JGR Space Physics

RESEARCH ARTICLE
10.1029/2022JA031006

**Key Points:**

- Dynamics of global geomagnetic connectivity during the ill-famed storms of February 2022 is reconstructed via mining archived in situ data
- A new hybrid model architecture is employed, combining flexible modular and high-resolution components with added interplanetary magnetic field penetration effect
- Deep magnetic depression and possibly a neutral line formed in the post-dusk sector around geosynchronous distance at the storm active phase

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

N. A. Tsyganenko,
nikolai.tsyganenko@gmail.com;
n.tsyganenko@spbu.ru

**Citation:**

Tsyganenko, N. A., Andreeva, V. A., Sitnov, M. I., & Stephens, G. K. (2022). Magnetosphere distortions during the "satellite killer" storm of February 3–4, 2022, as derived from a hybrid empirical model and archived data mining. *Journal of Geophysical Research: Space Physics*, *127*, e2022JA031006. https://doi.org/10.1029/2022JA031006

Received 12 SEP 2022
Accepted 29 NOV 2022

**Author Contributions:**

**Conceptualization:** N. A. Tsyganenko, M. I. Sitnov
**Data curation:** N. A. Tsyganenko, V. A. Andreeva
**Formal analysis:** N. A. Tsyganenko
**Funding acquisition:** N. A. Tsyganenko
**Investigation:** N. A. Tsyganenko
**Methodology:** N. A. Tsyganenko, M. I. Sitnov, G. K. Stephens
**Project Administration:** N. A. Tsyganenko
**Resources:** N. A. Tsyganenko
**Software:** N. A. Tsyganenko
**Supervision:** N. A. Tsyganenko



# Magnetosphere Distortions During the "Satellite Killer" Storm of February 3–4, 2022, as Derived From a Hybrid Empirical Model and Archived Data Mining

**N. A. Tsyganenko[1], V. A. Andreeva[1], M. I. Sitnov[2], and G. K. Stephens[2]**

[1]Saint-Petersburg State University, Institute and Department of Physics, Saint-Petersburg, Russia, [2]Applied Physics Laboratory, The Johns Hopkins University, Laurel, MD, USA

**Abstract** A pair of relatively mild geomagnetic storms of February 3–4, 2022, resulted in the loss of 38 Starlink satellites. In this work we reconstruct a sequence of magnetospheric configurations during that event, based on the largest ever archive of historical data, a new magnetic field model, and the nearest-neighbor method of mining the solar wind, magnetosphere, and ground activity data. The model is constructed as a combination of main modular framework, explicitly representing magnetic effects of principal current systems, and a high-resolution superstructure, based on expanding the field of equatorial currents into a sum of quasi-orthogonal harmonics, whose purpose is to correct residual inaccuracies of the modular component. In addition, the variable penetration of the interplanetary magnetic field (IMF) into the magnetosphere is taken into account in the form of a potential field, parameterized by the observed IMF components. Despite relatively modest intensity of the event in terms of Sym-H peak values, surprisingly strong transient distortions of the global magnetic field are found. Whereas under quiet pre-storm conditions the midnight segment of synchronous orbit and the dayside cusps mapped, respectively, to ∼67° and ∼77° of corrected latitude, at the peak of the storm their footpoints shifted equatorward to as low as ∼63° and ∼67°–68°, with formation of entangled flux-rope-like structures and magnetic neutral lines at distances $R \sim 12–15\ R_E$ on the nightside. At nearly the same time, a deep magnetic depression and, possibly, an island of reversed field polarity forms even closer, in the post-dusk sector of geosynchronous orbit.

**Plain Language Summary** In the beginning of February 2022, a relatively weak geomagnetic storm broke out, whose unexpected result was a premature demise of 38 Starlink satellites due to their abnormal drag in the upper atmosphere. This work analyses the concurrent distortions of the distant geomagnetic field, caused by the arrival of disturbed solar wind and interplanetary magnetic field. The storm-time deformations of the Earth's magnetosphere are reconstructed by using an advanced magnetic field model, a huge database of past spacecraft and ground-based observations covering the last 27 years, and novel methods of data mining. Strong distortions of the geomagnetic field during the storm culmination are revealed, that penetrated to unusually close distances up to the geosynchronous orbit.

## 1. Introduction

Quantitative reconstruction of the magnetic connectivity between the ionosphere and the vast volume of distant magnetosphere has traditionally been the principal goal of the data-based empirical modeling. Introduction of advanced mathematical architectures (N. A. Tsyganenko & Sitnov, 2007, henceforth TS07; N. A. Tsyganenko, Andreeva, Kubyshkina, et al., 2021, and references therein), the continuing growth of volume/quality of archived data, and the development of new methods of data ingestion into the models (Sitnov et al., 2008, 2020, and refs. therein) resulted in significant progress, making it eventually possible to model not only average magnetospheric states, quantified by ad hoc driving functions (e.g., N. A. Tsyganenko & Andreeva, 2015; henceforth TA15), but also to reproduce the dynamics of individual events (Sitnov et al., 2019; Stephens et al., 2019; Stephens & Sitnov, 2021; N. Tsyganenko, Andreeva, Kubyshkina, et al., 2021; N. A. Tsyganenko, Andreeva, & Sitnov, 2021; N. A. Tsyganenko, Andreeva, Sitnov, Stephens, et al., 2021). The main idea goes back to Sitnov et al. (2008); its essence is to abandon the one-time "universal" parameterization of a model by external drivers and/or ground indices, based on a single grand archive of data. Instead, the magnetosphere's dynamics is derived from a sequence of relatively small subsets, each of which is compiled (mined) from the grand set, based on the proximity of its data points to the modeled event trajectory in the hyperspace of magnetospheric state parameters and their temporal

**Validation:** N. A. Tsyganenko
**Visualization:** N. A. Tsyganenko
**Writing – original draft:** N. A. Tsyganenko
**Writing – review & editing:** N. A. Tsyganenko, V. A. Andreeva, M. I. Sitnov, G. K. Stephens

trends. Unifying that "nearest-neighbor" (NN) approach with sufficiently flexible high-resolution models made it possible to resolve previously unaccessible details of the spatial structure and time evolution of the magnetosphere during storms (e.g., Sitnov et al., 2018; N. A. Tsyganenko, Andreeva, & Sitnov, 2021) and even substorms (Stephens et al., 2019, 2022; N. A. Tsyganenko, Andreeva, Sitnov, Stephens, et al., 2021).

This work continues the above cited line of achievements by presenting results of a detailed study of the dynamics of the distant geomagnetic field during an interesting storm event of February 3–4, 2022, which resulted in unexpected premature demise of 38 Starlink satellites. Important innovations have been introduced, in particular, the "hybrid" fitting method, conceived earlier by N. A. Tsyganenko and Andreeva (2017). In that approach, a basic modular framework is complemented with a higher-resolution flexible superstructure, which serves a role of a next-order correction that reduces residual systematic biases of the modular component. Another new feature is that all three interplanetary magnetic field (IMF) components are now used in the IMF penetration module, previously limited to only a uniform fraction of the transverse component. Finally, the newest and the largest-ever grand data pool is used, covering the 27-year long interval from the beginning of 1995 to the end of 2021 (almost $2\frac{1}{2}$ solar cycles) and nearly twice larger than that employed in the most recent previous study of this kind (N. A. Tsyganenko, Andreeva, & Sitnov, 2021).

As shown in the present work, the above new techniques help reveal interesting details of the equatorial field structure at the nightside, such as the gradual formation and outward propagation of a relatively thin region of highly stretched field around geosynchronous orbit, accompanied with multiple $B_z$ reversals and neutral lines at larger distances during the long interval of southward IMF at the peak of the storm. The obtained results quantitatively demonstrate the high variability of magnetospheric magnetic configurations in response to the interplanetary driving, widely discussed in many previous studies (e.g., Kissinger et al., 2014; Pulkkinen et al., 2007, 2013; Runov et al., 2022).

The paper is structured as follows. The next Section 2 describes the event in question and the main principles/procedures of the grand data pool creation. Section 3 addresses details of generating the NN data subsets. Section 4 outlines the model's architecture and consists of two Sections 4.1 and 4.2, focused at its modular and high-resolution components, respectively. Section 5 presents the main results of this study in the form of field line plots for selected time moments of the storm (a full collection of 48 plots for February 3 and 4 at one-hour cadence is provided in Supporting Information S1). Sections 6 and 7 contain a discussion of results and a Summary.

## 2. The Double Storm of February 3–4, 2022

The event studied in this work started almost in line with a classic scenario of a CME-induced storm, but with one remarkable exception: there was no initial increase in the solar wind density nor pressure and, hence, no sudden commencement in the early hours of February 3 (DOY 34).

As shown in Figure 1, from the very beginning of the storm the magnetosphere was exposed to quite a large and positive azimuthal IMF $B_y$ (brown), which stayed around ∼+12 nT until UT = 07:30, when it started an almost linear excursion down to ∼−15 nT at UT = 10:30. The North-South IMF component $B_z$ (red) turned negative at UT = 02:00, after which it smoothly went down, having reached $B_z \sim -10$ nT by 07:30, nearly at the same time when IMF $B_y$ started its linear reversal. Around UT ∼ 07:50, IMF $B_z$ swiftly plunged further down to ∼−18 nT and stayed at that low level for nearly two hours, which drove the disturbance to its culmination at UT = 10:55, when the Sym-H index reached its peak value of −80 nT. The strong driving during that period resulted in a series of substorms, manifested in multiple negative peaks of the SuperMag auroral (SML) index, first of which occurred still in the very beginning of the storm at UT = 03:00. After UT = 10:00, IMF $B_z$ gradually relaxed to ∼−8 nT and then suddenly reversed to ∼+14 nT at UT ∼ 11:45, coincident with arrival of very dense solar wind, which resulted in a dramatic (∼10-fold) jump of the solar wind dynamic pressure. In spite of a relatively high speed (520 − 580 km/s) and pressure (∼3–10 nPa) during the rest of DOY 34, the disturbance gradually subsided due to small and intermittenty positive IMF $B_z$. The situation changed again at the beginning of the next day, February 4 (DOY 35), when a gradual increase of the solar wind density was followed by its second sudden jump to ∼15 $cm^{-3}$, accompanied by a temporary increase of the speed from 530 to 590 km/s during the first hours of DOY 35 and another long interval of southward IMF $B_z$, albeit not as large as on DOY 34. This resulted in the second wave of disturbance, manifested in the next round of Sym-H decrease, with peak values of −62 and −70 nT at 10:30 and 20:35 UT, respectively.

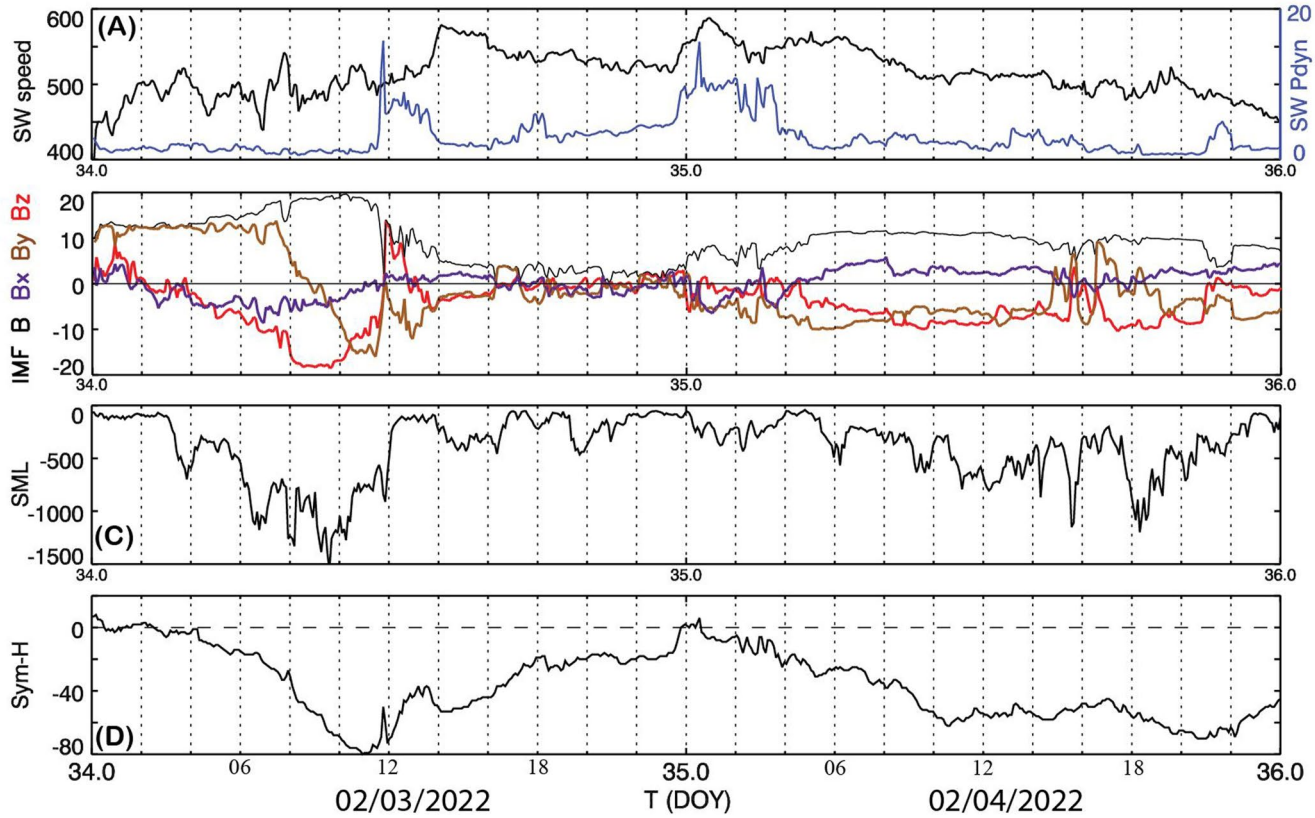


**Figure 1.** Variation of the interplanetary parameters and ground activity during the storm-time interval February 3–4, 2022. (a) The solar wind speed (black) and dynamic pressure (blue; separate scale on the right), (b) interplanetary magnetic field $B_x$ (violet), $B_y$ (brown), $B_z$ (red), full B (thin black), (c) SuperMag auroral index, (d) Sym-H index.

## 3. Data Sources and NN Selection Parameters/Procedures

As in all previous modeling studies, the data used in this work came from three sources: (a) interplanetary medium observations, collected mostly by the incoming solar wind monitors located around L1 point and stored in OMNI archive, (b) magnetospheric magnetic field data taken by both past and presently active missions, and (c) ground geomagnetic activity indices, also stored in the publicly available repositories. Also in line with the previously adopted procedures (e.g., N. A. Tsyganenko, 2013, Section 5), the interplanetary data employed in this work were downloaded and stored at 5-min resolution, assumed as an optimal tradeoff choice in view of inevitable timing errors, involved in the data propagation from the L1 point to the subsolar magnetopause, and additional uncertainties due to the delayed response of the inner and distant magnetosphere to the external drivers.

In regard to the magnetospheric space magnetometer data, all intermediate procedures involved in their processing (filtering of bad/biased or non-magnetospheric data, main geomagnetic field subtraction, averaging, etc.) were also covered in great detail in previous papers (N. A. Tsyganenko, 2013; N. Tsyganenko, Andreeva, Kubyshkina, et al., 2021). Likewise, all spacecraft missions whose data entered in the new database have already been overviewed in earlier publications (e.g., N. A. Tsyganenko, Andreeva, & Sitnov, 2021, Section 4). For that reason and in order to save page space, only a brief recounting of that material is given below.

Table 1 presents a list of satellites whose magnetometer data were used in this study, inner/outer coverage limits, total numbers of 5-min records, and updated ranges of covered time intervals. The inner distance limits $R_{min}$ were specified for each mission on the basis of results of visual day-by-day inspection of data plots. In most cases they were set at $R_{min} = 3.0\ R_E$; data at closer distances were excluded because of rapidly growing errors of the external field components, due to very large magnitude of the internal field at low altitudes. In other cases (Cluster 1, Van Allen, and Magnetospheric MultiScale) the near-perigee errors were found to be smaller, which allowed us to choose lower values of $R_{min}$. The values of the outer distance limit $R_{max}$ correspond either to the overall

**Table 1**
*Grand Data Set: Contributing Missions, Geocentric Distance Limits, Numbers of Records, Timespans*

| Mission | $R_{min}$ ($R_E$) | $R_{max}$ ($R_E$) | Number of records | Begin date (year/DOY) | End date (year/DOY) |
|---|---|---|---|---|---|
| Geotail | 8.08 | 31.00 | 841,088 | 1995/004 | 2021/363 |
| Polar | 3.01 | 9.61 | 871,250 | 1996/076 | 2008/108 |
| Cluster 1 | 2.17 | 22.01 | 921,311 | 2001/030 | 2020/152 |
| Cluster 2 | 3.00 | 22.16 | 925,291 | 2001/030 | 2020/152 |
| Cluster 3 | 3.00 | 22.16 | 834,097 | 2001/030 | 2020/152 |
| Cluster 4 | 3.00 | 22.16 | 941,476 | 2001/030 | 2020/152 |
| Themis A | 3.01 | 14.71 | 996,576 | 2007/060 | 2021/365 |
| Themis B | 3.01 | 31.00 | 94,646 | 2007/060 | 2010/202 |
| Themis C | 3.02 | 31.00 | 139,838 | 2007/060 | 2010/084 |
| Themis D | 3.01 | 15.36 | 987,475 | 2007/060 | 2021/365 |
| Themis E | 3.01 | 14.74 | 1,008,835 | 2007/060 | 2021/365 |
| Van Allen A | 2.01 | 5.80 | 472,481 | 2012/252 | 2019/243 |
| Van Allen A | 2.01 | 5.83 | 465,183 | 2012/252 | 2019/196 |
| MMS–1 | 2.02 | 29.19 | 307,513 | 2015/244 | 2021/363 |
| Total: 9,807,060 | | | | | |

Abbreviation: MMS, Magnetospheric MultiScale.

outer boundary of the modeling region $R = 31\ R_E$ (set in order to include the Geotail's apogee data since 1995), or to individual spacecraft apogees, if they fell short of that boundary.

A visual overview of the data coverage is presented in Figure 2 in the form of color-coded meridional and equatorial distributions of logarithm of columnar density of data points, integrated within $0.5 \times 0.5\ R_E$ bins of X–Z and X–Y (GSM). Peaks and sharp boundaries of the color density correspond to apogees of individual spacecraft that contributed to the grand database. A more detailed quantitative illustration of the radial distribution of individual mission data is given in Figure 3 as a histogram of the number of data records falling into consecutive $0.5\ R_E$ intervals of the geocentric distance. Figures 2 and 3 are similar in their format to Figure 39.1 of N. Tsyganenko, Andreeva, Kubyshkina, et al. (2021) and Figure 4 of N. A. Tsyganenko, Andreeva, and Sitnov (2021), respectively; the only but essential difference is almost twice larger number of records in the newly updated set, mainly due to added contributions from the recent Cluster 1–4 ($R_{max} \approx 22\ R_E$) and Themis A, D, and E ($R_{max} \approx 15\ R_E$) observations.

A standard approach used in older empirical models was to create a large single data file, whose each record included observed components of the magnetospheric magnetic field vector, concurrent values of interplanetary parameters and their geoeffective combinations, as well as the ground-based indices (all at the same 5-min resolution). That allowed us to directly relate the amplitudes of individual field sources with input parameters in a framework of a universal model code, driven by a limited set of concurrent observables.

The new NN approach is cardinally different in this respect: now the dynamical effect of external drivers enters into the model field not via analytical expansions of its coefficients, but through the procedure of selecting the most relevant data into a sequence of NN-subsets, generated by searching the grand data set in the $2N$-dimensional hyperspace of normalized solar-wind-magnetosphere state parameters $\langle G_i \rangle$ and their trends $\langle \dot{G}_i \rangle$ ($i = 1, \ldots, N$). This implies a largely different way of handling the data, according to which two separate grand sets are prepared: the first Set 1, containing only the state parameters, and the second independent Set 2 that includes all magnetospheric data records, each

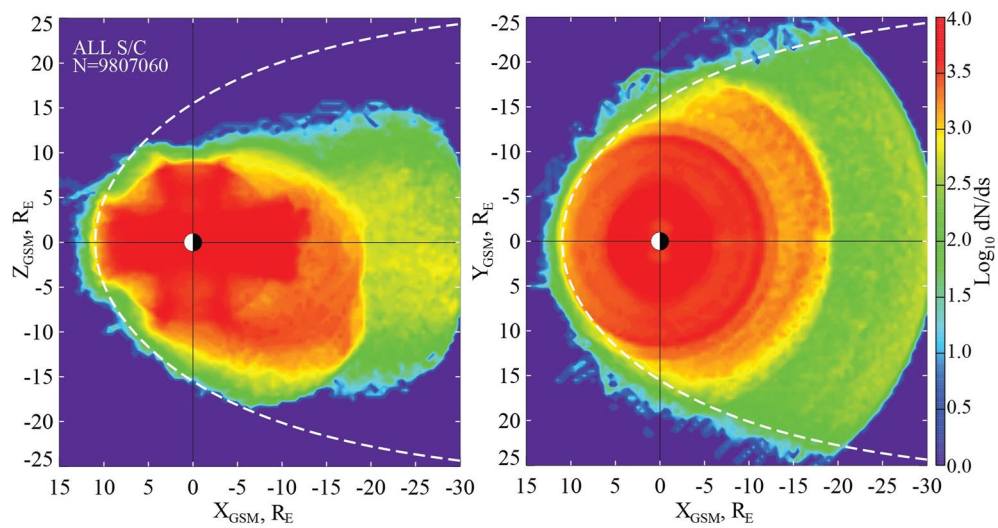


**Figure 2.** Spatial coverage of the new grand data set as viewed in meridional (left) and equatorial (right) projections. The color-coded quantity is the logarithm of column-integrated density of data records falling into $0.5 \times 0.5\ R_E$ X–Z and X–Y bins of GSM coordinates. Average magnetopause location is shown as white dashed line. Similar in format to Figure 39.1 in N. Tsyganenko, Andreeva, Kubyshkina, et al. (2021).

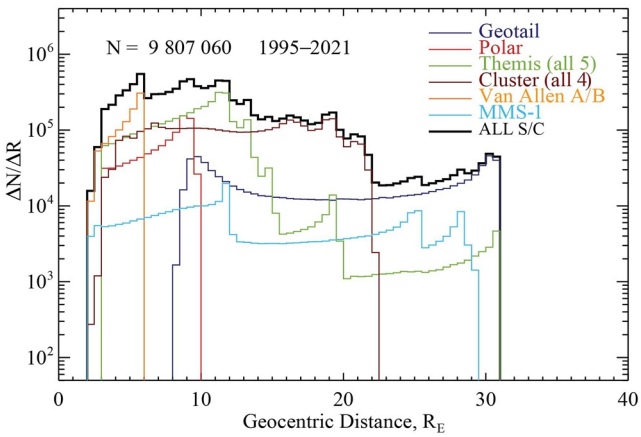


**Figure 3.** Histograms of the radial distribution of the grand set data, including both a cumulative plot for all missions (heavy black) and those for individual missions, as explained in the colored legend. Compare with Figure 4 in N. A. Tsyganenko, Andreeva, and Sitnov (2021) and Figure 1 of Stephens and Sitnov (2021).

of which is tagged with concurrent values of selection parameters stored in the Set 1. The essence of the NN selection procedure is then to jointly analyze both sets in such a way that, for any given time moment of interest, an iterative global search is performed over the Set 2, in order to define a subregion of the parametric space around the modeled query point $\langle \mathbf{G}^{(m)} \rangle$, containing a prescribed number of NN records. Once the NN subset is created, a fitting procedure is carried out, generating a set of best-fit model coefficients for that moment of interest.

The first and one of the most important questions is the choice of the NN selection parameters. At first sight, it would be natural to use as much available information as possible by including all principal interplanetary drivers, ground activity indices, and their temporal trends. However attractive it might seem, using all these parameters makes little sense: the growing number of selection criteria results in rapidly decreasing probability to find data records located sufficiently close to the query point in the parametric hyperspace. In essence, that fact is at the core of the so-called "dimensionality curse" problem (Gershenfeld, 2003, Ch.12.5), which effectively rules out using more than 5–7 selection parameters (Verleysen & François, 2005).

In this study, we set out to test the above constraint at its upper limit. Accordingly, the following parameters were used: Sym-$H_c$ index (standard Sym-H, corrected for the ram pressure effects), IMF $B_y$, IMF $B_z$, and the solar wind speed $V$. The four parameters were complemented by time derivatives, quantifying their temporal variations, which gave a total of eight NN selection parameters. Each parameter was normalized by its mean absolute value (see the metrics discussion below) and slide-averaged with a cosine mask over 4-hr intervals ($T = 4$), symmetrically centered on consecutive time moments $t_k$, following at a 5-min cadence:

$$\langle G_i \rangle (t_k) = \int_{-T/2}^{+T/2} G_i(t_k + \tau) \cos \frac{\pi \tau}{T} d\tau, \quad \langle \dot{G}_i \rangle (t_k) = \int_{-T/2}^{+T/2} \dot{G}_i(t_k + \tau) \cos \frac{\pi \tau}{T} d\tau \tag{1}$$

As in most of the previous models of this kind, effects of the solar wind ram pressure $P_{\rm dyn}$ were taken into account by (a) expanding the magnitude coefficients of the model field into binomials $a + b\sqrt{P_{\rm dyn}}$ and (b) proper spatial scaling of the model field in concert with the pressure-controlled size of the magnetopause.

A possibility to include the solar wind proton density $N_p$ and its time variation was also considered. Eventually, it was decided to leave it out, since the density effects are indirectly taken into account, once both the pressure and speed are included. In general, there is much room for further experiments with different sets of NN selection parameters, and the presently used variant is just one of many possible choices. For example, one may envision adding the auroral indices and their trends (e.g., Stephens et al., 2019; N. A. Tsyganenko, Andreeva, Sitnov, Stephens, et al., 2021), testing $VB_z$ or more sophisticated drivers like $\mathcal{N}$-index by Newell et al. (2007) or $\mathcal{B}$-index by Boynton et al. (2011), and so on. Here we note in passing that, even though the latter indices include IMF $B_y$ effects, they do not depend on its polarity, while our using the IMF $B_y$ in its original form distinguishes between situations with $B_y \gtrless 0$ and, hence, helps to more accurately select similar storm-time NN data from the grand archive.

Besides the "dimensionality curse", there is another complication, termed the "data paucity curse" (Sitnov et al., 2020). In essence, it stems from a well-known fact that the most powerful and practically most important natural events are not just rare, but their occurrence frequency rapidly decreases with the event magnitude. As a result, even the seemingly huge multi-year data archives may still contain too little information on the most violent space storms of greatest interest.

A partial remedy for both problems is to assign weights $W(\mathcal{R})$ to individual data records, such that their weighted contributions fall off with growing distance $\mathcal{R}$ from the query point in the parametric hyperspace. Various functional forms of $W(\mathcal{R})$ can be envisioned; here we chose a Gaussian weight function $W(\mathcal{R}) \propto \exp\left[-\alpha(\mathcal{R}/\mathcal{R}_c)^2\right]$,

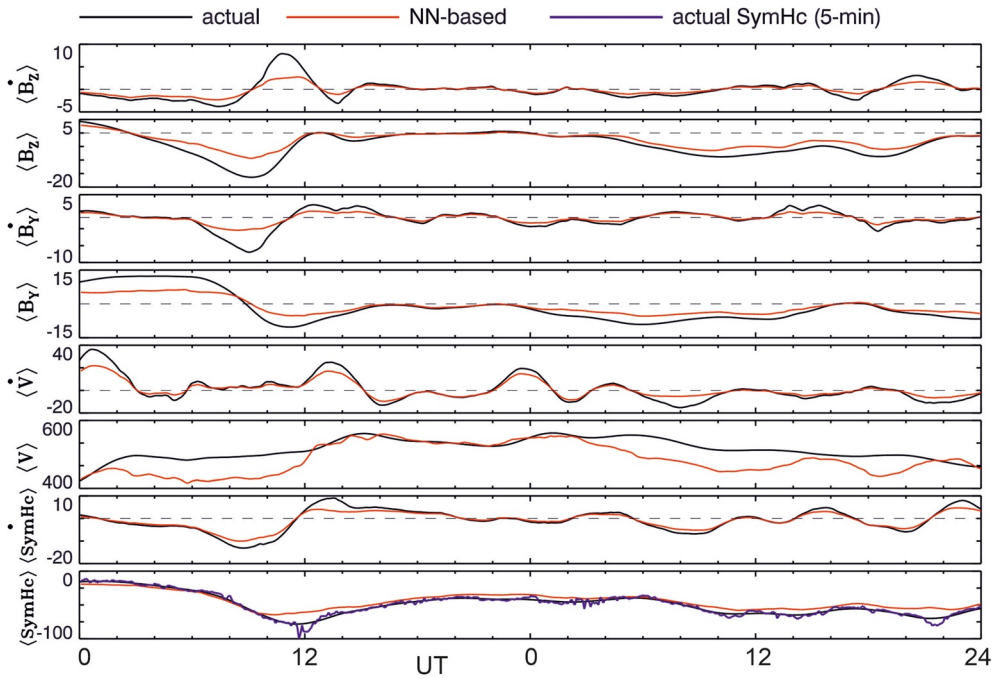


**Figure 4.** Variation of eight nearest-neighbor (NN) selection parameters over the 2-day interval February 3–4, 2022. Black and red curves correspond, respectively, to the actual parameter values during the event and their weighted average values calculated over NN subsets, compiled from the entire grand data pool. Time derivative units (dotted quantities) are in hours$^{-1}$.

also used by Sitnov et al. (2020). The critical distance $\mathcal{R}_c$ was assigned to be equal to the bounding radius of the NN subregion, calculated iteratively by the NN-selection algorithm based of a prescribed number of records $K$, sufficiently large to avoid instabilities due to overfitting (in this study, $K \approx 40{,}000$).

The factor $\alpha$ was introduced to fine tune the steepness of the weight function $W(\mathcal{R})$: larger/smaller values of $\alpha$ result in sharper/blunter $W$ profiles and, hence, better/poorer agreement between the NN parameter averages and the corresponding query point values. Based on many fitting experiments, a tradeoff value $\alpha = 4$ was assumed, providing, on the one hand, sufficiently stable sequences of model field configurations and, on the other, a reasonable agreement between the query point parameters and their NN-averaged values, as illustrated in Figure 4. The plot shows time variation of eight model state parameters during the 2-day interval of the storm, compared with their NN-average values calculated for each time moment over the corresponding NN subsets. There remains an unavoidable mismatch between the actual and NN-average parameters; as expected from the very principle of NN selection, the mismatch increases during active periods with negative IMF $B_z$ and decreases during activity lulls.

The above outlined procedure and its intricacies have already been described in detail in previous publications, from Sitnov et al. (2008) to more recent works (Sitnov et al., 2020; Stephens & Sitnov, 2021; N. A. Tsyganenko, Andreeva, & Sitnov, 2021). We briefly recapitulated it, not only to keep the paper more self-contained, but also to highlight an important new aspect, namely, the NN selection metrics. Aside from the Euclidian metrics commonly used in all previous studies:

$$\mathcal{R}_E = \left\{ \sum_{i=1}^{N} \left[ \langle G_i \rangle - \langle G_i \rangle^{(m)} \right]^2 + \left[ \langle \dot{G}_i \rangle - \langle \dot{G}_i \rangle^{(m)} \right]^2 \right\}^{1/2} \tag{2}$$

one may also consider the so-called "Manhattan" metric:

$$\mathcal{R}_M = \sum_{i=1}^{N} \left\{ \left| \langle G_i \rangle - \langle G_i \rangle^{(m)} \right| + \left| \left\langle \dot{G}_i \right\rangle - \left\langle \dot{G}_i \right\rangle^{(m)} \right| \right\} \tag{3}$$

or even more general case of $L_m$-norm metrics (Aggarwal et al., 2001):

$$\mathcal{R}_L = \left\{ \sum_{i=1}^{N} \left| \langle G_i \rangle - \langle G_i \rangle^{(m)} \right|^M + \left| \langle \dot{G}_i \rangle - \langle \dot{G}_i \rangle^{(m)} \right|^M \right\}^{1/M} \quad (4)$$

such that $M = 2$ and $M = 1$ correspond to Equations 2 and 3, respectively. As discussed by Aggarwal et al. (2001), one may go even further and explore the "fractional distance" metric, in which case the integer $M$ in Equation 4 is replaced by a floating point parameter $0 < f < 1$. In this work, based on comparing the results of using Euclidian, fractional, and Manhattan metrics, we finally chose the latter one, as it provided the most stable results, in line with Aggarwal et al. findings in regard to the dimensionality curse problem.

## 4. Model Architecture

As already noted in Section 1, the model used in this study is composed as a combination of (a) a base modular component, representing the magnetic field as a sum of contributions from several principal current systems, and (b) a more flexible higher-resolution component (superstructure), whose function is to correct systematic biases of the main modular part. Accordingly, the fitting algorithm operation was split into two phases: at the first phase, the modular component was fitted to the data, while all coefficients of the second component were kept at zero. At the second phase, all the obtained parameters of the modular component were frozen at their best-fit values, and the coefficients of the higher-resolution superstructure were fitted to the same data subset, to further reduce the r.m.s. Difference between the data and the first-approximation model field. In particular, the high-resolution part makes it possible to include smaller-scale radial variations of the magnetotail current, which increases the model's flexibility and allows us to reproduce such local details as the field depressions/peaks, reported in many studies of the stationary magnetospheric convection events (e.g., Kissinger et al., 2014; Sergeev et al., 1996).

In our first experiments with the hybrid approach (N. A. Tsyganenko & Andreeva, 2017), the base modular component was represented by the TA15 model, composed as a sum of contributions from the tail current, symmetric and partial ring currents (SRC/PRC), Region 1 field-aligned currents (R1 FAC), and the Earth's dipole shielding field. In its turn, the high-resolution part was based on the radial basis functions (RBFs; Andreeva & Tsyganenko, 2016). An undeniable advantage of the RBF models is their 3D structure, which allows to represent complex magnetic fields of a priori unknown geometry within a volume of interest, at least on a local scale. At the same time, the high flexibility and local nature of RBFs, combined with non-uniform data distribution may result in artificial spatial fluctuations of the model field in the regions of poor data coverage or gaps. In addition, a separate problem arises with large-scale global models, in which the total field is usually required to be confined (shielded) within a prescribed magnetopause. Although the contributions from each RBF source can in principle be individually shielded inside a given boundary, their large numbers (typically, up to several thousand) make the task computationally awkward/unfeasible. Since the foremost motivation of this study was to trace the magnetosphere distortion on a global scale, we eventually chose to represent the high-resolution component not by RBFs, but with sets of equatorial finite elements, first introduced in our 2007 publication (N. A. Tsyganenko & Sitnov, 2007) and then widely used in a number of more recent studies (e.g., Stephens et al., 2019; and refs. therein). A concise synopsis of both model components is outlined below.

### 4.1. Base Model (TA15)

The mathematical framework of the base model component was adopted with only minor changes from the TA15 work. The largest-scale field source in that model is the tail current disk (see Figure 9 in TA15), composed as a superposition of overlapping toroidal current rings with radially and azimuthally varying finite thickness $D = D(\rho, \phi)$. The disk current density steadily decreases tailward as $I(\rho) = I_0(R_N/\rho)^{\mu}$, where the exponent $\mu$ is a variable parameter defining the radial gradient of the tail lobe field. Two other nonlinear parameters of that module are (a) the geocentric distance $R_N$ to the inner edge of the current disk and (b) position of the disk center $X_c$ on the Sun-Earth line. Linear parameters include the current density $I_0$ at the inner edge of the disk, defining the overall magnitude of the tail field, and coefficients that quantify its dependence on the solar-wind pressure and the Earth's dipole tilt angle. Geometrical effects of the dipole tilt were taken into account by applying a 'bowl-shaped' deformation of the tail field; details of the numerical procedure can be found in an earlier publication (N. A. Tsyganenko, 2014, henceforth T14). As in all standard models of that kind, the disk magnetic field is shielded

inside a model magnetopause, which is effectively equivalent to the redirection of the extra-magnetospheric part of the disk current over the boundary (e.g., Sotirelis et al., 1994). Note that, while the original TA15 modules were shielded inside the solar wind pressure- and IMF Bz-driven magnetopause by Lin et al. (2010), in its present realization the IMF Bz effects were relegated to a flexible penetration term (see below). That allowed us to use a simpler fixed-shape variant of Lin et al. model with IMF $B_z = 0$, which not only simplified calculations but, most importantly, eliminated the problem of overrunning the model's validity limits during periods of anomalously large IMF $B_z$.

The SRC and PRC modules, originally derived in a semiempirical way by calculating the drift, magnetization, and field-aligned currents (see TA15, Section 4), were also transferred into the present model without modifications. Each one was parameterized by its magnitude coefficient, a nonlinear scale factor, and an "eversion" parameter, taking into account the day-night asymmetry of the source; the reader is again referred to the original TA15 paper for more details.

The R1 FAC field module was also adopted in almost the same form from the original TA15 model. It is based on a system of overlapping "wires" of finite radially varying thickness, extending from the ionosphere up to very large distances into the solar wind and deep magnetotail (see Section 5 and Figure 7 in T14). As in the case of the tail current, the model field of R1 FACs is shielded inside the magnetopause, which effectively creates a system of closure currents (e.g., Figure 4 in N. A. Tsyganenko, 2013). Three linear parameters of the R1 FAC module were the overall FAC magnitude and two coefficients, quantifying the modulation of its Northern and Southern parts by the dipole tilt angle. Two nonlinear parameters were (a) the noon $\theta_n$ and midnight $\theta_m$ colatitudes of the R1 FAC oval at ionospheric level, and (b) the radial eversion parameter $\nu$, controlling the distant shape of the FAC surface (see Equation 26 in T14 or Equation 19 in TA15 papers).

Finally, in the original prototype model (TA15, Section 4.5) the effect of partial IMF penetration inside the magnetosphere was included in the simplest form of a fixed fraction of the transverse component of external IMF. The penetrating part was assumed there uniform, normal to the Sun-Earth line, and proportional to the external $\mathbf{B}_{\perp}^{(\mathrm{IMF})}$ with a single penetration coefficient $\kappa$ and no $B_x$ component. In the present realization of the base model, a more sophisticated form of the IMF-related term is adopted, such that the penetrated field is split into two separate terms, driven, respectively, by $\mathbf{B}_{\perp}^{(\mathrm{IMF})} = B_{\mathrm{y}}^{(\mathrm{IMF})} + B_{\mathrm{z}}^{(\mathrm{IMF})}$ and $\mathbf{B}_{\mathrm{x}}^{(\mathrm{IMF})}$. Each penetrated term is a curl-free field of the form $\mathbf{B}_{\perp}^{(\mathrm{pen})} = -\nabla U_{\perp}$ and $\mathbf{B}_{\parallel}^{(\mathrm{pen})} = -\nabla U_{\parallel}$, where the potentials $U_{\perp}$ and $U_{\parallel}$ are represented by "box" harmonics $U_{\perp} \propto \exp(ax)\cosh(by)\sin(\gamma z)$ and $U_{\parallel} \propto \exp(ax)\cosh(by)\cos(\gamma z)$. Here the $y$- and $z$-axes are rotated around the Sun-Earth line, such that $z$-axis is aligned parallel to $\mathbf{B}_{\perp}^{(\mathrm{IMF})}$ and $y/z$-parities of the potentials match those of $\mathbf{B}_{\perp}^{(\mathrm{IMF})}$ and $\mathbf{B}_{\mathrm{x}}^{(\mathrm{IMF})}$. The penetrated field magnitudes are assumed as linear functions of the corresponding IMF components; the unknown proportionality coefficients and nonlinear scaling parameters are derived by fitting the entire model to consecutive NN data subsets.

### 4.2. Hybrid Superstructure (TS07 Expansion)

The TS07 hybrid field component is represented with a flexible superposition of contributions from an equatorial system of electric currents, spatially overlaid on the modular TA15 tail current and having the same x- and y-distribution of the current sheet thickness. The net field vector is a sum of three expansions

$$\mathbf{B}(\rho,\phi,z) = \sum_{m=1}^{M} a_m^{(s)}\,\mathbf{B}_m^{(s)} + \sum_{m=1}^{M}\sum_{n=1}^{N} a_{mn}^{(o)}\,\mathbf{B}_{mn}^{(o)} + \sum_{m=1}^{M}\sum_{n=1}^{N} a_{mn}^{(e)}\,\mathbf{B}_{mn}^{(e)} \tag{5}$$

where the first one is derived from axisymmetric vector potentials with different wave numbers $m$ (but, in its final form, not axisymmetric because of the varying sheet thickness and the shielding field asymmetry). The second and third double sums in Equation 5 include both radial and azimuthal Fourier harmonics and represent two magnetic field families, both of which are intrinsically asymmetric in longitude, but have different symmetry properties/parities with respect to the noon-midnight meridian plane, reflected in their superscripts $(o)$ and $(e)$. The upper summation limits in Equation 5 were set in this work at $M = 8$ and $N = 6$, while the basic radial scale distance entering in the vector potentials was assumed equal to $R_0 = 30\ R_E$, commensurate with the outer limit of data and the modeling region boundary. This defines the highest radial and azimuthal resolutions of the TS07 correction component on the order of $\delta\rho \sim R_0/M \sim 4R_E$ and $360/N \sim 60°$, respectively. Explicit equations for $\mathbf{B}_m^{(s)}$,

$\mathbf{B}_{mn}^{(o)}$, and $\mathbf{B}_{mn}^{(e)}$ in Equation 5 are rather long and page space consuming; the interested reader can find them in the original TS07 paper.

As in all previous studies that employed the TS07 equatorial sources, each term in the sums (Equation 5) was separately shielded within the common model magnetopause, using scalar potentials in the form of box harmonic expansions with required parity/symmetry properties, similar to those mentioned in Section 4. Finally, the same bowl-shaped transformation of the shielded TS07 field was applied, to take into account the dipole tilt-related distortion of the configuration.

## 5. Fitting Results

For the entire 2-day duration of the event, 48 NN data subsets were generated, each one containing nearly 40,000 data records and following at one-hour cadence of UT moments. Two-step hybrid fitting was carried out for each NN subset, and the obtained set of 48 model parameter files was used to calculate 48 consecutive configurations of field lines with starting footpoints in the noon-midnight meridian plane. Figures 5 and 6 display nine plots, following at two-hour intervals of UT and covering the most disturbed first 16 hr of the storm (DOY 34 and 35, respectively). A full set of 48 plots is provided separately in the SI attachment to the paper. Due to the large IMF $B_y$ (and, hence, its penetrated part) during the most of the event, some of the near-magnetopause and near-cusp field lines get swept in the dawn-dusk direction out of the plotting plane; in the top three panels that makes them appear as crossing the rest of the lines. For the same reason, in some panels there is an apparent gap between the field lines and the model magnetopause. To help visualize the highly variable distortions of the magnetic configurations in the course of the storm, the field lines with northern footpoint latitudes $\lambda_{SM} = 66°$, 68°, and 70° are colored, respectively, by red, blue, and magenta.

In the beginning of the event (UT = 00:00 and 02:00), one sees typical quiet-time configurations, with the red-colored lines ($\lambda_{SM} = 66°$) confined well inside the inner magnetosphere and the northern polar cusp mapped to $\lambda_{SM} \approx 76°$–77°. By UT = 04:00, IMF $B_z$ becomes slightly negative, resulting in already tangible response both in the midnight and noon sectors. By UT = 06:00, IMF $B_z$ falls to ∼−7 nT, the nightside line $\lambda_{SM} = 68°$ stretches tailward beyond 20 $R_E$, and the northern cusp footpoint shifts equatorward to $\lambda_{SM} = 73°$. The most dramatic changes are seen in the UT = 08:00 and 10:00 panels, corresponding to the negative peak of IMF $B_z \sim -18$ nT and indicating the emergence of a thin tongue-shaped stretching of the nightside field lines in the vicinity of synchronous orbit, which map as low as to $\lambda_{SM} = 63°$ at UT = 10:00. At that same moment, the line $\lambda_{SM} = 66°$ extends tailward all the way beyond the boundary of the modeling region, with formation of a pair of X- and O-type neutral lines at $R \sim 15$–20 $R_E$. The dayside cusp in that panel maps to unusually low latitude $\lambda_{SM} = 67°$–68°, significantly lower than typical values ($\gtrsim 70°$) obtained in most of previous statistical models, except for the case of much stronger X-class megastorm modeling (see N. A. Tsyganenko et al., 2003, Figures 9 and 11). The obtained local stretching of the magnetic field at synchronous distance was repeatedly reported and discussed in previous case (Kaufmann, 1987) and statistical (Pulkkinen, 2007) studies. In this work, we concentrate on the magnetic connectivity aspect of this issue by presenting the mapping results in their dynamics for the specific storm event.

The next panel (left in bottom row) is separated by only 2 hr (UT = 12:00) from the previous one, but shows a strikingly different, completely closed magnetosphere. According to the center plot in Figure 1, at that time IMF $B_z$ rapidly reversed its polarity and stayed large and positive for about 1 hr, which was promptly reflected in the magnetic configuration, especially on the dayside. The IMF jump coincided with arrival of very dense solar wind, which resulted in an overall strong compression of the magnetosphere.

The last two panels (UT = 14:00 and 16:00) correspond to the next southward excursion of IMF $B_z$. It was much weaker than that in the morning hours and was accompanied with a significant reduction of the solar wind pressure. Accordingly, the model responds with an overall expansion of the magnetopause and somewhat more stretched tail field on the nightside.

As already mentioned in Section 1, the next round of the storm on DOY 35 (February 4) proceeded rather differently from that on DOY 34. It began with an almost 4-hr-long period of enhanced solar wind density/pressure ($P_{dyn} \sim 10$ nPa) which resulted in a long and strong sudden commencement, manifested by a quick jump of Sym-H to zero in the end of DOY 34. The IMF behavior was also quite different from that in the first day: instead of a steady southward excursion, only a brief negative pulse of $B_z$ to ∼−6 nT occurred, followed by ∼4-hr period

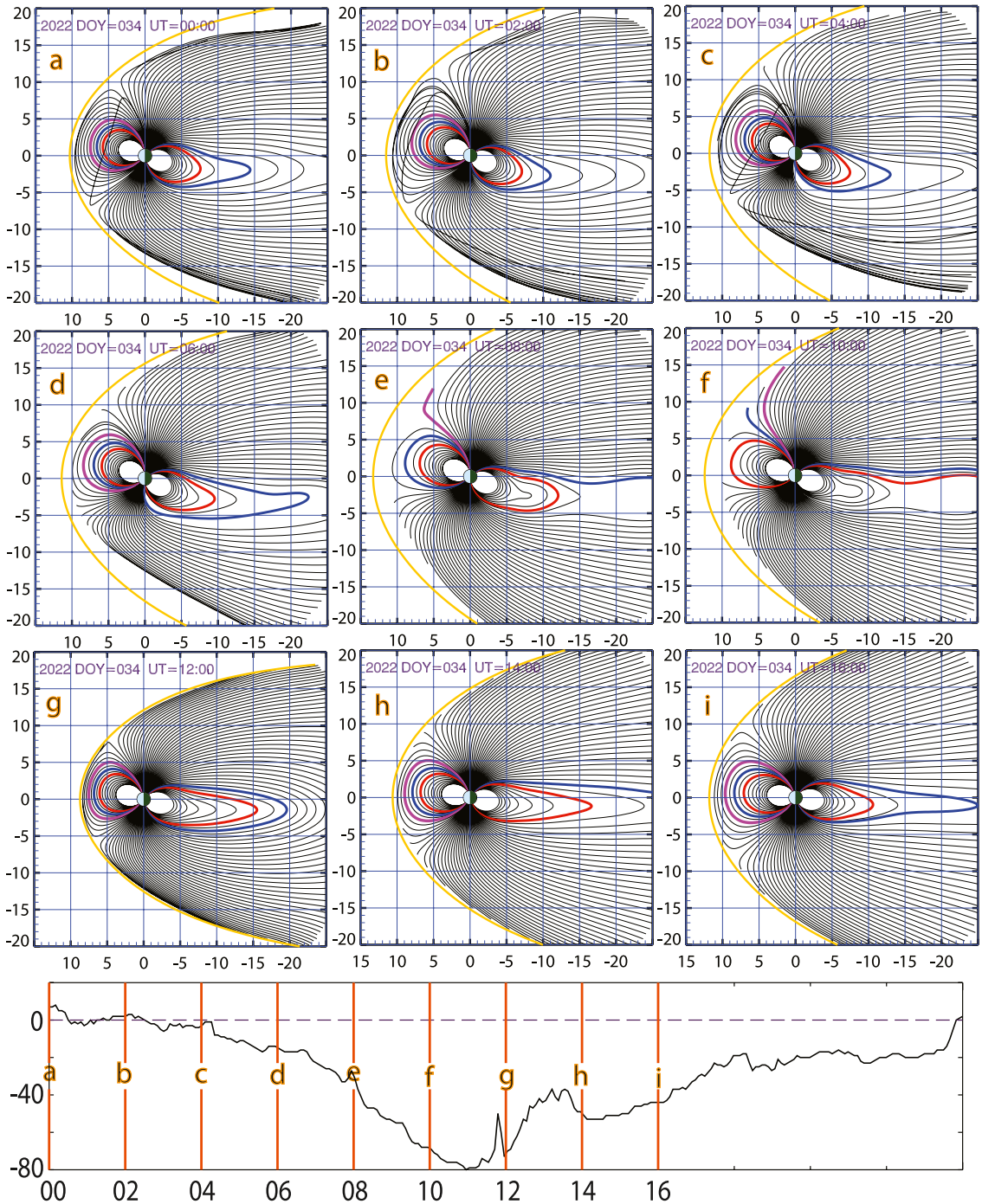


**Figure 5.** Magnetic field line configurations projected on the noon-midnight meridian plane for nine equidistant UT moments of the first day of the February 3–4, 2022 storm. The lines are plotted at 1° cadence of solar-magnetic footpoint latitude $\lambda$, starting from $\lambda = 59°$; those with $\lambda = 66°$, 68°, and 70° (dayside only) are highlighted by red, blue, and magenta, respectively; yellow contour shows position of the model magnetopause intersection with $Y_{GSM} = 0$ plane. Each panel is labeled by the corresponding date and UT. Bottom panel shows Sym-H index variation for the same day; orange bars labeled by letters a–i indicate time moments corresponding to the nine field line plots.

of fluctuations around zero (although with a tangible IMF $B_y \sim -5$ nT). After that, IMF $B_z$ gradually lowered to almost $\sim -10$ nT and stayed around that level (except a short fluctuation to positive $B_z$ at 16 UT) until ~21 UT.

During the DOY 35 the interplanetary driving was significantly weaker than on DOY 34, which resulted in a weaker magnetospheric reaction, both in terms of the Sym-H index and in the geomagnetic field line

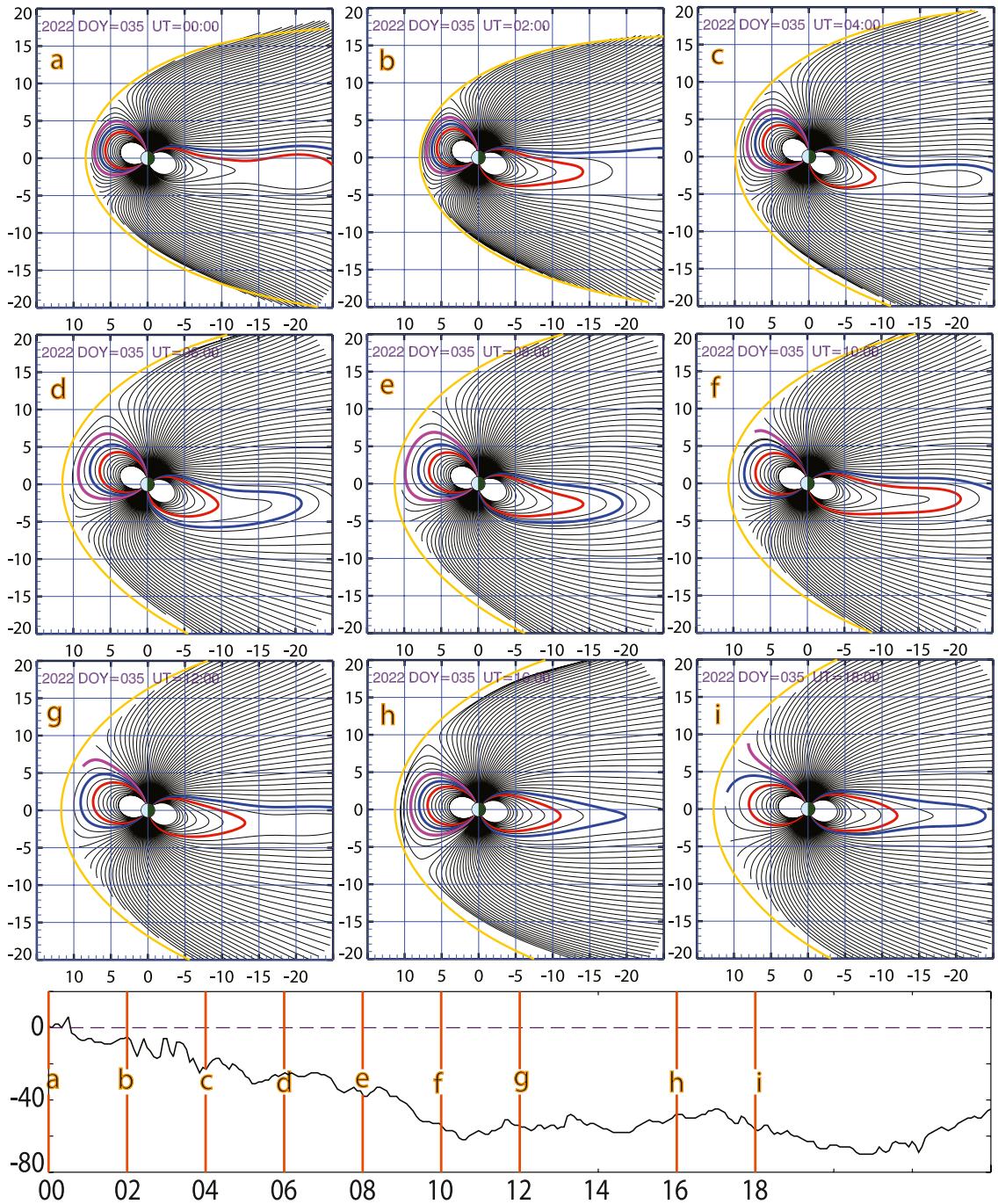


**Figure 6.** Same as in Figure 5, but for DOY 35 of 2022.

configurations. As shown in the next Figure 6, the largest nightside field distortion was observed in the very beginning of the day, obviously in response to the sharp increase in the solar wind pressure/density, coincident with the abrupt southward pulse of IMF $B_z$.

However, even though the 66° field line again distended down to $-25R_E$, as it did at 10:00 UT on DOY 34 (see in Figure 5), this time we do not see any tongue-like stretch near the synchronous distance, nor any unusual equatorward shift of the dayside cusp, which resided only slightly lower (at 74°) from its quiet-time location. Configurations in other panels also do not deviate too much from what one may expect for a moderately strong

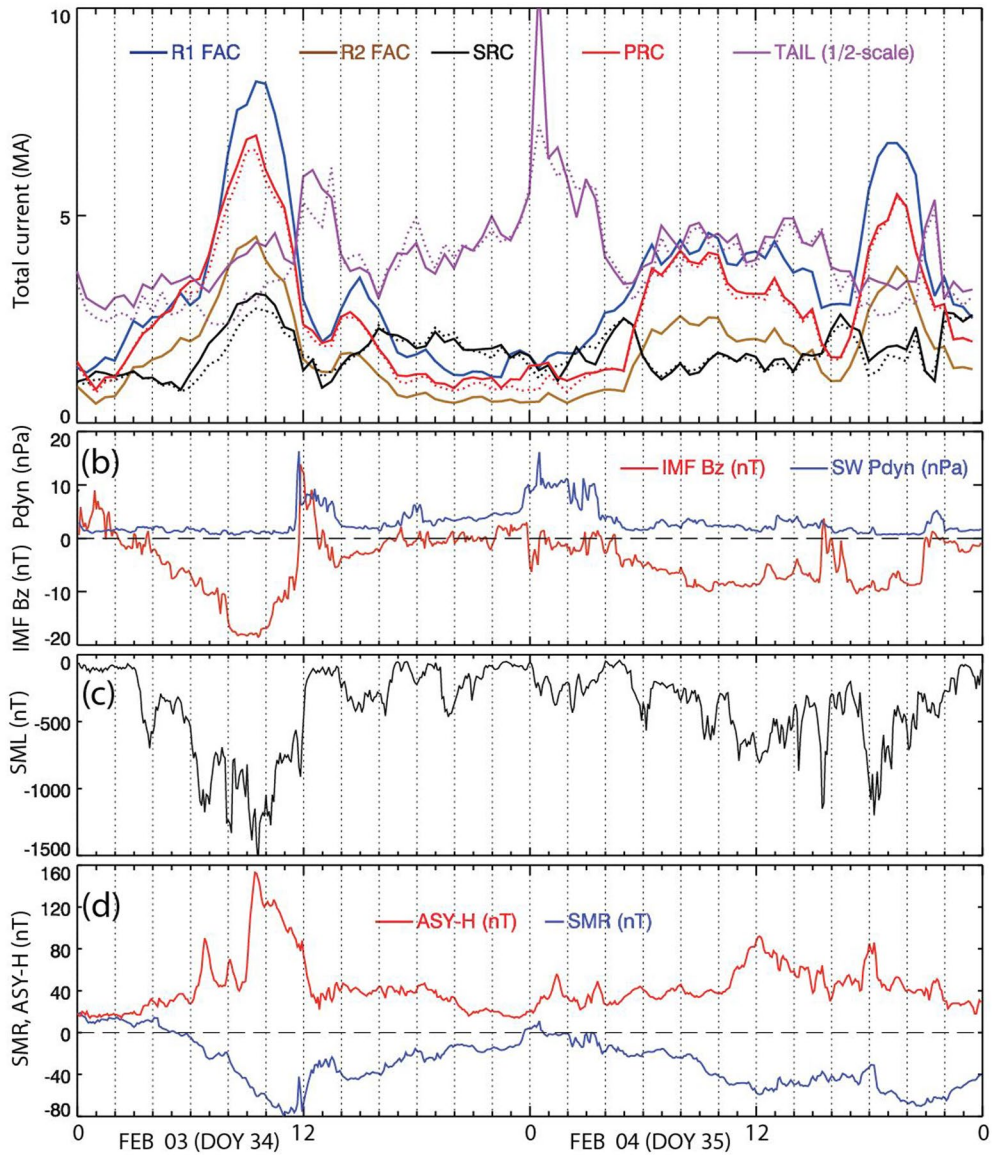


**Figure 7.** (a) Variation of the total electric current in the main TA15 current systems during the storm. R1 FAC, R2 FAC, partial (PRC), symmetric (SRC) ring, and tail currents are shown by blue, brown, red, black, and magenta, respectively. (b) Concurrent variation of interplanetary magnetic field $B_z$ (red) and solar wind pressure (blue). (c) Variation of the SuperMag auroral index. (d) SuperMag ring (blue) and ASY-H (red) indices.

disturbance, except maybe for the moments UT = 10:00 and 18:00, when IMF $B_z$ reached ∼−10 nT and prompted a significant equatorward shift of the polar cusps to 71° and 69°, respectively.

## 6. Discussion

The reconstructed dynamics of the storm-time magnetosphere raises a natural question on the relative role of different magnetospheric current systems in terms of their contributions to the obtained dramatic variability of magnetic configurations. The assumed structure of the model, based on the TA15 modules, allows to estimate their relative magnitudes and visualize their evolution during the event, as shown in Figure 7. Its top panel illustrates the variation of the total currents in megaamperes, calculated by integrating $\mathbf{j} = \nabla \times \mathbf{B}/\mu_0$ for each field module over the corresponding areas permeated by the current flow lines. Thus, the total R1 and R2 FACs were both calculated over a spherical cap of radius 1.12 $R_E$, corresponding to the altitude of Iridium satellites (780 km) within SM colatitude limit of 50°. The total SRC and PRC were calculated by integrating over rectangular areas

$2 \le \rho_{SM} \le 10\ R_E$, $|Z_{SM}| \le 5\ R_E$, lying either in the midnight meridian $\phi_{SM} = 180°$ (for SRC), or at $\phi_{SM}$, corresponding to best-fit values of the PRC peak SM longitude (usually around $\phi_{SM} \sim 150°$ in the pre-midnight sector). Dotted lines represent the currents calculated at the first phase of fitting the model, that is, before the TS07 hybrid correction. The other three panels reproduce the concurrent variations of the IMF $B_z$ and solar wind pressure (b), auroral SML index (c), and SuperMag ring (SMR) and ASY-H indices (d).

In the top panel one clearly sees three waves of growing magnitude of the field sources, of which the most outstanding are those for the R1 FAC (blue) and PRC (red) modules. Their rising slopes distinctly match three periods of falling SMR index, coincident with three long intervals of southward IMF. The total R2 FAC (brown) varies in close concert with R1 and PRC, and remains roughly twice smaller than the R1 FAC during the entire storm period. The latter significantly disagrees with a result by Coxon et al. (2014) based on AMPERE data, where the ratio R1/R2 was found to be no larger than 1.15. The most likely reasons are (a) a crude empirical geometry of the model R1/R2 FACs and (b) a significant spatial and temporal variability of the actual FAC magnetic effects, which results in the mixing of NN data in both geometric and parametric spaces, as manifested in the incomplete matching of the current and NN-averaged parameters in Figure 4.

The obtained model estimates of R1 and R2 total currents could in principle be compared against their values provided by the AMPERE online interface https://ampere.jhuapl.edu/ for the same time period. However, in view of possible ambiguities in separating the R1 and R2 contributions, we chose instead a more straightforward approach, based on direct comparison of the model output with net magnetic disturbances observed above the northern polar cap at Iridium altitudes of ~780 km and represented in the form of AMPERE summary plots. At the peak of the storm (UT = 09:00–10:00 of DOY 34), the model field due to combined effect of R1 and R2 at $X_{SM} = Y_{SM} = 0$, $Z_{SM} = 1.12 R_E$ was found equal to $\Delta B_x = 280$–$310$ nT, while the AMPERE plots for the same time interval (more exactly, 09:30–09:40) showed horizontal sunward-oriented vectors with magnitudes on the order ~250 − 270 nT. The most likely source of the 30–40 nT overshoot in the total model field is the above mentioned underestimate of the R2 FAC magnitude. In any case, the mismatch does not appear to be of serious concern, in view of the rather simplistic geometry of TA15 model FACs, not intended to represent their low-altitude effects, as well as the relative scarcity of high-latitude data and their total absence in the innermost magnetosphere at $R \le 2$–$3\ R_E$.

The PRC variation (red) is very similar to that of R1 and R2; the total current magnitude is roughly 60%–80% lower than that of R1, but significantly larger than that of R2. The latter result is consistent with the fact that the entire azimuthal PRC is by construction diverted into or out of the ionosphere via R2 FACs, being nearly equally divided between the Northern and Southern hemispheres. It is worth noting a close time coincidence of R1, R2, and PRC peaks with the peak of the ASY-H index in the bottom panel, indicating their close association with the development phase of the storm, when the dawn-dusk asymmetry of the inner magnetosphere is the largest and the external driving and internal convection reach their maximum.

The net model SRC magnitude (black trace in Figure 7) lies within 1–3 MA; its variation range through the storm is much less dramatic than that of other currents, such that the net SRC dominates over PRC only during the recovery phase of the first active period (between ~17:00 of DOY 34 and ~05:00 of DOY 35), in line with the concept of PRC as a principal source of the storm-time disturbance (Liemohn et al., 2001). During the active storm phases, one sees a significant effect of the TS07 correction, such that the corrected plot lies higher than that based on only TA15 model.

The tail current plot (magenta) was obtained by integrating dawn-dusk current density over midnight meridian plane in the range $|Z_{GSM}| \le 10\ R_E$, sufficiently wide to ignore the dipole tilt-related shifts of the current sheet. The integration limits were defined somewhat arbitrarily as $-30 \le X_{GSM} \le -7\ R_E$; because of such a wide radial span, the obtained total tail current is much larger than those of all other sources. For that reason and in order to keep all 5 plots in the same panel, the total tail current was divided by 2 and thus shown at 1/2 scale. As expected, its dependence on IMF $B_z$ is not as ordered as for the other sources; here the principal driving factor is the solar wind pressure, manifested in two strong peaks around 12:00 and 24:00 UT of DOY 34. Again, one sees a significant correction of the hybrid profile due to added TS07 expansions, in particular, during morning UT hours of DOY 34.

The latter raises a natural question of the importance of the TS07 correction in terms of magnetic field geometry. Figure 8 presents an example of comparing two noon-midnight configurations, corresponding to UT = 10:30 of

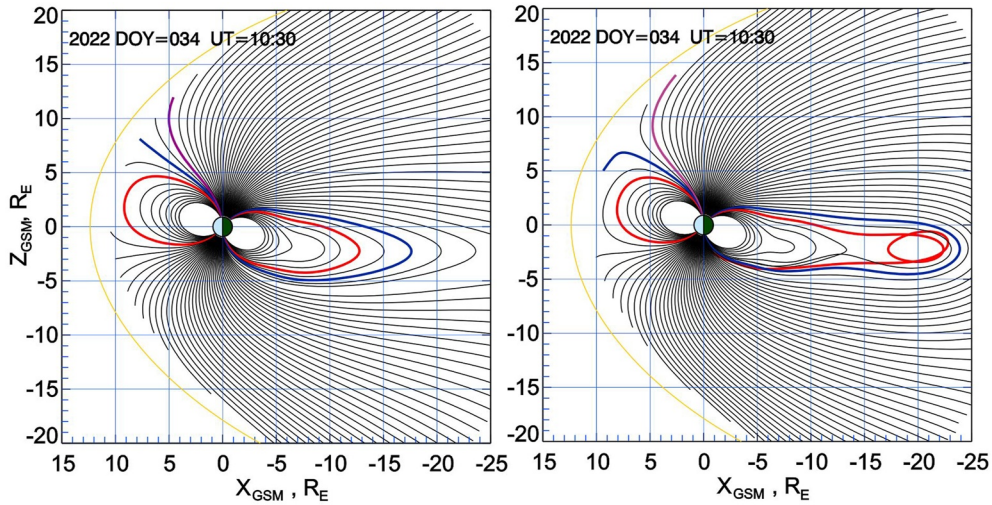


**Figure 8.** Compares field line configurations based on uncorrected (pure TA15; left panel) and full hybrid (with added TS07 contribution; right panel) for the period of peak storm intensity on February 3 (DOY 34).

DOY 34. The left panel shows field lines, plotted using the uncorrected (purely TA15) model, and the right panel displays field lines for the same UT moment but after having added the TS07 correction.

Despite a relatively small decrease in the residual field rms$|\langle \Delta \mathbf{B} |\rangle$ from 25.33 to 24.15 nT (with rms$\langle |\mathbf{B}| \rangle = 59.68$ nT), the corrected configuration looks significantly different: the distant field is much more stretched and includes a plasmoid-like feature with entangled spiraling of the magnetic field lines. Note in this regard that actual instantaneous field configurations at these distances are inherently very dynamic and diverse, while the NN subsets used in generating the model field by necessity include data taken rather far from the query points in the state space. For this reason, we prefer not to directly interpret such features in terms of plasmoids.

The next important question, prompted by the tongue-like shape of nightside field lines at $R \sim 6$–$8\ R_E$ in Figures 5e and 5f and 8, concerns the azimuthal distribution of the magnetic depression and the degree of its severity during the most active phase of the storm. It is well known since long ago that the largest storm-time distortion in the inner magnetosphere occurs in the post-dusk/pre-midnight sector (e.g., N. A. Tsyganenko et al., 2003, and references therein). To that end, we explored model distributions of the total near-equatorial magnetic field for the same sequence of UT moments as in Figure 5. More specifically, in order to take into account the varying geodipole tilt, the field was calculated not in the GSM equatorial plane, but on the bowl-shaped surface, corresponding to the center of the model tail current sheet and ring current. The result is presented in Figure 9 as a color-coded diagram of $B_z$ component of the total (i.e., model plus geodipole) magnetic field for the time moment UT = 10:00 of DOY = 34, corresponding to the meridional configuration in Figure 5f. Heavy blue isointensity contours delineate the regions of reversed $B_z < 0$ and show that, besides the wide area of negative $B_z$ at $R \sim 15$–$25\ R_E$, there is a narrow siecle-shaped region of strongly depressed field at much closer distances. In the immediate vicinity of geosynchronous orbit ($R \sim 7\ R_E$, MLT$\sim$21:00), the depression gets so strong that the total field drops to negative values.

As already noted, in spite of the highly flexible mathematical framework, the data mining and fitting procedures include many assumptions of technical nature, whose specific choice may affect the results of the modeling. For example, other forms of the weighting functions (both in geometric and parametric spaces) can be tested, in particular, anisotropic weighting instead of the simple isotropic Gaussians used in this study. Other options could be to try different combinations of NN selection variables, as well as different time integration limits $T$ in the convolutions Equation 1. Finally, the spatial resolution of TS07 expansions can be modified by changing the number of harmonics and/or their principal radial wave length $R_0$.

In view of that, one may pose a legitimate question on the reliability of the above results; in other words, to what extent the obtained configurations in Figures 5 and 6 are close to the actual ones? As already discussed earlier (e.g., N. A. Tsyganenko, Andreeva, Sitnov, Stephens, et al., 2021, Section 6), the commonly used validation

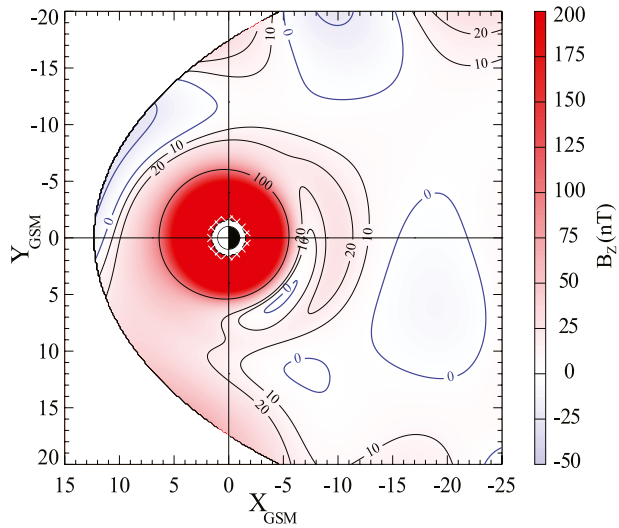


**Figure 9.** Distribution of the total near-equatorial $B_z$, obtained for UT = 10:00 of DOY = 34. Compare with panel f of Figure 5. Due to the rapidly growing contribution from the Earth's dipole, the colors are assumed to saturate at the upper limit $B_z$ = 200 nT. Blue contours delineate areas with reversed (negative) $B_z$ polarity.

techniques are hardly applicable here: unlike in the standard approach, our NN data mining is an example of the instance-based machine learning, which generates unique models, relevant to only specific time moments of interest. In many preliminary studies made in the course of this work, various modifications of the model and fitting procedures were tried, such as different numbers of NN subset records (from 15,000 to 80,000), averaging window lengths, other forms of the weighting function (e.g., an exponential instead of the Gaussian), other metrics, etc. In spite of all those changes, a remarkable degree of repeatability was found, in particular, unusually low latitudes of dayside cusp footpoints, persistent formation of neutral lines at $X \sim -15\ R_E$, and the extremely stretched nightside field around the synchronous distance during the storm culmination on DOY 34. This gives us sufficient confidence in the realism of the above described results and provides enough motivation for further research in this direction. In particular, there is an ample room to improve the model's spatial resolution by switching from the global to a local nightside field modeling at low/middle latitudes (N. A. Tsyganenko, Andreeva, & Sitnov, 2021), with help of a much more detailed representation of the correction component by RBFs instead of the TS07 expansions used in this work. Such studies extend well beyond the scope of this paper and are planned for the nearest future.

A fortunate possibility to locally validate this model opened up recently owing to the new GOES high-resolution data, made available by the NSSDC through the CDAWeb interface (https://cdaweb.gsfc.nasa.gov/cdaweb/). Figure 10 presents results of comparing the model output with GOES-16 and GOES-17 data taken during the 2-day period of interest.

While the agreement is quite satisfactory for GOES-16, in the case of GOES-17 one sees large negative excursions of the observed $B_y$, peaking at ~12:00 UT on February 3 and at ~10:00 on February 4. Both drops of $B_y$ (and somewhat weaker variations in $B_x$) are completely unaccounted for by the model. In both cases, the discrepancies

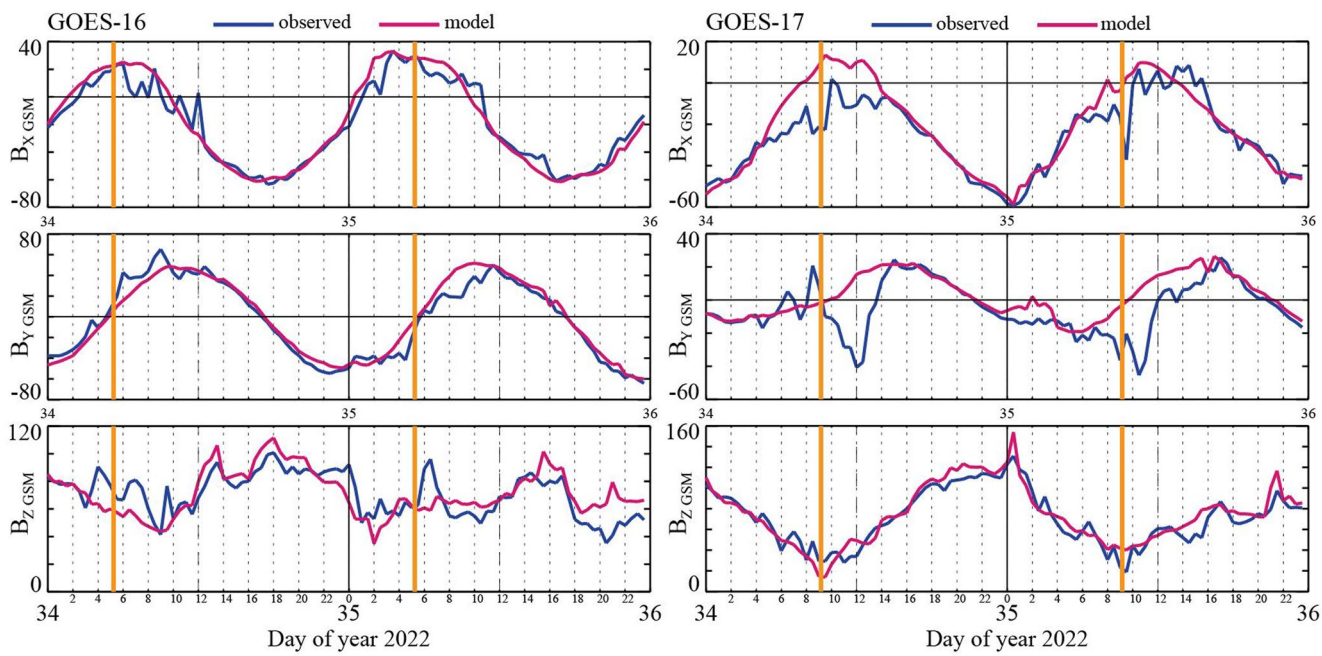


**Figure 10.** Variation of three GSM components of the total magnetic field along the orbits of GOES-16 (left) and GOES-17 (right) satellites, compared with this model prediction. The observed and model field variation is shown by blue and red lines, respectively. Yellow vertical bars correspond to the UT moments of the satellite passages through the midnight meridian.

fall mostly on the post-midnight sector and coincide in time with SML pulses (see Figure 1 or 7) and (on February 3) with the huge spike of the solar wind pressure. It is worth noting that there is virtually no trace of such a dramatic simultaneous variation in the $B_z$ component, which makes us conclude that the effect is caused by strong field-aligned currents flowing mostly parallel to the equatorial plane. Due to the lack of SCW-like structures in the base model and their absence in the high-resolution TS07 component, the observed jerks in $B_y$ and $B_x$ could not be captured by the model. This issue needs a further treatment, based on a fully 3D high-resolution representation, for example, using an RBF or cylindrical basis functions local modeling (N. Tsyganenko, Andreeva, Kubyshkina, et al., 2021; N. A. Tsyganenko, Andreeva, & Sitnov, 2021; N. A. Tsyganenko, Andreeva, Sitnov, Stephens, et al., 2021) and a more advanced NN selection procedure, explicitly including the auroral indices as indicators of substorm activations.

Note also that both the modeled and observed $B_z$ profiles along the GOES-17 orbit do not reach zero, which is most likely due to two factors. First, the GOES position is located northward from the dipole equator, that is, somewhat away from the field depression center. Second, as can be seen from the plot, by the time of disturbance peak at UT = 09:00 GOES-17 already crossed the midnight meridian and entered in the dawn sector, where the model $B_z$ depression gets significantly weaker (Figure 9).

As a final comment, we again emphasize that this work concentrates solely on the magnetic field distortions and does not touch upon the physical mechanism of magnetosphere-ionosphere-thermosphere coupling behind the Starlink failure (Dang et al., 2022). In this regard, a highly relevant reference is a recent paper by Angelopoulos et al. (2020), who showed that the near-Earth reconnection is far more geoefficient because the energy of the reconnected electrons grows as $V_A^2 \propto B_{\text{lobe}}^2$. In view of this, the obtained evidence of nightside X-line formation at unusually close distances may be quite telltale.

## 7. Summary

An empirical modeling of the magnetospheric magnetic field dynamics was carried out for the storm-time period of February 3–4, 2022, that received much attention in the media due to the unexpected loss of 38 Starlink satellites. The modeling is based on the new approach to mine large sets of historical data archives, combined with a hybrid method of the magnetic field representation, unifying a modular and high-resolution components into a single framework. In spite of a relatively mild storm intensity in terms of the ground low-latitude disturbance, the obtained magnetic configurations revealed a wide range of field line deformations induced by the dynamical external driving. In particular, an unusually enhanced distortion of the inner magnetosphere was obtained at the peak of the storm, with formation of a relatively thin area of strongly stretched and depressed tail-like field in the vicinity of synchronous orbit, probably indicating the formation of neutral lines at such unusually close distances. During the same period, the distant nightside field became extremely distended and formed magnetic neutral lines with entangled spiral structures as close as at $R \sim 12$–$15\ R_E$.

## Data Availability Statement

Geotail, Polar, Cluster, Van Allen Space Probes, and MMS data were downloaded from the NSSDC CDAWEB interface at https://cdaweb.gsfc.nasa.gov/cdaweb/index.html. THEMIS data were obtained from UC Berkeley repository (http://themis.ssl.berkeley.edu/index.shtml) and partly from the above NSSDC CDAWEB interface. High-resolution OMNI data were obtained from SPDF OMNIWEB at https://omniweb.gsfc.nasa.gov. SuperMag indices are available from https://supermag.jhuapl.edu. All data sets and software used to obtain the above described results are available from Zenodo repository: https://zenodo.org/record/7385214.

**Acknowledgments**
It is a great pleasure to acknowledge all teams and PIs of all experiments whose data contributed to this study. Specifically, Geotail MGF data were provided by S. Kokubun (STEL), and T. Nagai (Tokyo Institute of Technology). The data of Polar MGF experiment were made available online by the UCLA Polar team led by C. Russell. The magnetic field and ephemeris data of Cluster tetrahedron were made available via CDAWEB by A. Balogh and M. Tatrallyay. Van Allen Space Probes data were kindly provided by C. Kletzing. We acknowledge J. Burch, C. Russell, and W. Magnus for the use of Magnetospheric MultiScale data. We thank V. Angelopoulos, K.-H. Glassmeier, U. Auster, and W. Baumjohann for the use of THEMIS FGM data. Thanks are due to the AMPERE team and the AMPERE Science Center for providing open access to the Iridium Summary Plots. We also thank the SuperMag Project (J. Gjerloev) for providing the ground-based disturbance indices. We appreciate services of the NSSDC SPDF OMNIWEB facility (R. McGuire, N. Papitashvili) for the high-resolution interplanetary data and ground-based indices. This work has been funded by the Russian Foundation for Basic Research (RFBR) Grant 20-05-00218. MIS and GKS acknowledge the support of NASA Grant 80NSSC19K0074.

## References

Aggarwal, C. C., Hinneburg, A., & Keim, D. A. (2001). On the surprising behavior of distance metrics in high dimensional space. In J. Van den Bussche & V. Vianu (Eds.), *Database theory* (pp. 420–434). Springer. https://doi.org/10.1007/3-540-44503-x_27

Andreeva, V. A., & Tsyganenko, N. A. (2016). Reconstructing the magnetosphere from data using radial basis functions. *Journal of Geophysical Research: Space Physics*, *121*(3), 2249–2263. https://doi.org/10.1002/2015JA022242

Angelopoulos, V., Artemyev, A., Phan, T. D., & Miyashita, Y. (2020). Near-Earth magnetotail reconnection powers space storms. *Nature Physics*, *16*(3), 317–321. https://doi.org/10.1038/s41567-019-0749-4

Boynton, R. J., Balikhin, M. A., Billings, S. A., Wei, H. L., & Ganushkina, N. (2011). Using the NARMAX OLS-ERR algorithm to obtain the most influential coupling functions that affect the evolution of the magnetosphere. *Journal of Geophysical Research*, *116*(A5), A05218. https://doi.org/10.1029/2010JA015505

Coxon, J. C., Milan, S. E., Clausen, L. B. N., Anderson, B. J., & Korth, H. (2014). The magnitudes of the regions 1 and 2 Birkeland currents observed by AMPERE and their role in solar wind-magnetosphere-Ionosphere coupling. *Journal of Geophysical Research: Space Physics*, *119*(12), 9804–9815. https://doi.org/10.1002/2014JA020138

Dang, T., Li, X., Luo, B., Li, R., Zhang, B., Pham, K., et al. (2022). *Unveiling the space weather during the Starlink satellites destruction event on 4 February 2022* (Vol. 20. e2022SW003152). Space Weather. https://doi.org/10.1029/2022SW003152

Gershenfeld, N. (2003). *The nature of mathematical modeling*. Cambridge University Press.

Kaufmann, R. L. (1987). Substorm currents: Growth phase and onset. *Journal of Geophysical Research*, *92*(A7), 7471–7486. https://doi.org/10.1029/ja092ia07p07471

Kissinger, J., Kepko, L., Baker, D. N., Kanekal, S., Li, W., McPherron, R. L., & Angelopoulos, V. (2014). The importance of storm time steady magnetospheric convection in determining the final relativistic electron flux level. *Journal of Geophysical Research: Space Physics*, *119*(9), 7433–7443. https://doi.org/10.1002/2014JA019948

Liemohn, M. W., Kozyra, J. U., Thomsen, M. F., Roeder, J. L., Lu, G., Borovsky, J. E., & Cayton, T. E. (2001). Dominant role of the asymmetric ring current in producing the stormtime Dst. *Journal of Geophysical Research*, *106*(A6), 10883–10904. https://doi.org/10.1029/2000ja000326

Lin, R. L., Zhang, X. X., Liu, S. Q., Wang, Y. L., & Gong, J. C. (2010). A three-dimensional asymmetric magnetopause model. *Journal of Geophysical Research*, *115*(A4), A04207. https://doi.org/10.1029/2009JA014235

Newell, P. T., Sotirelis, T., Liou, K., Meng, C.-I., & Rich, F. J. (2007). A nearly universal solar wind—Magnetosphere coupling function inferred from 10 magnetospheric state variables. *Journal of Geophysical Research*, *112*(A1), A01206. https://doi.org/10.1029/2006JA012015

Pulkkinen, T. I., Partamies, N., Kissinger, J., McPherron, R. L., Glassmeier, K.-H., & Carlson, C. (2013). Plasma sheet magnetic fields and flows during steady magnetospheric convection events. *Journal of Geophysical Research: Space Physics*, *118*(10), 6136–6144. https://doi.org/10.1002/jgra.50574

Pulkkinen, T. I., Partamies, N., McPherron, R. L., Henderson, M., Reeves, G. D., Thomsen, M. F., & Singer, H. J. (2007). Comparative statistical analysis of storm time activations and sawtooth events. *Journal of Geophysical Research*, *112*(A1), A01205. https://doi.org/10.1029/2006JA012024

Runov, A., Angelopoulos, V., Weygand, J. M., Artemyev, A. V., Beyene, F., Sergeev, V., et al. (2022). Thin current sheet formation and reconnection at $X \sim -10\ R_E$ during the main phase of a magnetic storm. *Journal of Geophysical Research: Space Physics*, *127*(9), e2022JA030669. https://doi.org/10.1029/2022JA030669

Sergeev, V. A., Pellinen, R. J., & Pulkkinen, T. I. (1996). Steady magnetospheric convection: A review of recent results. *Space Science Reviews*, *75*(3–4), 551–604. https://doi.org/10.1007/bf00833344

Sitnov, M. I., Stephens, G. K., Gkioulidou, M., Merkin, V. G., Ukhorskiy, A. Y., Korth, H., et al. (2018). Empirical modeling of extreme events: Storm-time geomagnetic field, electric current, and pressure distributions. In *Extreme events in geospace* (pp. 259–279). Elsevier Inc. https://doi.org/10.1016/B978-0-12-812700-1.00011-X

Sitnov, M. I., Stephens, G. K., Tsyganenko, N. A., Korth, H., Roelof, E. C., Brandt, P. C., et al. (2020). Reconstruction of extreme geomagnetic storms: Breaking the data paucity curse. *Space Weather*, *18*(10), e2020SW002561. https://doi.org/10.1029/2020SW002561

Sitnov, M. I., Stephens, G. K., Tsyganenko, N. A., Miyashita, Y., Merkin, V. G., Motoba, T., et al. (2019). Signatures of nonideal plasma evolution during substorms obtained by mining multimission magnetometer data. *Journal of Geophysical Research: Space Physics*, *124*(11), 8427–8456. https://doi.org/10.1029/2019JA027037

Sitnov, M. I., Tsyganenko, N. A., Ukhorskiy, A. Y., & Brandt, P. C. (2008). Dynamical data-based modeling of the stormtime geomagnetic field with enhanced spatial resolution. *Journal of Geophysical Research*, *113*(A7), A07218. https://doi.org/10.1029/2007JA013003

Sotirelis, T., Tsyganenko, N. A., & Stern, D. P. (1994). Method for confining the magnetic field of the cross-tail current inside the magnetopause. *Journal of Geophysical Research*, *99*(A10), 19393–19402. https://doi.org/10.1029/94ja01656

Stephens, G. K., & Sitnov, M. I. (2021). Concurrent empirical magnetic reconstruction of storm and substorm spatial scales using data mining and virtual spacecraft. *Frontiers in Physiology*, *9*, 653111. https://doi.org/10.3389/fphy.2021.653111

Stephens, G. K., Sitnov, M. I., Korth, H., Tsyganenko, N. A., Ohtani, S., Gkioulidou, M., & Ukhorskiy, A. Y. (2019). Global empirical picture of magnetospheric substorms inferred from multimission magnetometer data. *Journal of Geophysical Research: Space Physics*, *124*(2), 1085–1110. https://doi.org/10.1029/2018JA025843

Stephens, G. K., Sitnov, M. I., Weigel, R. S., Turner, D. I., Tsyganenko, N. A., Rogers, A. J., et al. (2022). Global structure of magnetotail reconnection revealed by mining space magnetometer data. *Earth and Space Science Open Archive*. https://doi.org/10.1002/essoar.10511996.1

Tsyganenko, N., Andreeva, V., Kubyshkina, M., Sitnov, M., & Stephens, G. (2021). Data-based modeling of the Earth's magnetic field. In R. Maggiolo, N. André, H. Hasegawa, & D. Welling (Eds.), *Magnetospheres in the solar system, geophysical monograph 259* (pp. 617–633). Wiley & Sons, Inc. ch.39. https://doi.org/10.1002/9781119815624

Tsyganenko, N. A. (2013). Data-based modelling of the Earth's dynamic magnetosphere: A review. *Annales Geophysicae*, *31*(10), 1745–1772. https://doi.org/10.5194/angeo-31-1745-2013

Tsyganenko, N. A. (2014). Data-based modeling of the geomagnetosphere with an IMF-dependent magnetopause. *Journal of Geophysical Research: Space Physics*, *119*(1), 335–354. https://doi.org/10.1002/2013ja019346

Tsyganenko, N. A., & Andreeva, V. A. (2015). A forecasting model of the magnetosphere driven by an optimal solar wind coupling function. *Journal of Geophysical Research: Space Physics*, *120*(10), 8401–8425. https://doi.org/10.1002/2015ja021641

Tsyganenko, N. A., & Andreeva, V. A. (2017). A hybrid approach to empirical magnetosphere modeling. *Journal of Geophysical Research: Space Physics*, *122*(8), 8198–8213. https://doi.org/10.1002/2017JA024359

Tsyganenko, N. A., Andreeva, V. A., & Sitnov, M. I. (2021). Reconstruction of magnetospheric storm-time dynamics using cylindrical basis functions and multi-mission data mining. *Journal of Geophysical Research: Space Physics*, *126*(2), e2020JA028390. https://doi.org/10.1029/2020JA028390

Tsyganenko, N. A., Andreeva, V. A., Sitnov, M. I., Stephens, G. K., Gjerloev, J. W., Chu, X., & Troshichev, O. A. (2021). Reconstructing substorms via historical data mining: Is it really feasible? *Journal of Geophysical Research: Space Physics*, *126*(10), e2021JA029604. https://doi.org/10.1029/2021JA029604

Tsyganenko, N. A., Singer, H. J., & Kasper, J. C. (2003). Storm-time distortion of the inner magnetosphere: How severe can it get? *Journal of Geophysical Research*, *108*(A5), 1209. https://doi.org/10.1029/2002ja009808

Tsyganenko, N. A., & Sitnov, M. I. (2007). Magnetospheric configurations from a high-resolution data-based magnetic field model. *Journal of Geophysical Research*, *112*(A6), A06225. https://doi.org/10.1029/2007JA012260

Verleysen, M., & François, D. (2005). The curse of dimensionality in data mining and time series prediction. In J. Cabestany, A. Prieto, & F. Sandoval (Eds.), *Computational intelligence and bioinspired systems. IWANN 2005*, *Lecture notes in computer science* (Vol. 3512, pp. 758–770). Springer.