Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us
what having access to this work means to you and why it's important to you. Thank you.

Geomagnetic secular variation forecast using the NASA GEMS ensemble Kalman filter: A candidate model for IGRF 2020

Andrew Tangborn (Corresponding author), Joint Center for Earth Systems Technology, University of Maryland Baltimore County, Baltimore, MD, USA,tangborn@umbc.edu

Weijia Kuang, Geodesy and Geophysics Laboratory, Goddard Space Flight Center, Greenbelt, MD, USA, Weijia.Kuang-1@nasa.gov

Terence J. Sabaka, Geodesy and Geophysics Laboratory, Goddard Space Flight Center, Greenbelt, MD, USA, Terence.J.Sabaka@nasa.gov

Ce Yi, Geodesy and Geophysics Laboratory, Goddard Space Flight Center, Greenbelt, MD, USA, Ce.Yi@nasa.gov

Abstract

We have produced a 5 year mean secular variation (SV) of the geomagnetic field for the period 2020-2025. We use the NASA Geomagnetic Ensemble Modeling System (GEMS), which consists of the NASA Goddard geodynamo model and ensemble Kalman filter (EnKF) with 512 ensemble members. Geomagnetic field models are used as observations for the assimilation, including *gufm1* (1590-1960), CM4 (1961-2000) and CM6 (2001-2019). The forecast involves a bias correction scheme that assumes that the model bias changes on timescales much longer than the forecast period, so that they can be removed by successive forecasts. The algorithm was validated on the time period 2010-2015 by comparing with the 2015 IGRF before being applied to the 2020-2025 time period. This forecast has been submitted as a candidate model for IGRF 2025.

1. Introduction

It has long been observed that the Earth's main magnetic field changes slowly in time (Bullard et al 1950; Hide 1967; Bloxham & Gubbins 1985; Jackson et al 2000; Lund 2018). This change, called the secular variation (SV) originates dominantly from the the Earth's iron-rich fluid outer core that is in vigorous convection, driven by thermo-chemical buoyancy released from Earth's secular cooling and differentiation (geodynamo) (e.g. Lund and Olson 1987; Braginsky and Roberts 1996). Over the past 25 years, great progress has been made in numerically modeling the geodynamo (Glatzmaier and Roberts, 1995; Kagyama and Sato 1997; Kuang and Bloxham, 1997; Christensen et al 2001; Jones et al 2011; Matsui et al 2016). These models can reproduce much of the qualitative aspects of the Earth's magnetic field, including dipole dominance, westward drift and occasional reversals. Although it remains computationally prohibitive to simulate the geodynamo in the parameter regimes appropriate for the Earth's core, the asymptotic properties emerged from numerical solutions with wide range of dynamo parameter values open the door for quantitative applications of numerical models to the geomagnetic field (e.g. Christensen and Aubert 2006; Aubert et al 2017; Kuang et al 2017).

One such application is to use a numerical model to predict geomagnetic SV on time scales of several years and longer, if proper initialization is made using data assimilation techniques, as is done in numerical weather prediction (NWP). Forecast results from this geomagnetic data assimilation approach have already been used as predictive field candidates for previous IGRF models (Kuang, et al., 2010; Fournier, et al., 2015).

Obviously, this approach is different from other SV forecasts made without utilization of dynamo models. For example, Olsen, et al. (2010) modeled the time dependence of the Gauss coefficients with order 6 B-splines so that the predictive SV for 2010-2015 can be extrapolated from the coefficients derived from observations prior to 2010. While utilization of geodynamo models can take advantage of adding dynamically consistent motional induction in the outer core to make SV forecasts, it also carries additional complications, such as model bias that results from the mismatch between the parameter values used in the models and appropriate for the Earth's core, and propagation of observational errors in both time and space. These difficulties have been analyzed through assimilation experiments with simplified systems (Sun et al 2007; Fournier et al 2007), observation system simulation experiments (OSSEs) (Liu et al 2007; Aubert and Fournier, 2011; Fournier, 2013). It is expected that improvements to both modeling and assimilation techniques will make geomagnetic forecast a growing contributor of the IGRF candidate models.

In this paper, we report our mean SV forecast for the period from 2020 to 2025. This effort differs from our earlier work for IGRF candidate model (Kuang et al 2010) in two major areas: we assimilate the observed gauss coefficients from a single geomagnetic field model (Sabaka et al in this special issue) for the period from 1960 to 2019; and the error covariance in our assimilation is determined through a large ensemble of assimilation solutions, and is updated in each analysis cycle. The two improvements can help produce more accurate SV forecasts and uncertainty estimates, since, for example, the utilization of a single field model can avoid any potential disagreement in different field models that may lead to inconsistent observation error estimates, and because the time varying error covariance can provide more consistent analysis for each cycle.

This paper is organized as follows: the algorithm of our forecasts is given in Section 2; testing and validation are presented in Section 3, followed by the forecasts and discussion.

2. Geodynamo model and data assimilation algorith

The geomagnetic data assimilation system used for our SV forecasts is the NASA Geomagnetic Ensemble Modeling System (GEMS). It consists of a numerical geodynamo model (Kuang et al 1999, 2003), an ensemble Kalman filter (Evensen, 2009; Sun and Kuang 2015). The model solves a set of nonlinear magnetohydrodynamic equations in a spherical shell domain. Details of the model are given in Kuang and Bloxham (1999), Kuang and Chao (2003) and Jiang and Kuang (2008). The core state in this model is defined with the velocity $\mathbf{v}$, the magnetic field $\mathbf{B}$ and the relative density anomaly $\Delta\rho$. These state variables are described by by spherical harmonic expansions in the co-latitude θ and the longitude ϕ directions, with the spherical harmonic coefficients given on discrete radial grid points. If we denote by $\mathbf{x}$ the vector of these coefficients, then the dynamo system can be symbolically represented as

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{M}(\mathbf{x}) \tag{1}$$

where M includes all linear and nonlinear model operators. The GEMS system employs a sequential ensemble Kalman filter (EnKF), described below, which is the result of many years of development that started with a simple one dimensional system (Sun et al., 2007) and an optimal interpolation (OI) system (Kuang et al., 2008, 2009). These systems have shown how geomagnetic data assimilation can be used to estimate geodynamo model parameters and geomagnetic field model uncertainty (Tangborn and Kuang, 2015, 2018).

The observations used for GEMS are the gauss coefficients $\{g_l^m, h_l^m\}$ (where l and m are spherical harmonic degrees and orders) from geomagnetic field models. They represent only the internally generated magnetic field. In the outer core and electrically conducting lower mantle (called the D''-layer), the magnetic field can be decomposed into the toroidal ($\mathbf{B}_T$) and the poloidal ($\mathbf{B}_P$ components:

$$\mathbf{B} = \nabla \times (T_B)\hat{\mathbf{r}} + \nabla \times \nabla \times (P_B)\hat{\mathbf{r}} = \mathbf{B}_T + \mathbf{B}_P \tag{2}$$

where $\hat{\mathbf{r}}$ is the radial unit vector, and T_B, P_B are the toroidal and poloidal scalars and are described in GEMS as

$$\begin{bmatrix} P_b \\ T_b \end{bmatrix} = \sum_{m=0}^{M} \sum_{l=m}^{L} \begin{bmatrix} b_l^m(r,t) \\ j_l^m(r,t) \end{bmatrix} Y_l^m(\theta,\phi) \; + \; C.C. \,, \tag{3}$$

4

where $\{Y_l^m\}$ are the orthonormal spherical harmonic functions, and $C.C.$ denotes the complex conjugate part of the expansion (similar expansions are also made for $\mathbf{v}$ and $\Delta\rho$). At the top of the D''-layer, i.e. at $r = r_d$, the toroidal field vanishes, i.e. $j_l^m = 0$; but the poloidal field coefficients b_l^m can be matched with the observations

$$b_l^m(r_d) = \frac{r_s^2}{2l}\left(\frac{r_s}{r_d}\right)^l (g_l^m - ih_l^m) \qquad \text{for} \quad m \le l \le L_{obs}, \tag{4}$$

where r_s is the mean Earth surface radius, and L_{obs} is the highest degree of the observed geomagnetic field. In the rest of this paper, we denote the poloidal spectral coefficients in (4) that are derived from the field models by $(b_l^m)^o$ and the forecast coefficients $(b_l^m)^f$. The observations $\mathbf{y}^o = [(b_l^m)^o]$ are assimilated into the forecast $(\mathbf{x})^f$ at analysis time t_a to produce the analysis $\mathbf{x}^a$:

$$\mathbf{x}^a(t_a) = \mathbf{x}^f(t_a) + \mathbf{K}\left[\mathbf{y}^o - \mathbf{H}\cdot\mathbf{x}(t_a)^f\right] \tag{5}$$

where $\mathbf{H}$ is the observation operator matrix and $\mathbf{K}$ is the gain matrix, defined by:

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T\left[\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right]^{-1} \tag{6}$$

where $\mathbf{P}^f$ is the forecast error covariance and $\mathbf{R}$ is the observation error covariance. The analysis $\mathbf{x}^a$ is then used as the initial state of the dynamo system 1 for the next forecast. The forecast error covariance is computed from an ensemble of forecasts with N_{ens} members (Sun and Kuang 2015)

$$\mathbf{P}^f = \left\langle (\mathbf{x}^f - \mu_{\mathbf{x}}^f)(\mathbf{x}^f - \mu_{\mathbf{x}}^f)^T\right\rangle \tag{7}$$

where $\mu_{\mathbf{x}}^f$ is the ensemble mean forecast state. In our forecasts here, we take only the diagonal part of the observation error covariance $\mathbf{R}$ and is defined simply as the estimated variance of the field models, $(\sigma_l^m)^o$. The time dependent model for the observation error standard deviation is

$$(\sigma_l^m)^o = \alpha b_l^m(l)(l+1)/2 \tag{8}$$

where $\alpha = 0.1$ before 1900, $0.001 \le \alpha \le 0.1$ with an exponential decrease from 1900 to 2000, and

$\alpha = 0.001$ afterwards. This observation error model accounts for the increased accuracy of geomagnetic models at later times, as well as the higher relative errors for the higher degrees. It has been tested through a number of numerical experiments with the goal of minimizing forecast errors, and is therefore regarded as an appropriate estimate relative to our geodynamo forecasts.

The geomagnetic field models that are used to generate the gauss coefficients for our assimilation, and their maximum degree assimilated (L_{obs}) are

$$
\begin{array}{llll}
\text{Observations Period} & & L_{obs} \\
gufm1: & 1590 - 1960 & 4 \\
\text{CM4}: & 1960 - 2000 & 8 \\
\text{CM6}: & 2000 - 2019 & 13
\end{array}
\tag{9}
$$

Details of the first two models can be found in Jackson et. al. (2000) and Sabaka et al. (2004). Details of the last field model can be found in Sabaka et al (2020). We only summarize here that this model uses a "comprehensive inversion" technique which co-estimates parameters for the Earth's internal field, using estimates of systematic and random errors, so as to produce an optimal separation of the internal and external fields.

The ensemble is initialized in 1590 CE using a long free model run, and subsequent analyses are computed every 20 years until 2000 CE, when the analysis cycle is reduced to one year (the start of the forecast period using the CM6 model). $\mathbf{P}^f$ is updated with the forecasts at each analysis time t_a, but $\mathbf{H}$ is only updated if different L_{obs} is chosen for making the analysis. We use an ensemble of 512 for these experiments, and we have found this size to be essentially converged relative to smaller ensembles of 256 or 400. This means that an ensemble larger than 512 is not expected to produce appreciably more accurate forecasts.

Geomagnetic forecasts using a geodynamo model will naturally contain significant forecast errors due to mismatches in the dynamo parameter values and numerical approximations, and these can grow significantly during a multi-year forecast. But because these error sources vary on relatively long timescales, we can assume that the model error itself varies on time scales much longer than, e.g. the 5-year forecast period. This assumption allows us to implement a bias correction scheme for the SV prediction by producing a set of staggered forecasts, and then take differences between them to remove

the bias (Kuang et al 2010). The forecast of the observed poloidal field is defined as the projection of the state vector, $\mathbf{x}^f$ onto the observation space:

$$\mathbf{y}^f = \mathbf{H}\mathbf{x}^f \tag{10}$$

We define the forecast error, $\epsilon(t)$, as the difference between the forecast state, $\mathbf{y}^f(t)$, and the true state, $\mathbf{y}^t(f)$:

$$\mathbf{y}^f(t) = \mathbf{y}^t(t) + \epsilon(t). \tag{11}$$

The forecast error is a combination of model error and the growth of initial state errors, but we assume that model errors remain the largest component. Consider a forecast that is initialized from an analysis at time t_a and produces a forecast at time $t_a + \delta t$. We estimate the SV from the change in the geomagnetic field after time δt:

$$\mathbf{y}^f(t_a + \delta t) - \mathbf{y}^a(t_a) = \mathbf{y}(t_a + \delta t) - \mathbf{y}(t_a) + \epsilon(t_a + \delta t) - \epsilon(t_a) \approx (\dot{\mathbf{y}} + \dot{\epsilon})\delta t \tag{12}$$

where $\dot{\mathbf{y}}$ is the SV and $\dot{\epsilon}$ is the SV error growth rate. We can reduce this error term by making use of the assumption of slowly varying model error and considering staggered forecasts that begin at analysis times t_a and $\tilde{t}_a$, where $|t_a - \tilde{t}_a| << \delta t$. This gives us forecasts at times $t_a + \delta t$ and $\tilde{t}_a + \delta t$ in terms of the true state:

$$\begin{aligned} \mathbf{y}^f(t_a + \delta t) &= \mathbf{y}^t(t_a + \delta t) + \epsilon(t_a + \delta t) \\ \mathbf{y}^f(\tilde{t}_a + \delta t) &= \mathbf{y}^t(\tilde{t}_a + \delta t) + \epsilon(\tilde{t}_a + \delta t) \end{aligned} \tag{13}$$

The difference between these two sequences was shown to approximately remove the forecast error (Kuang et al. 2010), because the change in model error between $t_a + \delta t$ and $\tilde{t}_a + \delta t$ is small compared to the SV. So we can write an expression for the SV as

$$\tilde{\dot{\mathbf{y}}}^f(t) = \frac{1}{\tau_a}\left[\mathbf{H}\mathbf{x}^f(t_a + \delta t) - \mathbf{H}\mathbf{x}^f(\tilde{t}_a + \delta t)\right] \tag{14}$$

where $\tau_a = |t_a - \tilde{t}_a|$.

The details of the bias correction scheme can be found Kuang *et al.*, 2010.

3. Validation Experiments

We have validated the methodology by computing SV forecasts for the period 2009-2015. Two forecast experiments were conducted using the analysis from 2000 CE carried forward with one year forecast and analysis cycles. In the first experiment (E1), the assimilation is computed from 2000 to 2008; the analysis in 2008 is then used to make forecasts for the period 2009-2014. The second experiment (E2) is similar, but the assimilation is computed from 2000 to 2009; and the analysis in 2009 is used to make the forecast for the period 2010-2015. The final SV forecasts are then made with these two staggered forecasts:

$$\tilde{\dot{\mathbf{y}}}_{AB}(t_f) = \left[\mathbf{y}_2(t_f + \tau_a) - \mathbf{y}_1(t_f)\right]/\tau_a \tag{15}$$

where $\tau_a = 1$ year. The SV for the Gauss coefficients (g_n^m, h_n^m) are then computed as follows. We show the details for g_n^m, while h_n^m follow the same procedure. Note that this notation, traditionally used it geomagnetic field modeling, is equivalent to (g_l^m, h_l^m) which is generally used in geodynamo modeling. Because we are crossing between the two fields, we use these sets of notation interchangeably. In our assimilation, we utilize the scaled coefficients

$$G_n^m = g_n^m/g_1^0 \tag{16}$$

for the SV forecasts. We first compute G_n^m from our forecasts, and their time variation $\dot{G}_n^m$ via (15). The SV gauss coefficients can then be calculated via

$$\dot{g}_n^m = g_1^0 \dot{G}_n^m + G_n^m \dot{g}_1^0. \tag{17}$$

We then approximate the axial dipole coefficient g_1^0 in (17) with the values from the field models, e.g. CM6 (Sabaka, et al., 2020) as follows:

$$\dot{g}_1^0 = \left[(g_1^0)^{CM6}(2010) - (g_1^0)^{CM6}(2009)\right]/(1year), g_1^0(2009+t) = (g_1^0)^{CM6}(2009) + t\dot{g}_1^0. \tag{18}$$

The formulations (14)-(18) are the basis for the forecasts of the mean SV for the period 2010-2015 2015 SV which are then compared with the IGRF for each degree n and order m up to $n = 8$. In this case the years in E1 correspond to $\tilde{\tau}$ and those in E2 to τ. Some examples are shown in Figures 1-2 which

indicates the forecast SV, $\tilde{\mathbf{y}}^f$, for different n, m. The slope of each dashed line is the forecast while the solid line is the SV computed from CM6 coefficients. The coefficients shown here are representative of the range in accuracy obtained by the ensemble forecast and bias correction algorithm. The $(n = m = 1)$ and $(n = m = 2)$ coefficients in Figure 1 are typical of coefficients that predict the CM6 field model closely. The $(n = 3, m = 2)$ in Figure 2(a) is the least accurate case (for g_n^m), while $(n = 5, m = 3)$ has a typical mid-range accuracy for the SV forecasts.

These bias corrected SV forecasts are computed each year out to 2014, and the five year time average SV can be computed them. This is then used to determine a five year forecast of the Gauss coefficients in 2015 from

$$(g_n^m, h_n^m)_{2015} = (g_n^m, h_n^m)_{2010} + 5(\dot{g}_n^m, \dot{h}_n^m). \tag{19}$$

We compute the differences between this forecast and the CM6 coefficients in 2015 and plot them in terms of mean square fields (dashed line)(MSF), along side the CM6 MSF (solid line) in Figure 3. We also include the ensemble uncertainty from the 2015 forecast (dotted line). The relative differences are seen to vary from 10^{-7} for $n = 1$ to 10^{-2} for $n = 8$. The ensemble uncertainty, which does not contain information on forecast bias, slightly underestimates the errors for $n = 1 - 3$, and is very close at higher degrees. This is an indication that, after bias correction, most of the errors in the forecast that remain are random rather from model bias.

4. Mean SV for 2020-2025

We repeat the same method to make the SV forecast for the period from 2020 to 2025, except that we use three staggered forecasts. Starting from the analysis computed for 2000 CE, we have continued the assimilation and forecast (with the annual analysis cycle) forward in time for three sets of experiments, with final the analysis times in 2017, 2018 and 2019, and the forecasts are made in the next six years. The experiments, labeled A, B and C; are summarized as:

$$
\begin{array}{cccccccc}
 & t_a & t_{f_1} & t_{f_2} & t_{f_3} & t_{f_4} & t_{f_5} & t_{f_6} \\
Exp.\text{A} & 2017 & 2018 & 2019 & 2020 & 2021 & 2022 & 2023 \\
Exp.\text{B} & 2018 & 2019 & 2020 & 2021 & 2022 & 2023 & 2024 \\
Exp.\text{C} & 2019 & 2020 & 2021 & 2022 & 2023 & 2024 & 2025
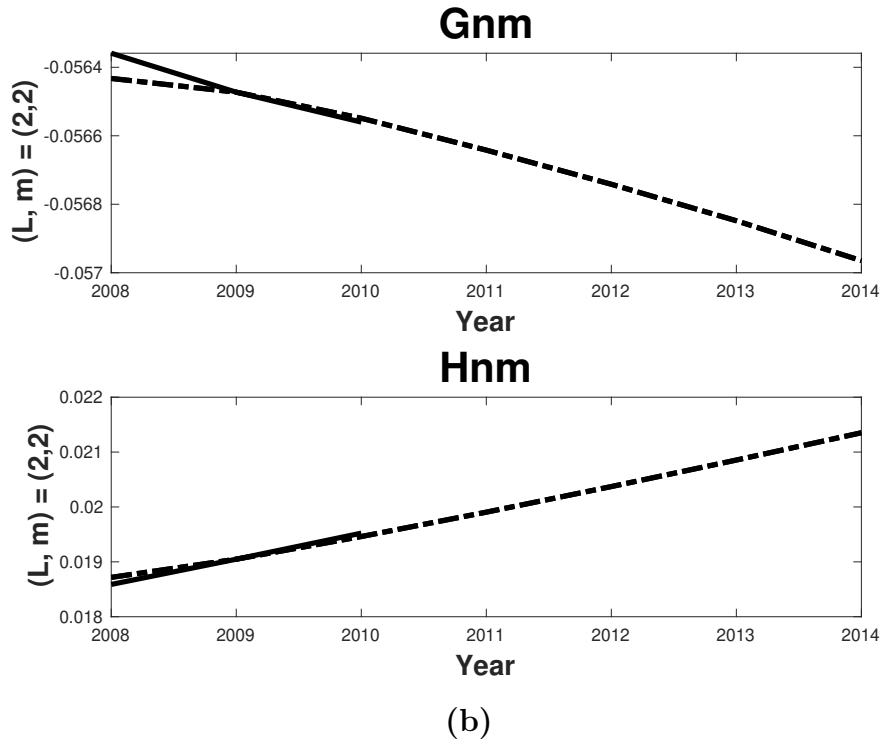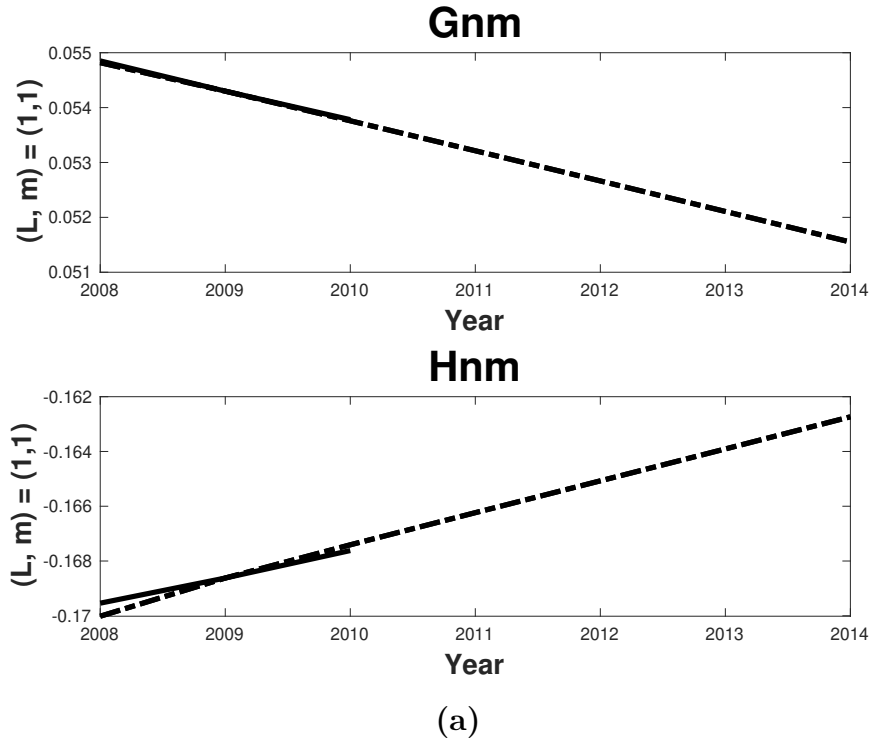\end{array}
\tag{20}
$$

Figure 1. SV validation forecasts (dashed line) and CM6 values (solid line) and for 2010-2015 for Gauss coefficients $n = 1, m = 1$ (a) and $n = 2, m = 2$ (b).
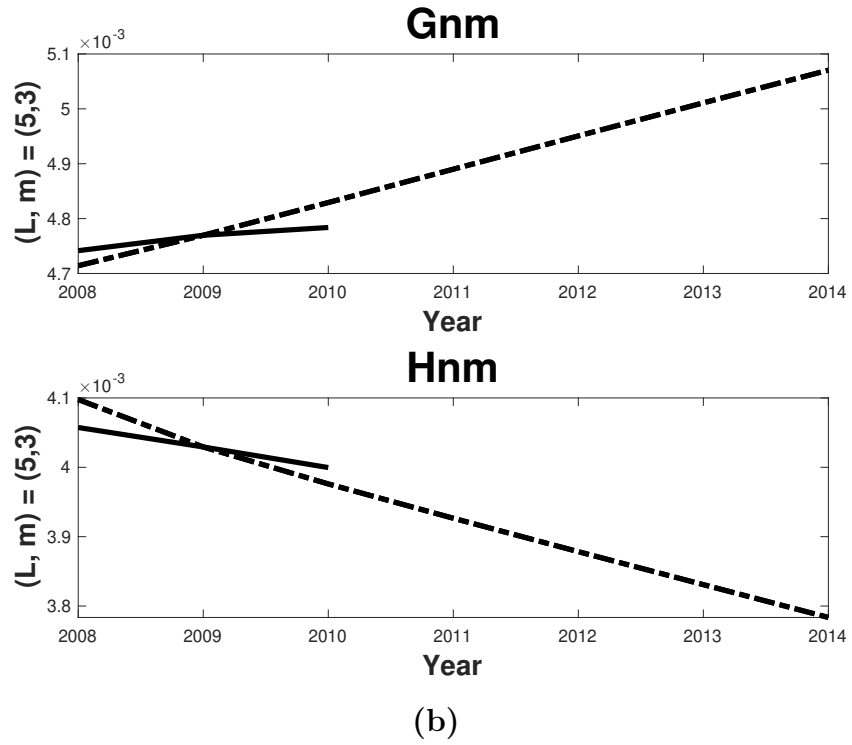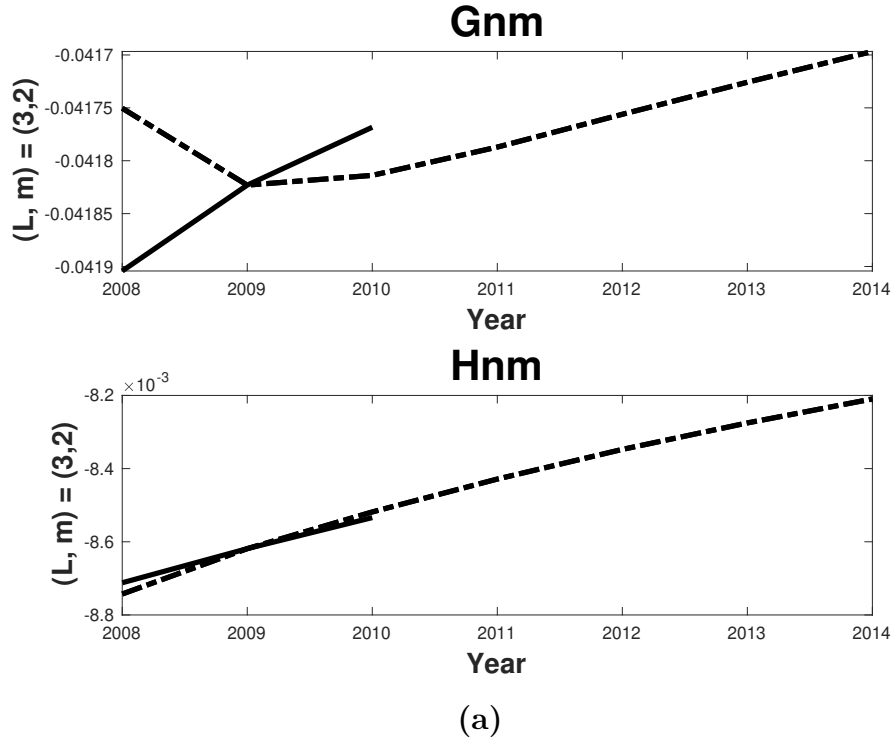
Figure 2. SV validation forecasts (dashed line) and CM6 values (solid line) for 2010-2015 for Gauss coefficients $L = n, m = 2$ and $n = 5, m = 3$.
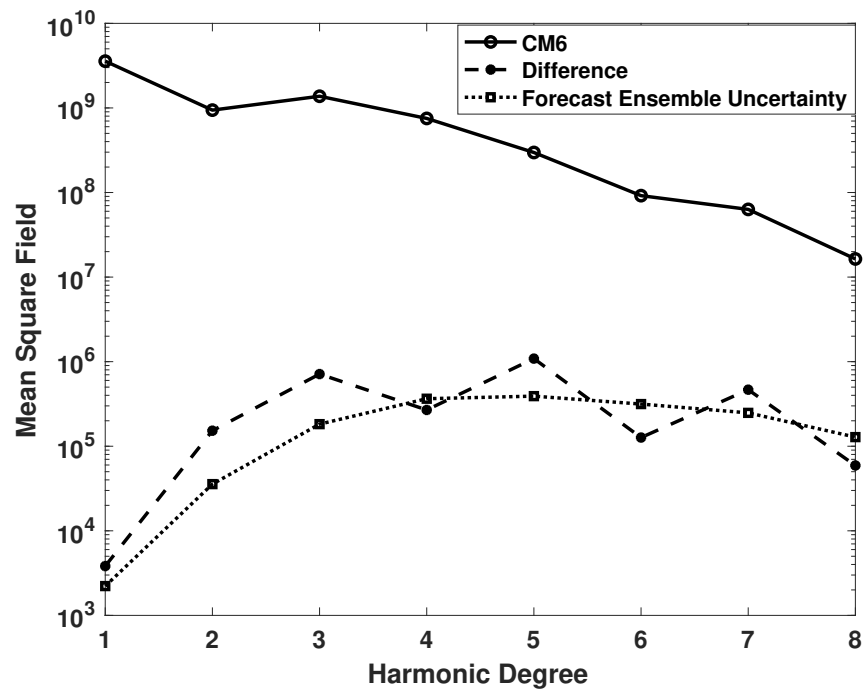
Figure 3. Mean square field in 2015 vs. harmonic degree for forecast (dash-dot), CM6 field (solid) and ensemble forecast uncertainty (dotted)

.

With these three sets of forecasts, we can use up to the following three different combinations of the differences between two staggered forecasts:

$$\tilde{\dot{\mathbf{y}}}_{BA} = (\mathbf{y}_B - \mathbf{y}_A)/\tau_a,$$

$$\tilde{\dot{\mathbf{y}}}_{CB} = (\mathbf{y}_C - \mathbf{y}_B)/\tau_a, \tag{21}$$

$$\tilde{\dot{\mathbf{y}}}_{CA} = (\mathbf{y}_C - \mathbf{y}_A)/2(\tau_a),$$

where again $\tau_a = 1$ year and the AC experiments are 2 years apart. The SV for the forecast Gauss coefficients computed as before, along with the SV for CM6.

The bias corrected SV forecast for 2025 will be a combination of the different forecasts, BA, CB or CA from (21). There are two main reasons for our last analysis ending at 2019: one is that the predictive field from CM6 for 2020 will be used as the initial condition for our forecast calibration; the other is the SV from CM6 for the period 2019-2020

$$(\dot{g}_n^m)^{CM6} = \left[(g_n^m)^{CM6}(2020) - (g_n^m)^{CM6}(2019) \right] /(1 year) \tag{22}$$

can be used to estimate our model bias. Some representative forecasts that demonstrate the bias correction approach and comparison to the CM6 2020 forecast are shown in Figures 4 - 6. Each figure consists of the three uncorrected forecasts for g_n^m and h_n^m, along with the 2017-2020 forecast from CM6, on top (A,B and C) and the bias corrected forecasts below (BA, CB and CA). So the corrected forecasts in each second set in each figure are combinations of the uncorrected forecasts, as described in (21). The slope of each line is the SV time varying SV, and we confirm the consistancy of each of the experiments with the CM6 SV in each case. For example, Figure 4(a) shows the forecasts for $(n, m) = (1, 1)$, and the SV for g_1^1 requires a negligble correction since the three experiments have essentially the same SV. The SV for h_1^1 shows a small change from the bias correction and the AB correction is used since it's slope is closest to the CM6 estimate.

The final calibrated mean SV for the period 2020-2025 is computed as the following: first, the optimal annual SV forecasts $\dot{g}_n^{m(f)}$ are made from our data assimilation system for 2020-2024, with the same procedure discussed in the previous section. Then, the time varying gauss coefficients are computed via

$$g_n^{m(f)}(2020 + k) = (g_n^m)^{CM6}(2020) + \sum_{n=0,k-1} \dot{g}_n^{m(f)}(2020 + n), \quad \text{for} \quad k = 1, 2, \cdot, 5. \tag{23}$$

13

Finally, the mean annual SV gauss coefficients are determined via

$$\overline{\dot{g}_n^{m(f)}} = \left[g_n^{m(f)}(2025) - g_n^{m(f)}(2020) \right]/5. \tag{24}$$

The uncertainties are determined by the standard deviation of the forecast ensemble. The final results are given in Table 1 along with the associated uncertainties, which are the ensemble standard deviations from the ensemble forecast.

5. Conclusions

We have used an EnKF to forecast the geomagnetic secular variation (Table 1) for the time period 2020-2025, using 400 ensemble members. The development of ensemble forecasting techniques has contributed to the SV forecasts, particularly due to the ability to estimate uncertainty, which is newly available with this geomagnetic forecast. The validation forecast for 2015 shows that the uncertainty estimate is in line with the differences with CM6 for that year, which also demonstrates that the bias correction scheme is removing a significant part of the systematic errors. We also use the assimilation of field models as a means to understand how estimates of field model errors impact the accuracy of geomagnetic forecasts from a geodynamo model. The 2025 SV forecast includes uncertainty estimates derived from the forecast ensemble spread. The large ensemble has placed a computational constraint on this forecast, because at present it is not computational feasible to run the assimilation system for 7000 years, as was done in [Kuang et al., 2010]. Future work will investigate the use of localization schemes that would enable the use of a much smaller ensemble that could be run for longer time periods.
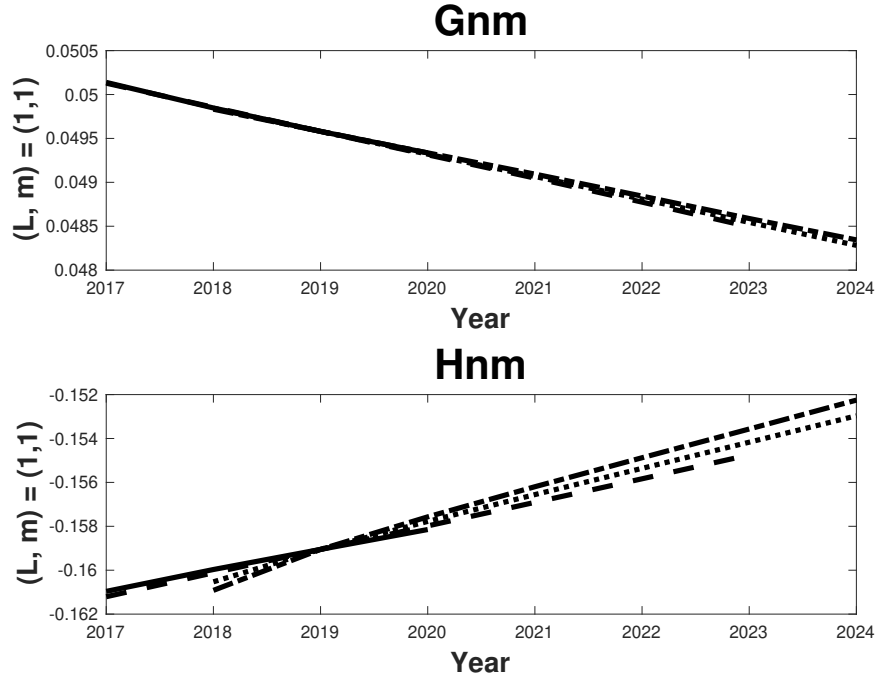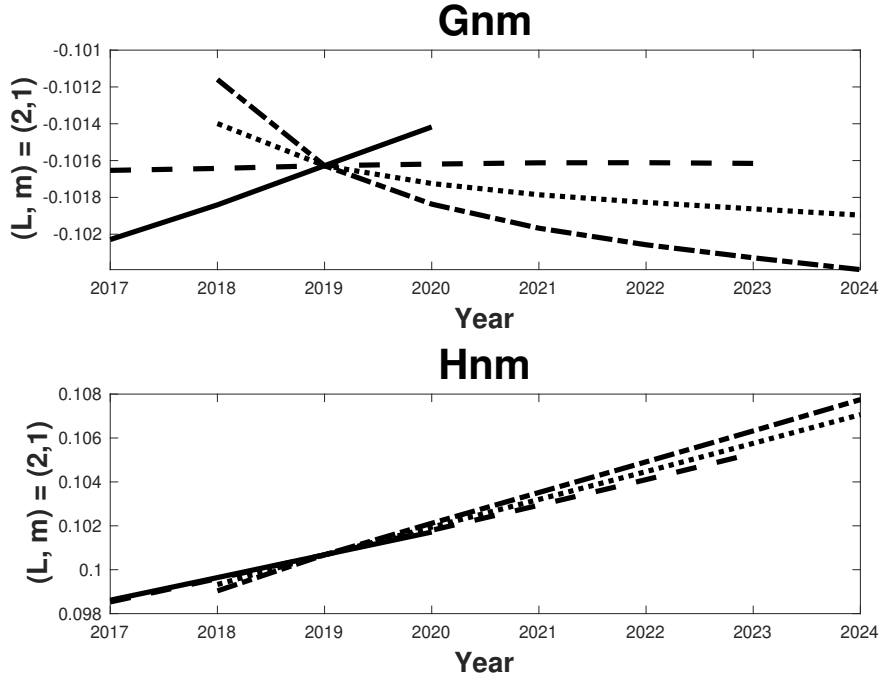
Figure 4. Forecasts of Gauss coefficients from analyses in 2017 (dashed line), 2018 (dash-dot) and 2019 (dotted), in the upper two plots. The grey line shows the 2020 forecast from the Goddard candidate model for IGRF-2025. Panel (a) shows the coefficients for $n = 1$ and $m = 1$, while panel (b) shows the coefficients for $n = 2$, $m = 1$. In each panel the upper figures show the raw (uncorrected) forecasts, while the lower panel shows the bias corrected forecasts from Eq. 21.

Figure 5. Forecasts of Gauss coefficients from analyses in 2017 (dashed line), 2018 (dash dot) and 2019 (dotted). The grey line shows the 2020 forecast from the Goddard candidate model for IGRF-2025. Panel (a) shows the coefficients for $n = 3$ and $m = 2$, while panel (b) shows the coefficients for $n = 3$, $m = 3$. In each panel the upper figures show the raw (uncorrected) forecasts, while the lower panel shows the bias corrected forecasts from Eq. 21.
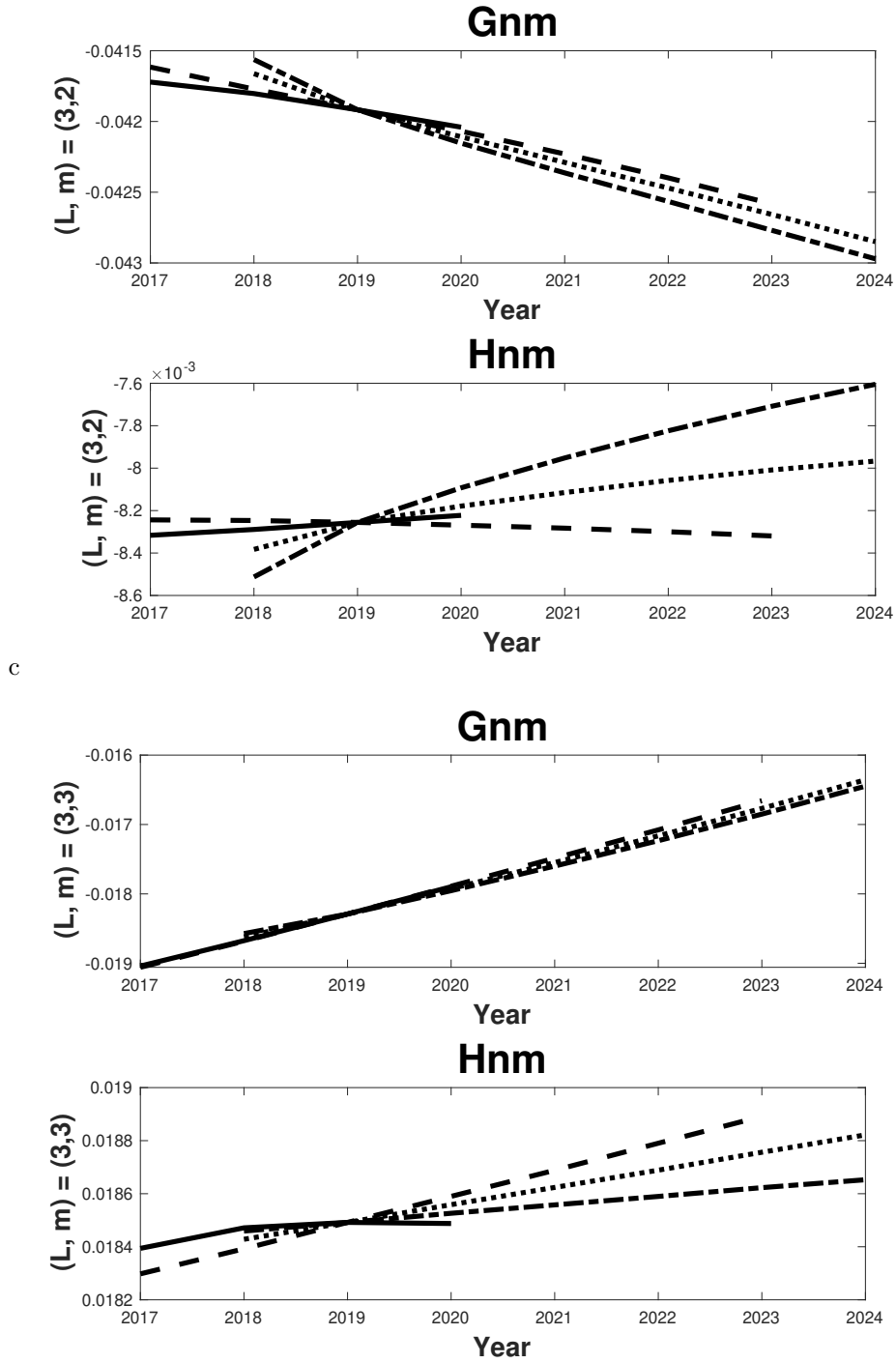
Figure 6. Forecasts of Gauss coefficients from analyses in 2017 (dashed line), 2018 (dash-dot) and 2019 (dotted). The grey line shows the 2020 forecast from the Goddard candidate model for IGRF-2025. Panel (a) shows the coefficients for $n = 6$ and $m = 4$, while panel (b) shows the coefficients for $n = 6$, $m = 5$. In each panel the upper figures show the raw (uncorrected) forecasts, while the lower panel shows the bias corrected forecasts from Eq. 21.
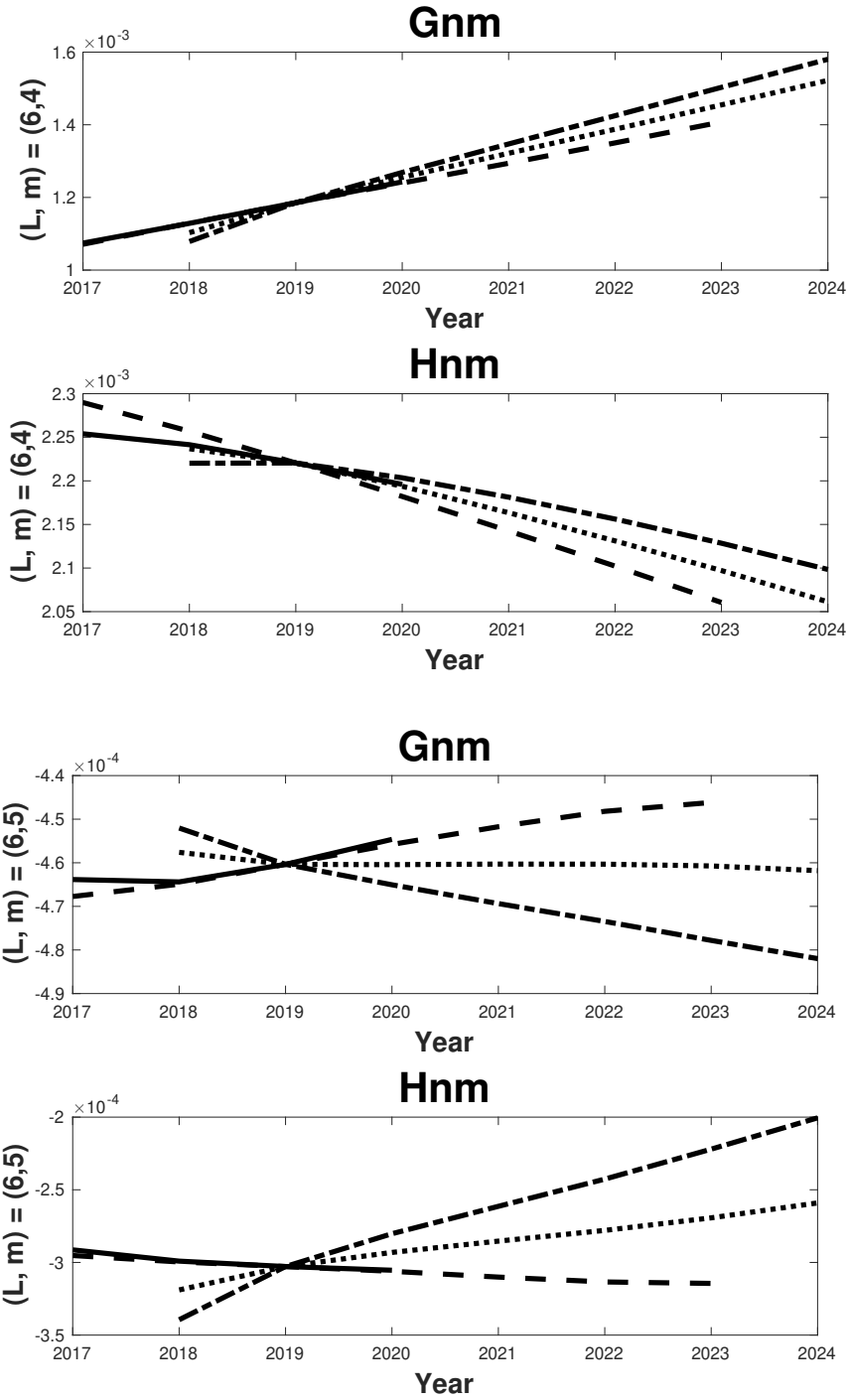
Table 1. Table of the mean SV gauss coefficients for g_n^m, h_n^m (with the unit nT/year). The axial dipole SV $\dot{g}_1^0$ is given by Sabaka et al (2020).

n	m	$\dot{g}$	$\dot{h}$	$\dot{g}$ uncert	$\dot{h}$ uncert
1	0	4.78	0.00	0.0056	0.0000
1	1	8.18	-32.26	0.0053	0.0084
2	0	-5.84	0.00	0.0054	0.0000
2	1	-0.50	-33.51	0.0073	0.0076
2	2	-3.87	-18.41	0.0042	0.0063
3	0	6.49	0.00	0.0057	0.0000
3	1	3.41	6.05	0.0038	0.0029
3	2	4.72	0.44	0.0042	0.0040
3	3	-10.90	-0.87	0.0034	0.0033
4	0	-1.17	0.00	0.0023	0.0000
4	1	-1.31	0.30	0.0029	0.0033
4	2	-6.34	6.82	0.0016	0.0022
4	3	4.91	4.66	0.0031	0.0030
4	4	-4.07	-5.10	0.0019	0.0018
5	0	-0.31	0.00	0.0028	0.0000
5	1	-0.17	0.42	0.0014	0.0013
5	2	-0.78	3.19	0.0016	0.0019
5	3	-0.08	0.46	0.0013	0.0012
5	4	2.57	3.20	0.0020	0.0012
5	5	1.24	-0.30	0.0010	0.0011
6	0	-0.60	0.00	0.0007	0.0000
6	1	-0.22	0.18	0.0010	0.0010
6	2	-0.94	-1.70	0.0008	0.0007
6	3	2.22	-0.65	0.0011	0.0010
6	4	-1.63	0.95	0.0008	0.0007
6	5	-0.11	0.08	0.0007	0.0006
6	6	0.87	0.82	0.0007	0.0007
7	0	-0.09	0.00	0.0006	0.0000
7	1	0.20	0.45	0.0005	0.0004
7	2	-0.28	0.49	0.0006	0.0005
7	3	0.91	-0.81	0.0004	0.0004
7	4	-0.55	0.05	0.0008	0.0006
7	5	-0.61	-1.01	0.0003	0.0003
7	6	-0.84	0.25	0.0004	0.0005
7	7	0.83	0.33	0.0004	0.0003
8	0	0.05	0.00	0.0003	0.0000
8	1	0.15	-0.36	0.0003	0.0003
8	2	-0.30	0.47	0.0003	0.0002
8	3	0.57	-0.02	0.0003	0.0004
8	4	-0.21	0.64	0.0003	0.0003
8	5	0.52	0.19	0.0003	0.0003
8	6	0.09	-0.32	0.0002	0.0002
8	7	-0.46	0.56	0.0003	0.0003
8	8	0.38	-0.08	0.0002	0.0002

List of abbreviations

CM4 - Comprehensive Model version 4

CM6 - Comprehensive Model version 6

EnKF - Ensemble Kalman Filter

GEMS - Geomagnetic Ensemble Modeling System

IGRF - International Geomagnetic Reference Field

NWP - Numerical Weather Prediction

OSSE - Observing System Simulation Experiment

SV - Secular variation

Availability of data and materials

All of the data presented in this paper is already publicly available (for example, the field models), or will be made available upon publication (the SV Gauss coefficients contained in this paper).

Competing Intrests

The authors declare that they have no competing interests.

Authors' contributions

Andrew Tangborn led the ensemble assimilations and validation tests for this work. Weijia Kuang led the validation methodology, GEMS/CM6 interface and the SV product for IGRF. Terence Sabaka provided all CM6 results used for this research, and interface validation. Ce Yi provided all technical support for

assimilation and analysis.

Data availability

The data (IGRF candidate Gauss coefficients) is included as a table in this paper.

References

Aubert, J.T. and A. Fournier (2011), Inferring internal properties of Earths core dynamics and their evolution from surface observations and a numerical geodynamo model, *Nonlin. Process. Geophys.*, 18, 657–674.

Aubert, J. T. Gastine and A. Fournier (2017). Spherical convective dynamos in the rapidly rotating asymptotic regime. *J. Fluid Mech.* 813, 558-593.

Bloxham, J. and D. Gubbins (1985). The secular variation of Earth's magnetic field. *Nature* 317, 777-781.

Braginsky, S. I. and P. H. Roberts (1996). Equations governing convection in Earth's core and the geodynamo. *Geophys. Astrophys. Fluid Dyn.* 79, 1-97.

Bullard, E.C., C. Freedman, H. Gellman and J. Nixon (1950). The westward drift of the Earth's magnetic field, *Phil. Trans. R. Soc. Lond.*, 243, 67-92.

Christensen, U.R., J. Aubert, P. Cardin, E. Dormy, S. Gibbons, G. Glatzmaier, E. Crote, Y. Honkura, C. Jones, M. Kono, M. Matsushima, A. Sakuraba, F. Takahashi, A. Tilgner, J. Wicht, K. ZHang (2001). A numerical dynamo benchmark. *Phy. Earth. Planet. Inte.* 123, 25-34.

Christensen, U.R. and J. Aubert (2006), Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields, *Geophys. J. Int.*, 166, 97-114.

Christensen U. R., J. Aubert and G. Hulot (2010). Conditions for Earth-like geodynamo models. *Earth Planet. Sci. Lett.* 296, 487-496.

Evensen, G. (2009), *Data assimilation: the ensemble Kalman filter*, Springer.

Fournier, A., Nerger, L. and Auber, J., (2013), An ensemble Kalman filter for the time-dependent analysis of the geomagnetic field, *Geochem. Geophys. Geosyst.*, 14, 4035–4043.

Fournier, A., J. Aubert and E. Thebaut (2015), A candidate secular variation model for IGRF-12 based

on Swarm data and inverse geodynamo modelling, *Earth, Planets and Space*, *67*, DOI 10.1186/s40623-015-0245-8.

Glatzmaier, G.A. and P.H. Roberts (1995), A three-dimensional convective dynamo solution with rotating and finitely conducting inner core and mantle, Phys. Earth Planet. Inter., 91, 63–75.

Hide, R. (1967). Motions of the Earth's core and mantle, and variations of main geomagnetic field. *Science* 157,55-58.

Jackson, A., A. T. Jonkers and M. R. Walker (2000). Four centuries of geomagnetic secular variation from historical records. *Phil. Trans. R. Soc. A* 358, 957-990.

Jones, C. A., P. Boronski, A. S. Brun, G. A. Glatzmaier, T. Gastine, M. S. Miesch, J. Wicht (2011). Anelastic convection-driven dynamo benchmarks. *Icarus* 216, 120-135.

Kuang, W.J. and J. Bloxham (1997), An Earth-like numerical dynamo model, Nature, 389, 371–374.

Kuang, W., A. Tangborn, W. Jiang, D. Liu, Z. Sun, J. Bloxham and Z. Wei (2008), MoSST-DAS: The first generation geomagnetic data assimilation framework, *Comm. Comp. Phys.*, *3*, 85-108.

Kuang, W., A. Tangborn and Z. Wei (2009), Constraining a numerical geodynamo model with 100 years of surface observations, **Geophys. J. Inter.**, 179, 1458-1468.

Kuang, W.,Z. Wei, R. Holme and A. Tangborn (2010), Prediction of geomagnetic field with data assimilation: a candidate secular variation model for IGRF-11, *Earth, Planets and Space*, 62, 775-785.

Kuang, W., B. F. Chao and J. Chen (2017). Decadal polar motion of the Earth excited by the convective outer core from geodynamo simulation. *J. Geophys. Res. SE*, 122, 8459-8473.

Lund, S. P. and P. Olson (1987). Historic and paleomagnetic secular variation and the Earth's core dynamo process. *Rev. Geophys.*, doi:10.1029/RG025i005p00917

Lund, S. P. (1987). A new view of long-term geomagnetic secular variation. *Front. Earth Sci.*, doi:10.3389/feart.2018.00040

Lund, S. P. and P. Olson (1987). Historic and paleomagnetic secular variation and the Earth's core dynamo process. *Rev. Geophys.*. doi:10.1029/RG025i005p00917

Matsui, H. and Co-authors (2016). Performance benchmarks for a next generation numerical dynamo

model, *Geochem., Geophys. and Geosys.*, https://doi.org/10.1002/2015GC006159.

Olsen, N., M. Mandea, T.J. Sabaka and L. T/offner-Clausen (2010), The CHAOS-3 geomagnetic field model and candidates for the 11th generation IGRF, *Earth, Planets and Space*, 62, 719-272.

Sabaka, T. J., N. Olsen, and M. E. Purucker (2004), Extending comprehensive models of the Earth's magnetic field with Ørsted and CHAMP data, *Geophys. J. Int.*, 159, 521–547.

Sabaka, T.J., L. T/offner-Clausen,N. Olsen and C.C. Finlay (2020), CM6: A Comprehensive Geomagnetic Field Model Derived From Both CHAMP and Swarm Satellite Observations, *submitted to this Issue*.

Sun, Z., A. Tangborn and W. Kuang (2007), Data assimilation in a sparsly observed one-dimensional modeled MHD system, **Nolinear Processes in Geophysics**, *14*, 181-192.

Tangborn, A. and W. Kuang (2015), Geodynamo model and error parameter estimation using geomagnetic data assimilation, *Geophys. J. Int. 200*, 664–675.

Tangborn, A. and W. Kuang (2018), Impact of archeomagnetic field model data on modern era geomagnetic forecasts, *Phys. Earth Planet. Int., 276, 2-9.*