Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **511,** 723–731 (2022) https://doi.org/10.1093/mnras/stac099
Advance Access publication 2022 January 17

Forecasting Geomagnetic activity (Dst Index) using the ensemble kalman filter

B. Nilam★ and S. Tulasi Ram★

Upper Atmosphere Sciences, Indian Institute of Geomagnetism, Navi, 410206, Mumbai, India

Accepted 2022 January 5. Received 2022 January 5; in original form 2021 November 9

ABSTRACT

A novel Ensemble Kalman Filter (EnKF) method is adopted to make reliable forecasts of geomagnetic activity (Dst index) for real-time applications. In this method, an educated estimate (forecast) of Disturbance storm time (Dst) is made based on the ring current dynamics like injection and decay rates. This estimated Dst is further updated with the assimilation of real-time Dst values or ΔH values from a single ground based magnetometer observations. The forecasted Dst values by this EnKF method are validated with the true Dst during the severely disturbed period that consists of 2 super (Dst< -250 nT), 8 intense (Dst< -100 nT), and 13 moderate geomagnetic storms (Dst< -50 nT). It is found that the EnKF method implemented in this work offers the best accurate forecast of Dst with a root mean square error (RMSE) and regression coefficient (R) of 4.3 nT and 0.99, respectively.

Key words: solar-terrestrial relation – solar wind – Planets and satellites: magnetic fields.

1 INTRODUCTION

Transient ejections from the Sun cause severe disturbances in the interplanetary space, and the charged particles can inject into the Earth's magnetosphere through the reconnection process during the southward interplanetary magnetic field (Perreault & Akasofu, 1978; Kan & Lee, 1979; Akasofu, 1980). The drift motion of trapped particles in the magnetosphere under the influence of gradient and curvature of the magnetic field, and the convection and corotation electric fields give rise to accumulation of plasma pressure, inducing a diamagnetic current. The westward current flows in the outer magnetosphere region and an eastward current flow in the inner magnetosphere. Though maximum intensities of the eastward and westward currents are nearly the same, the volume of the westward current region is much larger compared to the eastward current region. (Ebihara & Ejiri, 2000) This results in an induced southward magnetic field (opposite direction to that of Earth's magnetic field) causing a depression in the Geomagnetic field lasting for several hours to days, known as a geomagnetic storm (Ebihara & Ejiri, 2000 and 2003).

It has long been known that the horizontal component, H, of the geomagnetic field is depressed during periods of magnetic storms. The decrease in the H component of Earth's magnetic field at low and mid latitudes during a magnetic storm can be described by a uniform magnetic field antiparallel to the geomagnetic dipole axis, and the magnitude of this disturbance field varies with time due to changes in the ring current. The magnitude of the decrease in H represents the severity of disturbance and its recovery to pre-storm level takes several hours to days due to the gradual loss of ring current particles through the processes like energy-dependent charge exchange and scattering loss (Williams 1983; Daglis et al. 1999).

Disturbance storm time (Dst) index, expressed in nT, is a measure of geomagnetic activity used to indicate the intensity of geomagnetic storms and derived based on the hourly average value of the H component measured at four low latitude geomagnetic observatories Hermanus, Kakioka, Honolulu, and San Juan (Sugiura 1964). Observatories were chosen based on the quality of observation and for the reason that the locations of magnetic observatories are sufficiently far from the auroral and equatorial electrojets. The method for deriving the Dst index consists of removing secular and solar quiet (Sq) diurnal variations from observed H value for each observatory followed by the normalization by its geomagnetic latitude (refer Sugiura 1964 and Sugiua & Kamei 1991 for details). This index is routinely computed and published through World Data Centre for Geomagnetism (WDC), Kyoto University to monitor the geomagnetic activity continuously.

Nowadays, several space and ground based technological systems are susceptible to geomagnetic disturbances, therefore, predicting geomagnetic activity (Dst index) in advance is an important aspect of space weather studies. There are some empirical models for the estimation and/or prediction of Dst index based on the solar wind parameters and previous Dst values. Obrien & McPherron (2000) have assessed three models for the real time forecasting of Dst index. The first one is the very popular empirical model by Burton, McPherron & Russell (1975) which uses simple differential equation for the evolution of Dst index in terms of solar wind parameters and considers a constant decay rate of the ring current. Secondly, the Fenrich & Luhmann, (1998) have modified the Burton et al. (1975) model by incorporating different ring current decay rates for the main and recovery phases. The third model proposed by Obrien & McPherron, (2000) which uses the dawn–dusk solar wind electric field dependent ring current decay rate. These three models are based

★ E-mail: nilambhosale0@gmail.com (BN); tulasiram.s@iigm.res.in (STR)

on single step, first-order differential equation that provides evolution of Dst index with varying solar wind parameters. The predicted Dst values from these models generally found to be consistent when compared with observed Dst with an approximate root mean square error (RMSE) of about 20 nT during quiet and ∼40 nT during disturbed periods (Dst< -50 nT) Obrien & McPherron (2000).

Further, Newell et al., (2007) investigated correlation between several coupling functions (based on solar wind parameters) and geomagnetic indices. They have shown that the coupling functions E_{TL}Temerin & Li (2006), E_{SR} Scurry & Russel (1991), and E_{WV} Vasylunas et al. (1982) work better for Dst prediction among all coupling functions. Newell et al. (2007) proposed a new coupling function ($\frac{d\phi_{MP}}{dt}$) which gives a good correlation (R = -0.844) with Dst index when it has integrated over 72 h. One of the limitation of Newell function that it may not be suitable for real-time Dst forecasting when there are discontinuities in the solar wind observations.

Besides the above mentioned empirical models and coupling functions, there are several studies through theoretical and numerical simulation methods to predict the Dst under varying solar wind conditions (Murayama 1982; Billings et al. 1989; Raeder et al. 2001; Toffoletto et al. 2003; Tóth et al. 2005; Boynton et al. 2011). The Rice Convection Model (RCM) is a comprehensively used physical model consisting of plasma electrodynamics in the inner magnetosphere and coupling with the ionosphere (Harel et al. 1981; Wolf, Spiro & Rich 1991; Toffoletto et al. 2003 and 2015). Toffoletto et al. (2015) have further improved the RCM as a space weather tool to real time event simulation of different geomagnetic storms. Rastätter et al. (2013) have compared the predicted Dst results of various models for four different geomagnetic storms and found that the accuracy of different models varies from storm to storm and none of the models consistently works better for all the four events.

The real time data assimilation methods have been widely used in recent times for estimating the unknown state which evolves in time. The main idea of data assimilation models is to use the information obtained by observing a real time system while forecasting a state to improve the accuracy of the prediction. Kalman filter is one the most popular data assimilation method that predicts the state of a system by an educated guess followed by a correction on comparing with real-time measurement. The Ensemble Kalman Filter (EnKF) has the advantage of correcting the forecasted value in real time by propagating the uncertainties in time. In this paper, we present a novel method based on the EnKF to accurately forecast the Dst index in real time with the assimilation of some direct/indirect observations.

2 DATA

The interplanetary solar wind parameters like Solar wind velocity (V), dynamic pressure (P), and z-component of magnetic field (IMF Bz) at 1 h interval taken from NASA's OMNI Space Physics Data Facility (https://omniweb.gsfc.nasa.gov/ow_min.html) were considered as input parameters to make an educated guess of Dst by Kalman filter. We used Dst values provided by the WDC, Kyoto university to update, validate and investigate forecasted Dst values. Further, the ΔH values derived from a ground based magnetometer at low-latitude station, Alibag (Geog. Lat 18.6°N, Geog. Long 72.9°E, GeoMag. Lat 10.5°N) have been used as an indirect measurement for updating the estimated Dst in real time. The baseline for the observed horizontal (H) component at Alibag is computed by averaging midnight H value of five international quiet days in every month. The baseline is subtracted from the H-component variations in order to remove secular variations. Further, the Sq variations in H component are removed by subtracting mean diurnal variation of five international quiet days in each month. The H-component variations after the removal of baseline and Sq variations are referred as ΔH at Alibag.

3 METHODOLOGY

3.1 Kalman filter

The Kalman filter has long been used as the optimal solution to many data prediction problems (Pearson & Stear 1974; Galanis et al. 2006; Auger et al. 2013; Golestan, Guerrero & Vasquez 2018). The Kalman filter is capable of filtering out noisy data and at the same time, it also makes possible prediction of the next state of a system. It makes an accurate projection of the next state of the system using the knowledge of state dynamics and further updating it by assimilating some real-time observations. For example, if we know the Dst value at time k-1 (current state) and the knowledge of ring current injection and decay rates (state dynamics), then we can make an educated estimate of the Dst value at time k (next state). This estimate can be further updated by assimilating some direct or indirect observations in real time in order to improve the overall accuracy of the prediction.

3.1.1 Educated estimate of the state

General equations for the state dynamics of the Kalman filter are,

$$x'_k = F x_{k-1} + G\, u_k + w_k \tag{1}$$

$$P'_k = cov\,(F.x)_{k-1} + cov\,(w)_k \tag{2}$$

where,

x_{k-1}– State of the system at current state (k-1)
x'_k- Educated estimate of the next state of a system
F- State dynamic function
u_k—Input parameters
G—Input coefficients
w_k—Process noise
P'_k—Error covariance of the estimated state

The first term on the right-hand side of equation (1) represents the evolution of system from the states k-1 to k in the absence of external sources. The state dynamic function F controls this evolution. Further, there could also be some external inputs which can influence the system dynamics. The u_k and G in the second term represent the external input parameters and their coefficients, respectively. The third term, w_k represents the noise or uncertainty in the educated estimate. The main purpose of this term is to treat the unknown factors influences in the system. The equation (2) indeed represents the error covariance of the estimated state. The terms $cov(F.x)_{k-1}$ and $cov(w)_k$ in equation (2) indicate the covariances of state and process noise, respectively.

3.1.2 Updating the estimate by real-time measurements (Y_k)

The estimated state x'_k from equation (1) will be further updated with the help of some direct/indirect real-time observations at state k. Let Y_k represent observations at k, the measurement function can be given by,

$$Z_k = S_k\, Y_k + V_k \tag{3}$$

Where, S_k is the scaling factor that transforms the direct/indirect real time measurements into state equivalents for the correction of the estimated state. The measurement function 'Z_k' is not what we are trying to get out of the filter, but rather what we are able to measure in real time. The 'V_k' term can be used to include the errors in the measurements and accounts how good the equation (3) represent the state equivalents.

Now, the educated estimate x'_k and the error covariance P'_k can be updated appropriately based on the knowledge of the system and real-time measurements as,

$$x_k = x'_k + K_k \left(z_k - x'_k\right) \tag{4}$$

$$P_k = (1 - K_k S_k)\ P'_k \tag{5}$$

where, K_k is a Kalman gain, is a factor which makes a decision about the weights that should be given to the measurement and estimated state values. The Kalman gain is calculated from the errors in the estimated state and the measurement noise as

$$K_k = P'_k\ S_k / \left[cov\left(S_k . Y_k\right) + cov\ (V_k)\right] \tag{6}$$

The equations (4) and (5) give the updated state and the error covariance of the system with desirable gain. The updated state and error covariance are the outputs of the current iteration would become the previous state for the next iteration. This process of making the educated guess using equations (1) and (2) and updating/correcting those using equations (4) and (5) would continue iteratively.

3.2 Ensemble kalman filter

The EnKF originated as a version of the Kalman filter for large problems. It is a Monte Carlo based implementation of the Kalman filter for non-linear state estimation problem (Gillijns et al., 2006; Katzfuss, Stroud & Wikle 2016). In this method, the state and covariance matrices are replaced by the ensemble and sample covariances, respectively. The main idea of the EnKF is to propagate an ensemble of state instead of a single state. That means, instead of making a single forecast, a set of forecasts is produced. Generally, the forecasted value should fall within the predicted ensemble spread and the amount of spread should be related to the uncertainty or error of the

forecast. This approach can be used to make probabilistic forecasts of any dynamical system.

In the first step, a previous ensemble gets propagated forward in time and new states and/or ensemble is predicted by the following equation,

$$X_{n,\ i} = F(X_{n-1,\ i},\ W_{n-1,\ i}) \tag{7}$$

where, F represents, the non-linear state dynamic function and W_n is the process noise.

In the second step, the forecasted ensemble is updated by the measurement function as below,

$$Z_{n,\ i} = S_k\ (Y_{n,\ i},\ V_{n,\ i}) \tag{8}$$

where, $Z_{n,i}$ is the generated ensemble of the real time measurements $Y_{n,i}$ and Vn is the measurement noise and S_k is scaling factor. The expectation of the error in every ensemble mean of the states is,

$$E_n = (X_{n,i} - \overline{X_n}, \ldots\ldots\ldots\ X_{n,N} - \overline{X_n}$$

where, $\overline{X_n}$ is the ensemble mean of the states given as,

$$\overline{X_n} = \frac{1}{N}\sum_{i=1}^{N} X_{n,i}$$

The error covariance and Kalman Gain are calculated by,

$$P = \frac{1}{N-1} E_n E_n^T \tag{9}$$

$$K = P S^T \left(S P S^T + R\right)^{-1} \tag{10}$$

R is the variance of the measurement noise. Eventually, the updated ensemble calculated as,

$$X_{n,i}^{up} = X_{n,i} + K\ (Y_{n,i} + \overline{SX_n} \tag{11}$$

where, $X_{n,\ i}^{up}$ is the updated ensemble. These are the steps to predict and update the state ensemble of the EnKF. The first step adds the process noise/error and second step updates the ensemble that will be fed back as input for the next iteration.

3.3 Ensemble Kalman Filter to predict Dst index

In this section, we present the implementation of data assimilative EnKF method to predict the Dst index for real-time applications.

3.3.1 Educated estimate of Dst

In this EnKF method, the next state of Dst (Dst'_k) can be predicted from the knowledge of solar wind velocity (V), dynamic pressure (P), z-component interplanetary magnetic field (Bz) as,

$$Dst'_k = -a\ Dst_{k-1} + \left\{b\ E_y + c\sqrt{P}\right\}_k + d \tag{12}$$

The first term on the right hand of equation (12), *-aDst* indicate the ring current decay and the constant 'a' represents ring current decay constant. The second term on the right hand side of equation (12), *bEy* indicate the rate of ring current injection which is a function of interplanetary dawn-to-dusk electric field (Ey) given as -V × B_z, while the third term c$\sqrt{P}$ indicates the contributions due to the magnetopause currents. The contribution of magnetopause current to the H component is proportional to the$\sqrt{P}$ as confirmed by Siscoe et al. (1968) and Su & Konradi (1975). The equation (12) is similar to that given by Burton et al. (1975) except that the coefficients a, b, c, and d are dynamically derived here by a nonlinear regression on the previous 72 h values of the Dst and solar wind parameters. The equation (12) gives an educated estimate of the system (Dst index) based on the state dynamics like ring current decay and external contributions due to ring current injection and magnetopause currents.

Equation (12) can be re-written by adding some process noise to the system as,

$$Dst'_k = -a\ Dst_{k-1} + \left\{b\ E_y + c\sqrt{P}\right\}_k + d + W_{k,i}\ for\ i = 1,\ 2,\ 3\ldots\ldots\ldots n \tag{13}$$

where n is the ensemble size of the Dst'_k(estimated ensemble of the Dst) and $W_{k,\ i}$ is introduced to represent the process noise in the system. Equation (13) is used as the first step (educated estimate) of this EnKF method to make a forecast of Dst ensemble. The mean value of this ensemble gives the forecast of Dst value during the next state.

The EnKF method adopted in this work to predict the Dst (explained above) is summarized here with the help of a schematic as shown in Fig. 1. In the first step, an educated estimate of Dst of the next state (Dst'_k) is made using state dynamic function (F) and input (solar wind) parameters as given in equation (13). Here, the coefficients a, b, c, d in equation (13) are derived using the

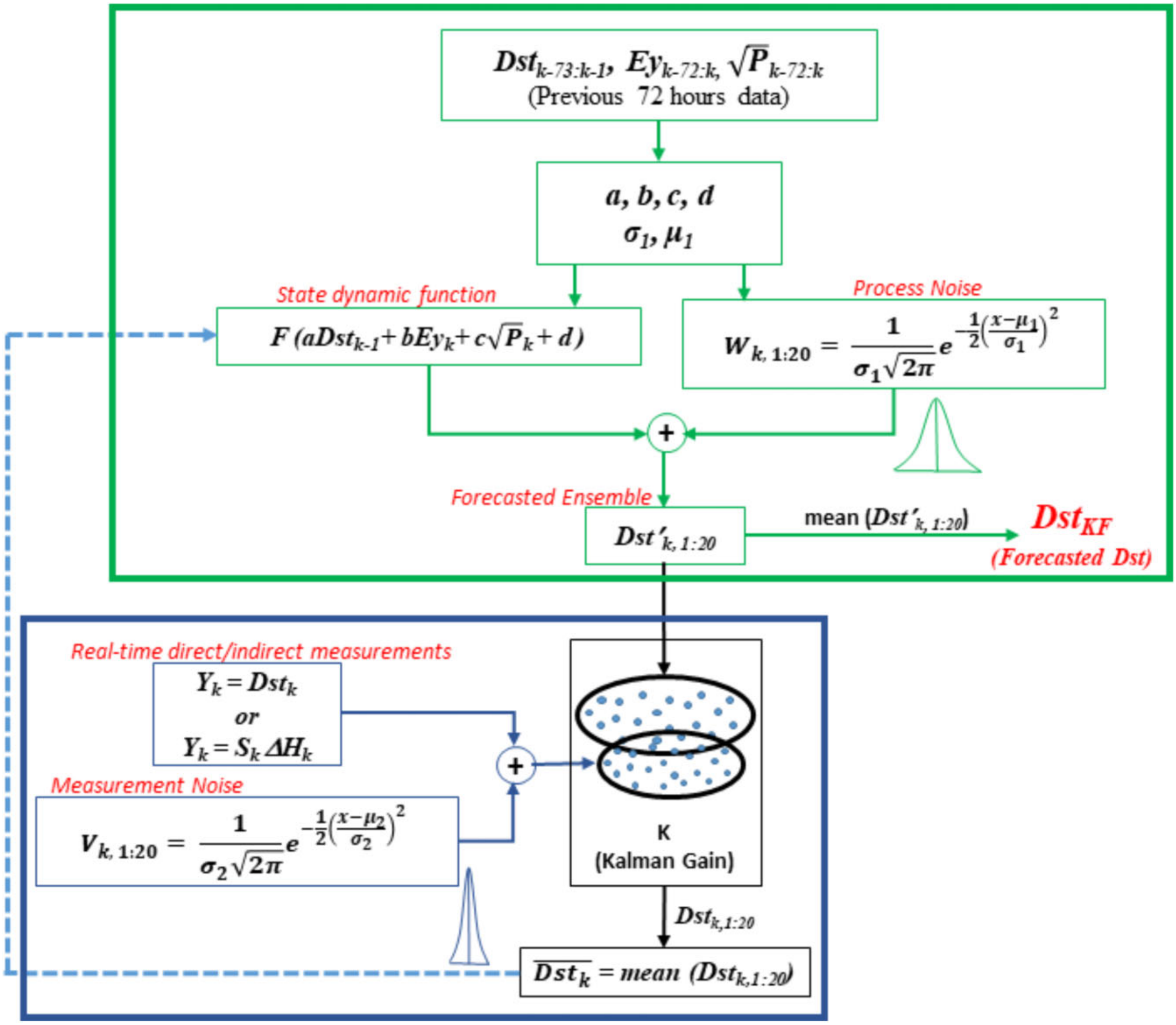

Figure 1. The architecture of EnKF method to forecast the Dst index

nonlinear regression fit of previous 72 h data of the Dst, Ey, and $\sqrt{P}$ (from k-72 to k-1). There could be some error in the estimated Dst'_k due to unknown sources in this process of estimation. To represent this process noise, a twenty point Gaussian noise ($W_{k,\,1:20}$) is generated with mean (μ_1) and standard deviation (σ_1) of errors derived from the previous 72 h values. This process noise ($W_{k,\,1:20}$) is added to Dst'_k in order to forecast a twenty point estimated ensemble $Dst'_{k,1:20}$. Here, the $Dst'_{k,\,1:20}$ would be the forecasted ensemble of Dst index based on the previous state (Dst_{k-1}), solar wind parameters and the process noise. The solar wind parameters usually measured at first Lagrangian point (L1-point) would give us a lead time of approximately ∼40 to 60 min before its effects can be seen on the ground. Therefore, for real time applications, this method can enable us to forecast the Dst at least ∼40–60 min in advance at the end of first step. Thus we call the $Dst'_{KF,1:20}$ as the forecasted ensemble of the Dst index and the mean of this ensemble would be the forecasted value of the next state of Dst (denoted as Dst_{KF}).

3.3.2 Updating the estimated Dst by direct measurements

In the second step, the above educated estimate of $Dst'_{k,1:20}$ can be further updated with some real-time direct/indirect measurements [equation (8)] to improve the overall accuracy. For this purpose, we have initially considered the observed true Dst values (direct measurement) with scaling factor set to unity ($S_k = 1$). It may be noted that, although one would not expect any significant error in the direct measurements (true Dst values), an ensemble of errors is required to be added in order to implement the EnKF algorithm. This ensemble of errors is usually to represent the errors due to the unknown factors in the measurement function. Hence, a twenty point Gaussian noise with zero mean ($\mu_2 = 0$) and very small standard deviation ($\sigma_2 = 1$ nT) is added to the true Dst and an ensemble of measurement values are generated ($Y_{k,\,1:20} = Dst_k + V_{k,1:20}$). Now, the forecasted ensemble $Dst'_{k,1:20}$ is updated with the ensemble of measurement values ($Y_{k,1:20}$) through the Kalman gain (K) using equations (10) and (11) to derive an updated ensemble of $Dst_{k,1:20}$ at state k. The mean of this updated ensemble ($\overline{Dst_k}$) is taken as the

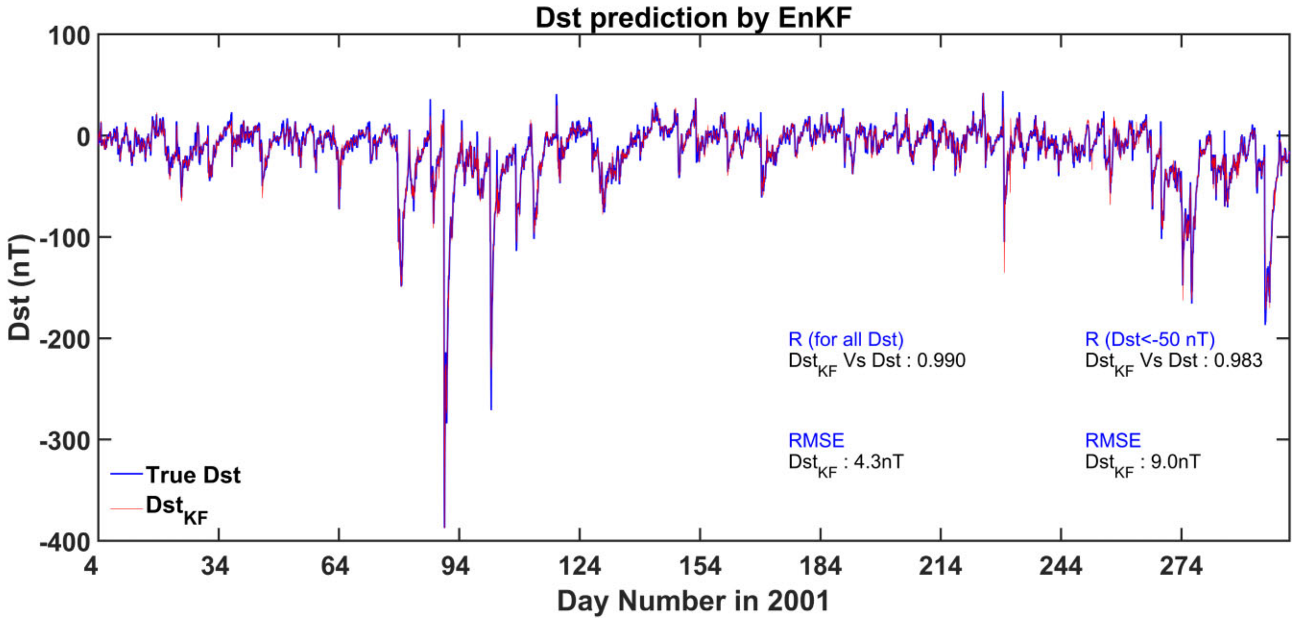

Figure 2. The comparison of EnKF forecasted Dst$_{KF}$ (red) with true Dst (blue) during day numbers 1 to 300 in the year 2001.

Table 1. Correlation coefficients and RMSE in the forecasted Dst by different models with respect to true Dst

Methods	R (all Dst values)	R(Dst ≤−50nT)	RMSE (all Dst values)	RMSE (Dst ≤ -50 nT)
Dst$_B$	0.986	0.981	5.4	10.6
Dst$_{FL}$	0.942	0.943	11.6	18.2
Dst$_{OM}$	0.922	0.934	12.2	18.6
Dst$_{KF\ (Dst)}$	0.990	0.983	4.3	9.0
Dst$_{KF\ (\Delta H)}$	0.943	0.934	10.1	19.0

updated forecast (or nowcast) which is fed again to the state dynamic function (F) for the next iteration and process continues iteratively (Fig. 1). The assimilation of measurement ensemble $Y_{k,\,1:20}$ through EnKF significantly improves the forecasted Dst during the next iteration.

4 RESULT AND DISCUSSION

In this section, we present the results of forecasted Dst from the EnKF method (Dst$_{KF}$) described above in comparison with the true Dst values provided by the WDC, Kyoto university. For this purpose, we have considered the Dst index data for the period from day number 1 to 300 in the year 2001. We have deliberately chosen this period when the geomagnetic activity is highly disturbed with multiple geomagnetic storms of different intensity levels in order to rigorously test this method under very complex conditions. Fig. 2 shows the forecasted Dst (red curve) by the EnKF in comparison with the true Dst (blue curve) from the WDC. One can clearly see from this figure that the forecasted Dst follow closely with the true Dst. The linear correlation coefficient (R) and RMSE of forecasted Dst with reference to the true Dst values are shown in the figure. When the total period (day number 1 to 300) is considered, the correlation coefficient between the forecasted Dst and true Dst is 0.998 and the RMSE in the forecasted Dst is 4.3 nT. This indicates that the EnKF method adopted here provides very accurate forecast of geomagnetic activity (Dst) with a very small error. In order to examine the correlation and RMSE during the highly disturbed conditions, we have further evaluated the correlations coefficients and RMSE during the periods when the Dst index is below -50 nT. During the period of highly disturbed conditions (Dst< -50 nT), the correlation coefficient (R) is 0.983 and RMSE is 9.0 nT. One may be noted that this period under consideration contain 2 super geomagnetic storms (Dst< -250 nT), 8 intense geomagnetic storms (Dst< -100 nT) and 13 moderate storms (Dst< -50 nT). Therefore, one can conclude that the EnKF method adopted here provides accurate forecast of Dst even during the highly disturbed and complex geomagnetic activity conditions.

There were some important empirical models available in the literature, for example, Burton et al. (1975), Fenrich & Luhmann (1998), Obrien & McPherron (2000) have given the simple differential equations which can predict next state of Dst from the previous Dst value and solar wind parameters (velocity, IMF Bz and pressure). The empirical relationships given in Burton et al. (1975), Fenrich & Luhmann, (1998), Obrien & McPherron (2000) are conceptually similar to that of equation (12) except that the coefficients vary in each [see Table 1 of Obrien & McPherron (2000)]. Fig. 3 shows the comparison of the forecasted Dst from EnKF (Dst$_{KF}$) with the other empirical models Dst$_B$, Dst$_{FL}$ and Dst$_{OM}$ given by Burton et al. (1975), Fenrich & Luhmann, (1998), Obrien & McPherron (2000), respectively, for the four major geomagnetic storms occurred during 2001. Different color curves indicate the forecasted Dst values by different methods and RMSEs of the different methods are indicated in the figure. It can be noticed from this figure that the EnKF method (Dst$_{KF}$) provides the best forecast of Dst with minimum RMSE compared to other methods during all the four geomagnetic storms of different intensity levels. Hence, one can conclude that the EnKF method adopted in this study provides the best forecast of Dst

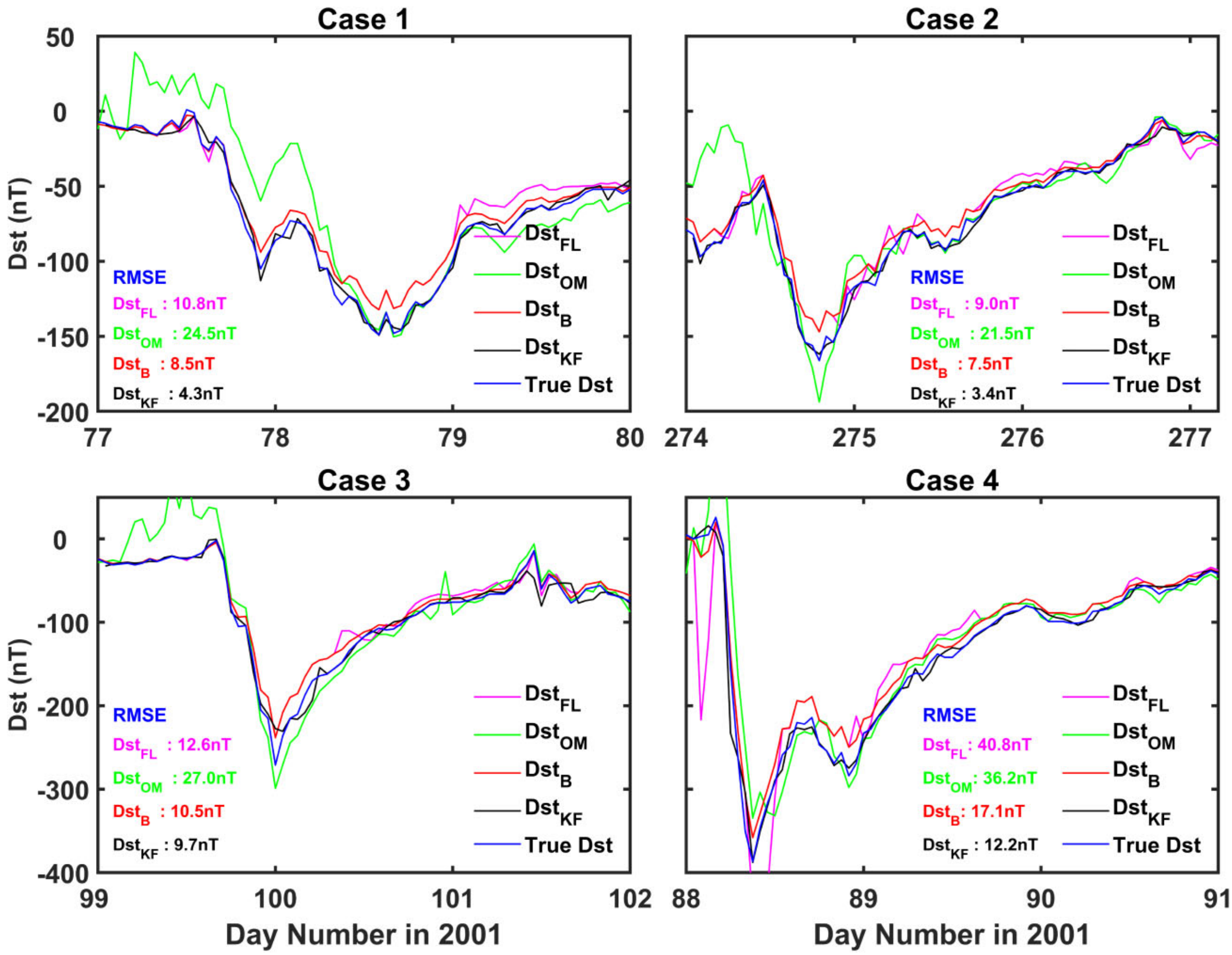

Figure 3. Comparison of forecasted Dst using different methods, Dst_B (red curve), Dst_{FL} (magenta curve), Dst_{OM} (green curve), and Dst_{KF} (black curve) with the true Dst (blue curve) during the four major geomagnetic storms occurred during 2001. (see text for more details)

even during the periods of highly disturbed geomagnetic activity by propagating an ensemble of uncertainties/errors in the forecast and by updating with the assimilation of real-time observations through the Kalman gain. It may also be noted that the RMSE of Dst_B is close to that of Dst_{KF} during these storms. Since the equation (12) used in this EnKF method to predict the next state (forecast) of Dst is similar to that of Burton's equation (except the coefficients), it is expected that the RMSE of Burton et al. (1975) method will be close to that of EnKF method and vice versa. However, from the close examination of Fig. 3, one can observe that the error in Burton's method is usually larger around the peak negative Dst and during the recovery phase of the storms. On the other hand, the error in the EnKF method is small throughout the storm period because of its real-time adaptability.

Now, we further evaluate the EnKF method when the direct measurements (true Dst) were not available for updating forecasts. In such scenario, we evaluate the possibility of assimilating the ΔH values (indirect measurements) from a single ground based low-latitude magnetometer observations. For this purpose, we have considered the ΔH variations measured from an Indian low-latitude station, Alibag (Geog. Lat. 18.6°N, Geog. Long. 72.9°E, GeoMag. Lat. 10.5°N). The ΔH values are derived after removing the baseline and solar-quiet current variations as explained earlier. It may be noted that ΔH values from the Alibag may differ from the longitudinally averaged Dst values due to the local time asymmetry in the ring current which can expected introduce significant error in the forecasted Dst by EnKF. Hence, in order to improve the accuracy of Dst forecast by Kalman filter, the indirect measurements (ΔH at Alibag) needs to be properly scaled to match with Dst through an appropriate scaling factor (S_k). For this purpose, we have examined the linear relationship between the true Dst and ΔH at Alibag at every one hour UT intervals as shown in the Fig. 4. The red straight lines in Fig. 4 indicate the linear fit between the true Dst and ΔH and the corresponding fit equation is provided in each panel. The value of slope indicates the linear relationship between the true Dst and ΔH at Alibag which founds to vary at different UT.

Fig. 5 shows the variation of slopes (black dots) derived from linear fits (Fig. 4) as a function of UT. The value of slope at any given time indicates relative variation in ΔH at Alibag with respect to variation in true Dst. For example, the value of slope less than unity indicates that the variation in ΔH is larger compared to the variation in Dst. In other words, the ΔH would exhibit larger decrease compared to the decrease in Dst in response to geomagnetic activity at times when the slope is less than unity. Similarly, the value of slope greater than unity indicates that the decrease in ΔH is smaller compared to the decrease in Dst. From Fig. 5, one can see that slopes exhibits a significant UT variation with a minimum around 14–16 UT (19.1–21.1 LT) and a maximum around 23–02 UT (04.1–07.1 LT). This indicates that the ΔH at Alibag exhibits a larger

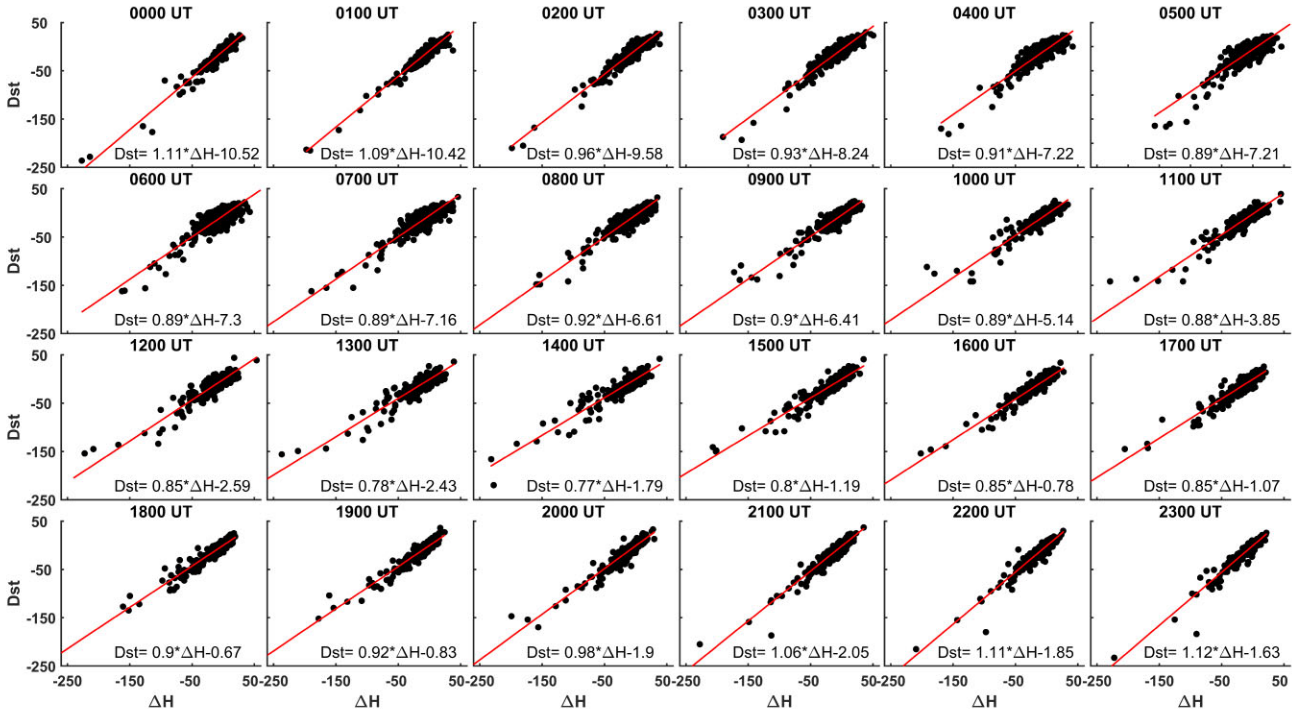

Figure 4. The linear regressions between the true Dst and ΔH at Alibag at every one hour UT intervals.

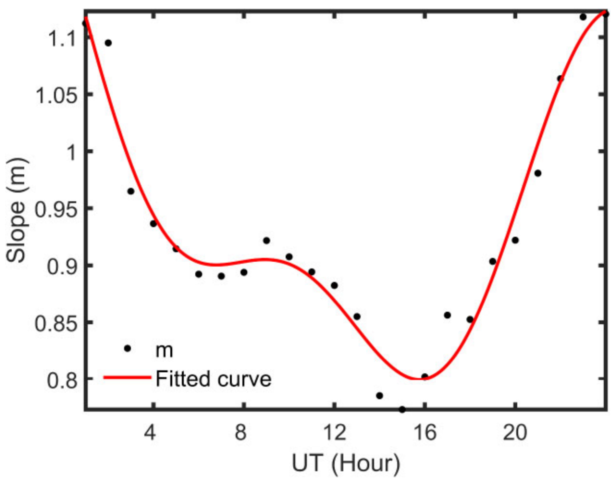

Figure 5. UT variation of derived slopes from the linear regression fit between true Dst and ΔH at Alibag (black dots). Red curve indicates the 3rd order sine series fit of derived slopes.

decrease (slope $<$ 1) compared to Dst during the local dusk hours and exhibits a smaller decrease (slope $>$ 1) compared to Dst during the local dawn hours. This dawn to dusk variation in the slopes can be attributed to the dawn-to-dusk asymmetry in the ring current and/or to the partial ring current (Fukushima and Kamide, 1973; Cummings, 1966). Recently, Ohtani (2021) has suggested the magnetosphere-ionosphere coupling processes that lead to the enhanced dawnside westward auroral electrojet also contributes to this dawn-to-dusk asymmetry of ring current. It is known that the partial ring current causes strong westward current around dusk sector leading to the larger decrease in ΔH compared to Dst. The minimum in slopes ($<$ 1) around 14 to 16 UT is consistent with the larger decrease in ΔH at Alibag due to partial ring current at local dusk hours.

From the results presented in Figs 4 and 5, it is understood that the linear relationship between the true Dst and ΔH at Alibag vary significantly with UT due to the dawn-to-dusk asymmetry in ring current and/or partial ring current. In order to incorporate the effects of partial ring current, we have modelled the UT variation of slopes using a third order sine series fit as given below

$$m = \sum_{i=1}^{n} a_i \quad \sin(b_i UT_{(0\ to\ 23)} + c_i) \tag{14}$$

where n = 3, $a_1 = 4.9884$, $a_2 = 4.1632$, $a_3 = 0.0729$, $b_1 = 0.0932$, $b_2 = 0.1097$, $b_3 = 0.46$, $c_1 = 0.6198$, $c_2 = 3.5653$ $c_3 = 3.0697$.

The red curve in Fig. 5 indicates the modelled UT variation of slopes between the true Dst and ΔH at Alibag. Then the measurement function Z_k in the EnKF method becomes,

$$Z_k = S_k\ .\ Y_k = \frac{m}{\cos\phi}.\Delta\mathrm{H} \tag{15}$$

where ϕ is Geomagnetic latitude of the Alibag and 'm' is a coefficient which varies with UT (Fig. 5 or equation 14) to account for the local time asymmetry in the ring current.

Fig. 6 shows the comparison of EnKF forecasted Dst with the true Dst during the day numbers 1 to 300 of the year 2001. It can be noticed from this figure that the forecasted Dst with the assimilation of ΔH at Alibag closely matches with the true Dst. It may be noted that the true Dst values were not used in the EnKF iterations here, instead the ΔH from a single low latitude station Alibag (indirect measurements) have been used for updating the forecasts. Nevertheless, the forecasted Dst values exhibit excellent similitude when compared with true Dst values. The linear correlation coefficients (R) and RMSE of forecasted Dst with reference to the true Dst values are given in the figure. When the total period is considered, the correlation coefficient (R) between the forecasted Dst and true Dst is 0.943 and the RMSE in the forecasted Dst is 10.1 nT. This indicates that the EnKF method adopted in this study

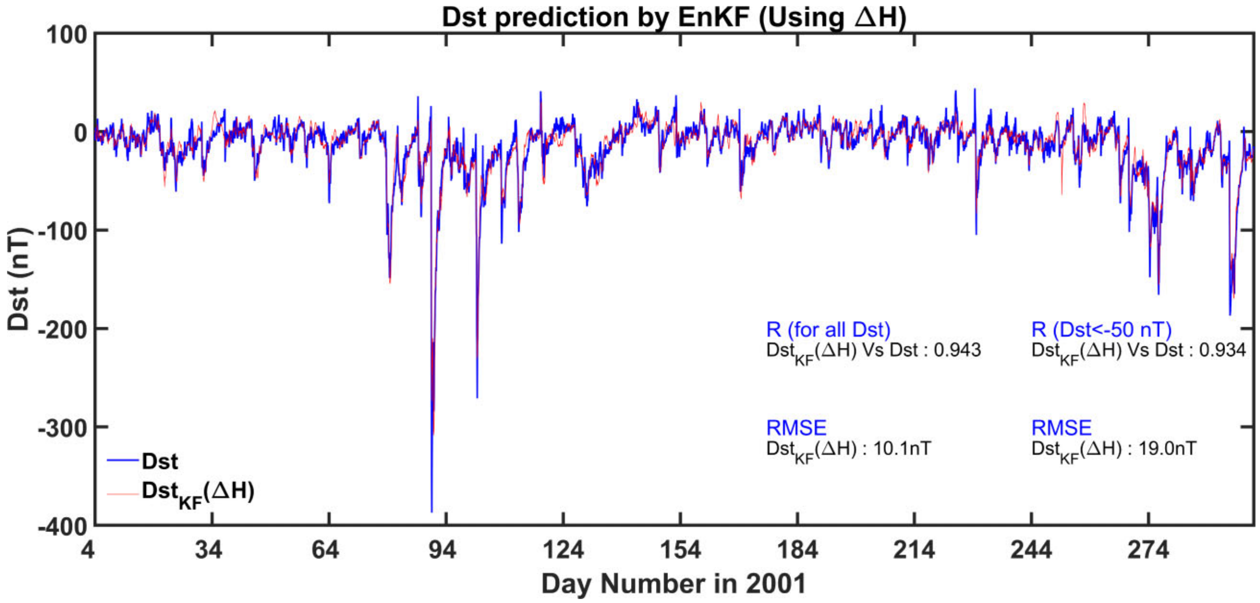

Figure 6. The comparison of EnKF forecasted Dst'_{KF} (blue) with true Dst (blue) during day numbers 1 to 300 in the year 2001 when the ΔH observations from Alibag are assimilated.

can accurately forecast geomagnetic activity (Dst) with a minimal error even with the assimilation of indirect measurements like ΔH from a single ground based magnetometer observations. During the period of highly disturbed conditions (Dst< -50 nT), the correlation coefficient (R) is slightly decreased to 0.934 and RMSE is slightly increased to 19 nT. Therefore, the EnKF method adopted in this study can provide accurate forecast of Dst even during the highly disturbed and complex geomagnetic activity conditions as shown in Fig. 6 with the assimilation of ΔH values from a single ground based magnetometer observations.

In summary, Table 1 provides the values of correlation coefficients and RMSE of forecasted Dst values by all methods (Dst_B, Dst_{FL}, Dst_{OM}, Dst_{KF} (Dst) & Dst_{KF} (ΔH)) with respect to true Dst when the entire period day numbers 1 to 300 were considered and for geomagnetically disturbed periods (Dst $\leq$ -50 nT). It can clearly noticed from this table that the EnKF method assimilated with the direct (Dst) measurement provides the best accurate forecast of Dst with minimum RMSE compared to other methods. Further, the EnKF method assimilated with indirect (ΔH) measurements from a single low-latitude observations also provides the reasonably accurate forecasts even during the highly disturbed geomagnetic activity conditions.

5 CONCLUSIONS

In this paper, we presented a novel EnKF method to forecast the geomagnetic activity, Dst-index. The EnKF makes an educated forecast of Dst using the information about the previous Dst value and state dynamics like ring current decay and injection rates based on the solar wind parameters. This forecasted Dst is further updated with the direct measurements like Dst or indirect measurements like ΔHfrom a ground based magnetometer at a single low-latitude station, Alibag. This method is thoroughly validated with the true Dst values provided by the WDC, Kyoto University under severely disturbed geomagnetic activity period in the year 2001 (day number 1 to 300) which consists of 2 superstorms, 8 intense storms, and nearly 13 moderate storms.

This EnKF method assimilated with true Dst found to provide very accurate forecasts on the next state of Dst compared to the other empirical models proposed by Burton et al. (1975), Fenrich & Luhmann (1998), Obrien & McPherron (2000) during the four geomagnetic storm of different intensities. Further, the EnKF method also provides accurate forecasts of Dst even with assimilation of ΔH measurements from a single low-latitude magnetometer observations. This EnKF is a very promising method to make reliable forecast of Dst in real time applications. In the future work, the forecasts can be further improved by considering the other factors like different loss mechanisms for ring current particles during the recovery phase of the geomagnetic storm and the effects of seasonal and diurnal variation in geomagnetic activity while making the educated estimates of Dst during the first step of EnKF.

ACKNOWLEDGEMENTS

This work is supported by Department of Science and Technology, Government of India. The open data policy of USA and WDC for the solar wind data and geomagnetic index (Dst) data is duly acknowledged.

DATA AVAILABILITY

The solar wind data underlying this article are publicly available at the open data policy of USA (https://omniweb.gsfc.nasa.gov/ow_min.html) and geomagnetic index (Dst) is available at WDC (http://wdc.kugi.kyoto-u.ac.jp/). The ground magnetometer observations at Alibag are available at http://wdciig.res.in/WebUI/Home.aspx.

REFERENCES

Akasofu S. -I., 1980, Planet. Space Sci., 28, 5
Auger F., Hilairet M., Guerrero J. M., Monmasson E., Orlowska-Kowalska T., Katsura S., 2013, IEEE Trans. Ind. Electron., 60, 12
Billings S. A., Chen S., Korenberg M. J., 1989, Int. J. Control, 499, 6
Boynton R. J., Balikhin M. A., Billings S. A., Sharma A. S., Amariute O. A., 2011, Ann. Geophys., 29, 6

Burton R. K., McPherron R. L., Russell C. T., 1975, J. Geophys. Res., 80, 31
Cummings W. D., 1966, J. Geophys. Res., 71, 19
Daglis I. A., Thorne R. M., Baumjohann W., Orsini S., 1999, Rev. Geophys., 37, 4
Ebihara Y., Ejiri M., 2000, J. Geophys. Res., 105, A7
Ebihara Y., Ejiri M., 2003, Space Sci. Rev., 105
Fenrich F. R., Luhmann J. G., 1998, Geophys. Res. Lett., 25, 15
Fukushima N., Kamide Y., 1973, Rev. Geophys. Space Sci., 11, 4
Galanis G., Louka P., Katsafados P., Pytharoulis I., Kallos G., 2006, Ann. Geophys., 24, 10
Gillijns S., Mendoza O. B., Chandrasekar J., De Moor B. L. R., Bernstein D. S., Ridley A., 2006, What is the ensemble Kalman filter and how well does it work?, American Control Conference, p. 6
Golestan S., Guerrero J. M., Vasquez J. C., 2018, IEEE Trans. Ind. Electron., 65, 12
Harel M., Wolf R. A., Reiff P. H., Spiro R. W., Burke W. J., Rich F. J., Smiddy M., 1981, J. Geophys. Res., 86, 2217
Kan J. R., Lee L. C., 1979, Geophys. Res. Lett., 6, 577
Katzfuss M., Stroud J. R., Wikle C. K., 2016, Am. Stat., 70, 4
Murayama T., 1982, Rev. Geophys., 20, 3
Newell P. T., Sotirelis T., Liou K., Meng C. - I., Rich F. J., 2007, J. Geophys. Res., 112, A01206
O'Brien T. P., McPherron R. L., 2000, J. Atmos. Sol. Terr. Phys., 62, 14
Ohtani S., 2021, J. Geophys. Res.: Space Phys., 126, 9
Pearson J. B., Stear E. B., 1974, IEEE Trans. Aerosp. Electron. Syst., AES-10, 3
Perreault P., Akasofu S. I., 1978, Geophys. J. Int., 54, 547
Raeder J. et al. 2001, J. Geophys. Res., 106, 381
Rastätter L. et al. 2013, Geospace environment modeling 2008–2009 challenge : Dst index, Space Weather, 11, 187
Scurry L., Russell C. T. 1991, J. Geophys. Res., 96, 9541
Siscoe G. L., Formisano V., Lazarus V., 1968, J. Geophys. Res., 73, 4869
Su S. - Y., Konradi A., 1975, J. Geophys. Res., 80, 195
Sugiura M., 1964, Ann. Int. Geophys. Year, 35, 9
Sugiura M., Kamei T., 1991, Equatorial Dst index 1957-1986, IAGA Bulletin, no. 40. IUGG, Paris
Temerin M., Li X., 2006, J. Geophys. Res., 111, A04221
Toffoletto F., 2015, Development of The RICE Convection Model as a Space Weather Tool, AFRL-RV-PS-TR-2015-0129, Accession Number: ADA624938
Toffoletto F., Sazykin S., Spiro R., Wolf R., 2003, Space Sci. Rev., 107, 175
Tóth G. et al. 2005, J. Geophys. Res: Space Phys., 110, 1
Vasyliunas V. M., Kan J. R., Siscoe G. L., Akasofu S. I., 1982, Planet. Space Sci., 30, 4
Williams D. J., 1983, Space Sci. Rev., 34, 223
Wolf R. A., Spiro R. W., Rich F. J., 1991, J. Atmos. Terr. Phys., 53, 817

This paper has been typeset from a TEX/LATEX file prepared by the author.