30 September 2021

AGU PUBLICATIONS
Space Weather Journal

Supporting Information for the submission

The polar cap (PC) index: PCS version based on Dome-C data

P. Stauning, pst@dmi.dk

Danish Meteorological Institute, Copenhagen, Denmark

Introduction. The present SI contribution presents basis for the potential use of magnetic data from Dome Concordia (Dome-C) observatory in Antarctica as a source for PCS index values in order to enhance the reliability and availability of PC indices to be used for solar-terrestrial sciences as well as for space weather monitoring applications. The description of the Dome-C-based PCS indices and the definition of reference levels and scaling parameters are very similar to the corresponding definitions and descriptions of Qaanaaq (THL)-based PCN indices or Vostok-based PCS indices available in Stauning (2016). A more detailed description of the index derivation methods may be found in the present Supporting Information (SI) file where the disagreements with features of the methodologies endorsed by the International Association for Geomagnetism and Aeronomy (IAGA) are also discussed. Such discussions may also be found, among others, in Stauning (2013, 2015, 2018a, 2020 and 2021a,b).

Contents of this file

1. Basic principles for calculation of Polar Cap indices.

The transpolar (noon to midnight) convection of plasma and magnetic fields driven by the interaction of the solar wind with the magnetosphere is associated with (equivalent Hall) electric currents in the upper atmosphere in opposite direction of the flow. These currents, in turn, induce magnetic variations at ground level (Troshichev et al., 1988, 2006; Vennerstrøm, 1991) from which the Polar Cap (PC) indices are derived.

The steps in the calculations of PC indices may be found elsewhere, for instance in Troshichev et al. (2006) or Stauning et al (2006), Stauning (2016, 2020, 2021c) but are summarized here for convenience. In order to focus on solar wind effects, the horizontal magnetic variations, $\Delta\boldsymbol{F} = \boldsymbol{F} - \boldsymbol{F}_{RL}$, of the recorded horizontal magnetic field vector series, $\boldsymbol{F}$, with respect to an undisturbed reference level, $\boldsymbol{F}_{RL}$, are projected to an "optimum direction" in space assumed perpendicular to the DP2 transpolar convection-related sunward currents. The optimum direction is characterized by its angle, φ, with the dawn-dusk meridian and defines the direction for positive values of the projected polar magnetic variations, ΔF_{PROJ}.

An important parameter for the interaction between the solar wind and the magnetosphere is the solar wind merging electric field, E_M, (also termed E_{KL}) formulated by Kan and Lee (1979):

$$E_M = V_{SW} \cdot (B_Y^2 + B_Z^2)^{½} \cdot \sin^2(\theta/2) \quad : \quad \theta = \arctan(B_Y/B_Z) \tag{1}$$

where V_{SW} is the solar wind velocity, B_Y and B_Z are Geocentric Solar-Magnetosphere (GSM) components of the Interplanetary Magnetic field (IMF), while θ is the polar angle of the transverse IMF vector. The merging electric field is supposed to control the rate of merging between solar wind and geospace magnetic fields at the front of the magnetosphere and thereby in control of the input of solar wind energy to the Earth's magnetosphere.

In consequence, the projected polar cap magnetic disturbances are assumed proportional to E_M:

$$\Delta F_{PROJ} = \alpha \cdot E_M + \beta \tag{2}$$

where α is the slope and β the intercept parameter named from a graphical display of the relation (2). The scaling parameters are derived from regression of past data of a lengthy epoch, preferably a full solar cycle. The timing between the series of ΔF_{PROJ} and E_M values should be adjusted for the propagation from space to ground.

The Polar Cap (PC) index is now defined by the inverse relation of Eq. 2, i.e.:

$$PC = (\Delta F_{PROJ} - \beta)/\alpha \quad (\approx E_M) \tag{3}$$

With the relation in Eq. 3, the ΔF_{PROJ} scalar values are scaled to make the PC index equal (on the average) to values of E_M in the solar wind. The scaling of the polar cap magnetic disturbances to a quantity in the solar wind removes (in principle) the dependence on the daily and seasonally varying ionospheric conductivities and other local conditions such as the location of the measuring polar magnetic observatory.

The projection angle for the projection of the horizontal magnetic variation vector, (ΔF_X, ΔF_Y), in the (rotating) observatory frame at longitude, λ, to the optimum direction, φ, in space is defined by:

$$V_{PROJ} = \text{Longitude}(\lambda) + \text{UTh} \cdot 15° + \text{optimum direction angle}(\varphi) \tag{4}$$

where UTh is the UT time at the observatory in hours.

Thus, the projected magnetic variations could be expressed by:

$$\Delta F_{PROJ} = \Delta F_X \cdot \sin(V_{PROJ}) \pm \Delta F_Y \cdot \cos(V_{PROJ}) \quad : \quad \text{(+ for southern, - for northern hemisphere)} \tag{5}$$

The propagation delay, τ, between the reference location in space for the solar wind data and the location for related effects at the polar cap, and the optimum angle, φ, are both estimated from searching optimum correlation between E_M and ΔF_{PROJ} (Troshichev et al., 2006; Stauning et al., 2006; Stauning, 2016). The correlation coefficient is usually around R=0.75 and the delay from Bow Shock Nose (BSN) to the polar cap is close to τ=20 min. regardless of the observatory positions in their daily rotation and vary little with seasonal and solar activity conditions.

The calibration parameters, the slope, α, and the intercept, β, are found by linear regression between delay-time adjusted samples of ΔF_{PROJ} and E_M for each moment of the day and year using an extended epoch of past data (Stauning et al., 2006; Stauning, 2016; Troshichev et al., 2006). The regression parameters and the optimum angle values are tabulated throughout the year at 1-min resolution. They are kept invariant over years.

During conditions where the IMF B_Z component is negative or just small, the forward convection (DP2) patterns prevail and generate positive ΔF_{PROJ} values. The slope parameter (α) is positive and the intercept term (β) is relatively small. Hence, the PC index values (cf. Eq. 3) are mostly positive. During positive (northward) IMF B_Z (NBZ) conditions, reverse convection patterns (DP3) may emerge and generate negative ΔF_{PROJ} values which, in turn, may generate negative PC index values.

The PCC (PC combined) indices defined in Stauning (2007) and used in Stauning et al. (2008) and Stauning, 2012, 2021c, 2021d) are derived from the mean of non-negative values of the PCN and PCS indices filling 0'es for negative index values:

$$\text{PCC} = (\text{PCN if} >0 \text{ or else } 0 + \text{PCS if} >0 \text{ or else } 0)\ /2. \tag{6}$$

Thus, the PCC index values are always non-negative like the merging electric field, E_M, used for the calibration of the individual polar cap indices. The rationale behind this formulation builds on a critical assessment of the consequences of negative index values. At negative PC index values in both hemispheres, the global magnetic activity goes low like the PCC index values. However, there could still be local magnetic activity such as upper atmosphere auroral heating and reverse transpolar convection. Positive PC index values in one hemisphere indicates unipolar solar wind energy entry and enable generation of global magnetic disturbances in agreement with the positive PCC index values even if the PC index for the other hemisphere is dominatingly negative.

Essential features of the calculation of PC index values are presented in further sections. The steps of index derivation procedures comprise:

- Preparation and control of space data for IMF B_Y and B_Z and V_{SW} values needed to generate E_M values forming the basis for the calibration of PC indices.
- Preparation of polar horizontal magnetic vector data series, $\boldsymbol{F}$. Quality control and definition of base-levels, $\boldsymbol{F}_{BL}$.
- Derivation of the undisturbed reference level $\boldsymbol{F}_{RL}$ (including Quiet Day Curve, QDC) for calculations of the magnetic variations used for calculations of index values in definitive (post event) or real-time versions.
- Parallel calculations of delay (τ) and optimum angle values (φ) by optimizing the correlation between E_M and the projected polar magnetic variations, ΔF_{PROJ}, in their definitive versions.
- Regression of ΔF_{PROJ} on E_M in their definitive versions to derive slope (α) and intercept (β) scaling parameter values.
- Calculation and quality control of definitive PC index series for space science investigations.
- Derivation and validation of real-time PC index values for space weather monitoring applications.

The calibration parameters (φ, α, β) are derived from space and ground data from a reference epoch which for Dome-C considered here comprises the interval from 2009 to 2020. The basic 1-min polar magnetic data have been provided by the INTERMAGNET data service (https://intermagnet.org) while the space data are provided by the OMNIweb data service (https://omniweb.gsfc.nasa.gov) based on contributions from the ACE, WIND, and Geotail space missions. By appropriate time-shifting of the measurements, the data in the OMNI files have been referred to the magnetospheric bow shock nose (BSN) located at a distance of approximately 12 earth radii in front of the Earth towards the Sun.

In order to enhance the reliability and quality of the statistical processing, all calibration parameters are in the first step derived as mean hourly values for each calendar month and in the next step interpolated to generate specific values for each moment of the year. They are held constant over years.

2. Space data for generation of merging electric field values.

An example of IMF GSM B_X, B_Y, and B_Z components, and the solar wind velocity, V_{SW}. throughout 2015 is shown in in the top fields of Fig. 1 based on data from the OMNIweb data service (https://omniweb.gsfc.nasa.gov). Values of the derived merging electric field, E_M, are shown in the second-lowest field. The slowly varying curves (in red lines) superimposed on the fast field variations (in blue lines) indicate Gaussian-smoothed values. All space parameters are time-shifted from the satellite positions to the reference Bow Shock Nose (BSN) positions.

Note in Fig. 1 the systematic modulation of IMF B_X and B_Y intensities in opposite phases with a period of mostly around 27 days, i.e., the solar rotation period. This is an indication of the solar wind sector structure, where the general solar magnetic field has consistently organized structures through considerable parts of the rotating Sun's circumference. In Fig. 1 for 2015, the structure indicates a two-sector mode through most months.

The solar wind velocity, V_{SW}, as well as the solar F10.7 cm index considered a proxy for the solar ionizing radiation also display structured intensities in part related to the solar rotation.

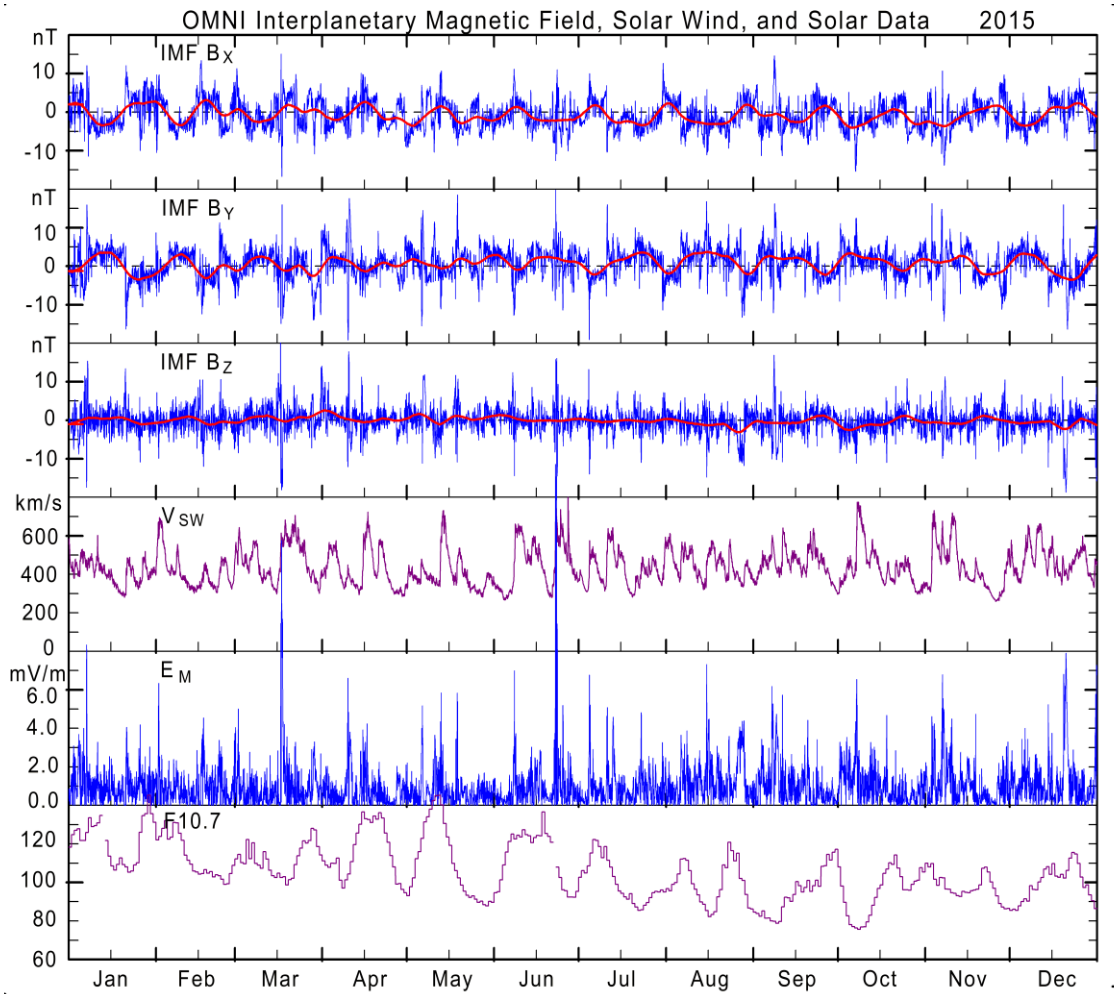

Fig. 1. OMNI solar and solar wind data. From top: interplanetary magnetic field (IMF) components B_X, B_Y, B_Z, in blue line with their smoothed values in red line; solar wind velocity (V_{SW}), merging electric field (E_M), and F10.7 cm solar index. All space data are shifted to bow shock nose (BSN).

The recurrent sector structure features for IMF B_Y are further illustrated in Fig. 2 that displays the smoothed IMF B_Y values from 1998 throughout 2019 against time of year. The larger B_Y amplitudes are generally associated with the two-sector structures reflecting the solar 27 days rotation period. Fig. 2 is an updated version from Stauning (2013b)

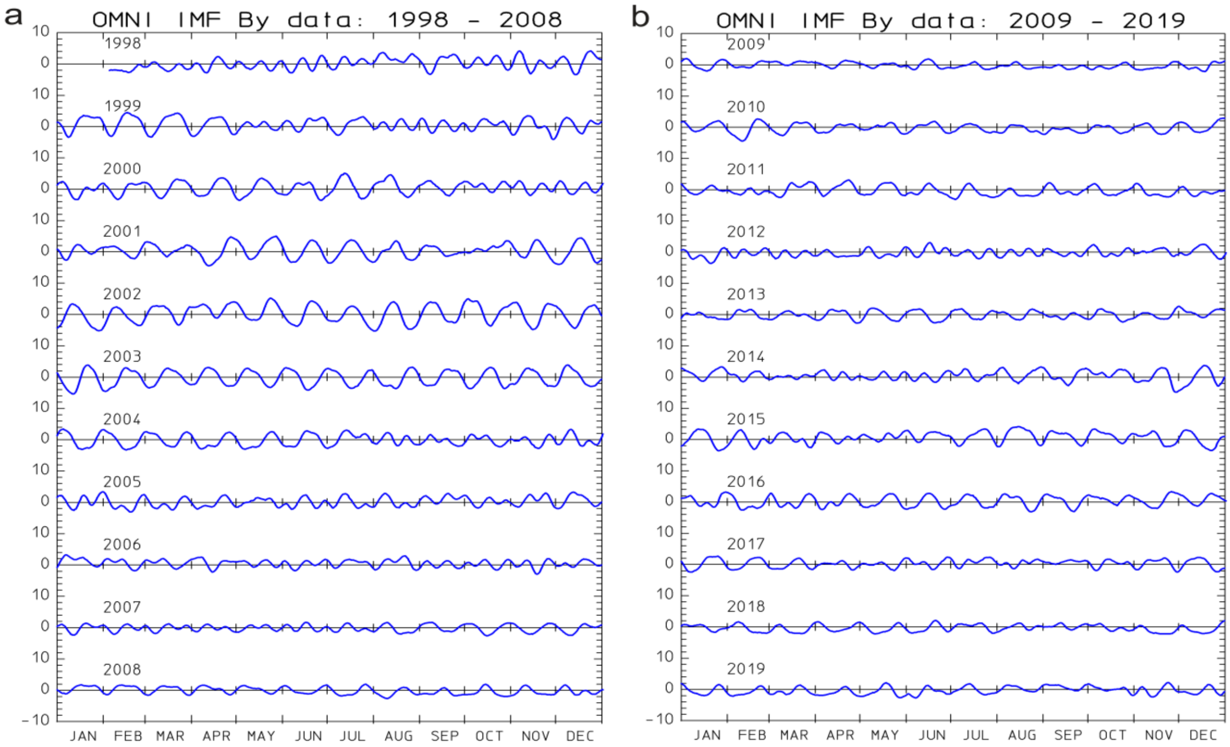

Fig. 2. Recurrence features (sector structure) for IMF B_Y. The IMF B_X data display corresponding features (in antiphase). Updated from Stauning (2013b).

3. Handling of geomagnetic observations.

The magnetic data used for the standard PCN indices are collected from Qaanaaq observatory in Greenland operated by the Danish Meteorological Institute (DMI) while the Danish Space Research Institute (DTU Space) operates the magnetic instruments and the data collection and processing. Data for the standard PCS indices are collected from Vostok observatory operated by the Arctic and Antarctic Research Institute (AARI) in St. Petersburg while data for en alternative PCS index are collected from French-Italian Dome Concordia (Dome-C) observatory (Chambodut et al., 2009; Di Mauro et al., 2014).

Essential geographic coordinates and geomagnetic parameters for 2021 based on the NASA VITMO application are listed in Table A1 for the standard and alternative observatories considered here and for further observatories (ALE and EUR) that would provide optimum conditions for PCN index calculations if data quality permits.

Table 1. Geographic and geomagnetic parameters at 100 km of altitude for selected stations.

Observatory	Station	Latitude	Longitude	CGMlat	CGMlon	LT=00	MLT=00
Name	Acr.	Deg.	Deg.	Deg.	Deg.	UThrs	UThrs
Qaanaaq	THL	77.47	290.77	83.86	23.86	4.62	3.60
Thule AB	TAB	76.54	291.18	83.00	22.65	4.59	3.68
Alert	ALE	82.50	297.65	87.02	70.10	4.16	0.14

Eureka	EUR	80.00	274.10	86.95	343.57	5.73	6.00
Resolute Bay	RES	74.68	265.10	81.97	327.82	6.33	6.88
Dome-C	DMC	-75.25	124.17	-89.31	44.52	15.72	1.77
Vostok	VOS	-78.46	106.84	-84.04	56.64	16.88	0.95

The magnetic data supplied from INTERMAGNET (https://intermagnet.org) are examined prior to their use in PC index calculations. It is of major importance that the base level values are correctly adjusted. In order to disclose possible problems, the monthly average X- and Y-component values are inspected. These values are derived as the means for all hours of the recordings from the 5 quietest (QQ) days each month defined by the International Service for Geomagnetic Indices (ISGI). Figs. 3a,b display the average values for the observed X and Y components from Qaanaaq (THL) and Vostok (VOS).

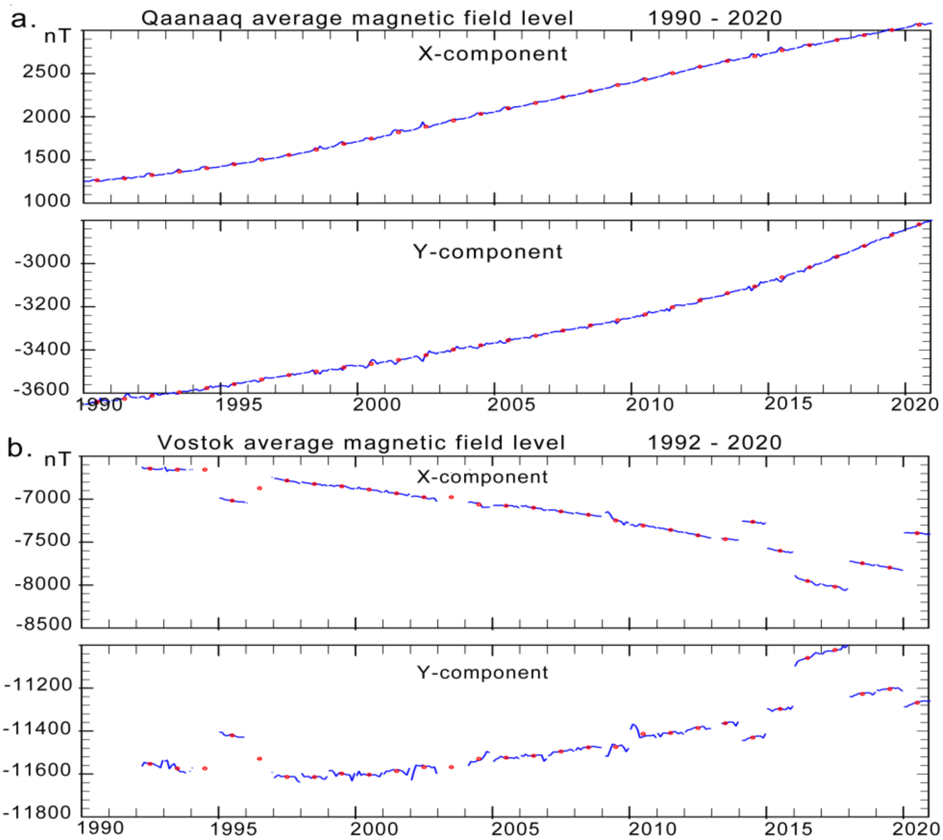

Fig. 3. Monthly (blue line) and yearly (red dots) average X- and Y-component values compiled throughout all hours of the 5 quietest days each month (http://isgi.unistra.fr). (a) Qaanaaq (THL). (b) Vostok (VOS). (data from https://intermagnet.org).

Corresponding average data from Dome-C observatory are displayed in Fig. 4a. There is a base level problem during 2016-2017 for Dome-C data. However, the adjustments are simple and the data quality is otherwise good. The monthly and yearly average data values after level correction are displayed in Fig. 4b.

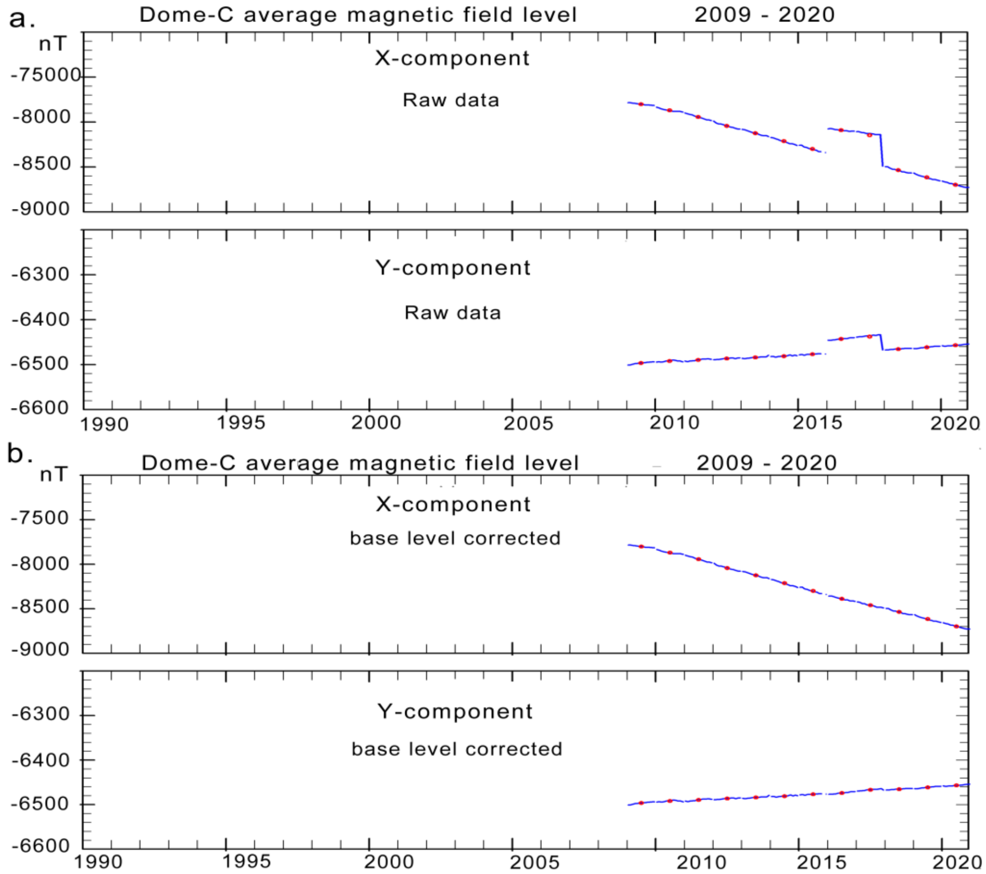

Fig. 4. Monthly (blue line) and yearly (red dots) average X- and Y-component values compiled throughout all hours of the 5 quietest days each month. (a) Dome-C measurements. (data from https://intermagnet.org). (b) Dome-C data with base level corrections.

In order to detect further data quality problems, all data are displayed in plots of the type shown in Figs. 5a,b of hourly values for a sequence of 2 years at a time. Here, the base levels have been adjusted and the regular secular variations have been removed. Thus, the amplitudes counted here from the 0 nT base line enter the calculations of PC indices after removal of the quiet daily variation ($\boldsymbol{F}_{QDC}$). The base level values, X_{BL} and Y_{BL} (after corrections), are noted in the displays.

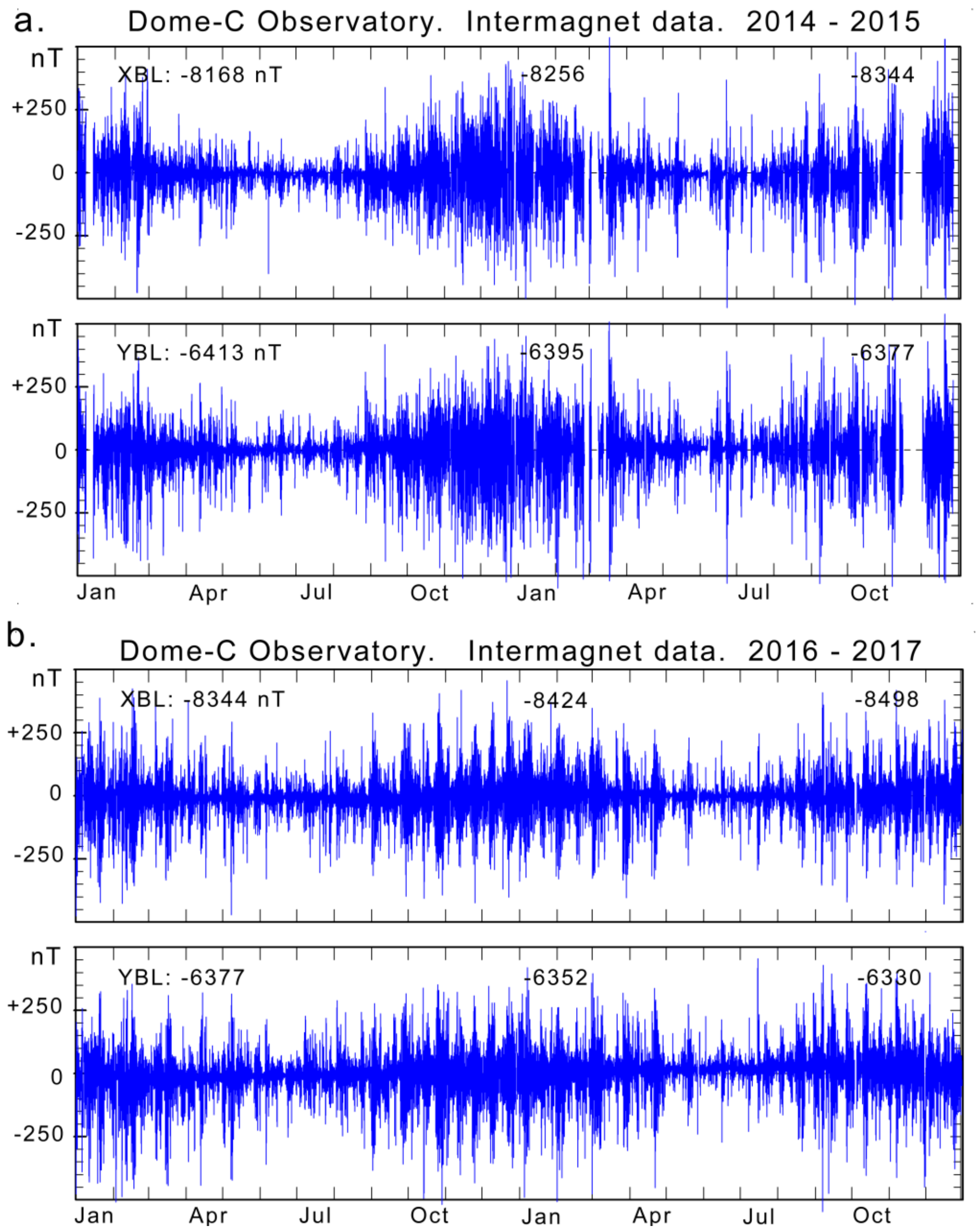

Fig. 5. Dome-C magnetic data. X- and Y-component displayed by 1-h average values. Base levels during 2016-2017 corrected. (a) data 2014-2015. (b) data 2016-2017. (data from https://intermagnet.org)

4 Reference level (QDC) for PC index calculations in the SRW version.

The definition of reference levels, $\boldsymbol{F}_{\mathrm{RL}}$, to be used for calculation of the polar magnetic variations needed for PC index calculations differs among the PC index versions. In the version developed by Vennerstrøm (1991), just the secularly varying base level, $\boldsymbol{F}_{\mathrm{BL}}$, was used as reference level.

$$\boldsymbol{F}_{\mathrm{RL}} = \boldsymbol{F}_{\mathrm{BL}} \qquad \text{(Vennerstrøm, 1991)} \qquad (7)$$

This level does not reflect the daily magnetic variations during undisturbed conditions. However, the calibration parameters, notably the intercept coefficient, reflect the undisturbed daily variation averaged over the epoch used for the regression.

In the version developed at AARI, the varying level on "*extremely quiescent days*" (Troshichev et al., 2006) was used as the PC index reference level. This level could be considered built from a quiet day curve (QDC), $\boldsymbol{F}_{\mathrm{QDC}}$, added on top of the base level, $\boldsymbol{F}_{\mathrm{BL}}$. Thus in vector formulation:

$$\boldsymbol{F}_{\mathrm{RL}} = \boldsymbol{F}_{\mathrm{BL}} + \boldsymbol{F}_{\mathrm{QDC}} \qquad \text{(AARI, Troshichev et al., 2006)} \qquad (8)$$

Extremely quiescent days are particularly rare at polar latitudes. Therefore, the concept was broadened to imply the generation of QDC values from quiet segments of nearby days within 30 days at a time (Troshichev et al, 2006; Janzhura and Troshichev, 2008). The use of a basic interval close to the solar rotation period (~27.4 days) with equal weight on each day's quiet samples removes most solar rotation effects on the QDCs by adding equal amounts of oppositely shifted solar or solar sector contributions.

In order to restore the more lengthy solar rotation contributions in the QDCs, Janzhura and Troshichev (2011) introduced solar sector ($\Delta\boldsymbol{F}_{\mathrm{SS}}$) terms ($\Delta H_{\mathrm{SS}}$,$\Delta D_{\mathrm{SS}}$) derived as the differences between the daily median component values and their average values. The terms were smoothed over 7 days with the day of interest at the middle. Further, the QDC values were calculated from the data less the SS-terms. The reference level was then formed by the sum of the baseline, the SS-terms, and the (30 days) QDC values. In vector notations:

$$\boldsymbol{F}_{\mathrm{RL}} = \boldsymbol{F}_{\mathrm{BL}} + \Delta\boldsymbol{F}_{\mathrm{SS}} + \boldsymbol{F}_{\mathrm{QDC}} \qquad \text{(AARI, Janzhura and Troshichev, 2011)} \qquad (9)$$

The procedure used for the IAGA-endorsed version described in Matzka and Troshichev (2014) uses the 7-days smoothed median value $\boldsymbol{F}_{\mathrm{SS}}$ and the $\boldsymbol{F}_{\mathrm{QDC}}$ values derived from the data less the median values in the reference level construction:

$$\boldsymbol{F}_{\mathrm{RL}} = \boldsymbol{F}_{\mathrm{SS}} + \boldsymbol{F}_{\mathrm{QDC}} \qquad \text{(IAGA, Matzka and Troshichev, 2014)} \qquad (10)$$

Actually, as explained in Stauning (2020), the two expressions define the same reference level quantity since $\boldsymbol{F}_{\mathrm{SS}} = \boldsymbol{F}_{\mathrm{BL}} + \Delta\boldsymbol{F}_{\mathrm{SS}}$.

The definition of the reference level is one of the issues that distinguish the PC index version presented in Stauning (2016) and used in the present work from the reference level definition in the IAGA-endorsed PC index versions. The reference level construction used here (Eq. 8) is based on the formulation in Troshichev et al. (2006) but uses the "solar rotation weighted" (SRW) reference level construction published in Stauning (2011) instead of the 30-days equal weight QDC methods detailed in Janzhura and Troshichev (2008) with the added SS-term from Janzhura and Troshichev (2011) or Troshichev and Janzhura (2012).

As formulated in Stauning (2011, 2018b,c, 2020, 2021c), the essential point for the SRW method is deriving the reference level from quiet samples collected on nearby days at conditions otherwise as close as possible to those prevailing at the day of interest. The factors of primary importance are:

(i) Sample "quietness"

(ii) Separation of the date of samples from the QDC date

(iii) Solar wind conditions (particularly IMF B_{Y} and V_{SW})

(iv) Solar UV and X-ray ionizing radiation (F10.7 cm solar flux index, Ly-α solar radiation)

For these factors weight functions are defined to optimize the selection of samples for the QDC. For each hour of the day, observed hourly average values at corresponding hours within an extended interval (±40 days) are multiplied by the relevant weights, added and then divided by the sum of weights to provide the hourly QDC value as shown in Eq. 11. Subsequently, the hourly QDC values are smoothed to remove irregular fluctuations and interpolated to provide any more detailed resolution as required.

$$X_{\mathrm{QDC}} = \Sigma\,(X_{\mathrm{OBS}}\cdot\mathrm{WF}) / \Sigma\,\mathrm{WF} \quad \text{and} \quad Y_{\mathrm{QDC}} = \Sigma(\,Y_{\mathrm{OBS}}\cdot\mathrm{WF}) / \Sigma\,\mathrm{WF} \qquad (11)$$

The weight function (i) for sample quietness is determined from the variability of 1-min data values within the hour much like the technique used in Troshichev et al. (2006) and detailed in Janzhura and Troshichev (2008). Two parameters are calculated on a vector basis. One is the maximum time derivative used to indicate the smoothness within the sample hour. The other is the average variance to define the slope of data values. Both parameters need to take small values for the hourly sample to be considered "quiet" (flat and featureless display). The parameters are independent on data representation in (X,Y) or (H,D) components.

For an estimate of further weight functions (ii) to (iv), the factors of importance were subjected to auto-covariance analyses vs. separation between the date of interest and the dates of the samples to be included in the construction of the QDC values. The auto-covariance values normalized by the variances should take large values to meet the condition that the quiet samples used to build the QDCs must represent conditions close to those prevailing at the day of interest. The auto-covariance results from the epoch (2009-2020) used here for definition of the scaling parameters are illustrated in Fig. 6 (similar to Fig. 3 of Stauning, 2011).

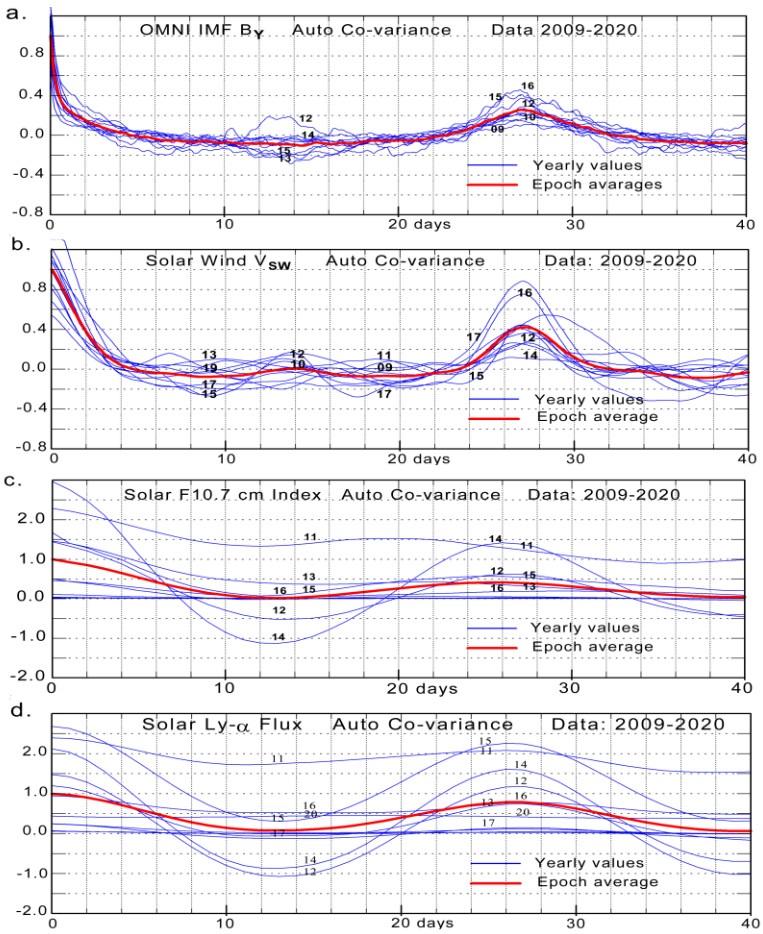

Fig. 6. Display of autocovariance values vs. shift in days. (a) IMF B_Y (OMNI), (b) Vsw, (c) F10.7 20 cm flux, (d) Ly-α flux. Data displayed throughout the years 2009-2020. Thin (blue) lines display auto covariance for one year, thick (red) lines displays mean auto covariance through 12 years (one solar cycle). Last two digits of the year are noted at the curves (similar to Fig. 3 of Stauning, 2011).

Details of the auto-covariance analysis are provided in Stauning (2011). The main results are, as expected, high autocorrelation values at nearby dates and also high values at dates displaced one full solar rotation of 27.4 days from the day of interest where the solar illumination and the solar wind conditions are similar, on a statistical basis, to the prevailing conditions. In between, at half a solar rotation, mixed IMF B_Y auto-covariance results were found. In a few cases a local maximum was seen indicating the occurrences of 4-sector solar wind structures. In most cases the autocorrelation functions have minima at half a solar rotation indicating 2-sector structures or just weak or mixed sector structures. The autocorrelation for solar wind velocities (V_{SW}) and solar illumination (F10.7 and Ly-α) gave unequivocal maxima at separations of a full solar rotation period (~27 days) and small or even strongly negative values at half a solar rotation.

For the solar rotation weighting a squared cosine function was selected to provide unity weights at the QDC date (zero separation), and at dates separated by 27.4 days (L_{SR}), and zero weight at half a solar rotation period when the opposite face of the Sun is directed toward the Earth. For these cases the recurrence features of solar UV illumination and solar wind intensity are absent while the solar wind sector effect, most likely, is in the opposite direction (at 2-sector structures) or weak (at multi-sector structure).

For the date separation, exponential weight factors functions were selected. The combined solar rotation and date difference weight function, WF_{DR}, is defined in Eq. 12:

$$WF_{DR} = WF_{SR} \cdot WF_{DD} = \cos^2(\pi \cdot X_{DD}/L_{SR}) \cdot \exp(-X_{DD}^2/R_{DD}^2) \qquad (12)$$

With R_{DD}=40 days, the final weight factor function, WF_{DR}, for sample separation, X_{DD}, has a central maximum holding 50% of the total weights and two secondary maxima located a solar rotation period (27.4 days) before and after the QDC day holding weights corresponding to 25% of the total weight each. The total span of samples included in the QDC construction is set to ±40 days to encompass all three weight maxima. The separation weight factors displayed in Fig. 7 have been pre-calculated and tabulated (for details see Stauning, 2011).

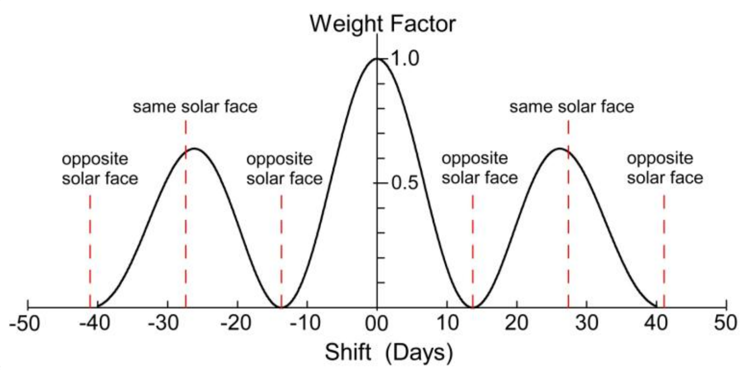

Fig. 7. Display of combined date difference and solar rotation weight factors vs. date shift. (from Stauning, 2011).

As data are collected, the quietness weight factor could be calculated promptly for each hour of recordings along with the hourly averages of each component. The three values are stored. The quietness weight factors are common for the two horizontal components and independent of their representation in (X,Y) or (H,D) coordinates.

Thus, at any time after initial 40 days of data collection, the relevant real-time QDC could be calculated and after further 40 days of initial data collection the final QDCs could be calculated for any day in the past on the provision that the basic data are final. The hourly component averages and their quietness weight factors are fetched from their stored values and their separation weight factors are found from the tabulated values. For each UT hour of the day, the hourly average component values within ±40 days are multiplied by the weight factors and summed up. The weight factors are summed up. The sum of weighted component hourly average values divided by the sum of weights defines for each hour the QDC value according to Eq. 12. The hourly sums of weights are quality factors for which alert limits could be set to caution against invalid values. The hourly QDC values are smoothed to remove fluctuations and then interpolated to provide the desired time resolution. The derived QDCs are routinely displayed in yearly plots like Fig. 8a.

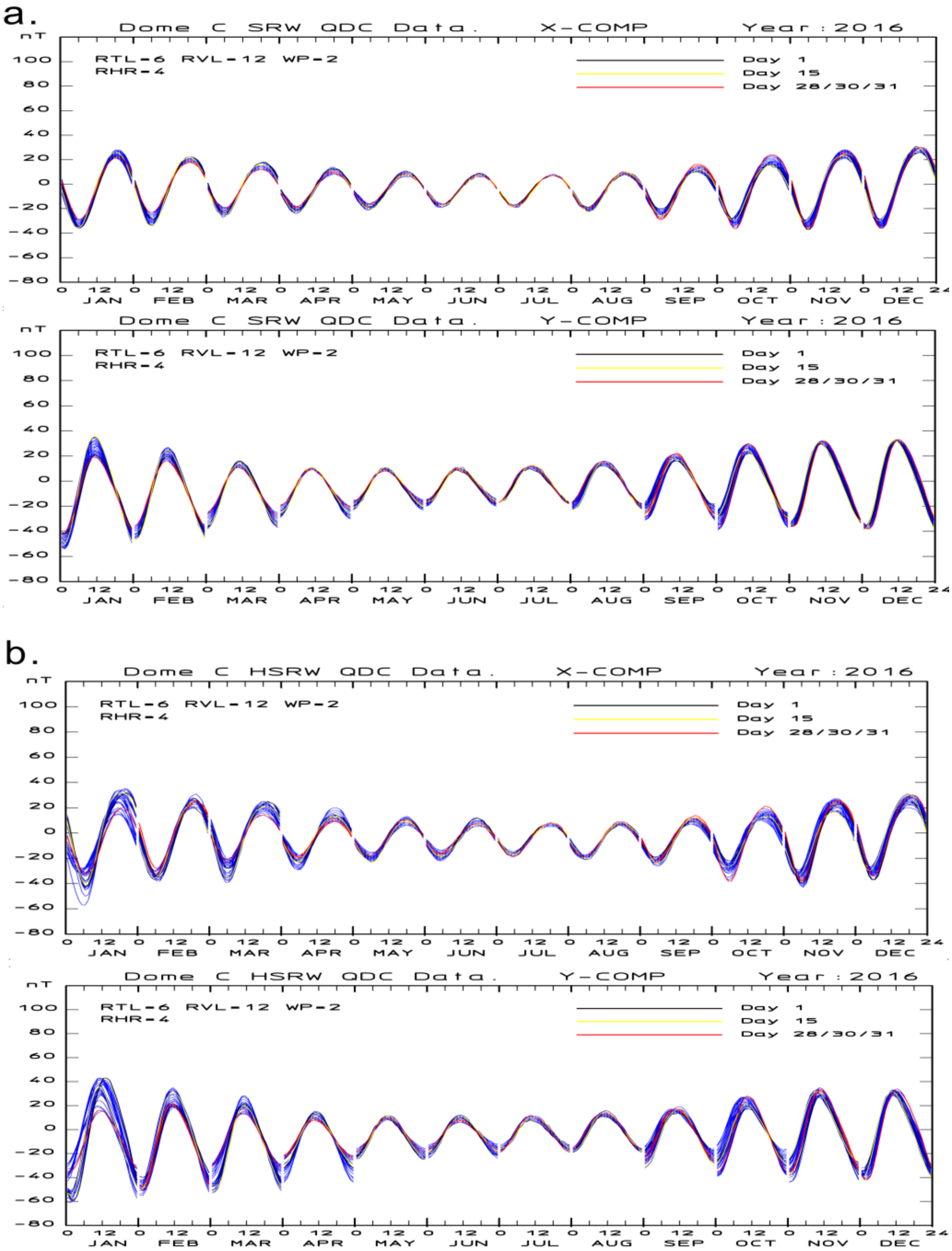

Fig. 8. One year's (2016) QDC values for Dome-C (DMC). The monthly assembly of daily QDCs is displayed in blue lines. The QDC values on day 1, 15, and the last day of the month are superposed in black,

yellow, and red lines, respectively. (a) Display of (post-event) SRW X- and Y-components. (b) Display of (simulated real-time) HSRW X- and Y-components.

In these diagrams for the magnetic data from Dome-C (DMC) there is a QDC curve for each day of the year. For one month at a time, the daily QDC curves are drawn on top of each other in blue line. For day 1 (in black line), day 15 (yellow), and last day of the month (in red line) the QDCs are re-drawn on top of the other QDCs. Going from the black through the yellow to the red curves provides an impression of the development of the QDCs throughout the month. The seasonal variations are very distinct with amplitude maxima at local summer. Most of the additional variability in the QDCs is caused by the IMF B_Y–related solar sector effects which are taken into account this way.

The weighting over ±40 days makes the determination of the final QDC fairly insensitive to intervals of missing data. Thus, the weighting technique allows calculations of real-time QDCs with reduced accuracy from past data collected within -40 to 0 days half interval solar rotation weighted (HSRW) QDCs by simply ignoring the not yet available post-event samples without changing the ±40 days' calculation scheme. As further data arrives, then the QDCs could be gradually improved to be completed after passing +40 days with respect to the day of interest. Thus, there are seamless transitions between real-time and post-event QDC values. An example of HSRW QDCs for 2016 is displayed in Fig. 8b.

Detailed displays of the relations between the observed values and the derived QDCs are provided in Fig. 9 with data from Dome-C for January and July, 2016. Note how accurately the variations in QDC levels and amplitudes make the QDCs match the relevant variations in the geomagnetic data during quiet intervals in spite of the otherwise very disturbed conditions.

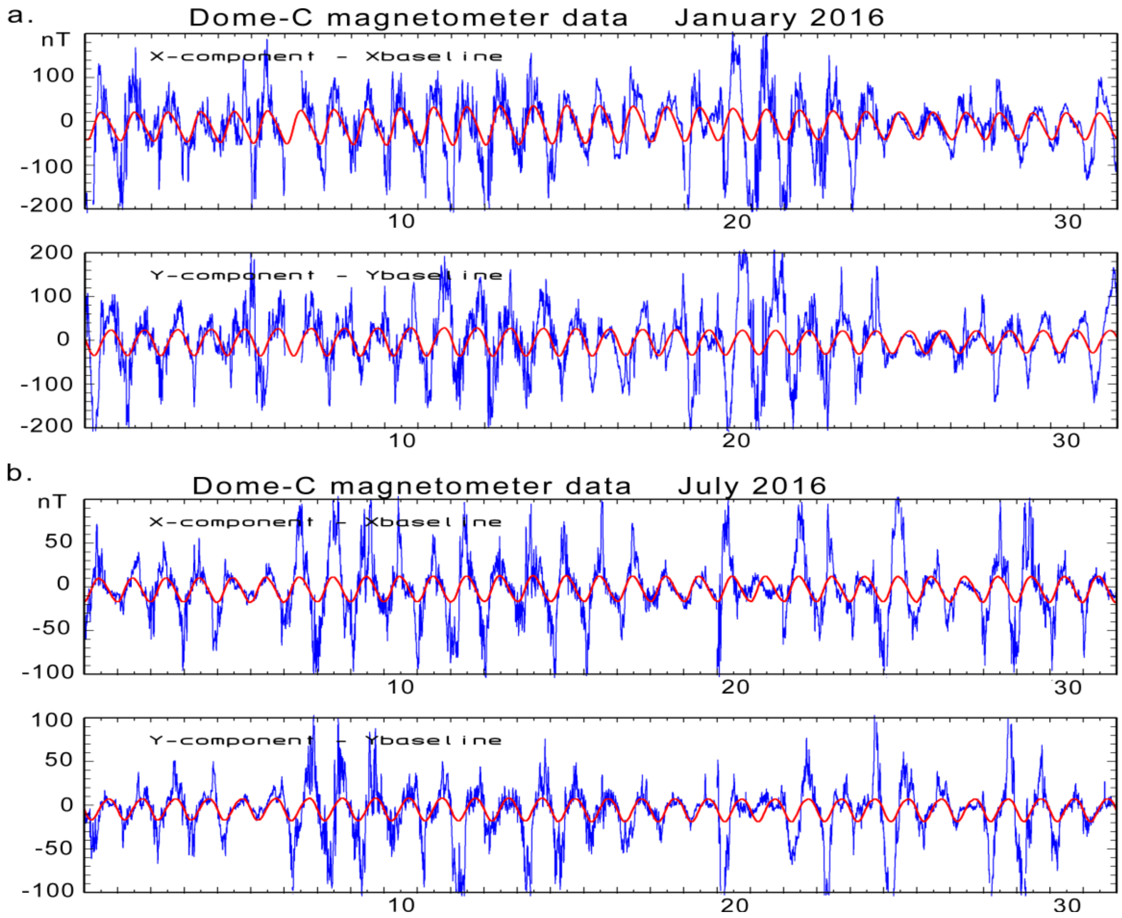

Fig. 9. Recorded data from Dome-C (blue line) and derived QDCs (red) for (a) January and (b) July 2016. Different scales are used to accommodate seasonal effects.

5. BSN to Polar Cap delays and optimum direction angle calculations.

The correlation between the horizontal disturbance vector $\Delta\boldsymbol{F}$ (corrected for the quiet daily variations) and the merging electric field, E_M, could be increased by projecting $\Delta\boldsymbol{F}$ to a specific direction, the so-called "optimum direction" considered to be perpendicular to the dominant DP2 forward convection (equivalent) currents. The optimum direction in space is characterized by its angle, φ, with the dawn-dusk meridian and varies slowly with local time and season. The optimum direction values are specific for each moment of the year and for each observatory.

Values of the optimum direction angle are calculated from analyses to find the maximum correlation between the reference level-corrected geomagnetic variations measured in the polar cap and the solar wind merging electric field values derived from interplanetary spacecraft data.

In order to correlate the satellite data with polar ground-based magnetic data it is important to adjust the relative timing of samples. The satellite data are first shifted from the satellite position to the reference location at the bow shock nose (BSN) at appr. 12 R_E in front of the Earth. In a rough estimate this time shift equals the difference in the X coordinates between the actual satellite position and the BSN location in a Geocentric Solar-Ecliptic (GSE) coordinate system divided by the solar wind velocity V_X. The OMNI data files are merged from best available interplanetary satellite measurements and referenced to the BSN position by careful modelling of the timing. For ACE satellite data the time shifts are on the order of 1 hour (cf., https://omniweb.gsfc.nasa.gov).

Next, the satellite data are referred to the polar regions by imposing a shift corresponding to an anticipated delay, τ, between the estimated time for solar wind parameters at BSN and the time of the resulting effects on the polar ionospheric convection as observed through the geomagnetic recordings. The delay is varied such that values of E_M at time t are correlated with values of ΔF_{PROJ} at time, $t + \tau$, looking for the maximum correlation coefficient to define τ .

5.1. Optimum angle calculations for Dome-C.

At the correlation studies by Stauning (2016) using 5-min samples, the best correlations between OMNI BSN values of E_M and Qaanaaq ΔF_{PROJ} series were obtained for delays close to 20 min. The delays were generally a little larger during night hours (01-07 UT) and smaller at morning and daytime hours (07-19 UT) than the average values. However, the variations were rather small, ranging from a minimum value at 17.0 minutes (winter day) to maximum at 20.4 minutes (summer night). The average optimum correlation coefficient was 0.717 while average optimum delay was 18.8 min (epoch 1997-2009). In further calculations a fixed value of τ = 20 min delay is kept throughout further epochs and for all polar stations including Qaanaaq, Vostok and Dome-C.

With the delay fixed, the optimum direction angles are now derived by the method defined in Stauning (2016). For each calendar month and each UT hour of the day and with steps of 10° in the optimum direction angle through all possible directions, the disturbance vectors, $\Delta \boldsymbol{F}$**,** are projected to the optimum direction according to Eqs. 4 and 5, and the correlations between the projected magnetic disturbances and the solar wind merging electric fields are calculated. The correlation coefficients (R) are calculated as function of the optimum direction angle using textbook's product-momentum formula.

$$R = \frac{N\,\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[N\,\Sigma X^2 - (\Sigma X)^2][N\,\Sigma Y^2 - (\Sigma Y)^2]}} \qquad (13)$$

where X = E_M, Y = ΔF_{PROJ} while the summations are extended over all relevant 5-min samples throughout the data interval years exempting NBZ cases, where IMF $B_Z>|B_Y|+3$ nT .

Among the calculated values of the correlation coefficients derived through all steps in optimum direction angle, the maximum value is found. Based on the direction angle for this maximum value along with the angles for the preceding and the following values of the correlation coefficient, a parabolic function is then adapted to determine the precise value of the optimum direction angle at the top of the parabola and the corresponding maximum correlation coefficient for the calendar month and UT hour in question.

In order to make the values generally representative some averaging and smoothing is necessary. In the present version, the values are exposed to bivariate Gaussian smoothing over months and UT hours by weighted averaging. The exponents used in the smoothing exponential weight functions characterize the degree of smoothing and are stored with the resulting optimum direction values.

$$WF = \exp\{ - (H - H0)^2/HR^2 - (M - M0)^2/MD^2 \} \qquad (14)$$

where H is the variable UT hour, H0 is the selected UT hour while HR is the half-width of the Gaussian weight function for the time-of-day. Correspondingly, M is the variable month, M0 the selected month and MD the half-width of the Gaussian weight function for months. The summations involved in the averaging are extended to twice the width of the Gaussian. Assuming cyclic variations, provisions are made for summation beyond the 24 hours of a day and 12 months of a year. The values used here are HR=4 hours and MD=2 months.

In order to avoid that the bivariate Gaussian smoothing reduces the amplitudes of the daily and monthly variations, the series, X(N), are exposed to a "peak amplitude enhancement" by applying the modification

$$X_M(N) = X(N) - A \cdot \{X(N-1) + X(N+1) - 2 \cdot X(N)\} \qquad (15)$$

With A=0.25, this modification enhances the peak values for a sine-like data distribution by typically a few per cent, which balances the reduction imposed by the Gaussian smoothing or interpolation process.

As a precaution against unfortunate data sections such as invalid data throughout a year, the total epoch was subdivided in 3 intervals of 4 years each (2009-12, 2013-16, and 2017-20). The optimum angles were calculated for each interval and the results averaged after inspection of the intermediate results.

The results for the mean daily variations in the optimum angles within each month of the year are displayed in the 12 monthly sections of Fig. 10. The influence from including QDC correction in the processing of the magnetic data has been examined. In the presently used program (DMI2016 version) to derive the optimum direction angles, the QDC correction is invoked in a single command line and can easily be switched on or off. The QDC correction was found to have negligible effects on the optimum direction angles. Correspondingly, the effects from the screening against NBZ samples were examined. Fig. 10 displays the optimum angle values derived without QDC and without NBZ samples (blue line), with QDC and without NBZ samples (magenta), and with QDC and with NBZ samples (red line). The differences between the three cases are small. The curves with QDC correction and without NBZ samples display the preferred set of optimum angle values.

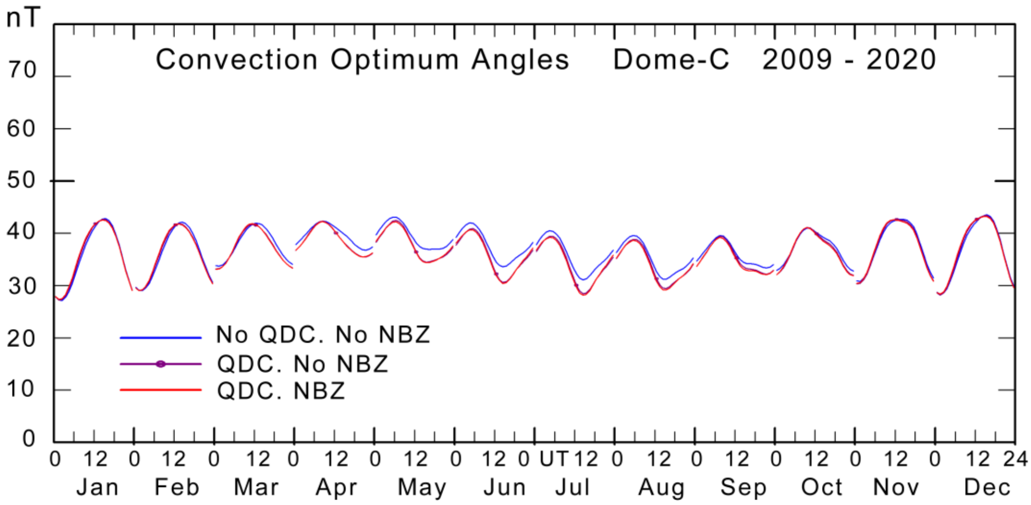

Fig. 10. Monthly mean daily variation in optimum angles for Dome-C for each month of the year. Angles have been derived by using DMI2016 methods without QDC and without NBZ samples (blue line), with QDC and without NBZ (magenta), with QDC and with NBZ samples (red).

The smoothed monthly mean daily variation in the optimum angles are converted into series of hourly values for each day of the year by continued application of bivariate Gaussian smoothing. The hourly values are converted into any more detailed resolution by simple parabolic interpolation. The details of the bivariate Gaussian smoothing or interpolation, the peak amplitude enhancement, and the interpolation techniques are demonstrated in Stauning (2016). The selection of parameters involved in smoothing or interpolation is kept with the resulting values.

6. Calculations of slope and intercept

Recalling that we are searching for proxy values based on polar magnetic disturbances to represent the solar wind "merging" electric field ($E_M = E_{KL} = V_{SW}\, B_T \sin^2(\theta/2)$), the general assumption is that there is a (statistical) linear relation between the polar magnetic variations, ΔF_{PROJ}, and the solar wind electric field, E_M, and that this relation can be inverted and used to define a polar cap (PC) index by equivalence (cf. Eqs. 1-3).

6.1. QDC and NBZ effects on calculations of slope and intercept.

Contrary to the calculation of the optimum direction, the QDC issue has a considerable importance for the calculations of slope and intercept parameters. Figs. 11a,b provide sketches of the consequences of including the projected QDC value (Fq) at the regression. The sketches demonstrate that the slope remain the same ($\alpha 2 = \alpha 1$) while the intercept is modified by the amount Fq (i.e., $\beta 2 = \beta 1 - F$q).

The question whether reverse convection cases should be included in the data base used for the regression is also important. The sketches in Figs. 11c,d illustrate that the large negative reverse convection, ΔF_{PROJ}, samples combined with small positive E_M values occurring during NBZ conditions make the regression slope steeper ($\alpha 3 > \alpha 2$) while the intercept value gets more negative ($\beta 3 < \beta 2$).

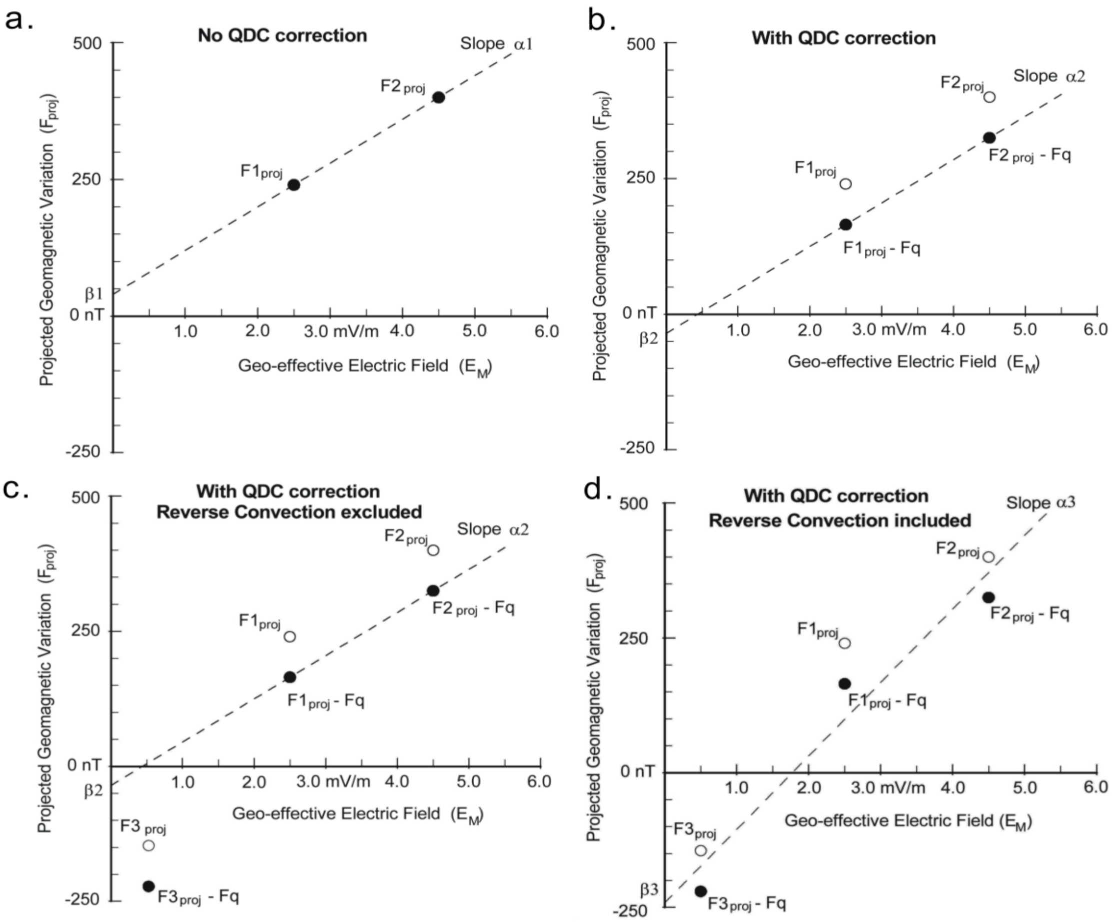

Fig. 11. (a) Regression without QDC involvment. (b) Regression with QDC involved (parallel displacement by Fq). (c) Regression with QDC and without NBZ samples. (d) Regression with QDC and with NBZ sample ($F3_{PROJ}$) included. (after Stauning, 2013).

The preferences in the present version are using QDC involvement and omission of NBZ samples in the calculation of index calibration parameters. However, the different options are looked at.

6.2. Slope and intercept regression calculations

Basis for the regression is the above-mentioned assumption of a linear relation between the merging electric field, E_M, and the projected (baseline and QDC corrected) magnetic variation, ΔF_{PROJ}, as expressed in Eq. 2 from which average values of the slope, α, and the intercept parameter, β ,should be derived by proper statistical methods from a comprehensive and representative data base.

Similar to the optimum direction angles (φ) the regression coefficients are derived as series of mean hourly values for each calendar month. To solve for the coefficients in the linear relation in Eq. 2 ($\Delta F_{PROJ} = \alpha\, E_M + \beta$), statistical text-books provide the least squares regression formulas:

Slope: $$\alpha = \frac{N\,\Sigma XY - (\Sigma X)(\Sigma Y)}{N\,\Sigma X^2 - (\Sigma X)^2} \qquad (16)$$

Intercept: $$\beta = \frac{(\Sigma Y)(\Sigma X^2) - (\Sigma X)(\Sigma XY)}{N\,\Sigma X^2 - (\Sigma X)^2} \qquad (17)$$

In these regression formulas, the merging electric field (E_M) is parameter X while the projected magnetic disturbance (ΔF_{PROJ}) is parameter Y. For each calendar month of the year the hourly values of α and β are formed by processing all corresponding 5-min values of E_M (t-20 min) and ΔF_{PROJ} (t) throughout that hour of all days of the month and all years of the selected epoch. In the first step the epoch from 2009 to 2020 are divided in three sets of 4 years each as done for the optimum angles. The three subsets are subsequently inspected and then averaged.

In order to avoid reverse convection cases in the data base used for calculations of PC index coefficients, a combination of limits on actual IMF values and projected magnetic variations is used. For the IMF it is required that IMF B_Z < | IMF B_Y | + 3.0 nT. This condition excludes cases where strong northward B_Z is the dominant IMF component. A further condition imposed on the selection of data requires that the projected magnetic variation, ΔF_{PROJ}, is larger than the value corresponding to PC = -2 mV/m (≈ -50 nT). This condition ensures that cases with strong reverse convection, which may continue for a while after the driving northward IMF parameter has been reduced or has changed polarity, are also omitted.

6.3. Regression slope and intercept for Dome-C for epoch 2009-2020.

The raw (non-smoothed) values of the slopes and intercept coefficients derived from using Eqs. 16 and 17 are exposed to bivariate Gaussian smoothing over months and UT hours by weighted averaging (Stauning, 2016). The resulting slope and intercept values for epoch 2009-2020 are presented in Figs. 12a,b in the format corresponding to Fig. 10. Each of the 12 monthly sections presents the mean hourly parameter variation for the month.

The monthly mean hourly values of the slopes and intercepts are converted into series of hourly values for each day of the year by Gaussian bivariate weight function interpolation with peak amplitude enhancements, corresponding to the handling of the optimum angle parameter. For finer resolutions, e.g., 5-min or 1-min samples, simple parabolic or linear interpolations are used. (Stauning, 2016).

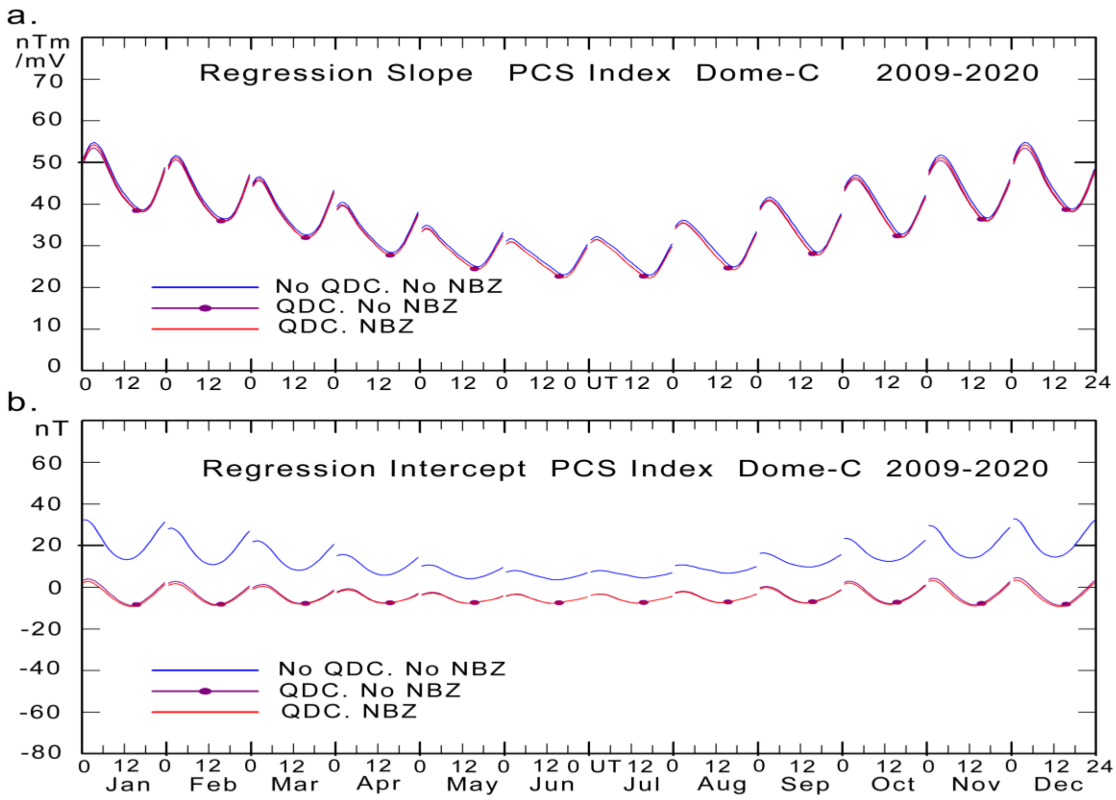

Fig. 12. PCS slope and intercept values derived by regression of ΔF_{PROJ} on E_M. Data from Dome-C (DMC) for epoch 2009-2020. Data processed without QDC involvement and without NBZ samples are displayed in blue line; data with QDC and without NBZ samples in magenta line with dots; data with QDC and including NBZ samples in red line.

It is seen from Fig. 12 as anticipated from the sketches in Fig. 11 that the slope values are little affected whether the data are handled with or without QDC. The intercept values without QDC involvement (blue line) are increased by an amount representing the projected QDC contribution while including the NBZ samples (red line) has no significant effects or slope or intercept. Due to its proximity to the magnetic pole the amount and the strength of reverse convection events are minimal at Dome-C which makes the station an ideal location for supply of data for PCS calculations.

The monthly mean hourly values of the calibration parameters shown in Figs. 10 and 12 may be interpolated to provide finer resolution as described above and converted into tables of parameters for each 5-min (or 1-min) interval of the year. The calibration parameters are not invariant to general changes in solar activity or to secular variations in local polar magnetic configuration. However, they are kept invariant over years unless a new index version is implemented.

6.4. Calculation of PC index values post event and in real time.

With the DMI methods (Stauning, 2016), the scaling parameters, (φ, α, β), are derived as monthly mean hourly values and then interpolated to provide tables at finer resolution as required. With the optimum angle values displayed in Figs. 8, the slope and intercept values displayed in Fig.10, and the QDC values derived by the solar rotation weighted (SRW) method described in section 5 (cf., Figs. 6 and 7), it is now possible to calculate PCS index values vs. UT time and date. The magnetic variations are derived from the observed values by subtracting base line and QDC values according to Eq. 8. The projection angles are derived from Eq. 4 using the tabulated optimum angles (φ). The projection is accomplished by Eq. 5. The slope and intercept values, α and β are fetched from their tabulated values to be used in Eq. 3 defining PC index values. These steps are illustrated in Fig. 13.

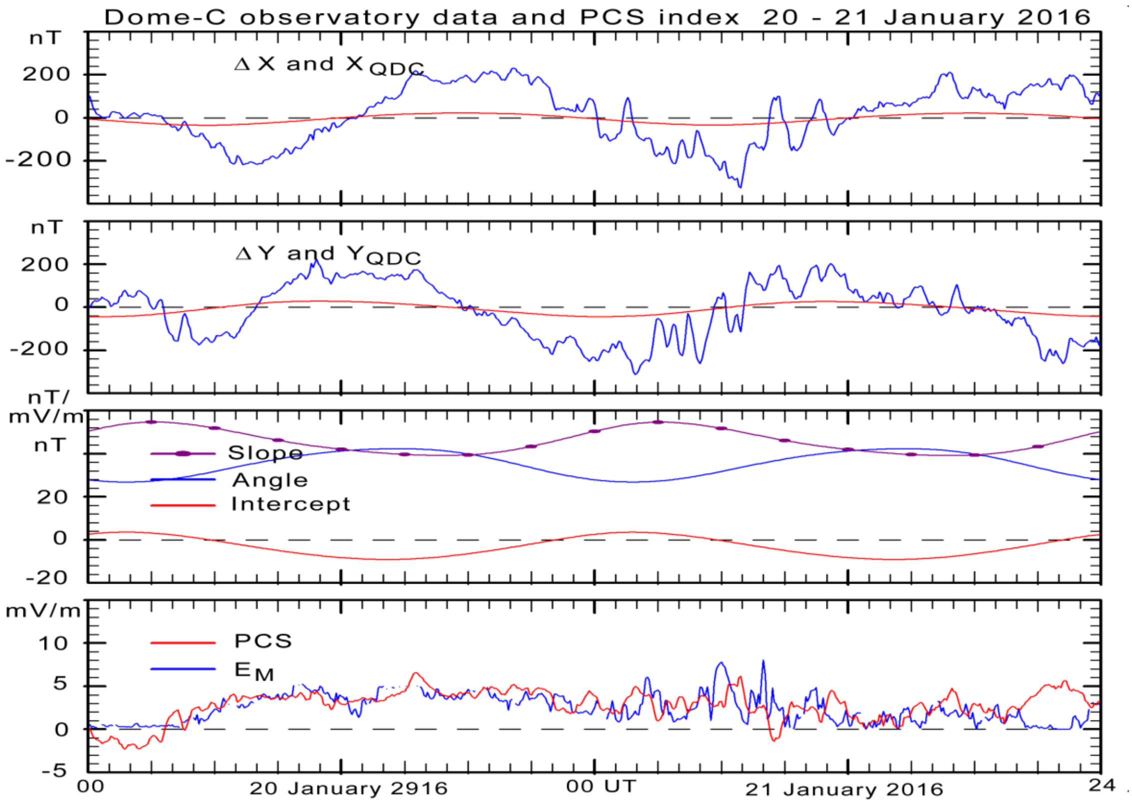

Fig. 13. Example of PC index calculations for 20-21 January 2016. From top of the diagram the X-component (blue line) and X_{QDC}-component (red), the Y- and Y_{QDC}-components, the slope (magenta with dots, intercept (red) and optimum angle (blue) scaling parameters. In the bottom field the PCS values (in red line), and the relevant time-shifted merging electric field, E_M (for illustration). All data are 5-min samples.

For real-time applications the critical issue is defining the undisturbed reference level. For the present approach the QDC values are derived by the HSRW method using quiet samples collected from past data only during the interval from -40 to 0 days. Examples are provided in Figs. 6b. A detailed description of methods for calculations of current QDC values and PC indices in real-time may be found in the appendix to Stauning (2018c).

7. Assessments of PC index quality.

For a geophysical index offered to the international scientific community and important space weather services, the quality of the post event (definitive) as well as the real-time index values is of utmost importance. In spite of this (seemingly) obvious ascertainment, little efforts have been provided on this issue at past and present PC index versions.

The main quality principles were formulated in Troshichev et al. (1988).

"- PC index in any UT time should be determined by the polar cap magnetic disturbance value related to influence of the geoeffective solar wind, and therefore

- the magnetic disturbance vector δF should be counted from level of the quiet geomagnetic field to eliminate variations unrelated to the solar wind fluctuations;

- PC index should correspond to the value of the interplanetary electric field E_{KL} (E_M) impacting the magnetosphere, irrespective of UT time, season and point of observation."

The reference levels advocated here are by their definition (cf. section 5) based on quiet (the quietest) geomagnetic samples and thus they comply with the quality requirements.

The relations between the PC indices and the solar wind are illustrated by the correlations between 15-min average values of Dome-C-based PCS index values (PCD) and values of the merging electric field shifted by 20 min displayed in Fig. 14. The quarterly mean correlation coefficients between 15-min E_M values and PCS values based on Dome-C data are displayed in heavy magenta line while the corresponding correlation coefficients for Vostok-based PCS values are displayed in red line and the coefficients for Qaanaaq (THL)-based PCN values are shown in blue line.

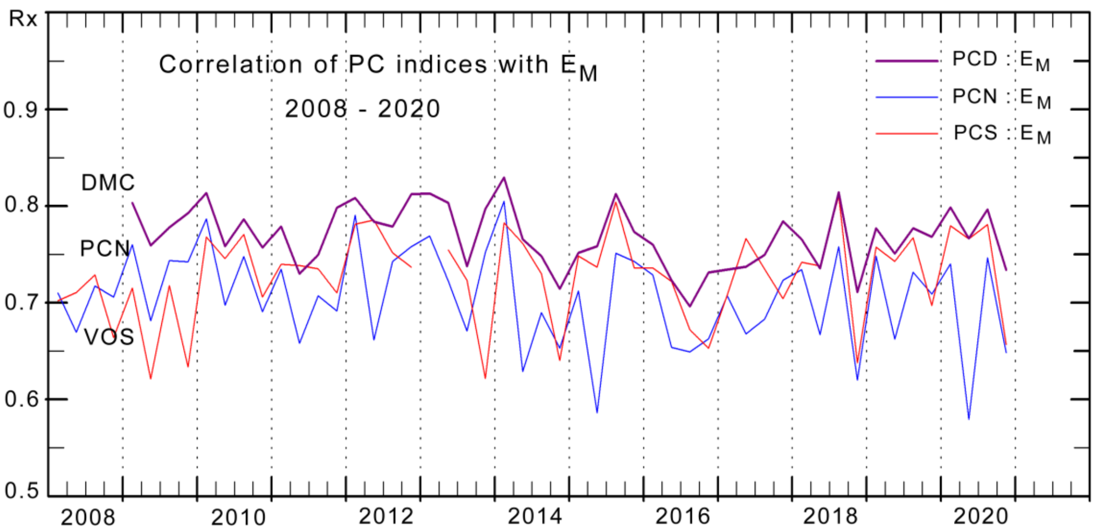

Fig. 14. Quarterly means of coefficients for the correlation between 15-min averages of the merging electric field, E_M, and Dome-C-based PCS values (PCD) in heavy magenta line and corresponding coefficients for Vostok-based PCS values (red line) and Qaanaaq-based PCN values (blue line).

With a single exception in 2017, the correlation between E_M and Dome-C based PCS index values seen in Fig. 14 is higher – at times much higher – than the correlation between E_M and the Vostok-based PCS values and consistently much higher than the correlation between E_M and the Qaanaaq (THL)-based PCN values throughout the epoch (2009-2020).

The correlation between E_M and PCS values based on Dome-C data is close to the corresponding values for PCS indices based on Vostok data throughout the local winter season (May-September) but much higher at local summer (October-April). The correlation coefficients between E_M and

Qaanaaq-based PCN index values are much lower than either E_M - PCS correlations during most of the year, particularly during northern summer.

The seasonal variations in the correlation between the merging electric field and the Dome-based PCS (PCD), the Vostk-based PCS, and the Qaanaaq THL) based PCN values are displayed in Fig. 15 by the monthly mean correlation coefficients for 15-min samples averaged over the epoch 2009-2020. The line types are the same as those used in Fig. 14. The order of southern months has been rearranged to make seasons match.

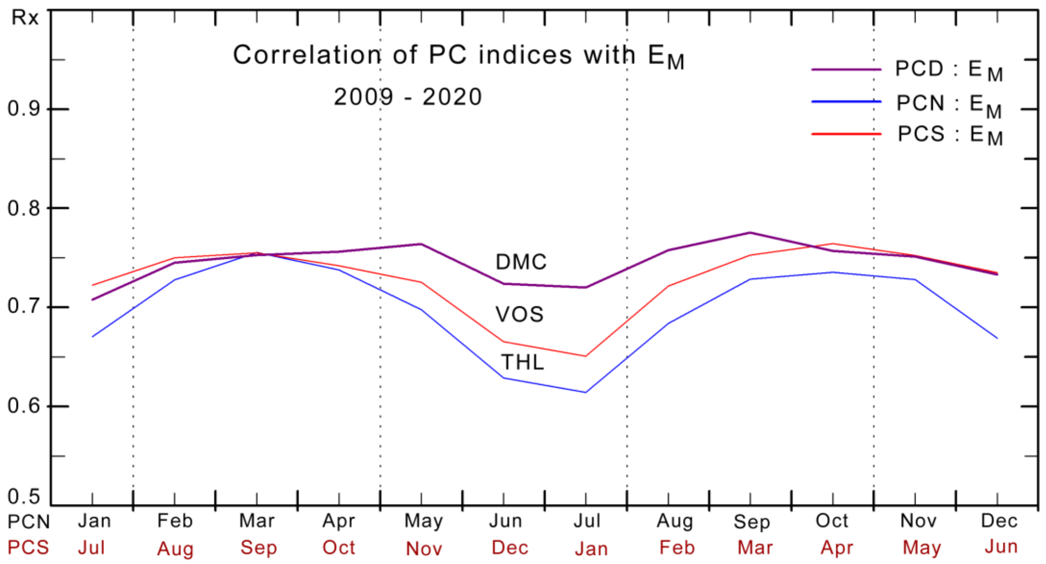

Fig. 15. Monthly means of coefficients for the correlation between 15-min averages of E_M and Dome-C-based PCS values (PCD) in heavy magenta line. Corresponding coefficients for Vostok-based PCS values in red line and Qaanaaq-based PCN values in blue line. The order of southern months has been rearranged.

It is seen from Fig. 15 that the coefficients for the correlation between E_M and PCS values based on Dome-C data are close to the corresponding values for PCS indices based on Vostok data throughout the local winter months (April-September) but much higher at local summer (October-March). The correlation coefficients between E_M and Qaanaaq-based PCN index values are much lower than either E_M - PCS correlations during most of the year.

The main reason for the low correlations during local summer months is the increased occurrences and enhanced intensities of reverse convection events compared to conditions at (local) winter. In terms of location, such reverse convection events are particularly frequent and intense midway between the Cusp region at the dayside and the geomagnetic pole. Thus, they are less frequent and intense at Vostok compared to Qaanaaq and furthermore less frequent at Dome-C compared to Vostok due to the closer proximity to the (southern) geomagnetic pole (cf. Table 1).

The quality of the Dome-C-based PCS index compared to further PC index versions could be verified by contrasting the different versions which in most cases should provide the same index values. There could be deviations in response to IMF B_Y–related effects that may act in opposite directions at the opposite polar caps. NBZ cases may also generate large hemispherical differences. During NBZ conditions, the PC index values are often strongly negative in one hemispherical version while the corresponding index values for the opposite hemisphere are just small. Such features are readily seen in composite plots such as Fig. 16 and distinguished from erroneous values.

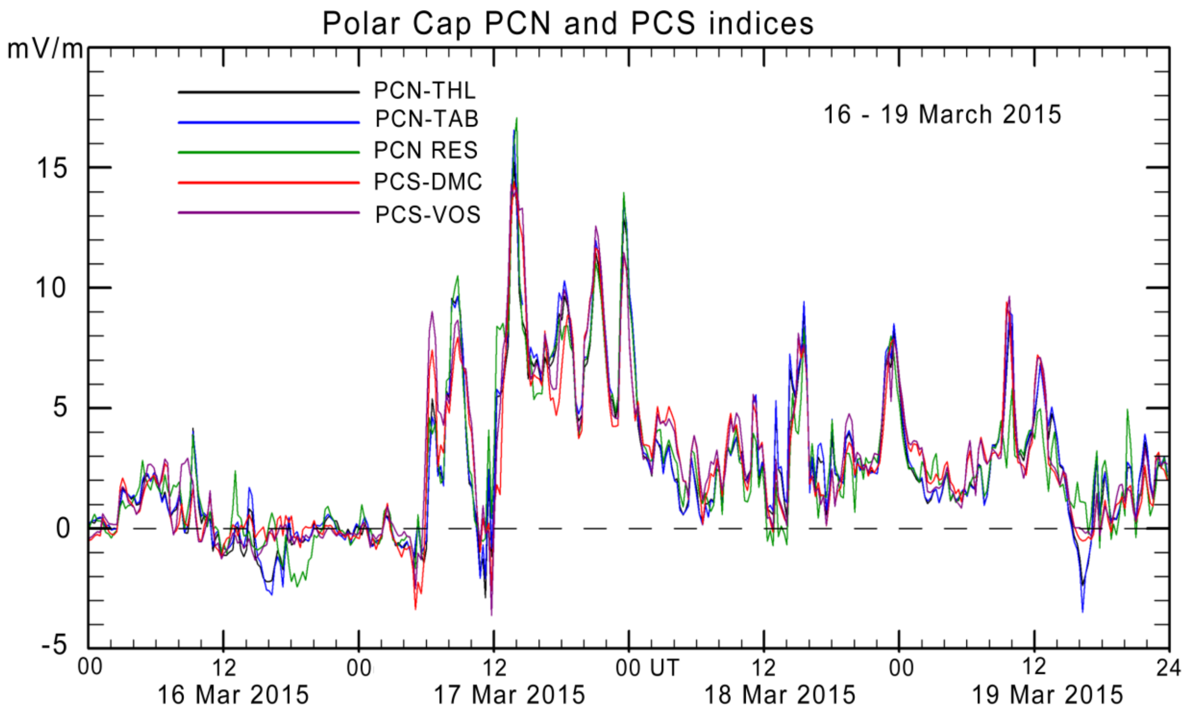

Fig. 16. PC indices in different versions for the strong magnetic storm on 22-25 June 2015.

Fig. 16 displays a fair agreement between indices for positive index values, in particular between Vostok-based and Dome-C-based PCS indices. At times the (northern) PCN values take large negative excursions while the corresponding (southern) PCS indices are just small.

A convenient method to detect irregular indices is by inspecting monthly diagrams as the example shown in Fig. 17. In these diagrams one should look for agreement between amplitudes of E_M and positive PC index values while negative PC index values should be related to small E_M values. Another feature to observe is the consistency between the component values and their respective QDC values. The QDCs should agree with the recordings at low activity levels and be in-phase during larger disturbances while turning out-of phase during reverse convection cases.

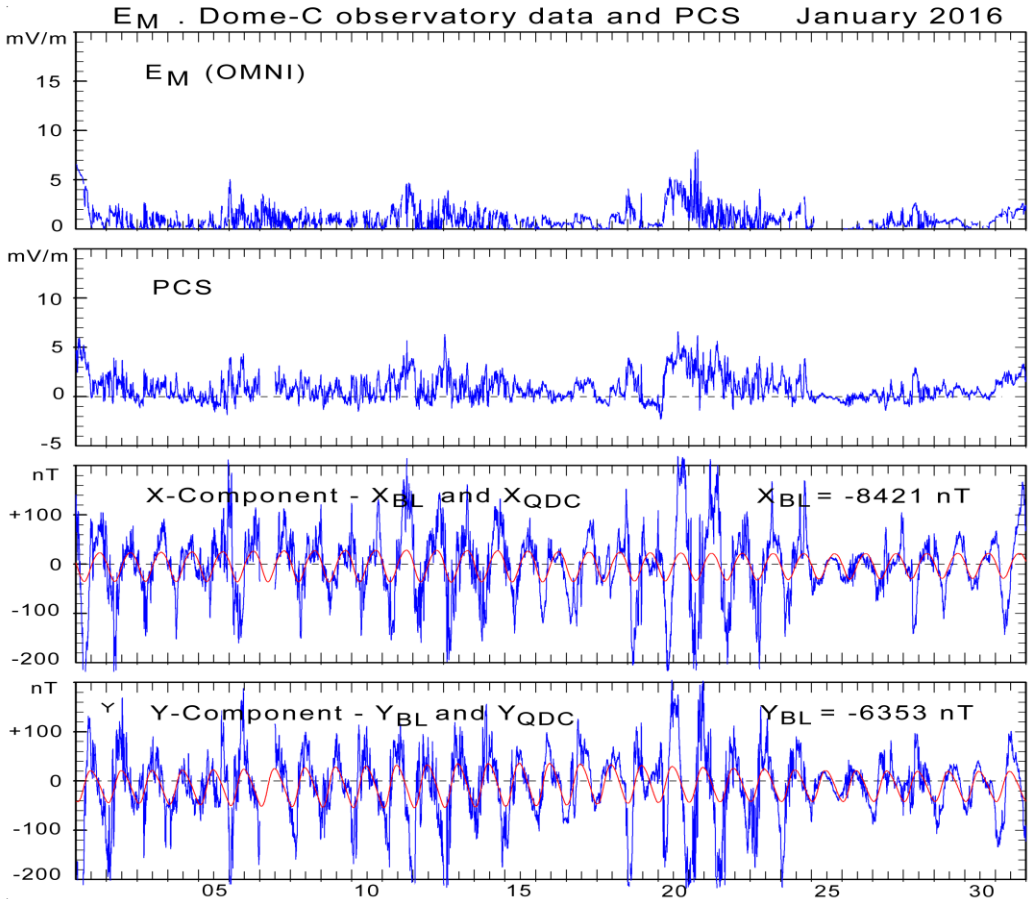

Fig. 17. Quality assessment diagram. From top of the diagram display of E_M, PCS, *X*- and *X*-QDC, and *Y*- *Y*-QDC components. The QDC values are displayed in red line. Other values in blue line.

One might expect that such quality control measures were implemented by the index providers (AARI and DTU Space) in respect for potential users in the scientific community. However, neither the index providers nor the IAGA-supported index supplier (ISGI) appear having implemented supervision of the index quality.

A striking example of invalid PCS index values is displayed in Fig. 18 with indices for 27-30 June 2011 for Qaanaaq (THL), Vostok (VOS) and Dome-C (DMC) in the versions (DMI) defined in the present work and PCN and PCS index values in the IAGA-supported versions.

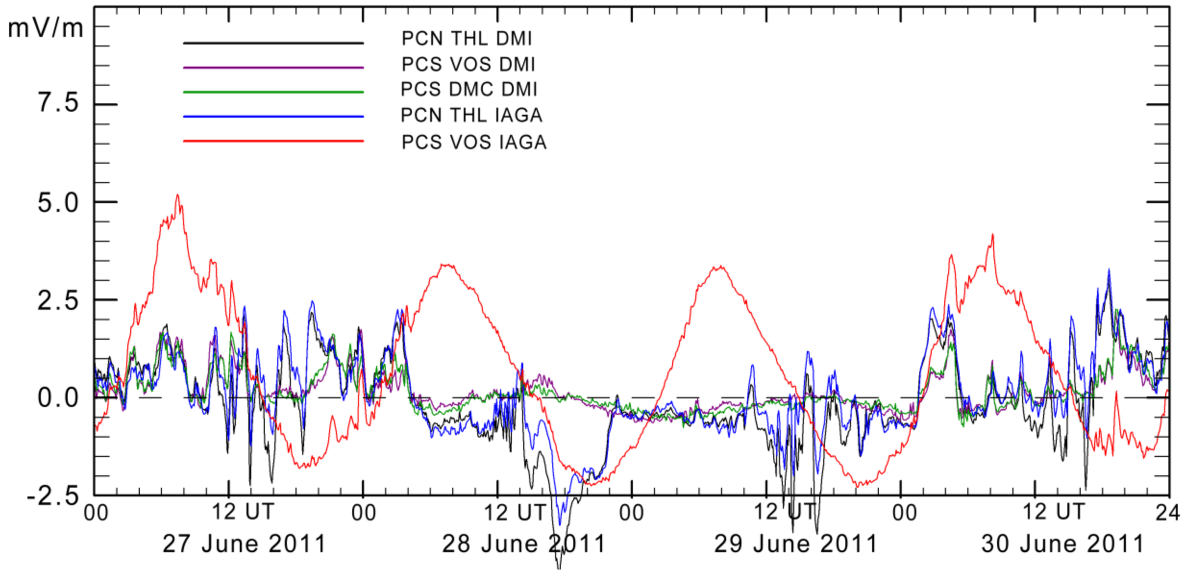

Fig. 18. PCN and PCS index values for 27-30 June 2011 in DMI2016 versions based on data from Qaanaaq (THL) in black line, from Vostok (magenta), and from Dome-C (green). PCN and PCS index values in IAGA-supported versions based on data from Qaanaaq (blue line) and Vostok (red line).

It is readily seen that the daily excursions between -2 and +4 mV/m (magnetic storm level) in the IAGA PCS values (red line) must be in error when compared to the other index values recorded on these rather quiet days. In passing it might be noted that the Vostok-based PCS indices (magenta line) agree well with the Dome-C-based PCS index values (green) in the DMI versions.

These erroneous Vostok-based PCS values are easily detected in comparative plots of PCS values from the available sources, that is, Dome-C and Vostok, in the DMI2016 version and in the IAGA2014 version. Fig. 19 presents an example for 1-30 June 2011.

Values of the merging electric fields, E_M, have been added in the top field of Fig. 19. These data are added here for information but not strictly necessary for a basic quality control. The invalid Vostok-based PCS indices are seen in the bottom field.

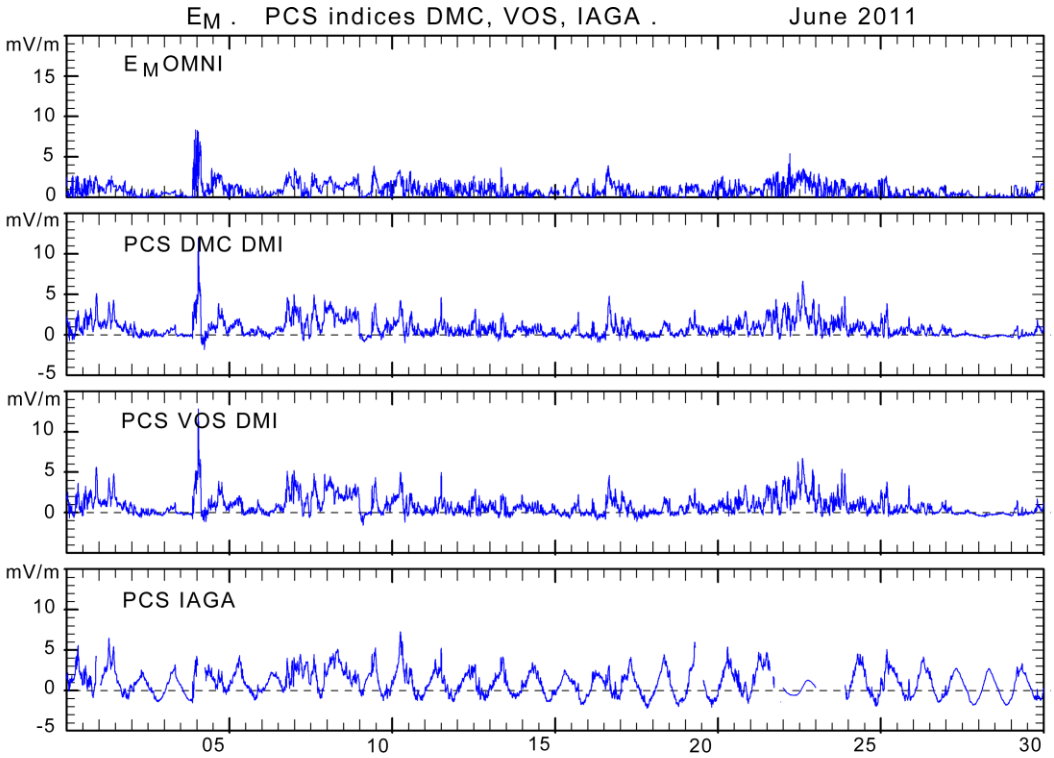

Fig. 19. Quality control diagram. From top: merging electric field (E_M), Dome-C-based PCS and Vostok-based PCS indices in DMI versions, and Vostok-based PCS index values in the IAGA version.

The PCN and PCS index values in the IAGA-supported versions displayed in Figs. 18 and 19 were downloaded in September 2021 from the "final" version link at the AARI web site https://pcindex.org and confirmed by the identical index data downloaded also in September 2021 from the IAGA-supported ISGI web service at (http://isgi.unistra.fr .

7. Summary of differences between the IAGA-endorsed and the present index methods

The definitions and the descriptions of the IAGA-endorsed PC index derivation methods are (in principle) available at Matzka (2014). The IAGA endorsement by Resolution #3 (2013) rely on the recommendations by the IAGA Task Force (Menvielle et al., 2013) that, in turn, refers to the publications Troshichev et al. (2006), Janzhura and Troshichev (2008), and Janzhura and Troshichev (2011). Troshichev et al. (2006) describes the general principles for PC index derivation that are also applied in the present work. The reference levels described there and in Janzhura and Troshichev (2008) are built from quiet samples only while the reference level construction in Janzhura and Troshichev (2011), Troshichev (2011, 2017), or Troshichev and Janzhura (2012) implies adding a solar sector term derived from smoothed daily median values.

Specific issues where the methods defined in the present work differ from the IAGA-endorsed methods comprise:

(1) The present work based on "DMI2016" methods from Stauning (2016) attempts to avoid DP3 (NBZ) convection samples in the calculations of scaling parameters by the requirement IMF $B_Z < |B_Y| + 3$ nT on the solar wind conditions and $\Delta F_{PROJ} > \Delta F_{MIN}$ (-2 mV/m) (~-50 nT) on the ground-

based data. The “IAGA2014” methods described in Matzka (2014) and Nielsen and Willer (2019) do not discriminate between DP2 and DP3 convection cases but use all available samples in the correlation and regression-based calculations of scaling parameters.

(2) With the method presented here, the magnetic data are inspected at the initial processing step looking for irregularities. In cases of irregularities the base levels are corrected to remove other than the regular secular variations. It has not been possible to obtain information on corresponding handling of the basic magnetic data in the IAGA-endorsed processing system.

(3) The reference levels defined here use the sum of the base line and the quiet day variation (QDC) derived from the quietest samples within –40 to +40 days intervals weighted with solar rotation phase and differences between QDC and sample dates (Stauning, 2011) while the reference levels in IAGA2014 use 30-days quiet samples with equal weights and additional solar sector terms derived as daily median values smoothed over 7 days (Janzhura and Trochichev, 2011); Troshichev and Janzhura, 2012; Matzka and Troshichev, 2014). The solar sector terms are not quiet but vary with the amplitudes of disturbances, particularly the IMF B_Y levels. Furthermore, using the daily median values, strongly affected by IMF B_Y-related effects at the noon Cusp region, in the reference levels applied throughout all hours (all observatory positions in their daily rotation) may generate unfounded index modifications at different hours (see comments by Stauning, 2013a, 2015, 2020, 2021a).

(4) The correlation and regression calculations defined here use in each step all available 5-min samples from each hour of each calendar month throughout the 12 years calibration epoch to derive mean hourly values for each calendar month (i.e., 12 s/h*30 d/m*12 y/epoch=4320 5-min samples) to derive each element of the arrays of hourly values for each calendar month (288 values of optimum angles, slopes or intercepts. These values are interpolated to provide the desired time resolution.

For IAGA2014, each step uses all 5-min samples for each 5-min moment of each month of each year (i.e., 30 samples) to calculate each element of the calibration parameter arrays derived for each 5 min of each month of each year providing 288*12*12=41472 values of optimum angles, slopes or intercepts. These values are smoothed and interpolated to provide final values in the desired time resolution.

The total number of available 5-min samples throughout the reference epoch is the same (~1261440 5-min samples) and the correlation and regression results would be the same if fluctuations were absent or all interim processes were linear and unlimited. However, any smoothing whether “box”, “Gaussian”, or “lowess” - type used in the IAGA-endorsed index procedure (Matzka, 2014) may generate systematic differences in the end results. Thus, the question is whether interpolation among a few values derived from strictly linear processing of many samples or smoothing of many values from strictly linear processing of few samples is the better way to ensure proper statistical handling.

The differences between the two methods are particularly evident when the basic data material has large fluctuations such as the optimum angle values at local winter time. The differences between optimum angles derived by the different methodologies are discussed in Stauning (2021b). For the PCN versions the differences may amount to 20-30° while for the PCS versions, the differences may amount to 30-40° in the optimum angles notably at local winter where the initial values are most fluctuating. The problem is not mentioned in available documentation from the index providers (e.g., Troshichev, 2011, 2017; Matzka and Troshichev, 2014) and suggestions to look into this issue have been ignored.

Concluding remarks.

The polar cap indices provide potentially very useful assets for investigations of solar wind – magnetosphere interactions and geomagnetic disturbances and for monitoring of space weather conditions. The pioneering efforts by Dr. Troshichev and his colleagues (published, e.g. in Troshichev and Andrezen, 1985; Troshichev et al. 1988) are duly acknowledged. However, the derivation procedures and index calculations have been haunted by errors and mistakes.

In a number of publications (e.g., Stauning, 2013a,b, 2015, 2018a,b, 2020, and 2021b) the faulty PC index features have been criticized and close examinations of the IAGA-recommended PC index versions have been suggested with little success apart from prompting the development of independent index versions (Stauning et al., 2006; Stauning, 2016, 2018c).

The submitted work (including the present SI file) provides coherent definitions and detailed descriptions of all steps involved in the generation of Polar Cap (PC) index scaling parameters and index values in their post-event and real-time versions.

It is disappointing that IAGA in spite of its "Criteria for endorsement of indices" (2009) upon endorsing the current "official" PC index versions by its Resolution #3 (2013) has failed to request comprehensive documentation of derivation procedures, proper validation of methods, and effective quality control of published index series supplied to the international scientific community.

Data availability:

Near real-time (prompt) PC index values and archived PCN and PCS index series derived by the IAGA-endorsed procedures are available through AARI and ISGI web sites. Archived PCN and PCS data used in the paper were downloaded from the "final" version link at https://pcindex.org and from http://isgi.unistra.fr in September 2021 unless otherwise noted.

Space data from the WIND, ACE, and GeoTail missions for deriving E_M and IMF B_Y values have been obtained from OMNIweb space data service at https://omniweb.gsfc.nasa.gov .

Geomagnetic data from Qaanaaq, Vostok and Dome-C were provided from the INTERMAGNET data service web portal at https://intermagnet.org .

The observatory in Qaanaaq is managed by the Danish Meteorological Institute, while the magnetometer there is operated by DTU Space, Denmark. The Vostok observatory is operated by the Arctic and Antarctic Research Institute in St. Petersburg, Russia. The Dome-C observatory is managed by Ecole et Observatoire des Sciences de la Terre (https://eost.unistra.fr) (France) and Istituto Nazionale di Geofisica e Vulcanologia (https://ingv.it) (Italy).

The "DMI2016" PC index version is documented in the report DMI SR-16-22 (Stauning, 2016) available at the web site: https://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2016/SR-16-22-PCindex.pdf

Acknowledgments. The staffs at the observatories in Qaanaaq (Thule), Vostok, and Concordia and their supporting institutes are gratefully acknowledged for providing high-quality geomagnetic data for this study. The space data contributions managed through OMNIweb data center from the ACE, GeoTail, and WIND spacecraft missions are gratefully acknowledged. The efficient provision of geomagnetic data from the INTERMAGNET data service centre, and the excellent performance of the PC index portals are greatly appreciated. The author gratefully acknowledges the collaboration and many rewarding discussions in the past with Drs. O. A. Troshichev and A. S. Janzhura at the Arctic and Antarctic Research Institute in St. Petersburg, Russia.

References

Chambodut A, Di Mauro D, Schott JJ, Bordais P, Agnoletto L, Di Felice P. (2009). Three years continous record of the Earth's magnetic field at Concordia station (Dome-C, Antarctica). *Annals of Geophysics* **52** 15-24. https://doi.org/10.4401/ag-4569 .

Di Mauro D, Cafarella L, Lepidi S, Pietrolungo M, Alfonsi L, Chambodut A. (2014). Geomagnetic polar observatories: the role of Concordia station at Dome C, Antarctica, *Annals of Geophysics* **57**, 6, G0656. https://doi.org/10.4401/ag-6605 .

IAGA Resolution no. 3 (2013). https://www.iaga-aiga.org/resolutions.

IAGA Index endorsement criteria (2009): https://www.ngdc.noaa.gov/IAGA/vdat/GeomagneticIndices/Criteria_for_Endorsement.pdf

Janzhura, A.S. and Troshichev, O.A. (2008). Determination of the running quiet daily geomagnetic variation, J Atmos Solar-Terr Phys 70, 962–972, https://doi.org/10.1016/j.jastp.2007.11.004

Janzhura, A.S. and Troshichev, O.A. (2011). Identification of the IMF sector structure in near-real time by ground magnetic data. *Annales Geophysicae* **29** 1491-1500. https://doi.org/10.5194/angeo-29-1491-2011 .

Kan, J. R. & Lee, L. C. (1979). Energy coupling function and solar wind-magnetosphere dynamo. *Geophysical Research Letter*, 6 (7)**,** 577- 580. https://doi.org/10.1029/GL006i007p00577 .

Matzka, J. (2014). PC index description and software. https://doi.org/10.11581/DTU:00000057

Matzka, J. and Troshichev, O.A. (2014). PC index description main document and references. http://isgi.unistra.fr/Documents/References/PC_index_description_main_document.pdf

Menvielle, M., McCready, H., Demetrescu, C., 2013**.** Recommendation by the Task Force: Fully recommend endorsement of the PC index. IAGA_documentation_20130225.pdf.

Nielsen, J. B. & Willer, A. N. (2019). Restructuring and harmonizing the code used to calculate the Definitive Polar Cap Index. *Report from DTU Space*. https://tinyurl.com/sx3g5t5

Stauning, P. **(**2007). A new index for the interplanetary merging electric field and the global geomagnetic activity: Application of the unified Polar Cap (PCN and PCS) indices. *AGU Space Weather*, *5*, S09001. https://doi.org/10.1029/2007SW000311.

Stauning, P. (2011). Determination of the quiet daily geomagnetic variations for polar regions. *J. Atmos Solar-Terr Phys*, *73(16)*, 2314-2330. https://doi.org/10.1016/j.jastp.2011.07.004 .

Stauning, P. (2012). The Polar Cap PC Indices: Relations to Solar Wind and Global Disturbances, Exploring the Solar Wind, Marian Lazar (Ed.). *InTech Publ*. ISBN: 978-953-51-0339-4. https://doi.org/10.5772/37359

Stauning, P. (2013a). Comments on quiet daily variation derivation in "Identification of the IMF sector structure in near-real time by ground magnetic data" by Janzhura and Troshichev (2011). *Annales Geophysicae*, 31, 1221-1225. https://doi.org/10.5194/angeo-31-1221-2013 .

Stauning, P. (2013b). The Polar Cap index: A critical review of methods and a new approach, *J. Geophys. Res. Space Physics*, 118, 5021-5038. https://doi.org/10.1002/jgra.50462

Stauning, P. (2015). A critical note on the IAGA-endorsed Polar Cap index procedure: effects of solar wind sector structure and reverse polar convection. *Annales Geophysicae*, *33***,** 1443-1455. https://doi.org/10.5194/angeo-33-1443-2015 .

Stauning, P. (2016). The Polar Cap (PC) Index.: Derivation Procedures and Quality Control. *DMI Scientific Report SR-16-22*. Available at: https://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2016/SR-16-22-PCindex.pdf .

Stauning, P. (2018a). A critical note on the IAGA-endorsed Polar Cap (PC) indices: excessive excursions in the real-time index values. *Ann Geophys 36*, 621–631. https://doi.org/10.5194/angeo-36-621-2018 .

Stauning, P. (2018b). Multi-station basis for Polar Cap (PC) indices: ensuring credibility and operational reliability. *J Space Weather Space Clim 8*, A07. https://doi.org/10.1051/swsc/2017036 .

Stauning, P. (2018c). Reliable Polar Cap (PC) indices for space weather monitoring and forecast, *J Space Weather Space Clim 8***,** A49. https://doi.org/10.1051/swsc/2018031

Stauning, P. (2020). The Polar Cap (PC) index: invalid index series and a different approach. *Space Weather* 16, e2020SW002442. https://doi.org/10.1029/2020SW002442

Stauning, P. (2021a). Comment on "Identification of the IMF sector structure in near-real time by ground magnetic data" by Janzhura and Troshichev (2011). *Ann. Geophys*., 39, 369–377. https://doi.org/10.5194/angeo-39-369-2021

Stauning, P., 2021b. Invalid Polar Cap (PC) indices: Erroneous scaling parameters, J. Geophys. Res. Space Phys. https://doi.org/10.1029/2020SW002442 .

Stauning, P. (2021c). The Polar Cap (PC) index combination, PCC: relations to solar wind properties and global magnetic disturbances. *J Space Weather Space Clim* https://doi.org/10.1051/swsc/2020074

Stauning, P., 2021d. Transpolar convection and magnetospheric ring currents: real-time applications of Polar Cap (PC) indices. Space Weather. https://doi.org/10.1029/2020SW002072

Stauning, P., Troshichev, O. A., & Janzhura, A. S. (2006). Polar Cap (PC) index. Unified PC-N (North) index procedures and quality. *DMI Scientific Report, SR-06-04*. (revised 2007 version available at https://www.dmi.dk/fileadmin/Rapporter/SR/sr06-04.pdf).

Stauning, P., Troshichev, O. A., & Janzhura, A. S. (2008). The Polar Cap (PC) index: Relations to solar wind parameters and global activity levels, *J Atmos Solar-Terr Phys 70(18)*, 2246-2261. https://doi.org/10.1016/j.jastp.2008.09.028 .

Troshichev, O. A. (2011). Polar Cap (PC) Index. Available at: https://pcindex.org (see Supported materials).

Troshichev, OA (2017). Polar cap magnetic activity (PC index) and space weather monitoring, *Édition universitaires européennes*, ISBN: 978-3-8381-8012-0.

Troshichev, O.A., Andrezen, V.G., 1985. The relationship between interplanetary quantities and magnetic activity in the southern polar cap, Planet. Space Sci., 33, 415-419.

Troshichev, O. A. and Janzhura, A. S. (2012): Space Weather monitoring by ground–based means, Springer Praxis Books, Heidelberg, ISBN 978–3–642–16802–4, https://doi.org/10.1007/978–3–642–16803–1 .

Troshichev, O. A., Andrezen, V. G., Vennerstrøm, S. & Friis-Christensen, E. (1988). Magnetic activity in the polar cap – A new index. *Journal of Planetary and Space Sciences 36(11)*, 1095-1102. https://doi.org/10.1016/0032-0633(88)90063-3

Troshichev, O. A., Janzhura, A. S., Stauning, P. (2006): Unified PCN and PCS indices: method of calculation, physical sense and dependence on the IMF azimuthal and northward components, J. Geophys. Res., 111, A05208, https://doi.org/10.1029/2005JA011402, 2006.

Troshichev, O. A., Janzhura, A. S., Stauning, P. (2009): Correction to "Unified PCN and PCS indices: Method of calculation, physical sense, and dependence on the IMF azimuthal and northward components, J. Geophys. Res., 111, A05208", J. Geophys. Res. 114, A11202, https://doi.org/10.1029/2009JA014937, 2009.

Vennerstrøm, S. (1991). The geomagnetic activity index PC, PhD Thesis, Scientific Report 91-3, Danish Meteorological Institute, 105 pp. https://www.dmi.dk/fileadmin/user_upload/Rapporter/SR/1991/sr91-3.pdf

30 September 2021

Peter Stauning. pst@dmi.dk