EGU General Assembly 2020

Extraction of the daily quiet variation from the geomagnetic field observations with the principal component analysis

Anna Morozova, Rania Rebbah

[1] Centre for Earth and Space Research of University of Coimbra, Department of Physics, University of Coimbra, 3000-456 Coimbra, Portugal
[2] Department of Physics, University of Coimbra, 3000-456 Coimbra, Portugal

Abstract

- Geomagnetic field (GMF) variations from external sources are classified as regular (diurnal) or occurring during periods of disturbances.
- The most significant regular variations are the quiet solar daily variation (Sq) and the disturbance daily variation (S_D).
- These variations have well recognized daily cycles and need to be accounted for before the analysis of the disturbed field.
- Preliminary analysis of the GMF variations shows that the principal component analysis (PCA) is a useful tool for extraction of regular variations of GMF; however the requirements to the data set length, geomagnetic activity level etc. need to be established.
- Here we present preliminary results of the PCA-based Sq and S_D extraction procedure based on the analysis of the Coimbra Geomagnetic Observatory (COI) measurements of the geomagnetic field components H, X, Y and Z between 2007 and 2017.

Data

- H, X, Y and Z components of the geomagnetic field
- Measured at the Coimbra Geomagnetic Observatory (COI), Portugal
  - 40° 13´ N, 8° 25´ W, 99 m asl
- Hourly series
- Only December months from 2007 to 2017

- The month-long hourly series of each component was analyzed
  - for individual month of each of 11 years
  - for all 11 years together

- Due to the location of the COI observatory H ≈ X (since D ≈ -4°)

- COI is located near or slightly north to the mean Sq vortex focus position for European sector (≤ 40° N) (*e.g., Yamazaki and Maute, 2017*)

Methods for Sq & S_D extraction

1. Standard approach using quietest days of a month

2. Principal component analysis (PCA)

- Correlation analysis
  - Similarities between series were analyzed using the correlation coefficients (r) and their statistical significances (p value)
  - Statistical significance (p value) was estimated using the Monte Carlo approach with artificial series constructed by the “phase randomization procedure” (*Ebisuzaki, 1997*).

Method 1:
Sq & S_D – standard approach

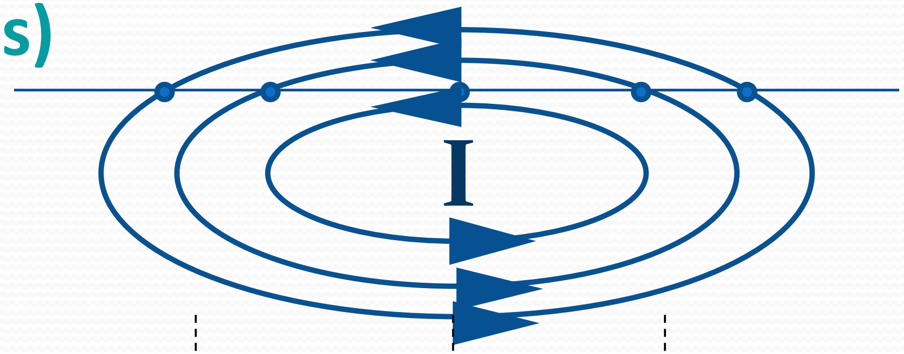
- "daily quiet" (Sq):
  - calculated as the mean daily variation of the 5 most quiet days of a month
    - international quiet days – IQD, estimated by the GFZ-Potsdam from Kp
    - local quiet days – LQD, estimated from the local K-index
  - ionospheric origin
    - Source: electric current vortex in the sunlit hemisphere
  - contamination from magnetospheric currents (mostly in polar regions)

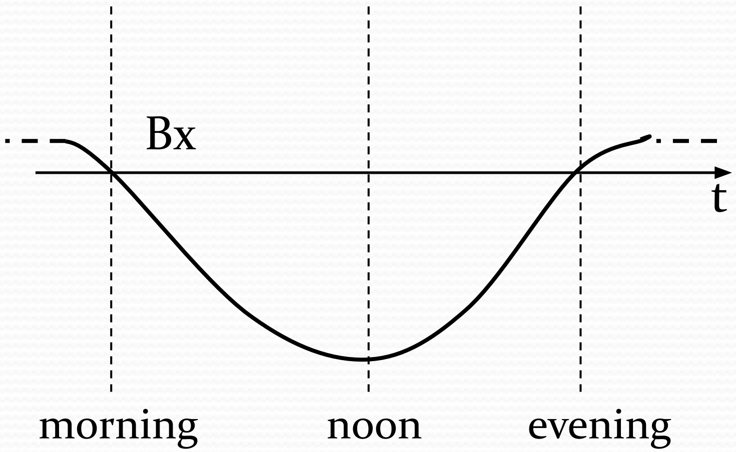
- "daily disturbed" (S_D):
  - calculated as the mean daily variation of all days of the month (S) minus Sq
  - the name comes from the similarity of shapes of the S_D and Dst variations
  - magnetospheric origin

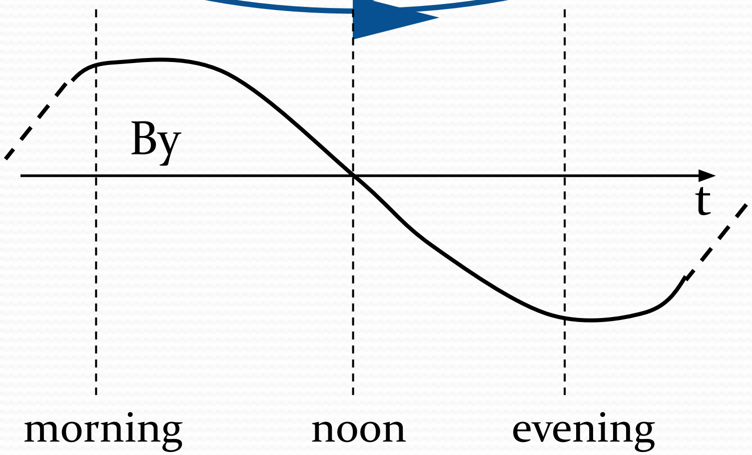
Method 1:
Problems of the standard approach

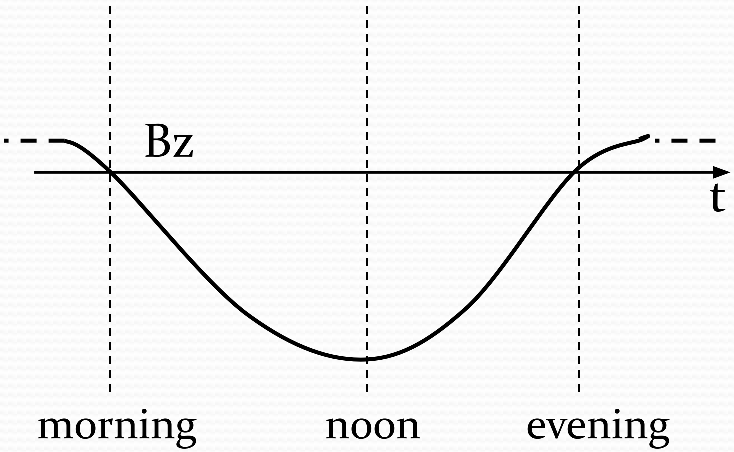
- IQD are days that are only **relatively quiet** comparing to others days of a month
- They can be disturbed on the **absolute** scale
- Final IQD definition is lagged by 1-2 yr
- Observations for certain IQD day at a particular observatory can be missing
- There is a single curve for all days of a month without accounting for variability
  - in the ionosphere and magnetosphere,
  - for the position of the Sq-generating vortex
  - for the shape of the Sq-generating vortex
- A number of studies *(Xu and Kamide, 2004; Chen et al., 2007; Yamazaki et al., 2016*) showed the need for methods of Sq (and S_D) extraction which take into account day-to-day variability of the ionospheric conditions.

Sq “ideal” shape for a mid-latitudinal station
(north of the Sq vortex focus)



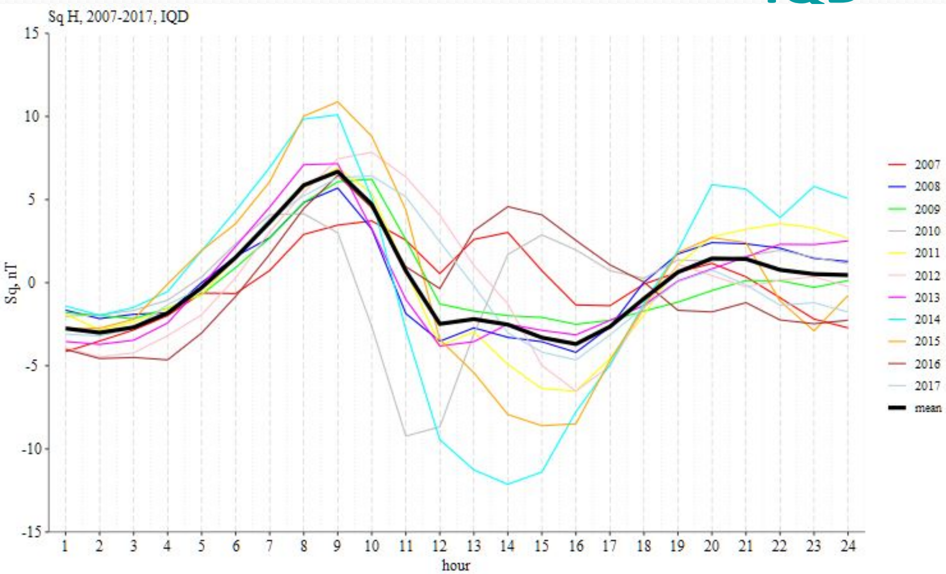


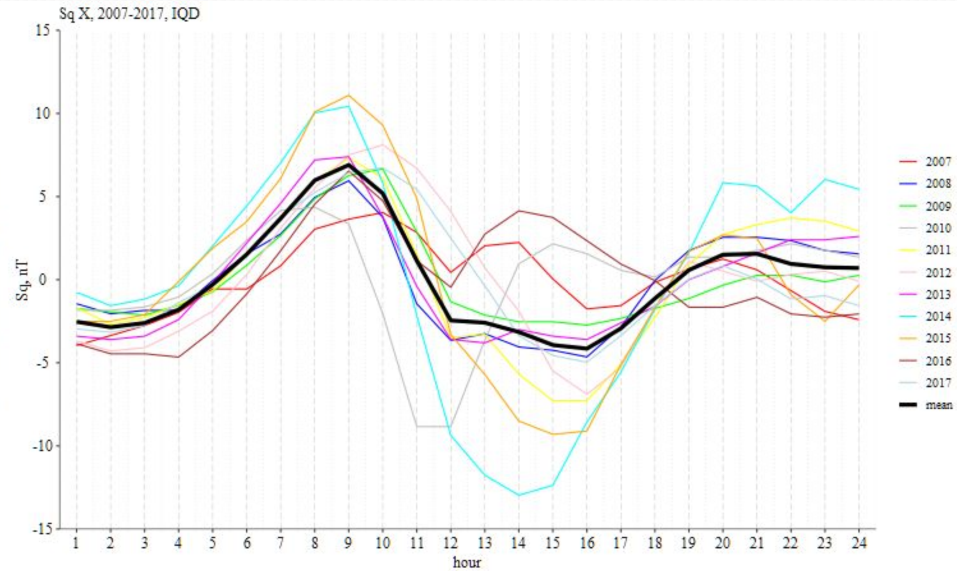


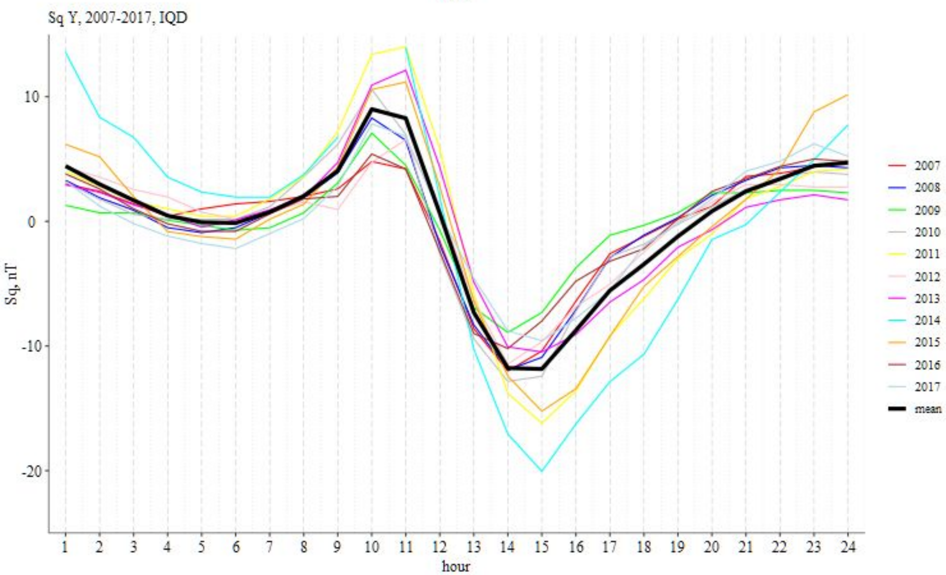


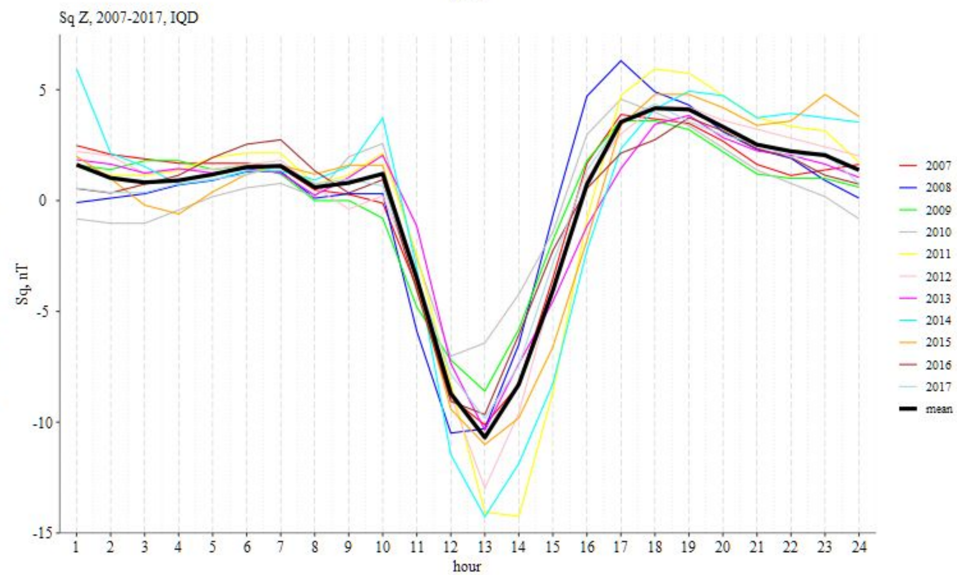
adapted from
Chapman and Bartels (1940)

COI data: Sq_{IQD} – individual December series





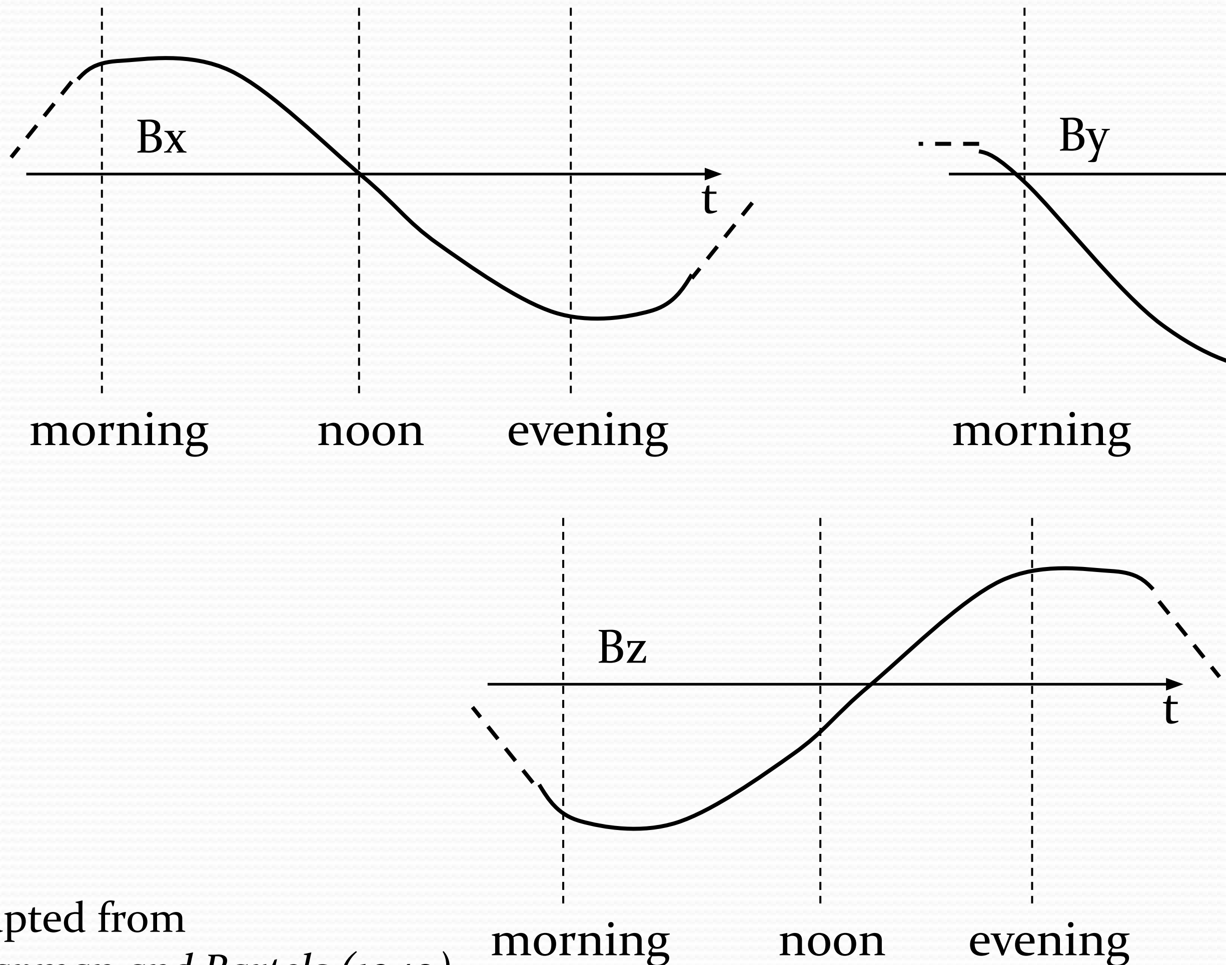




COI data: Sq_{IQD} – individual December series

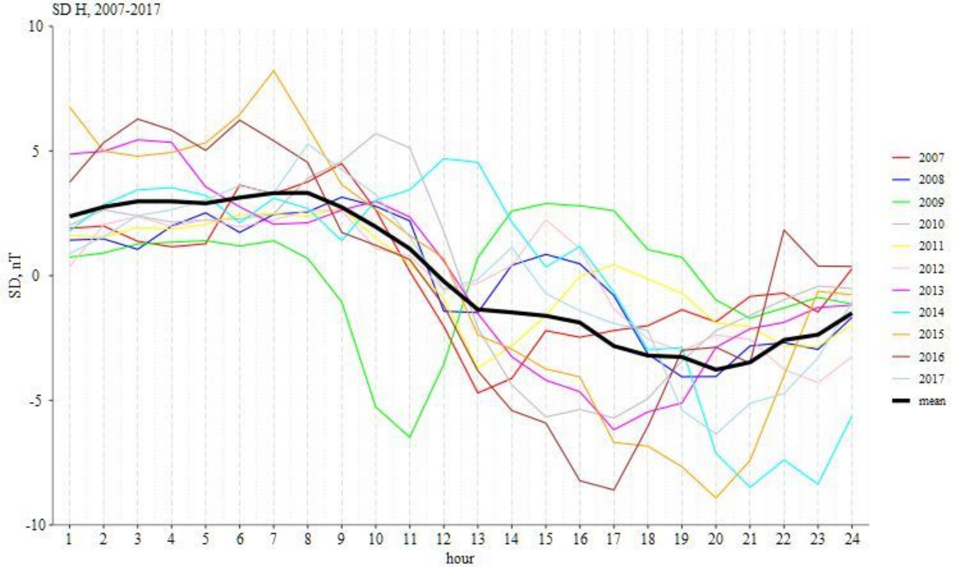
- H ≈ X components
  - Mean Sq is far from the “ideal Sq” for a station located north of the Sq vortex focus, i.e.
    - either there is contamination by disturbances
    - or for most of these IQD days COI was located near the Sq vortex centre
  - High months-to-month variability of the Sq_{IQD} shape:
    - the shapes of Sq_{IQD} for December of 2010, 2011, 2014, 2015 are similar to the “ideal Sq”
    - the shapes of Sq_{IQD} for December of 2008, 2012, 2013, 2017 are close to the “ideal Sq”
    - the shapes of Sq_{IQD} for December of 2007, 2009, 2016 are strongly affected by disturbances/Sq vortex shape and position
- Y, Z components
  - Both mean Sq and Sq for individual months are similar to the “ideal Sq”
  - Low month-to-month variability of the Sq_{IQD} shape

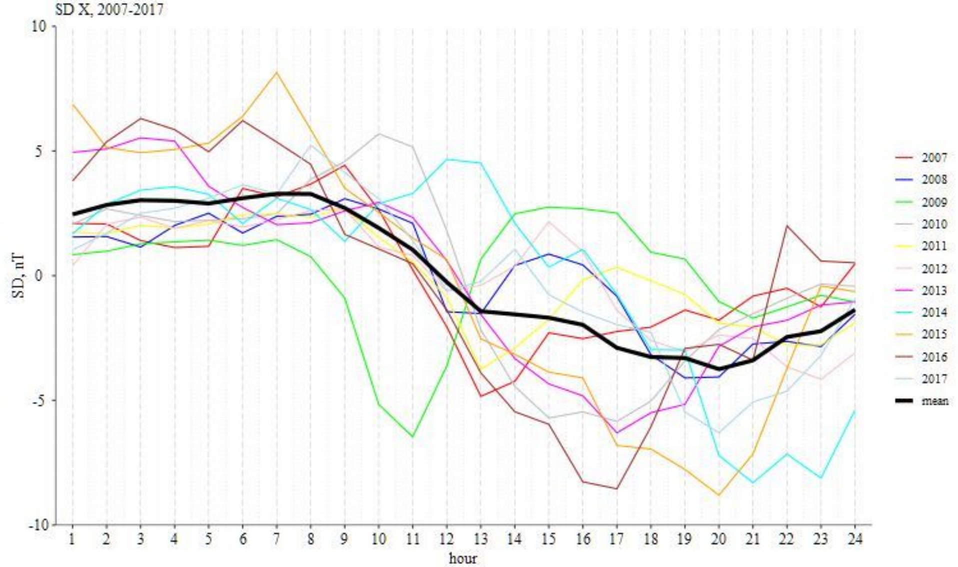
S_D "ideal" shape for a mid-latitudinal station

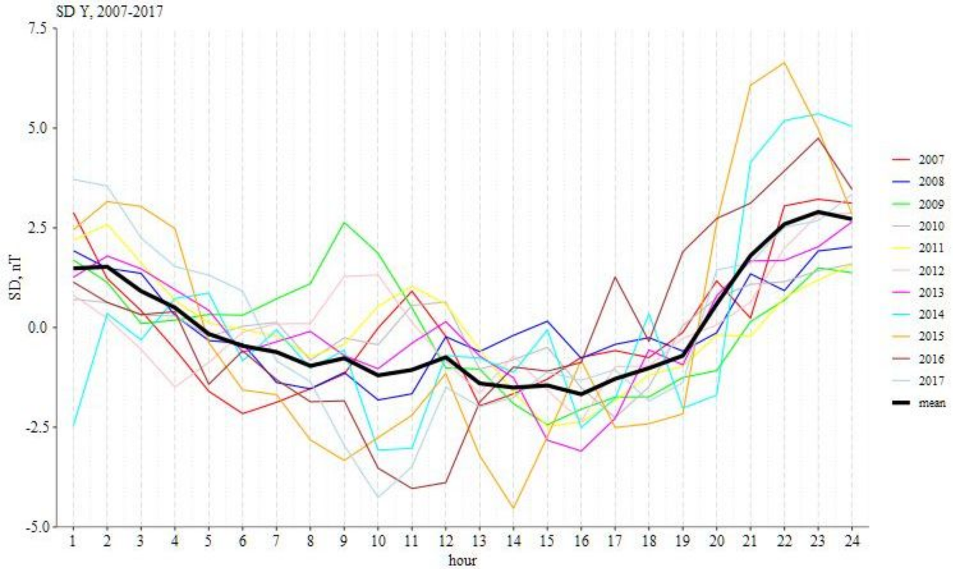


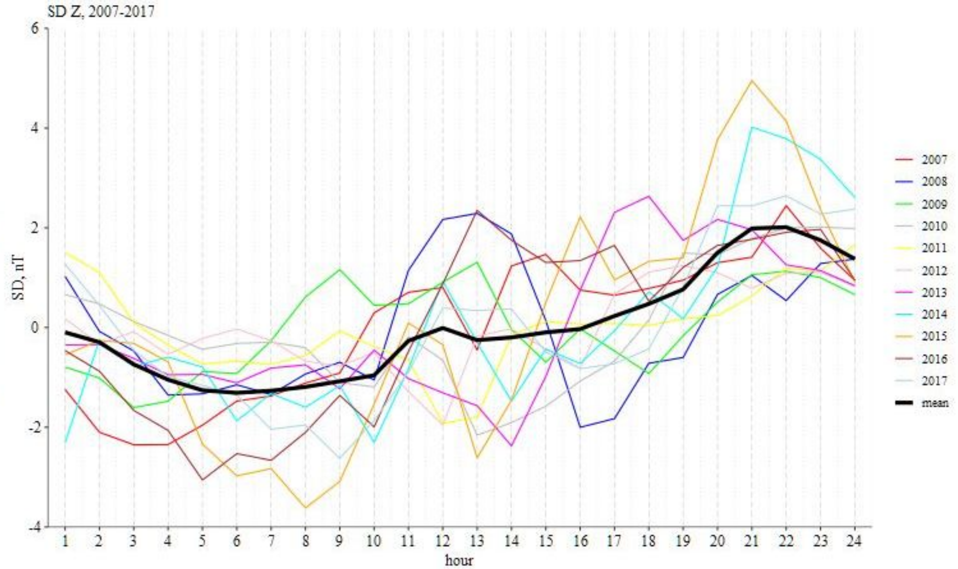
adapted from
Chapman and Bartels (1940)

COI data: S_D – individual December series









COI data: S_D – individual December series

- H ≈ X components
  - Mean S_D is similar to the “ideal S_D”
  - The shapes of S_D for individual months can deviate from the “ideal S_D”, sometimes significantly (e.g., December 2007)
  - High month-to-month variability of the S_D shape

- Y, Z components
  - Mean S_D are similar to the “ideal S_D”
  - The shapes of S_D for individual months can deviate from the “ideal S_D” shape
  - Moderate month-to-month variability of the S_D shape

Method 2:
Principal components analysis (PCA)

- Previous studies show (*Xu and Kamide, 2004; Chen et al., 2007*) that the principal component analysis (PCA) is a useful tool for the extraction of regular variations of GMF.

- PCA is a widely used method to extract independent modes of variability when a number of series of the same parameter of, e.g., different stations or days is used.

Principal components analysis (PCA)

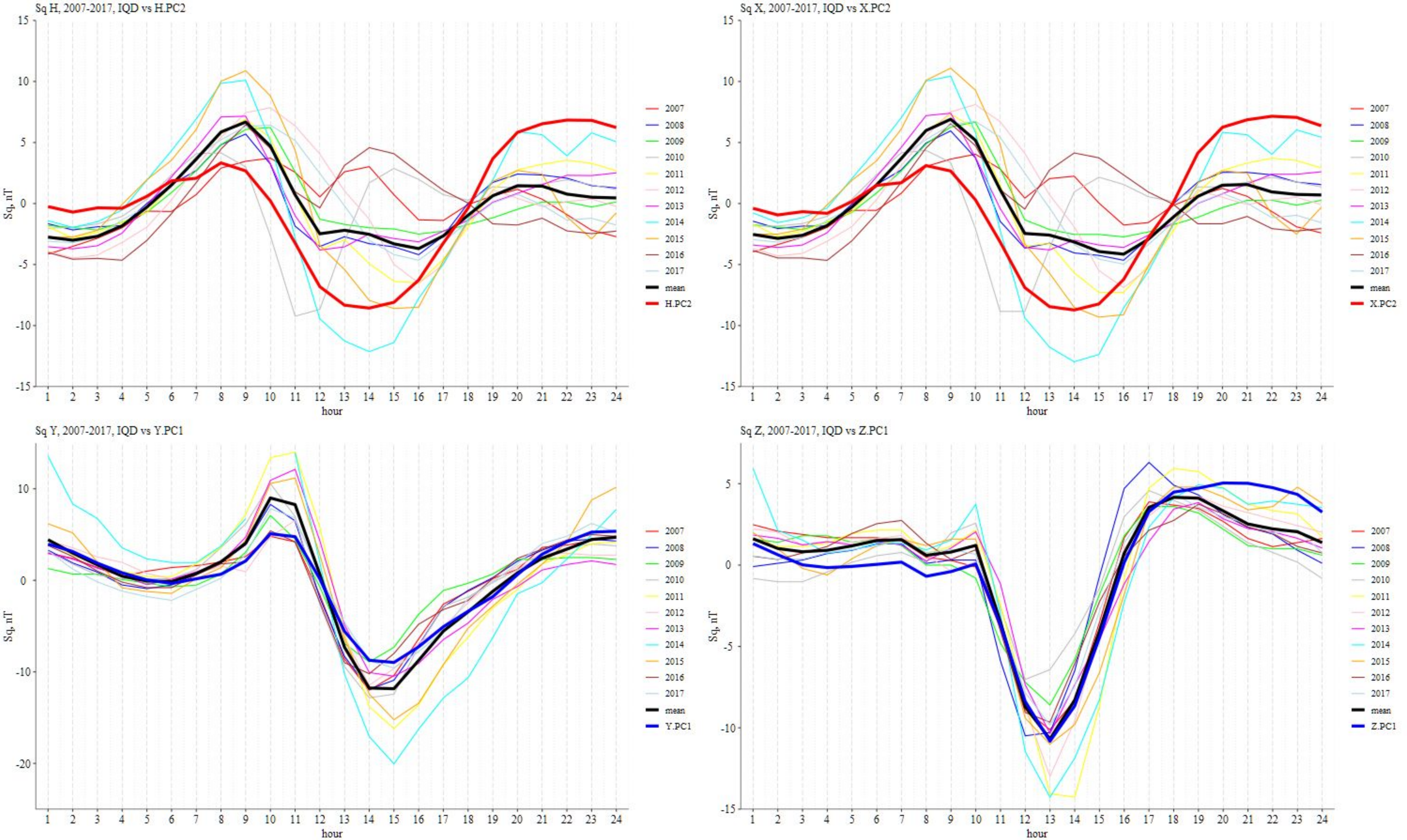
- Input data ⇒ covariance matrix ⇒ eigenvalues & eigenvectors.
- Eigenvalues ⇒ explained variances of the extracted modes
- Eigenvectors ⇒ principal component (PC) & empirical orthogonal function (EOF).
- PCs = daily variations of different types
- EOFs = amplitudes of daily variations (PCs) for each of the analyzed days
- PC# & EOF# ⇒ mode#

- PCA input matrix for COI data:
  - each column contains 24 observations (every 1 h)
  - number of columns:
    - 31 for an individual December (PCA for an individual month)
    - 31*11 for all 11 Decembers together (PCA for 2007-2017)

PCA results:
Sq
all Decembers 2007-2017

- Each of the following plots shows
  - Sq_{IQD} calculated for each of 11 Decembers – colored thin lines
  - Sq_{IQD} calculated for December of all 11 year – black thick line
  - Sq_{PCA}: PC1 (Y & Z) and PC2 (H & X) obtained for the whole data set (11 years) – blue and red thick lines, respectively

COI PCA: Sq_{IQD} vs Sq_{PCA} – all Decembers 2007-2017



PCA results: explained variances

all Decembers 2007-2017

Components	PC1	Identified as…	PC2	Identified as…
H	54%	S_D	18%	Sq
X	54%	S_D	19%	Sq
Y	67%	Sq	12%	S_D?
Z	71%	Sq	10%	S_D?

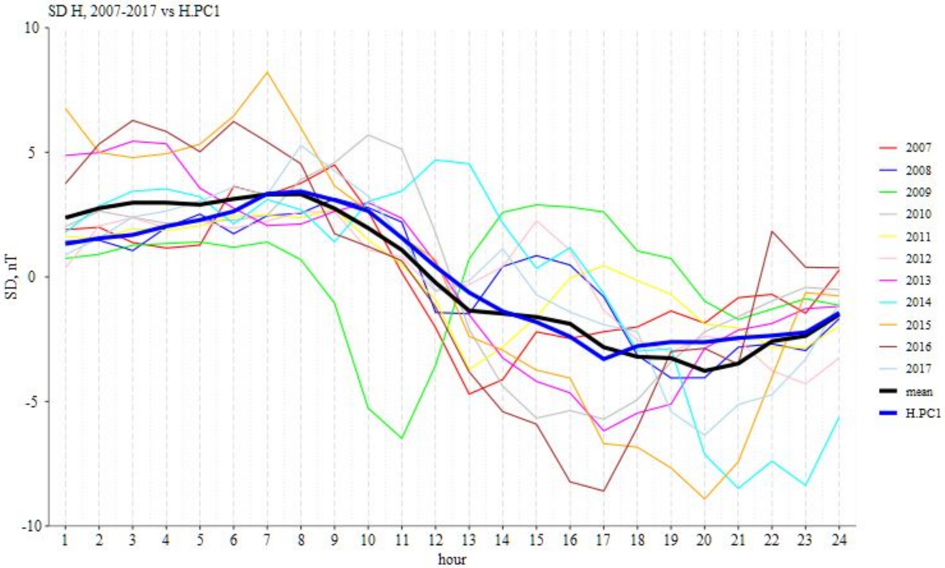
COI PCA: Sq_{PCA} - all Decembers 2007-2017

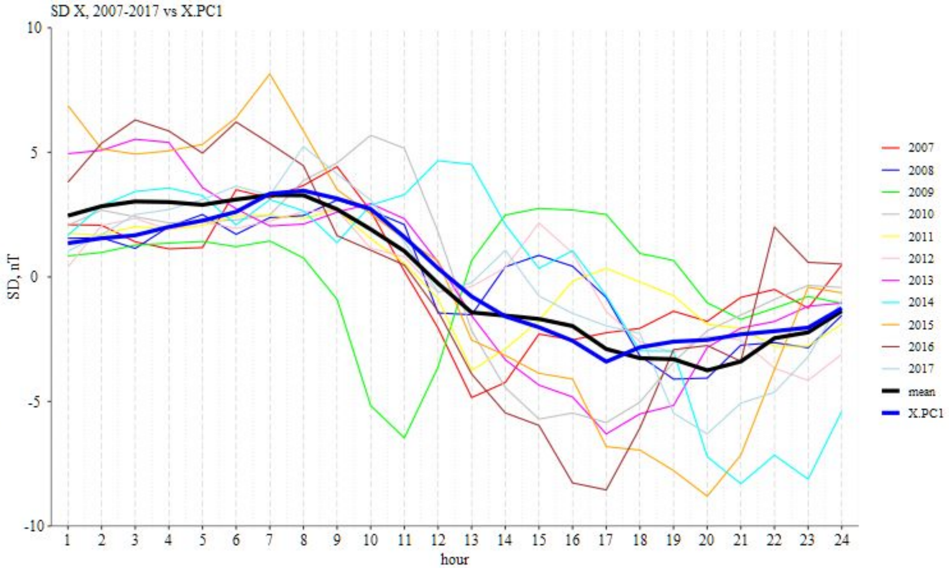
- H ≈ X components
  - Sq_{PCA} is identified as PC2 and is similar to the “ideal Sq” without notable contamination by disturbances
  - $Sq_{PCA} \neq Sq_{IQD}$
- Y, Z components
  - Sq_{PCA} is identified as PC1 and is similar to the “ideal Sq” without notable contamination by disturbances
  - $Sq_{PCA} = Sq_{IQD}$

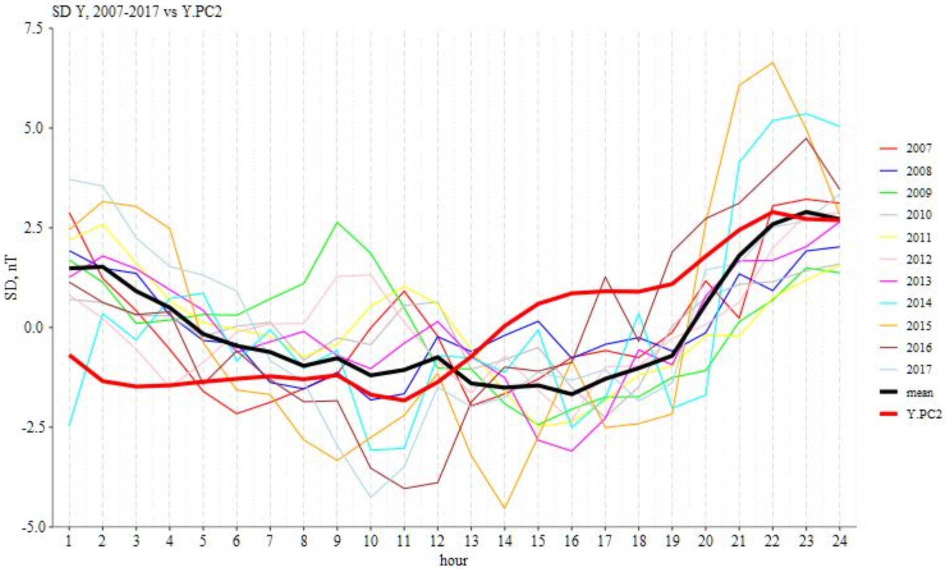
PCA results:
S_D
all Decembers 2007-2017

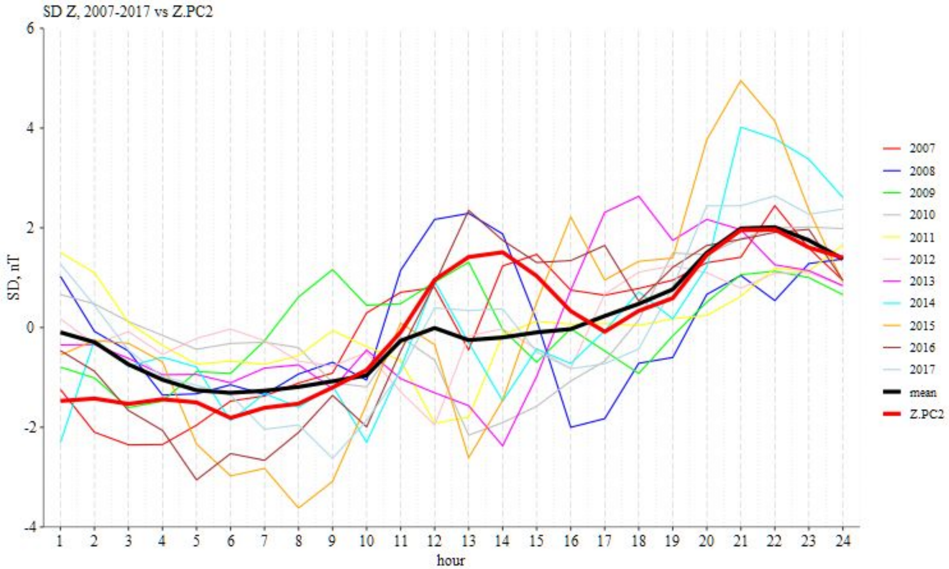
- Each of the following plots shows
  - $S_{D\ IQD}$ calculated for each of 11 Decembers – colored thin lines
  - $S_{D\ IQD}$ calculated for December of all 11 year – black thick line
  - $S_{D\ PCA}$: PC2 (Y & Z) and PC1 (H & X) obtained for the whole data set (11 years) – red and blue thick lines, respectively

COI PCA : S_D vs $S_{D\ PCA}$ - all Decembers 2007-2017









COI PCA : $S_{D\ PCA}$ - all Decembers 2007-2017

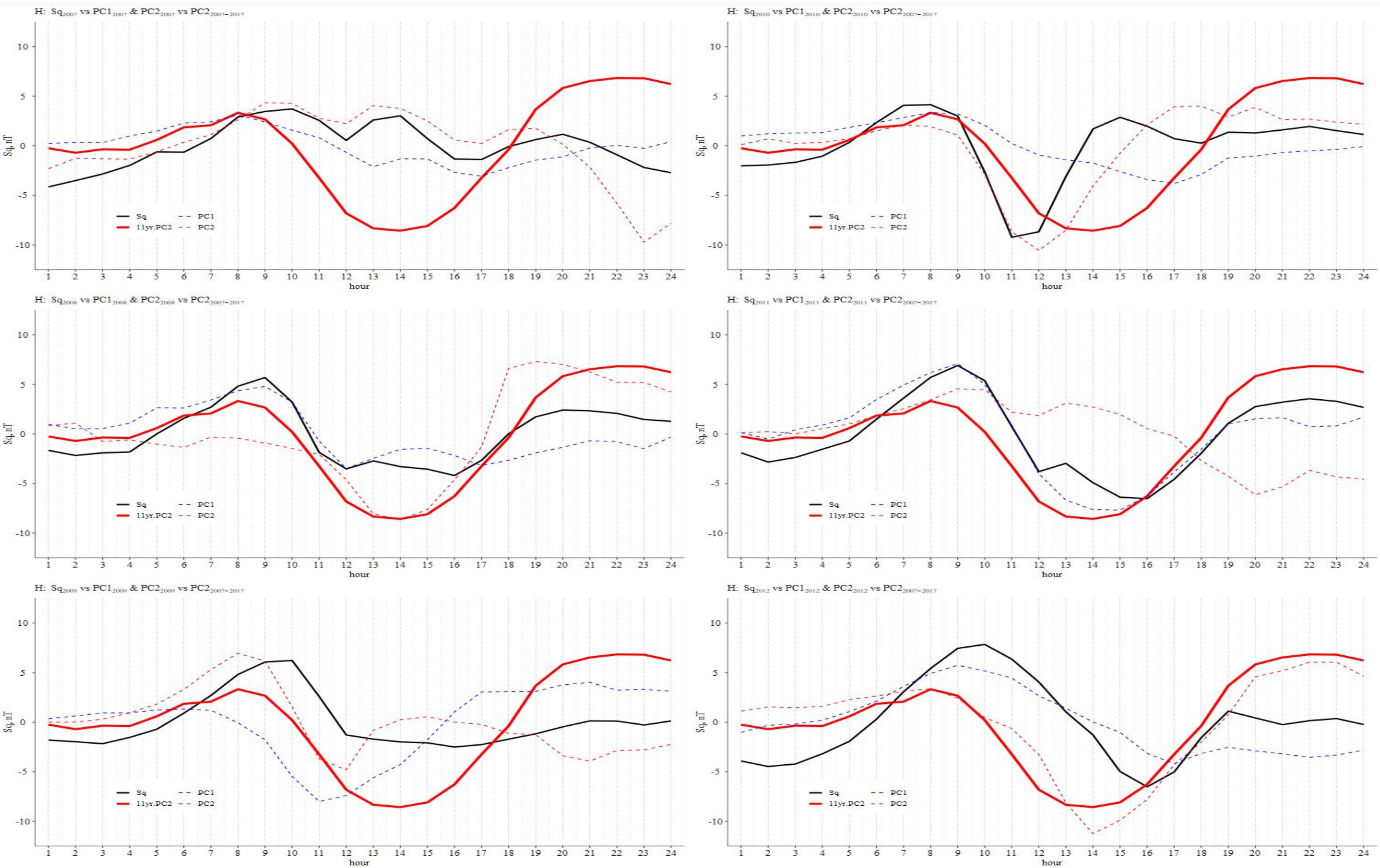
- H ≈ X components
  - $S_{D\ PCA}$ is identified as PC1 and is similar to the "ideal S_D"
  - $S_{D\ PCA} \approx S_D$
- Y, Z components
  - $S_{D\ PCA}$ is identified as PC2 and is similar to the "ideal S_D"
  - $S_{D\ PCA}$ is similar to S_D

COI PCA: Sq_{IQD} vs Sq_{PCA} individual Decembers 2007-2012 vs all Decembers (only H component)

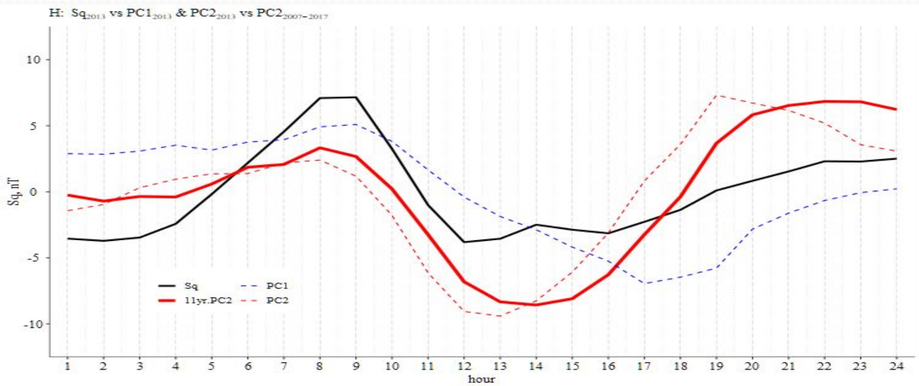
- To test the effect of the data set length on the quality of the PCA-based method of the Sq extraction we applied PCA to the 1-month long data sets of 11 individual Decembers (H component only).

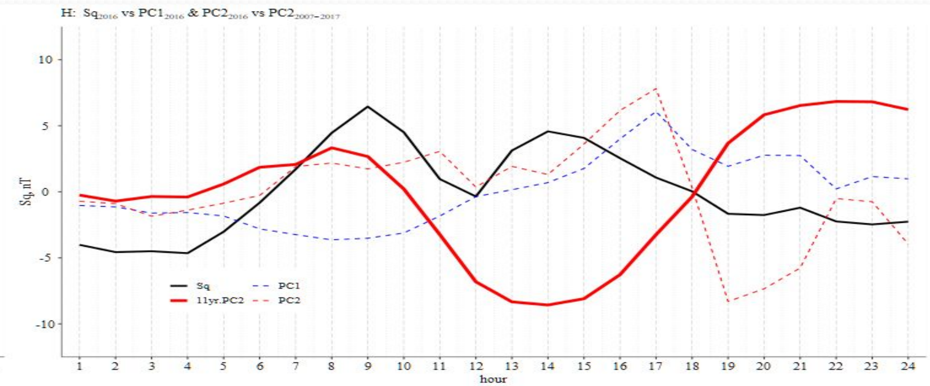
- Each of the following plots shows
  - Sq_{IQD} calculated for December of this year – black thick line
  - PC2 obtained on the whole data set (11 years) – red thick line
  - PC1 and PC2 obtained for this particular month – blue and red dashed lines

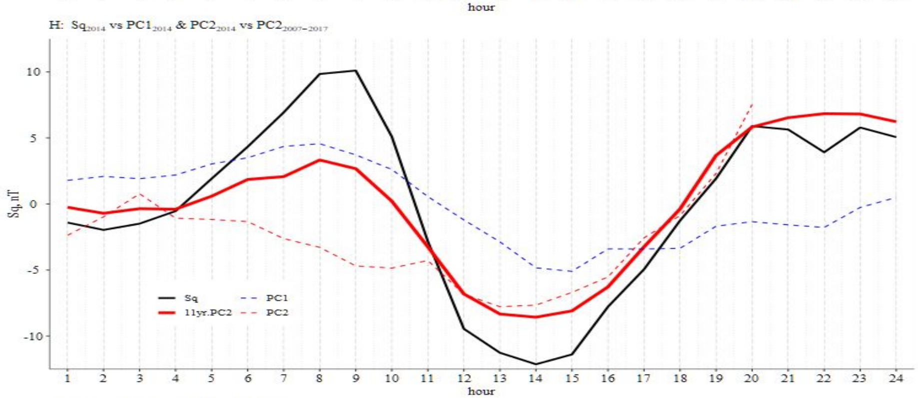
COI PCA: Sq_{IQD} vs Sq_{PCA}
individual Decembers 2007-2012 vs all Decembers

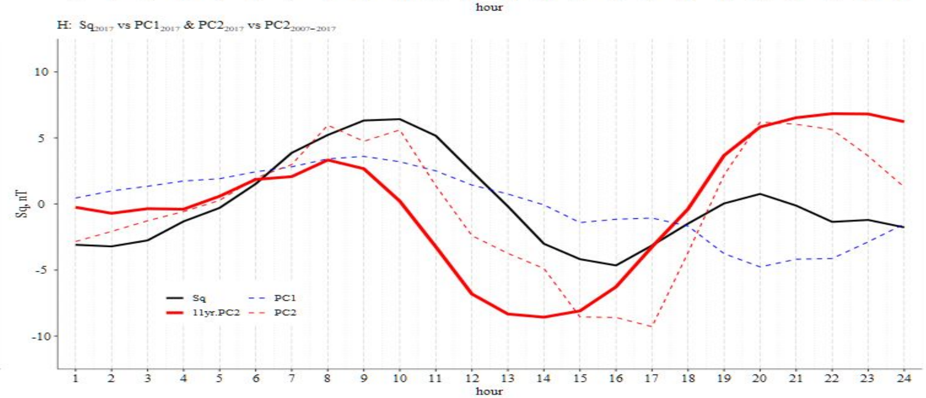


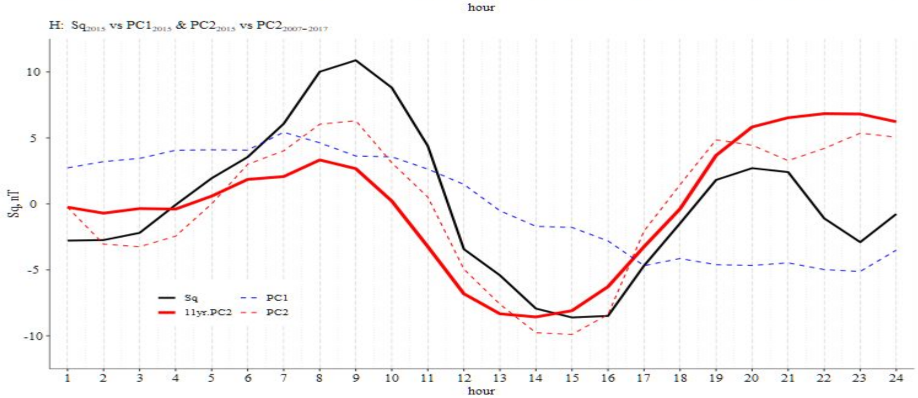
COI PCA: Sq_{IQD} vs Sq_{PCA}
individual Decembers 2013-2017 vs all Decembers











PCA results: explained variances

individual Decembers from 2007 to 2017

Time interval	PC1	Identified as…	PC2	Identified as…
December 2007	47%	S_D?	21%	?
December 2008	49%	S_D	**17%**	**Sq**
December 2009	**39%**	**Sq??**	28%	?
December 2010	56%	S_D	**21%**	**Sq**
December 2011	**54%**	**Sq**	17%	S_D?
December 2012	54%	S_D	**22%**	**Sq**
December 2013	57%	S_D	**23%**	**Sq**
December 2014	46%	S_D	**25%**	**Sq??**
December 2015	78%	S_D?	**11%**	**Sq**
December 2016	59%	?	12%	?
December 2017	55%	S_D?	**18%**	**Sq**
Decembers 2007-2017	54%	S_D	**18%**	**Sq**

COI PCA: Sq_{PCA} & $S_{D\ PCA}$ - individual Decembers

- For 9 out of 11 analyzed individual months PCA extract daily variation that can be identified as Sq
  - For 7 out of 11 analyzed months Sq_{PCA} is identified as PC2
  - For 2 out of 11 analyzed months Sq_{PCA} is identified as PC1

- For 9 out of 11 analyzed individual months PCA extract daily variation that can be identified as S_D
  - For 8 out of 11 analyzed months $S_{D\ PCA}$ is identified as PC1
  - For 1 out of 11 analyzed months $S_{D\ PCA}$ is identified as PC2

Sq_{IQD} vs Sq_{PCA} individual Decembers 2013-2017 vs all Decembers

- To compare IQD-based and PCA-based Sq curves for individual Decembers and for the whole data set we calculated correlation coefficients between:
  - Sq_{IQD} for individual December and PCs obtained for the whole data set ($PC2_{11}$)
  - Sq_{IQD} for an individual December (PCi_1: $PC1_1$ or $PC2_1$)
  - $PC2_{11}$ and PCi_1 which is identified as Sq

- In the following Table the values in parentheses are p-values.
- Only p-values < 0.15 are shown.

Correlation coefficients

Sq_{IQD} vs Sq_{PCA} , individual Decembers vs all Decembers

Time interval	Sq_{IQD} vs $PC2_{11}$	Sq_{IQD} vs PCi_1	i	$PC2_{11}$ vs PCi_1	i
December 2007	0.21	0.72 (0.08)	**2**	0.60	**2**
December 2008	**0.80 (0.04)**	**0.65 (0.003)** / 0.56	**1** / 2	0.87 (0.04)	**2**
December 2009	0.37	0.52	**2**	0.67	1
December 2010	0.39	0.79 (0.002)	**2**	0.64	**2**
December 2011	**0.80 (0.07)**	0.90 (0.005)	1	0.76 (0.14)	1
December 2012	0.24	0.72 (0.05)	1	0.94 (0.02)	**2**
December 2013	**0.67 (0.12)**	0.52	**2**	0.90 (0.05)	**2**
December 2014	**0.90 (0.03)**	**0.69 (0.07)** / 0.54	**1** / 2	0.83 (0.04)	**2**
December 2015	**0.57**	0.47 **0.79 (0.1)**	1 / **2**	0.91 (0.002)	**2**
December 2016	0.39	0.55	**2**	0.61 (0.04)	**2**
December 2017	0.25	0.53 / **0.67**	1 / **2**	0.80 (0.07)	**2**

Sq_{IQD} vs Sq_{PCA} individual Decembers 2013-2017 vs all Decembers

- Sq_{IQD} is highly correlated with Sq_{PCA} for those years when its shape is very similar to the "ideal Sq" shape:
  - 2008, 2011,2013, 2014, 2015 (compare to slide # 9)
  - Exceptions: 2010 & 2017 – years when the time of the daily minimum is shifted to the earlier /later hours (respectfully) resulting in low correlation coefficients

- For 7 out of 11 analyzed months Sq_{IQD} is highly correlated with Sq_{PCA} = $PC2_1$ for this particular month
- For 9 out of 11 analyzed individual months $PC2_{11}$ is highly correlated with $PC2_1$

Conclusions

- Preliminary results show that PCA can be successfully used for extraction of the Sq and S_D variations from the observations of the geomagnetic field.
- We analyzed H, X, Y and Z components for December months measured at the Coimbra Geomagnetic Observatory (COI) from 2007 to 2017.
- The PCA-based Sq and S_D curves were compared with the standard ones obtained using 5 IQD per month.
- PCA was applied to data sets of different length:
  - either 1 month-long data set for one of the analyzed years
  - or data series for the same month but from all years combined together.

Conclusions

- For most of the analyzed years
  - PC1 was identified as
    - S_D variation for H and X components and
    - Sq variations for Y and Z components.
  - PC2 was identified as
    - Sq variation for H and X components
    - S_D variations for Y and Z components.
- The PCA of the longer series (data for the same month but from different years combined together) produces more reliable results.

References

- Ebisuzaki, W. (1997), A method to estimate the statistical significance of a correlation when the data are serially correlated, J. Clim., 10 (9), 2147-2153.
- Chapman, S. and Bartels, J. (1940) Geomagnetism. Oxford University Press, Oxford
- Chen, G.X., Xu, W.Y., Du, A.M., Wu, Y.Y., Chen, B. and Liu, X.C. (2007): Statistical characteristics of the day-to-day variability in the geomagnetic Sq field. Journal of Geophysical Research: Space Physics, 112, A6, doi:10.1029/2006JA012059.
- Xu, W.Y. and Kamide, Y. (2004): Decomposition of daily geomagnetic variations by using method of natural orthogonal component. Journal of Geophysical Research: Space Physics, 109, A5, doi:10.1029/2003JA010216.
- Yamazaki, Y., Häusler, K. and Wild, J.A. (2016): Day-to-day variability of midlatitude ionospheric currents due to magnetospheric and lower atmospheric forcing. Journal of Geophysical Research: Space Physics, 121, 7, 7067-7086.
- Yamazaki, Y., Maute, A. (2017): Sq and EEJ—A Review on the Daily Variation of the Geomagnetic Field Caused by Ionospheric Dynamo Currents. Space Sci Rev 206, 299–405, https://doi.org/10.1007/s11214-016-0282-z.

Acknowledgement

- This study is funded by national funds through FCT (Foundation for Science and Technology, I.P.), under the project MAG-GIC: PTDC/CTA -GEO/31744/2017.
- CITEUC is funded by National Funds through FCT - Foundation for Science and Technology (project: UID/MULTI/00611/2019) and FEDER – European Regional Development Fund through COMPETE 2020 – Operational Programme Competitiveness and Internationalization (project: POCI-01-0145-FEDER-006922).