ is being updated. Preliminary tests suggest this addition can effectively smooth SQ in the neighborhood of the current observation, but further testing is needed before this can be put into production. There is also recent work that extends the concept of exponential smoothing to a “general seasonality” (Hanzák, 2012). The idea here is to define a set of periodic, real-valued basis functions with associated scale parameters whose amplitudes update according to exponential smoothing. This begins to resemble a real-time, adaptive variant of more traditional methods used to estimate SQ and would be most valuable when input observations happen at arbitrary times, not as part of a regular time series.

There are certainly additional improvements that could be made because exponential smoothing belongs to a general family of forecast tools that has been built up via decades of advances by mathematicians and scientists working across many disciplines. However, ours is a deliberately simple algorithm written with emphasis on clarity, robustness, and efficiency. Any improvements are likely to be at the software level and designed to either increase speed or address specific operational issues.

The direct demand for SV, SQ, and DIST time series is indicated by the popularity of the SuperMAG project, but these time series are intended primarily as inputs to existing and planned real-time geomagnetic data products at the USGS. The most immediate is the USGS’ real-time, 1-minute resolution Dst (disturbance storm time) index, which is only slightly more sophisticated than a weighted average of four low-latitude observatories’ magnetically northward DIST. Dst is a proxy for the intensity of the magnetospheric ring current, and is the most widely used indicator of geomagnetic storm intensity.

In addition to Dst, the so-called K indices from the observatories (Bartels and others, 1939; Menvielle and others, 1995) also require removal of SV and SQ before binning, transforming DIST to a quasi-logarithmic scale, and merging into Kp, a geomagnetic planetary index used to indicate a general geospace disturbance (Menvielle and Berthelier, 1991). In addition, the official AE (auroral electrojet) (Davis and Sugiura, 1966) index does not generally remove SQ because it is a nonlinear combination of multiple observatories’ data that does not exhibit much of a repeatable diurnal signal. A future, real-time AE may emulate the official version by simply combining SQ and DIST, or possibly taking the sensible step of leaving out SQ in order to provide a more robust measure of high-latitude magnetic activity.

In addition to their contribution to geomagnetic indices, SQ and (or) DIST can be inverted to obtain equivalent ionospheric currents. These currents can then be used to calculate magnetic disturbance between observatories, thus providing a physics-inspired geomagnetic field interpolation method. The USGS is transitioning a research software package on the basis of spherical elementary

currents (Amm and Viljanen, 1999; Pulkinnen and others, 2003) into a real-time gridded map of magnetic disturbance throughout the continental United States. Eventually, this map will be combined with gridded maps of surface impedances to predict geoelectric field variations across the contiguous United States.

Summary and Conclusions

The theoretical basis for Holt-Winters exponential smoothing algorithm with “seasonal” corrections was summarized, implemented as a Python software package, and validated using both synthetic and real-world geomagnetic time series as input. The results indicate that this algorithm can effectively decompose real-time geomagnetic data into a secularly varying baseline, a periodic solar quiet signal, and a remainder that constitutes the aperiodic, mostly unpredictable, magnetic disturbance of greatest concern as a space weather hazard. This algorithm is certainly not the first to do this for geomagnetic time series, but it is the first to be designed from the beginning with real-time data processing in mind. The algorithm is (1) time-causal, so it does not rely on knowledge of future observations to estimate the current state; (2) recursive, so it does not require a tremendous amount of data to be ingested for each desired output; and (3) adaptive, so it adjusts to evolving background conditions with an effective “memory” that can be configured at runtime. Finally, the algorithm is very computationally efficient, with a simple but rigorous underlying mathematical framework that allows its performance to be extrapolated beyond just the most recent observations.

References Cited

- Amm, O., and Viljanen, A., 1999, Ionospheric disturbance magnetic field continuation from the ground to the ionosphere using spherical elementary current systems: *Earth Planets Space*, v. 51, p. 431–440, accessed March 17, 2017, at <https://www.terrapub.co.jp/journals/EPS/pdf/5106/51060431.pdf>.
- Archibald, B.C., and Koehler, A.B., 2003, Normalization of seasonal factors in Winters' methods: *International Journal of Forecasting*, v. 19, no. 1, p. 143–148, accessed March 17, 2017, at [https://doi.org/10.1016/S0169-2070\(01\)00117-0](https://doi.org/10.1016/S0169-2070(01)00117-0).
- Bartels, J., Heck, N.H., and Johnston, H.F., 1939, The three-hour-range index measuring geomagnetic activity: *Terrestrial Magnetism and Atmospheric Electricity*, v. 44, no. 4, p. 411–454, accessed March 17, 2017, at <https://doi.org/10.1029/TE044i004p00411>.
- Brown, R.G., 1956, *Exponential smoothing for predicting demand*: Cambridge, Mass., Arthur D. Little, Inc., 15 p., accessed March 17, 2017, at <https://www.industrydocumentslibrary.ucsf.edu/tobacco/docs/#id=jzlc0130>.
- Brown, R.G., 1963, *Smoothing forecasting and prediction of discrete time series*: Englewood Cliffs, N.J., Prentice-Hall, 468 p.
- Davis, T.N., and Sugiura, M., 1966, Auroral electrojet activity index AE and its universal time variations: *Journal of Geophysical Research*, v. 71, no. 3, p. 785–801, accessed March 17, 2017 at <https://doi.org/10.1029/JZ071i003p00785>.
- Efron, B., and Tibshirani, R.J., 1986, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy: *Statistical Science*, v. 1, no. 1, p. 54–75, accessed March 17, 2017, at <https://doi.org/10.1214/ss/1177013815>.
- Gannon, J.L., Love, J.J., Friberg, P.A., Stewart, D.C., and Lisowski, S.W., 2011, *U.S. Geological Survey Near Real-Time Dst Index: USGS Open-File Report 2011–1030*, 10 p., accessed March 17, 2017, at <https://pubs.usgs.gov/of/2011/1030/>.
- Gardner, E.S., 2006, Exponential smoothing—The state of the art, Part II: *International Journal of Forecasting*, v. 22, no. 4, p. 637–666, accessed March 17, 2017, at <https://doi.org/10.1016/j.ijforecast.2006.03.005>.
- Gjerloev, J.W., 2012, The SuperMAG data processing technique: *Journal of Geophysical Research*, v. 117, no. A9, 19 p., accessed March 17, 2017, at <https://doi.org/10.1029/2012JA017683>.
- Hanzák, T., 2012, Holt-Winters method with general seasonality: *Kybernetika*, v. 48, no. 1, p. 1–15, accessed March 17, 2017, at <http://dml.cz/dmlcz/142058>.
- Holt, C.C., 1957, Forecasting seasonals and trends by exponentially weighted averages: *International Journal of Forecasting*, v. 20, no. 1, p. 5–10, accessed March 17, 2017, at <https://doi.org/10.1016/j.ijforecast.2003.09.015>.
- Hyndman, R.J., and Athanasopoulos, G., 2013, *Forecasting—Principles and practice*: OTexts, accessed September 1, 2016 at <http://www.otexts.org/fpp/>.
- Hyndman, R.J., Koehler, A.B., Ord, J.K., and Snyder, R.D., 2005, Prediction intervals for exponential smoothing using two new classes of state space models: *Journal of Forecasting*, v. 24, no. 1, p. 17–37, accessed March 17, 2017, at <https://doi.org/10.1002/for.938>.
- Love, J.J., and Chuilliat, A., 2013, An international network of magnetic observatories: *Eos*, v. 94, no. 42, p. 373–384, accessed March 17, 2017, at <https://doi.org/10.1002/2013EO420001>.
- Love, J.J., and Gannon, J.L., 2009, Revised Dst and the epicycles of magnetic disturbance—1958–2007: *Annales Geophysicae*, v. 27, no. 8, p. 3101–3131, accessed March 17, 2017, at <https://doi.org/10.5194/angeo-27-3101-2009>.
- Love, J.J., and Rigler, E.J., 2014, The magnetic tides of Honolulu: *Geophysical Journal International*, v. 197, p. 1335–1353, accessed March 17, 2017, at <https://doi.org/10.1093/gji/ggu090>.

- Love, J.J., Rigler, E.J., Pulkkinen, A., and Riley, P., 2015, On the lognormality of historical magnetic storm intensity statistics—Implications for extreme-event probabilities: *Geophysical Research Letters*, v. 42, p. 6544–6553, accessed March 17, 2017, at <https://doi.org/10.1002/2015GL064842>.
- McCollough, J.P., Young, S.L., Rigler, E.J., and Simpson, H.A., 2015, Real-time validation of the Dst Predictor Model: Air Force Research Laboratory Technical Report, AFRL-RV-PS-TR-2015-0073, 17 p.
- Menvielle, M., and Berthelier, A., 1991, The K-derived planetary indices—Description and availability: *Reviews of Geophysics*, v. 29, no. 3, p. 415–432, accessed March 17, 2017, at <https://doi.org/10.1029/91RG00994>.
- Menvielle, M., Papitashvili, N., Häkkinen, L., and Sucksdorff, C., 1995, Computer production of K indices—Review and comparison of methods: *Geophysical Journal International*, v. 123, p. 866–886, accessed March 17, 2017, at <https://doi.org/10.1111/j.1365-246X.1995.tb06895.x>.
- Pulkkinen, A., Amm, O., and Viljanen, A., and BEAR Working Group, 2003, Ionospheric equivalent current distributions determined with the method of spherical elementary current systems: *Journal of Geophysical Research*, v. 108, no. A2, 9 p., accessed March 17, 2017, at <https://doi.org/10.1029/2001JA005085>.
- Sugiura, M., 1964, Hourly values of equatorial Dst index for the IGY: *Annals of the International Geophysical Year*, v. 35, p. 9–45.
- Winters, P. R., 1960, Forecasting sales by exponentially weighted moving averages: *Management Science*, v. 6, p. 324–342, accessed March 17, 2017, at <https://doi.org/10.1287/mnsc.6.3.324>.

Glossary

Magnetic disturbance. The aperiodic, short-term variations that remain after subtracting secular variation and solar quiet signals from geomagnetic measurements.

Secular variation. The slowly evolving geomagnetic baseline about which shorter-term periodic and aperiodic geomagnetic variations happen. Ideally, it is only the contribution from Earth's internal main field, but in practice, it often includes the effects of external currents systems or even instrumental artifacts.

Solar quiet. The periodic, mostly daily, variation of geomagnetic data that may be considered a magnetic "tide." At lower to middle geomagnetic latitudes, it is primarily caused by the so-called solar quiet ionospheric current system, but the term is commonly used to describe periodic variations seen at higher latitudes as well.

Appendix 1. Pseudocode.

Algorithm 1. Holt-Winters Exponential Smoother.

Parameters: $m \in \mathbb{R}^1$,
 $\alpha \in \mathbb{R}^1, \beta \in \mathbb{R}^1, \gamma \in \mathbb{R}^1, \phi \in \mathbb{R}^1$,
 $h \in \mathbb{R}^1, zThresh \in \mathbb{R}^1$

Inputs: $yObs \in \mathbb{R}^N$,
 $yHat0 \in \mathbb{R}^h, \ell0 \in \mathbb{R}^1, b0 \in \mathbb{R}^1, s0 \in \mathbb{R}^m$,
 $\sigma0 \in \mathbb{R}^{h+1}$

Outputs: $yHatOut \in \mathbb{R}^N, sOut \in \mathbb{R}^N, \sigmaOut \in \mathbb{R}^N$,
 $yHat0 \in \mathbb{R}^h, \ell0 \in \mathbb{R}^1, b0 \in \mathbb{R}^1, s0 \in \mathbb{R}^m$,
 $\sigma0 \in \mathbb{R}^{h+1}$

Assumes: All arrays use zero-based indexing

```
# allocate arrays
1.  $yHat \in \mathbb{R}^{N+h}$ 
2.  $s \in \mathbb{R}^{N+m}$ 
3.  $r \in \mathbb{R}^{N+1}$ 
4.  $\sigma \in \mathbb{R}^{N+h+1}$ 

# initialize arrays
5.  $yHat[0..h-1] \leftarrow yHat0$ 
6.  $s[0..m-1] \leftarrow s0$ 
7.  $\sigma[0..h] \leftarrow \sigma0$ 

# initialize scalars
8.  $\phiHminus1 \leftarrow 0$ 
9.  $sumC2\_H \leftarrow 1$ 
10. FOR  $i=\{1..h\}$  DO
11.    $\phiHminus1 \leftarrow \phiHminus1 + \phi**(i-1)$ 
12.    $sumC2\_H \leftarrow sumC2\_H + (\alpha * (1 + \phiHminus1 * \beta) +$   

    $\gamma * (1 \text{ IF } (i \% m == 0) \text{ ELSE } 0)**2$ 
13. ENDFOR
14.  $\phiJminus1 \leftarrow \phiHminus1$ 
15.  $sumC2 \leftarrow sumC2\_H$ 
16.  $j \leftarrow h$ 
```

Algorithm 1. Holt-Winters Exponential Smoother.

```
# main loop
17. FOR  $i=\{0 \dots N-1\}$  DO
    # reestimate error variance
18.  $\sigma^2 \leftarrow \sigma[i]**2$  IF  $j == h$ 

    # predict  $h$ -step error standard deviation
19.  $\sigma[i+h+1] \leftarrow \sqrt{\sigma^2 * \text{sumC2}}$ 

    # predict  $h$ -steps ahead
20.  $\hat{y}[i+h] \leftarrow \ell + \phi_{Hminus1} * b + s[i + h \% m]$ 

    # determine prediction residual
21.  $et \leftarrow y_{Obs}[i] - \hat{y}[i]$ 

    # propagate smoothed variables if no valid observation,
    # or if error threshold is exceeded
22. IF ( $\text{isnan}(et)$  OR  $\text{abs}(et) > zThresh * \sigma[i]$ ) THEN
    # propagate seasonal corrections
23.  $r[i+1] \leftarrow r[i]$ 
24.  $s[i+m] \leftarrow s[i]$ 

    # propagate level and slope
25.  $\ell \leftarrow \ell + \phi * b$ 
26.  $b \leftarrow \phi * b$ 

    # propagate/evolve error standard deviation if no valid obs.
27. IF ( $\text{isnan}(et)$ ) THEN
    #  $\phi_{Hminus1}$ ,  $\text{sumC2}$  and  $j$  will be reset with next valid obs.
28.  $\phi_{Hminus1} \leftarrow \phi_{Hminus1} + \phi ** j$ 
29.  $\text{sumC2} \leftarrow \text{sumC2} + (\alpha * (1 + \phi_{Hminus1} * \beta)) +$ 
     $\gamma * (1 \text{ IF } (j \% m == 0) \text{ ELSE } 0) ** 2$ 
30.  $j \leftarrow j + 1$ 
    # update error standard deviation even if error threshold exceeded
31. ELSE
    # this is identical to  $\sigma$  update below
32.  $\sigma[i+1] \leftarrow \alpha * \text{abs}(et) + (1 - \alpha) * \sigma[i]$ 
33. ENDIF
```

Algorithm 1. Holt-Winters Exponential Smoother.

```
# update smoothed variables if valid observations
34. ELSE
    # update seasonal corrections
35.    $r[i+1] \leftarrow \gamma * (1 - \alpha) * et / m$ 
36.    $s[i+m] \leftarrow s[i] + \gamma * (1 - \alpha) * et$ 

    # update level and slope
37.    $\ell \leftarrow \ell + \phi * b + \alpha * et$ 
38.    $b \leftarrow \phi * b + \alpha * \beta * et$ 

    # this is identical to sigma update above
39.    $\sigma[i+1] \leftarrow \alpha * \text{abs}(et) + (1 - \alpha) * \sigma[i]$ 

    # reset prediction interval variables
40.    $\text{sumC2} \leftarrow \text{sumC2} - H$ 
41.    $\phi J_{\text{minus}1} \leftarrow \phi H_{\text{minus}1}$ 
42.    $j \leftarrow h$ 
43. ENDIF
44. ENDFOR

# adjust seasonal corrections so they would have had a zero mean when applied
# (this cumulative summation could have been done easily inside the main loop,
# but it might be made more efficient with optimized library functions)
45. FOR  $i = \{1 \dots N\}$  DO

46.
47.    $r[i] \leftarrow r[i-1] + r[i]$ 
48.    $s[i-1] \leftarrow s[i-1] - r[i-1]$ 
49. ENDFOR

# adjust the seasonal corrections to be passed back as s0
50. FOR  $i = \{N \dots N+m-1\}$  DO
51.    $s[i] \leftarrow s[i] - r[N]$ 
52. ENDFOR

# main output variables; all synchronous with input yObs
53.  $y\text{HatOut} \leftarrow y\text{Hat}[0 \dots N-1]$ 
54.  $s\text{Out} \leftarrow s[0 \dots N-1]$ 
55.  $\sigma\text{Out} \leftarrow \sigma[1 \dots N]$ 

# initial variables for subsequent calls
56.  $y\text{Hat0} \leftarrow y\text{Hat}[N \dots N+h-1]$ 
57.  $s0 \leftarrow s[N \dots N+h+m-1]$ 
58.  $\ell0 \leftarrow \ell + r[N]$ 
59.  $b0 \leftarrow b$ 
60.  $\sigma0 \leftarrow \sigma[N \dots N+h+1-1]$ 
```
