northumbria UNIVERSITY NEWCASTLE

UNIVERSITY OF BRADFORD
MAKING KNOWLEDGE WORK

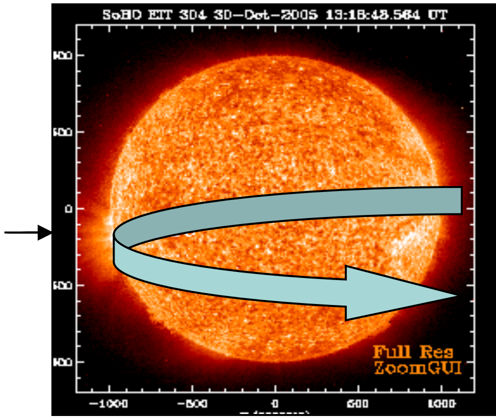
Principal Component Analysis of solar magnetic field and prediction of solar activity on a millennium timescale

V.V. Zharkova[1], Shepherd S.J. [2] , Popova E.[3] and S.I. Zharkov[4]

Zharkov et al., 2008, Solar Phys., 248
Zharkova et al., MNRAS, 2012
Popova et al, AnnGeo, 2013
Shepherd et al, ApJ, 2014
Zharkova et al, Nature Sci. Rep., 2015

THE UNIVERSITY OF HULL

Automated data analysis



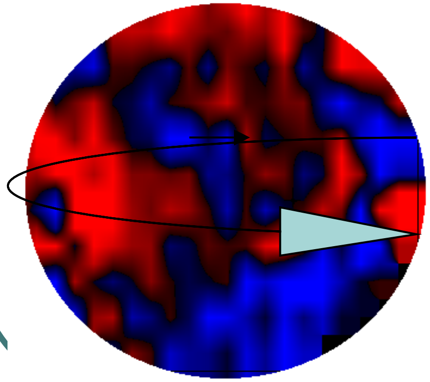
- **High resolution MF (MDI) – EGSO project (2002-2005)**

  – sunspot magnetic field – SMF

- **Automated detection -**

Solar Feature Catalogues:

Zharkova et al., 2005, Sol Phys, 228, 365



- **Low resolution MF (WSO) – synoptic maps of SBMF**

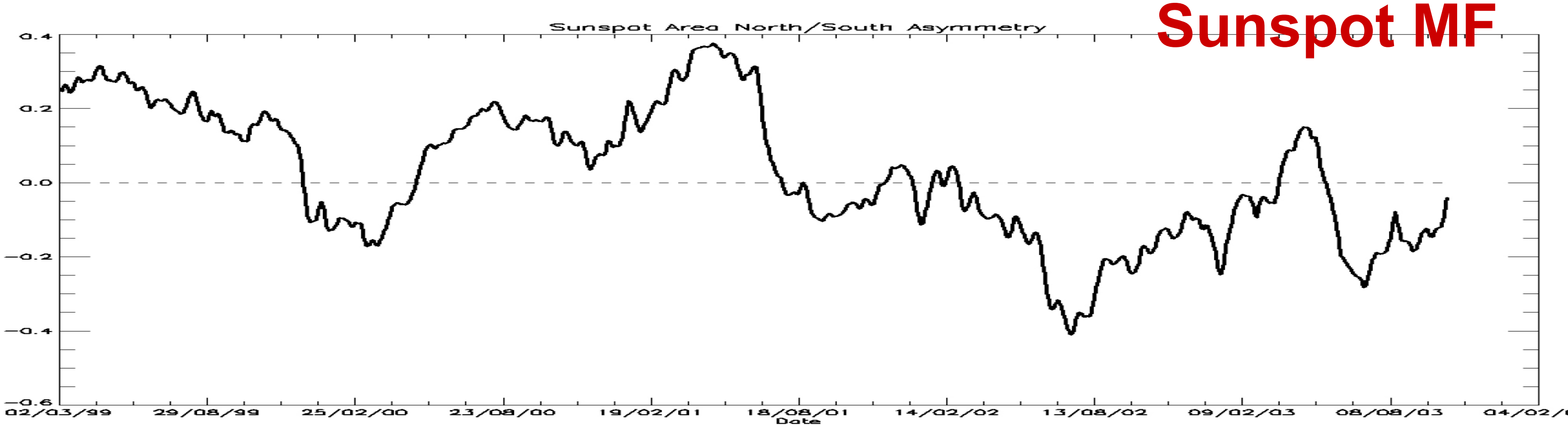
Solar Feature Catalogues - EGSO

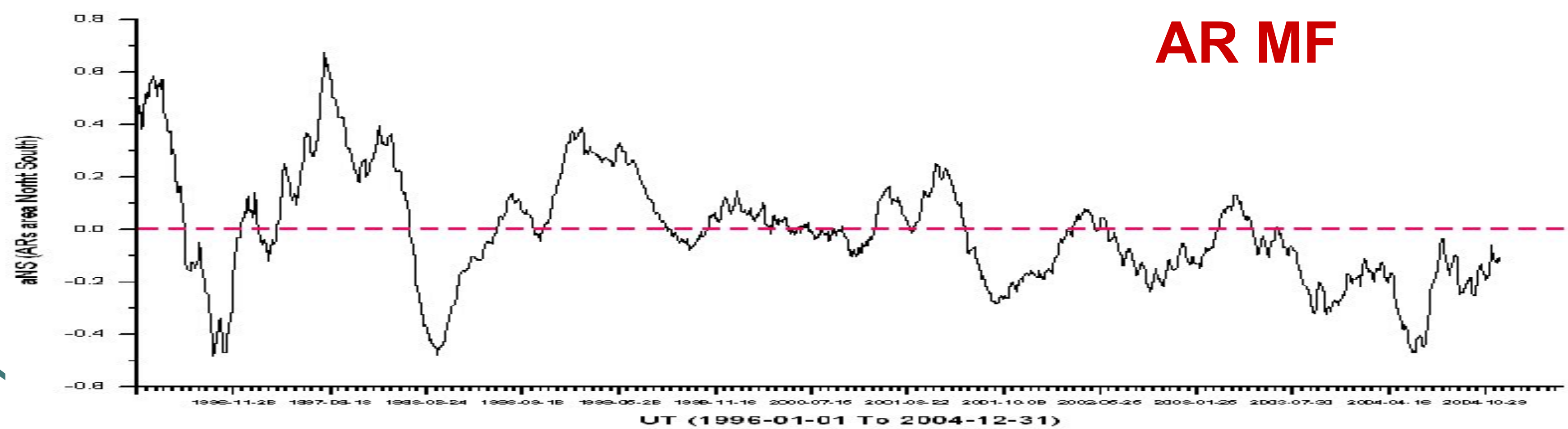
Zharkova et al., 2005, Sol Phys, 228, 365

http://solar.inf.brad.ac.uk

- Sunspot Catalogue (from 1996-05-19 19:08:35 to 2010-05-31 19:51:32)
  - Automated feature detection and data extraction
  - About 40000 observation processed
  - ~370,000 sunspots and 100 000 ARs stored and processed
- Sunspot Catalogue (SOHO MDI) (*Zharkov et al., 2005)*
  - Space Observations, Accuracy & Image Quality
  - Synoptic Continuum images every 6 hour
  - LOS Magnetogram Data
- AR Catalogue (Meudon+MDI) (*Benkhalil et al., 2006)*
  - Meudon Ca II K3 images
  - Meudon H-alpha images
  - MDI LOS Magnetograms
- Filaments and prominences (Meudon) (*Fuller et al., 2005)*
- Meudon H-alpha images

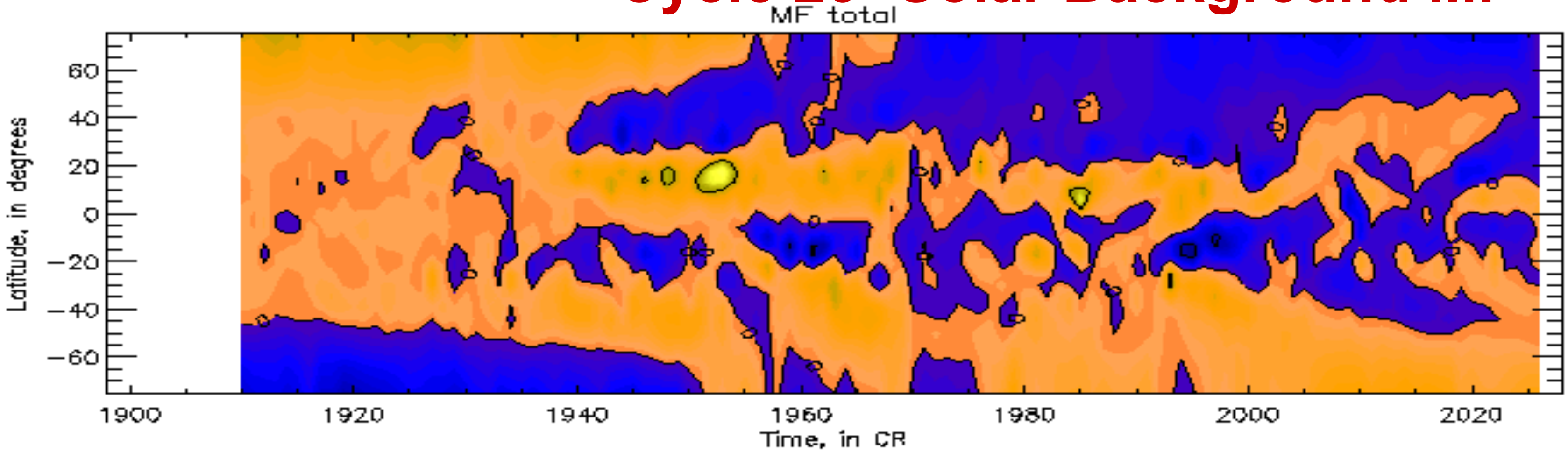
North South Asymmetry (averaged by 170 days) (N-S)/(N+S) (ss –top, ar –bottom) – Zharkov and Zharkova ‘06

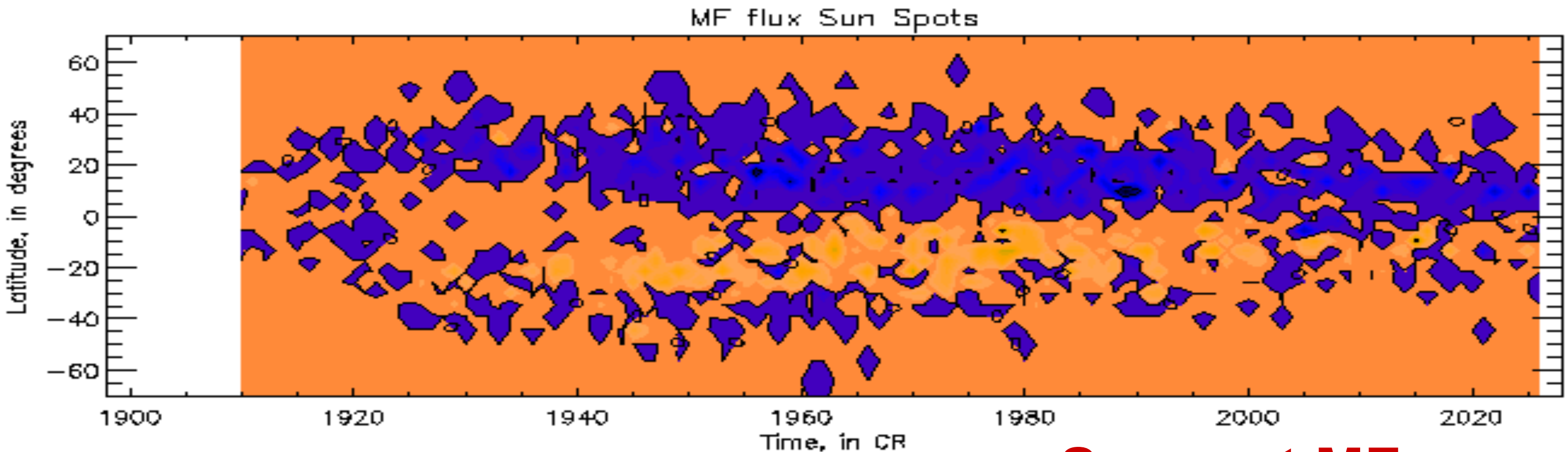




SBMF (top) and sunspot MF (bottom)– phase between them is π~11y (Zharkov et al, 2008, Stix 1976)

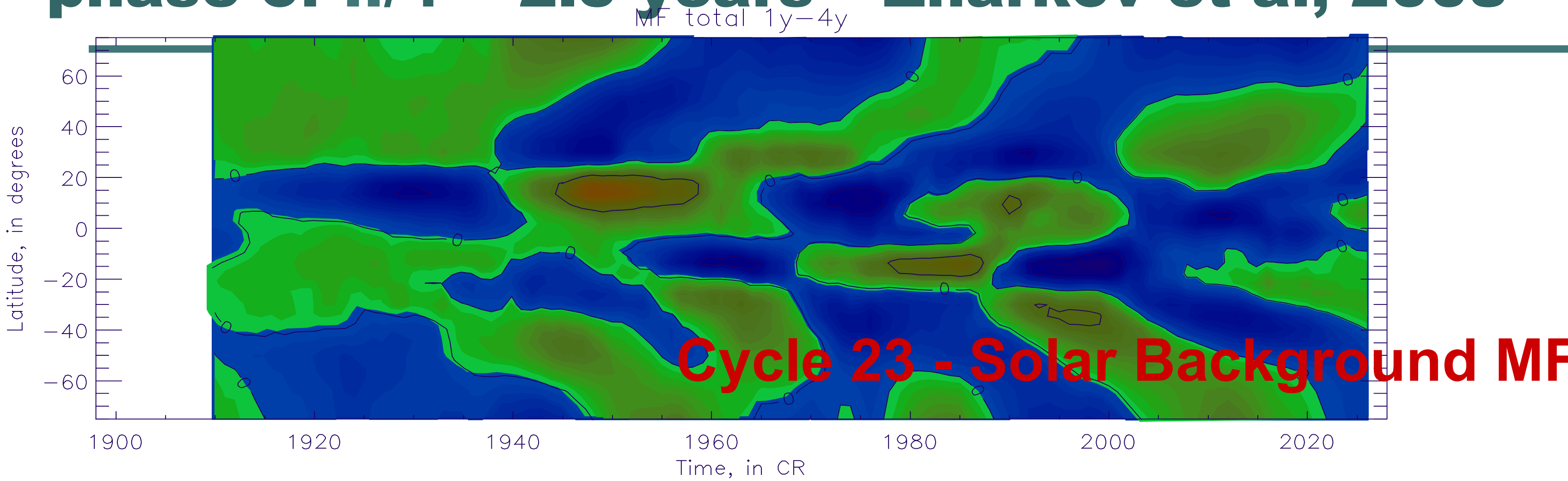
Cycle 23 -Solar Background MF

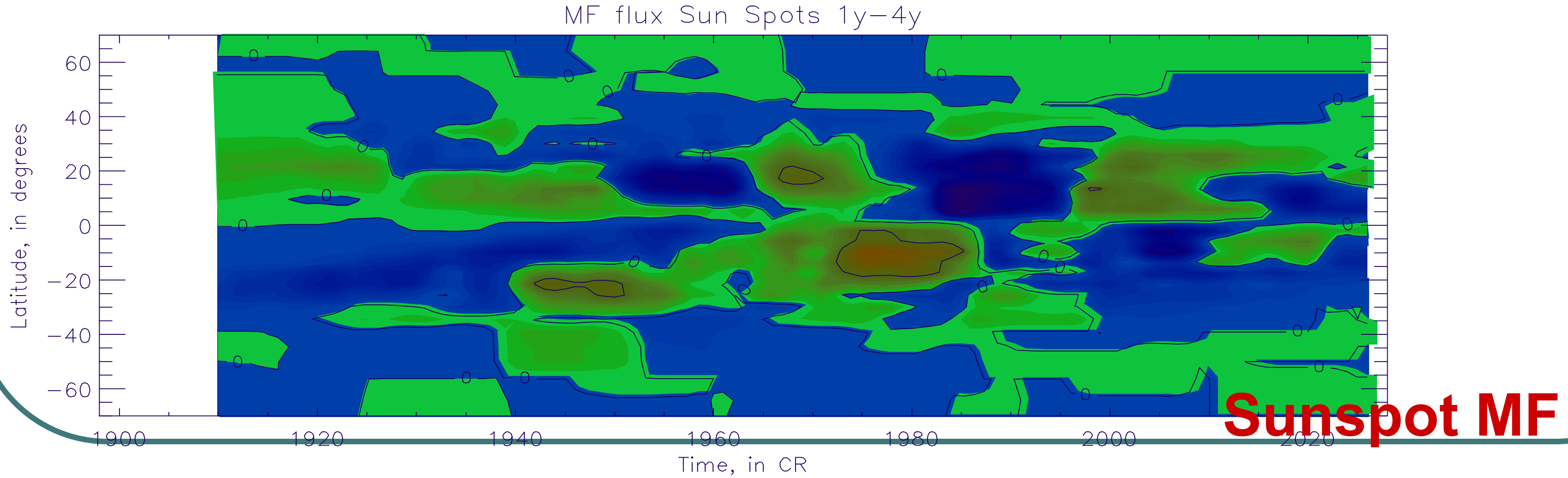




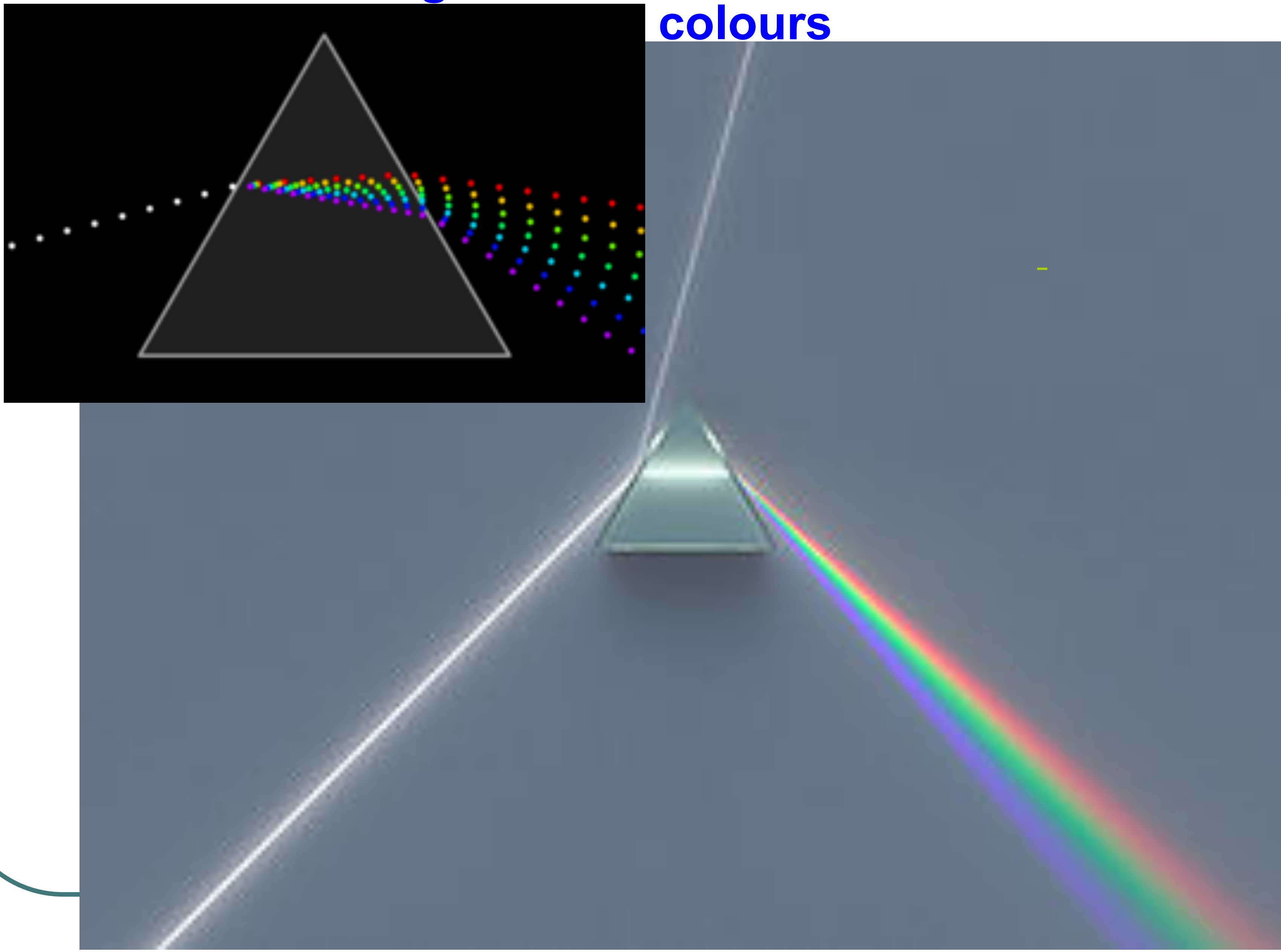
Sunspot MF

2. 1y-4y residuals for BMF (top) and excess SMF (bottom) reveal additional phase of π/4 ~ 2.5 years – Zharkov et al, 2008





White light refraction into waves of different colours



PCA acts as a prism for magnetic waves
Philosophy of PCA

- Introduced by Pearson (1901) and Hotelling (1933) to describe the variation in a set of multivariate data in terms of a set of uncorrelated variables
- We typically have a data matrix of n observations on p correlated variables $x_1, x_2, \ldots x_p$
- PCA looks for a transformation of the x_i into p new variables y_i that are uncorrelated

To reduce dimension – weighted average

- Weighted average based on some criterion. Which one? -→
- Looking for a transformation of the data matrix $\mathbf{X}$ $(n \times p)$ such that

$$Y = \boldsymbol{\delta}^T \boldsymbol{X} = \delta_1 X_1 + \delta_2 X_2 + .. + \delta_p X_p$$

- Where $\boldsymbol{\delta} = (\delta_1, \delta_2, .., \delta_p)^T$ is a column vector of weights with

$$\delta_1^2 + \delta_2^2 + .. + \delta_p^2 = 1$$

One good criterion

- Maximize the variance of the projection of the observations on the Y variables
- Find $\boldsymbol{\delta}$ *so that*

$$Var(\delta^T \mathbf{X}) = \delta^T \mathbf{Var(X)}\, \delta \quad \textit{is maximal}$$

- The matrix **C=Var(X)** is the covariance matrix of the X_i variables

Calculating eignevalues and eigenvectors

- The eigenvalues λ_i are found by solving the equation

$$det(C-\lambda I)=0$$

- Eigenvectors are columns of the matrix A such that

$$C=A\ D\ A^T$$

- Where

$$D=\begin{pmatrix} \lambda_1\ 0 \ldots\ldots.0 \\ 0 \quad \lambda_2 \ldots\ldots.0 \\ 0 \\ 0 \ldots\ldots\ldots\ldots.\lambda_p \end{pmatrix}$$

So PCA gives

- New variables Y_i that are linear combination of the original variables (x_i):
- $Y_i = a_{i1}x_1 + a_{i2}x_2 + \ldots a_{ip}x_p$; $i=1..p$
- The new variables Y_i are derived in decreasing order of importance;
- they are called 'principal components'

Interpretation of PCA

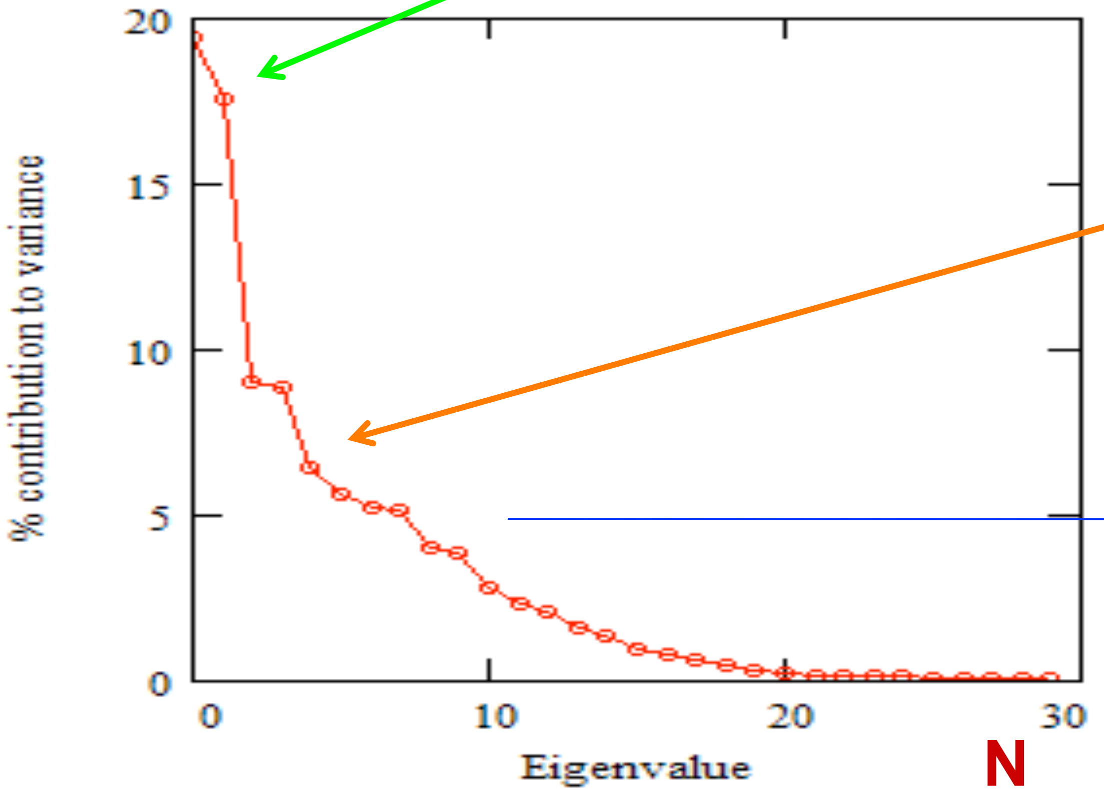
- The new variables (PCs) have a variance equal to their corresponding eigenvalue

$$Var(Y_i) = \lambda_i \text{ for all } i=1\ldots p$$

- Small $\lambda_i \Leftrightarrow$ small variance $\Leftrightarrow$ data change little in the direction of component Y_i
- The relative variance explained by each PC is given by $\lambda_i / \Sigma \lambda_i$
- ***PCs can be assigned to separate physical processes***

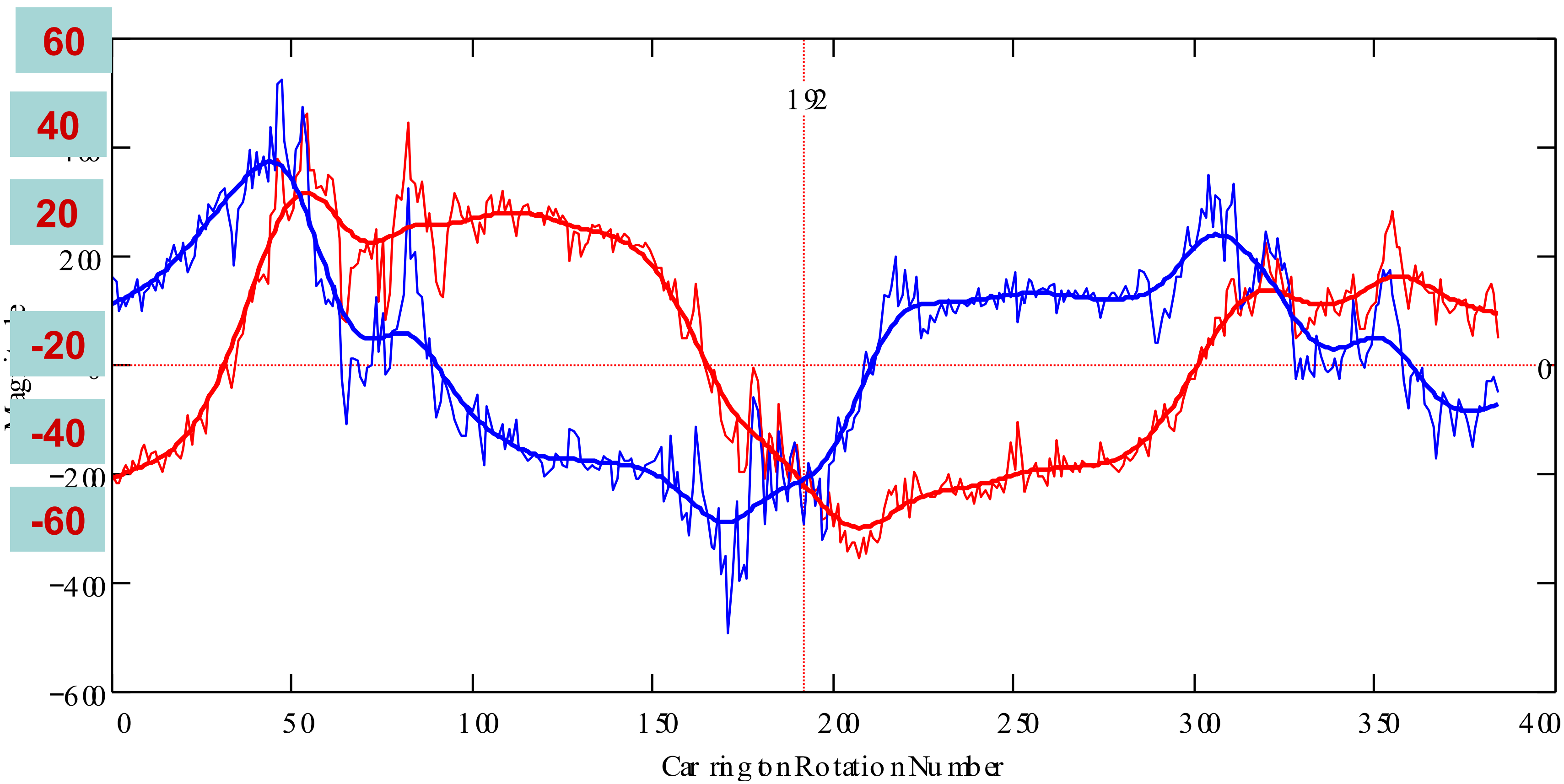
SBMF results: Scree plot- Eigenvalues vs variances

- 2 main eigenvalues covering 40% of variance – dipole source

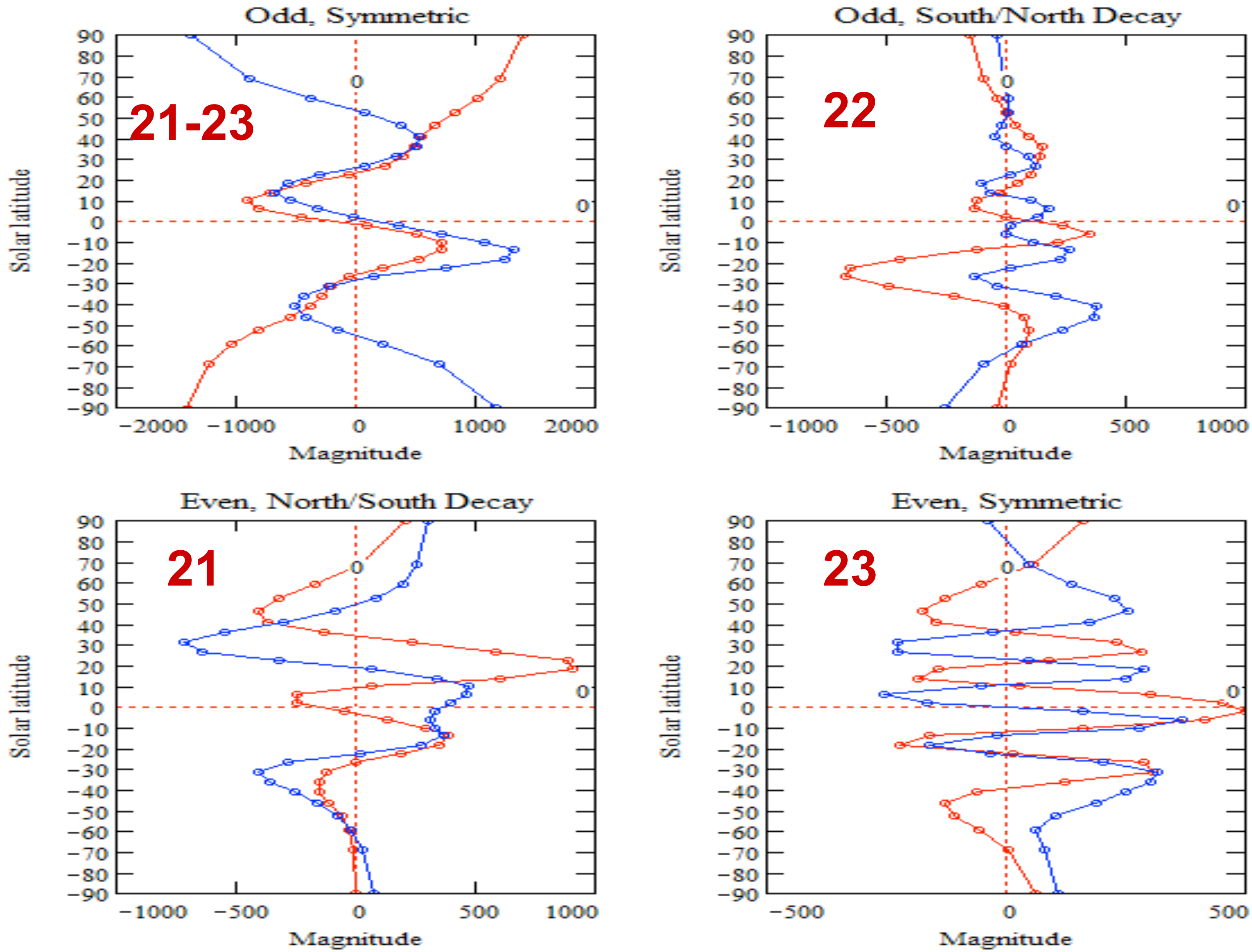


45% of variance other sources (quadruple

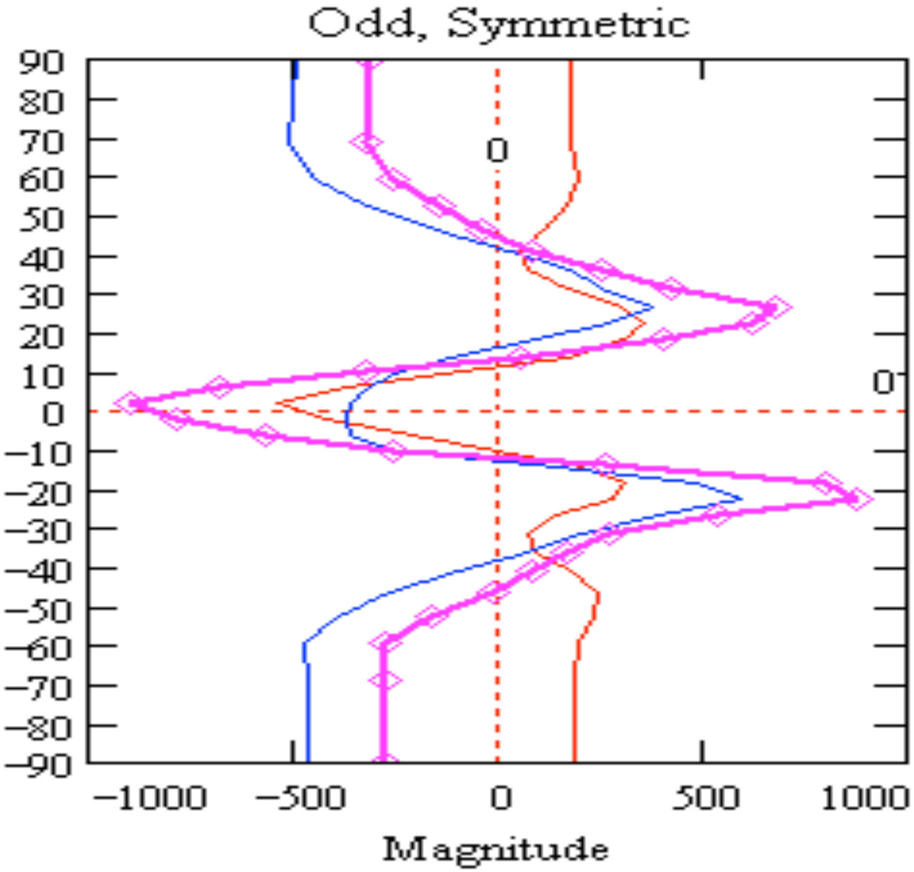
Two magnetic waves of the opposite polarities extracted in SMF with PCA (Zharkova et al., 2012, MNRAS)



PCs and ICs for 4 largest pairs of eigenvalues (Zharkova et al, 2012, MNRAS)



Derivatives of 2 main EOFs



From top left figure above

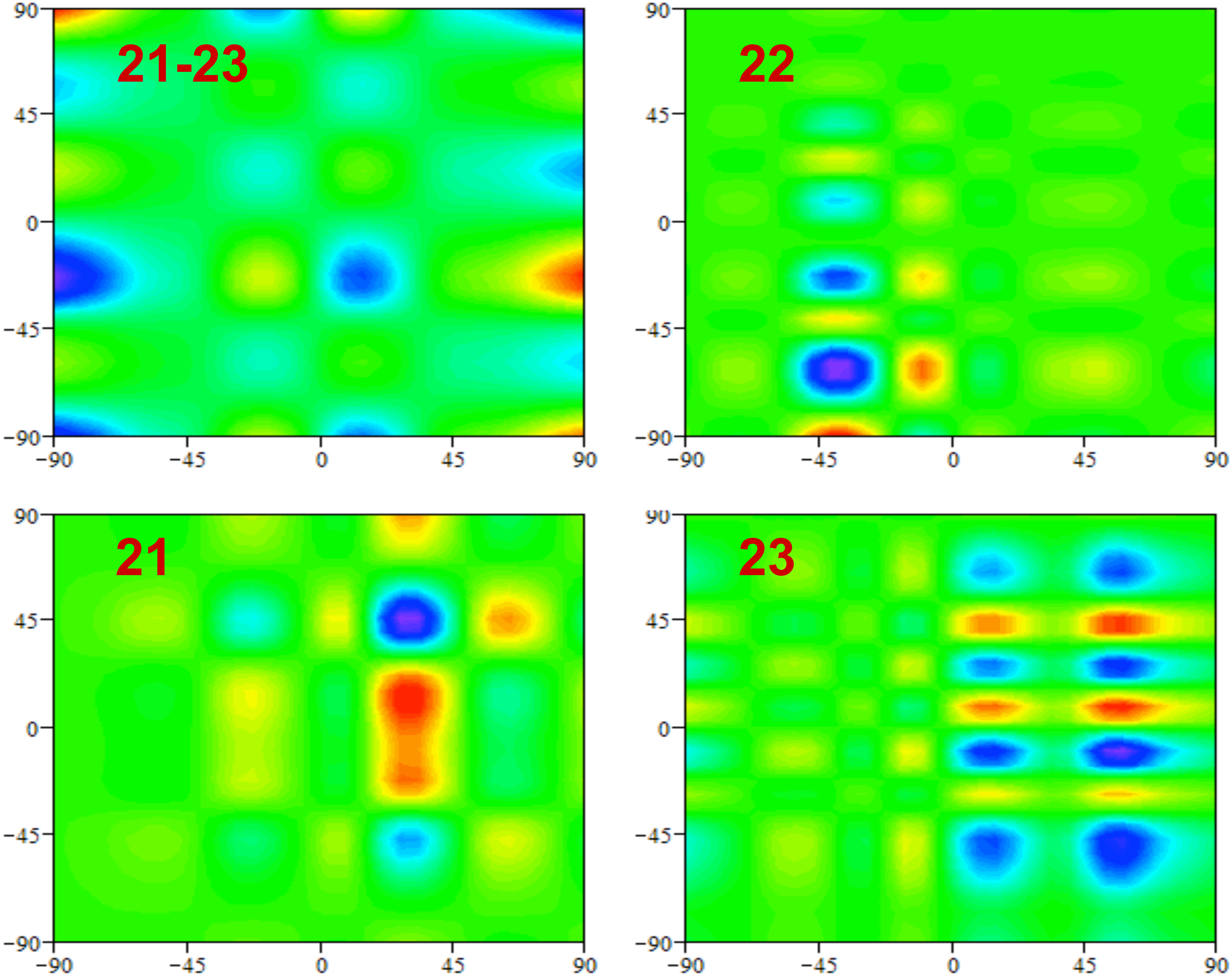
component 1

component 2

resultant

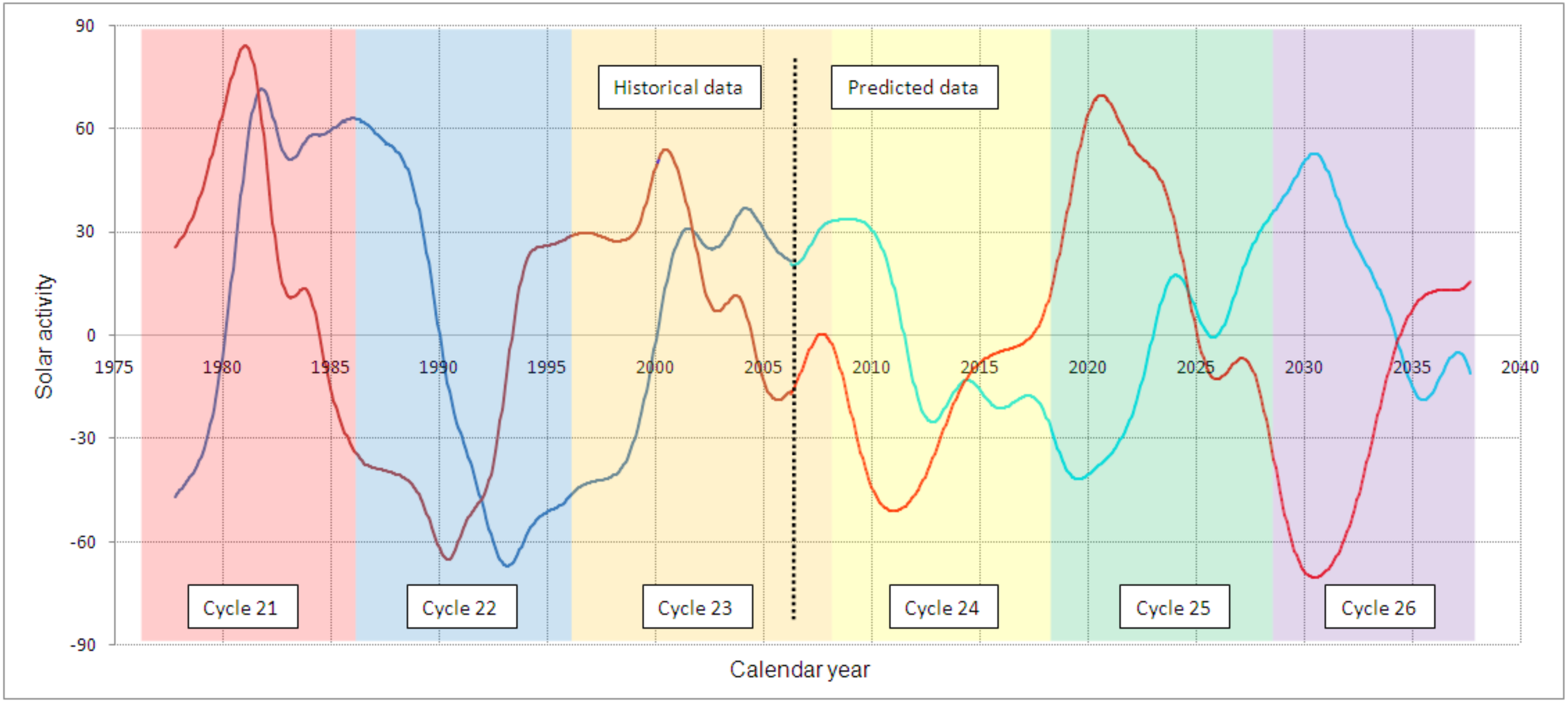
These latitudinal ICs were modeled with the updated 2-layers Parker's dynamo fitting amplitudes, phase shifts and number of equator crossings Popova et al, 2013

Cross-correlation of 8 EOFs

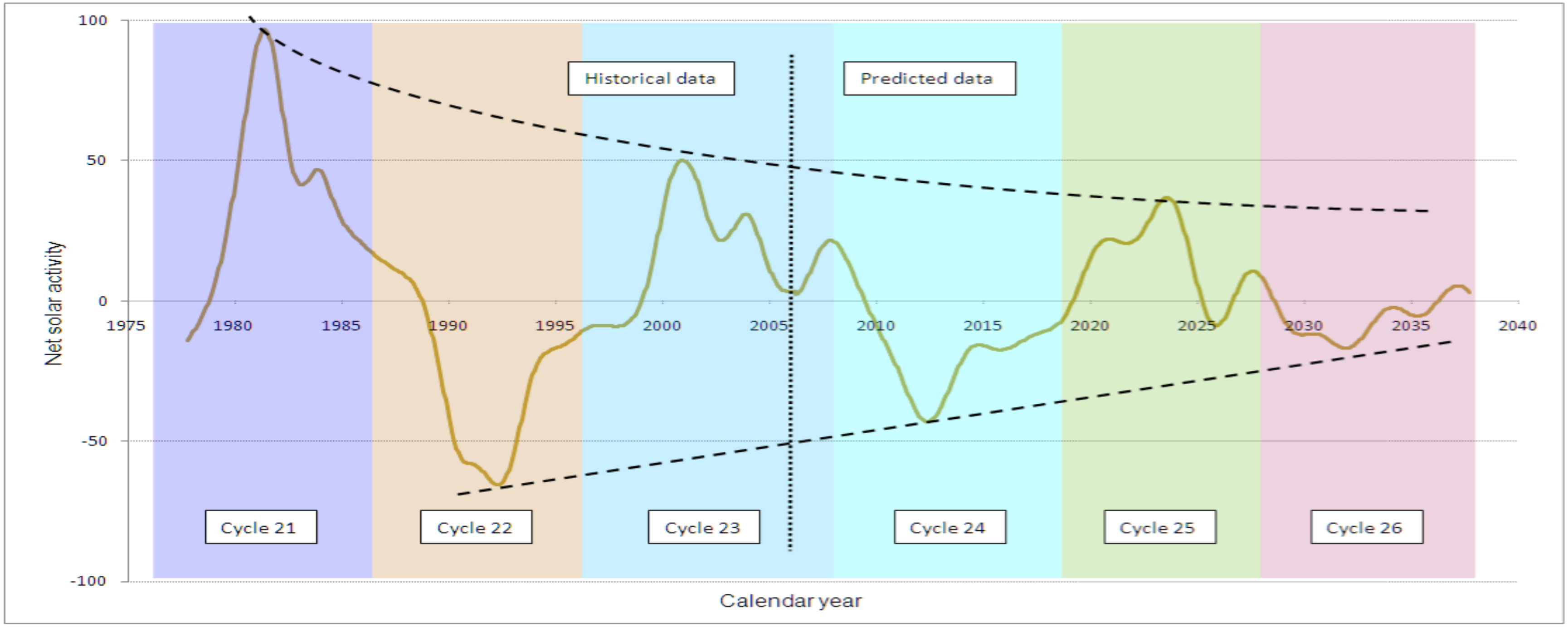


Classification of and predicted principal components

(Zharkova et al, 2012, Shepherd et al, 2014)



Summary curve



Mathematical laws from PCs: Symbolic regression -Hamiltonian approach ~~Schmidt & Lipson, 2009, Science)~~

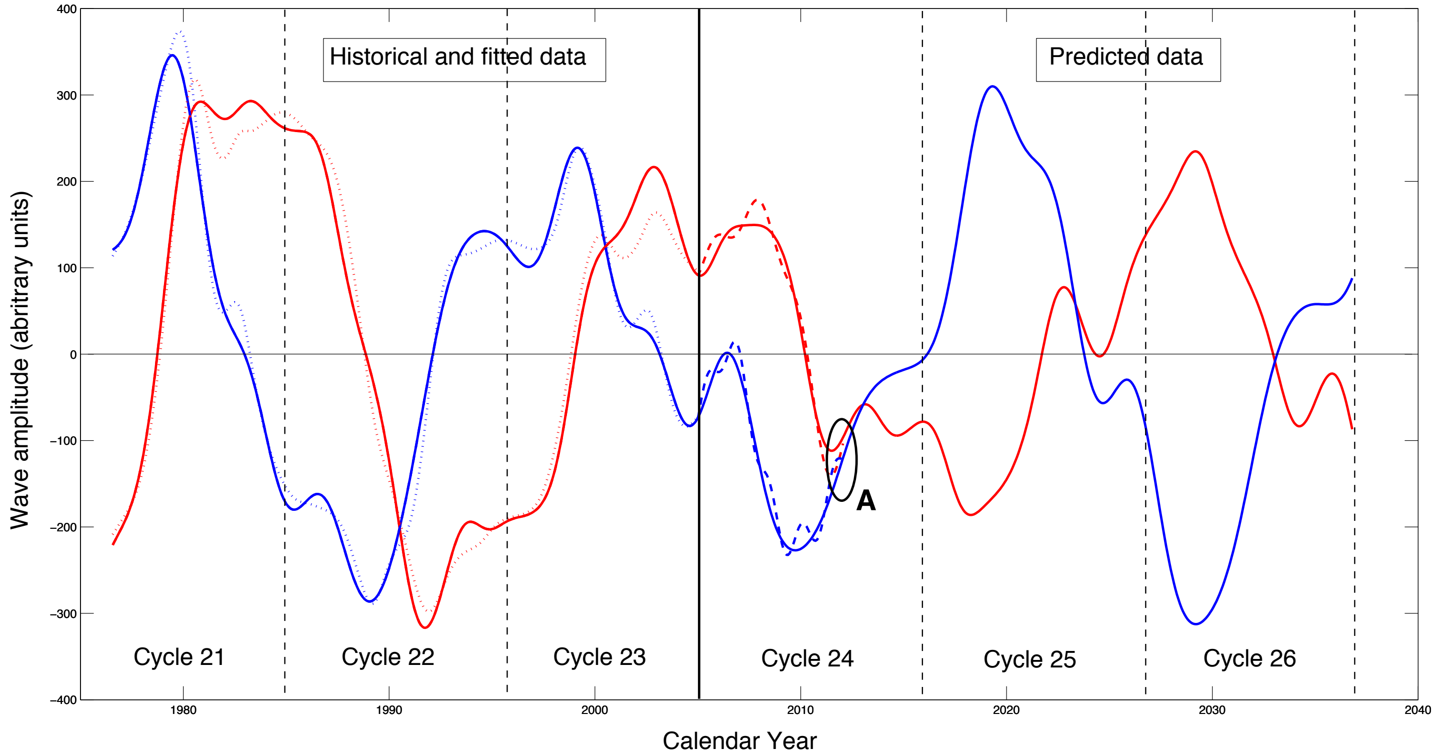
- Mathematical law for the first principal component:

$$F_1(t) = \Sigma_{k=1,..,5} A_k \cos(\omega_{k,1} t + \phi_{k,1})\cos(B_{k,1} \cos(\omega_{k,1} t + \phi_{k,1})$$

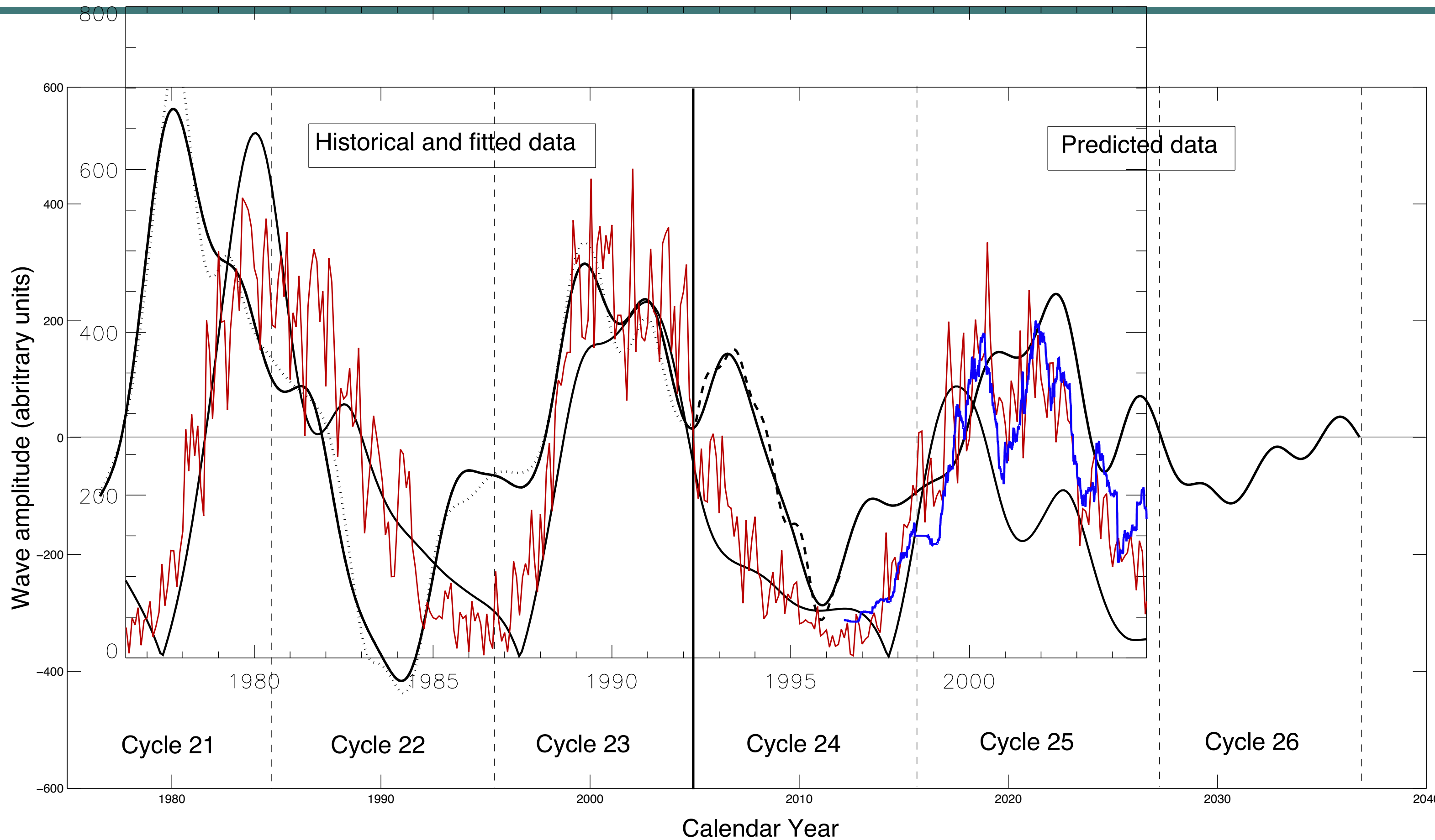
- Mathematical law for the second principal component:

$$F_2(t) = \Sigma_{k=1,..,5} A_k \cos(\omega_{k,2} t + \phi_{k,2})\cos(B_{k,2} \cos(\omega_{k,2} t + \phi_{k,2})$$

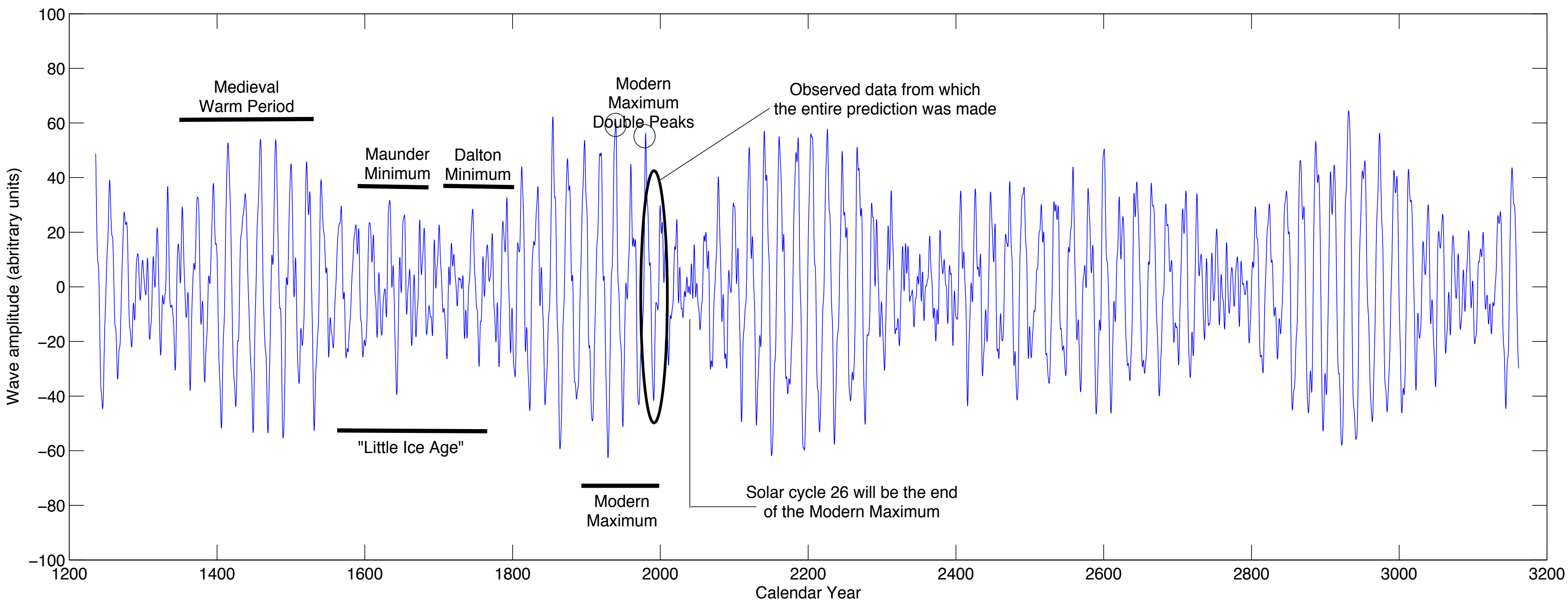
Fitting the PCs to measured in cycle 24 and prediction to 25-26



Summary PC and modulus summary PC vs sunspot data
(Shepherd et al, 2014, ApJ)

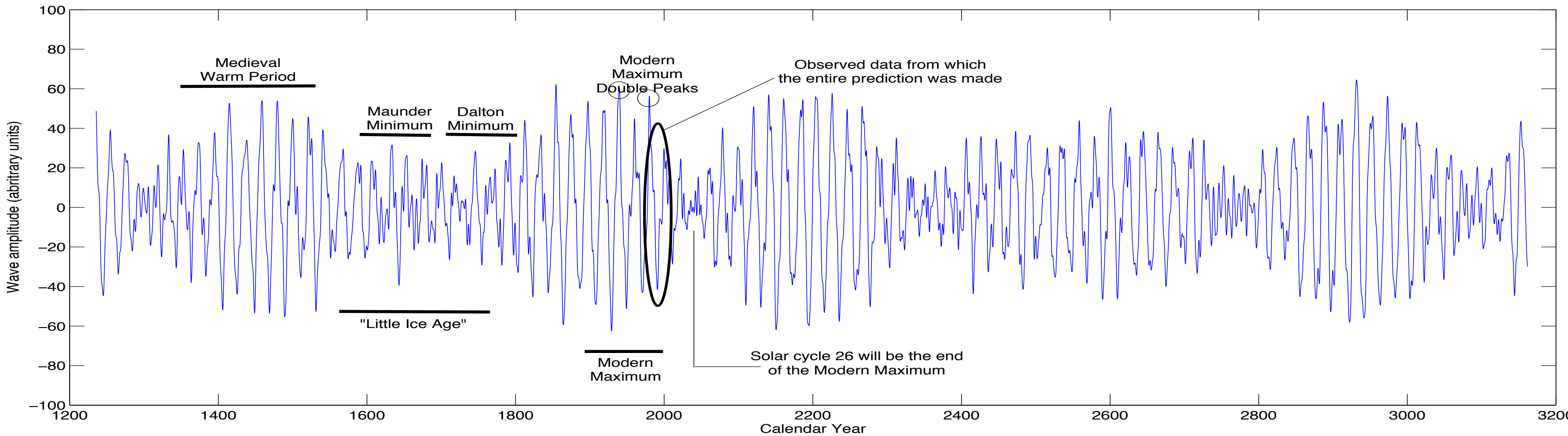


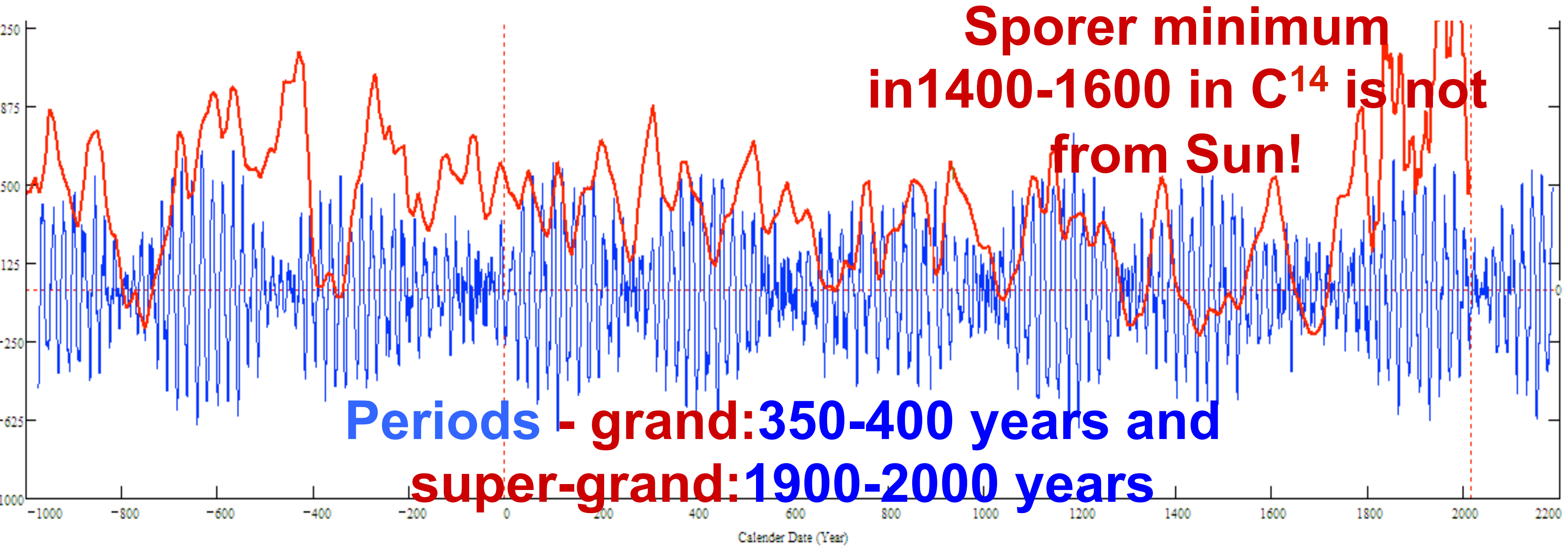
Predicted summary curve on the millennium timescale



Periods - grand:350-400 years and super-grand:1900-2000 years

Updated curve for 3000 years (blue) versus a curve by Usoskin et al. (red)

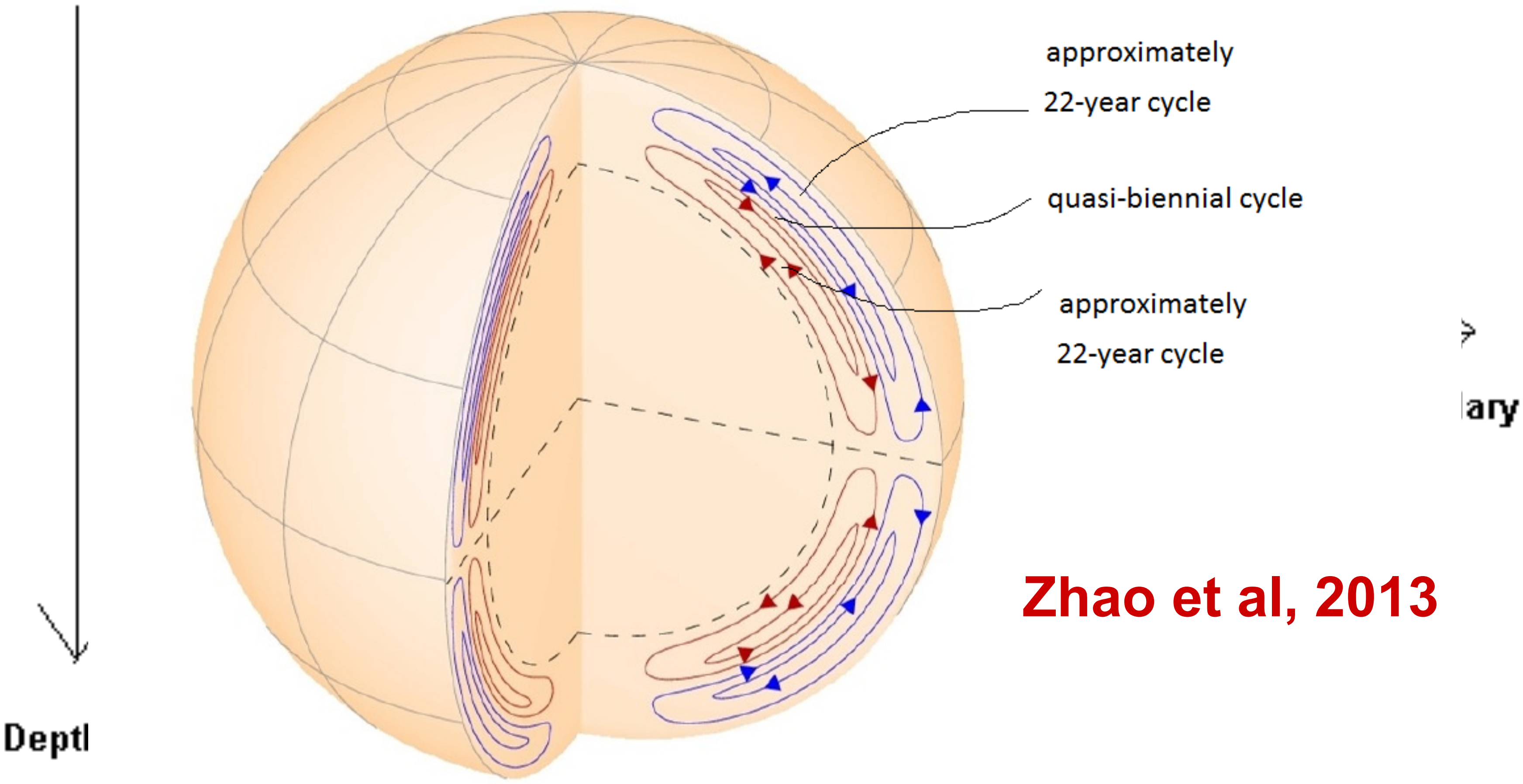




2 layer dynamo model explaining some PCA features

- Dynamo model was not even considered yet while we did PCA in 2010-12 and SEA 2014
- Started discussing possible mechanisms since the end of 2010
- In 2011 we considered 2 layer Parker's model (1993) with meridional circulation
- 2013 – first model paper appeared in Annals in Geophysics (Popova et al, 2013)

Dynamo in Two-Layer Medium



Zhao et al, 2013

We included the meridional flows in each layer:

$$\frac{\partial B}{\partial t} + \frac{\partial (VB)}{\partial \theta} = \beta \Delta B, \qquad \frac{\partial A}{\partial t} + V\frac{\partial A}{\partial \theta} = \alpha B + \beta \Delta A, \tag{2.3}$$

$$\frac{\partial b}{\partial t} + \frac{\partial (vb)}{\partial \theta} = D\cos\theta \frac{\partial a}{\partial \theta} + \Delta b, \qquad \frac{\partial a}{\partial t} + v\frac{\partial a}{\partial \theta} = \Delta a\,, \tag{2.4}$$

here $V(\theta), v(\theta)$ are the meridional flows in the respective layers.

Following Parker we prescribe $r = 0$ for the radial boundary between two layers and use boundary conditions:

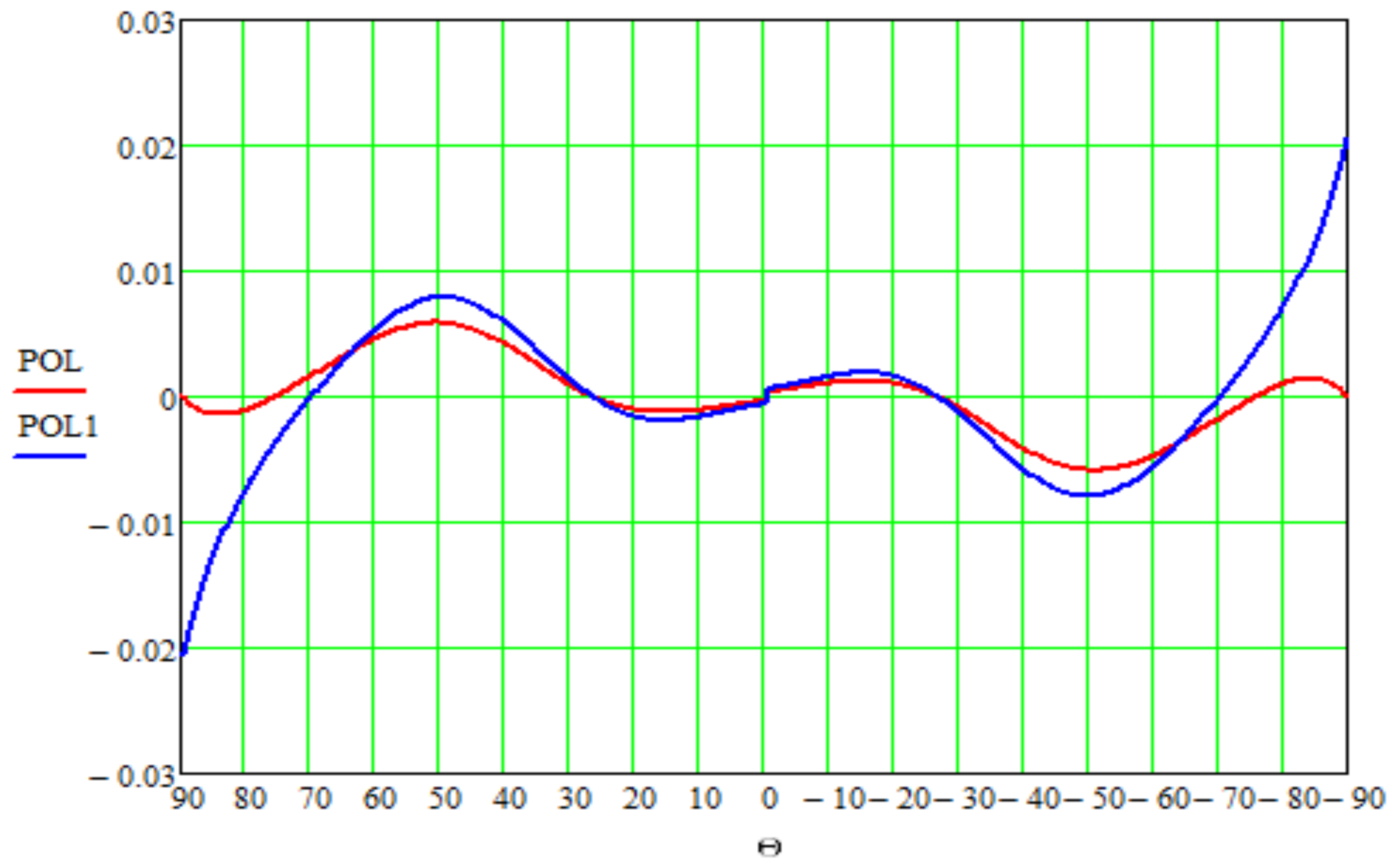
$$b = B, \qquad a = A, \qquad \frac{\partial b}{\partial r} = \beta\frac{\partial B}{\partial r}, \qquad \frac{\partial a}{\partial r} = \frac{\partial A}{\partial r}. \tag{2.5}$$

In view of the symmetry conditions $\alpha(-\theta) = -\alpha(\theta), V(-\theta) = -V(\theta)$ the above system of equations can be considered in only one (e.g., the northern) hemisphere using anti-symmetry (dipolar symmetry) or symmetry (quadrupolar symmetry) conditions at the equator.

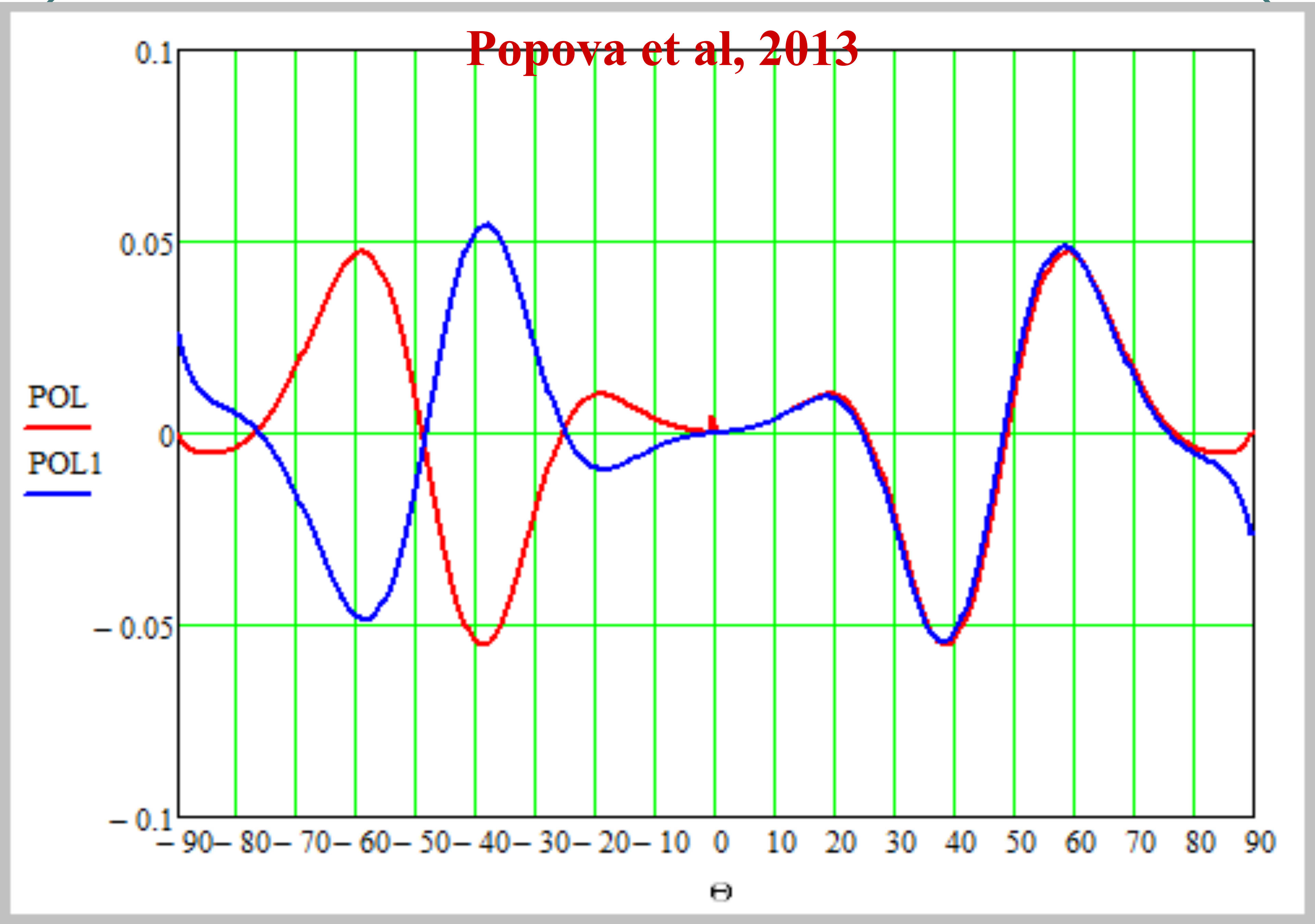
We obtained Hamilton-Jacobi equation for eqs. (2.3) and (2.4) by a method similar to the method described in Popova et al. (2010).

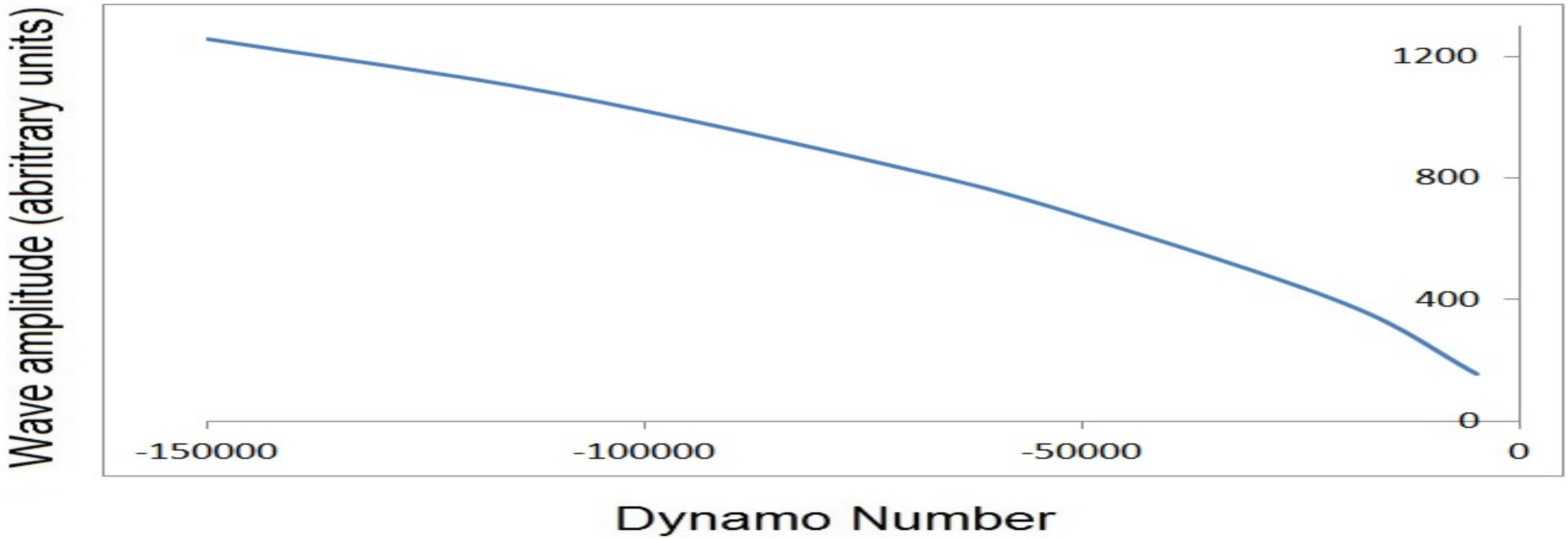
Two dipole components with different boundary conditions of BMF

Popova et al, 2013

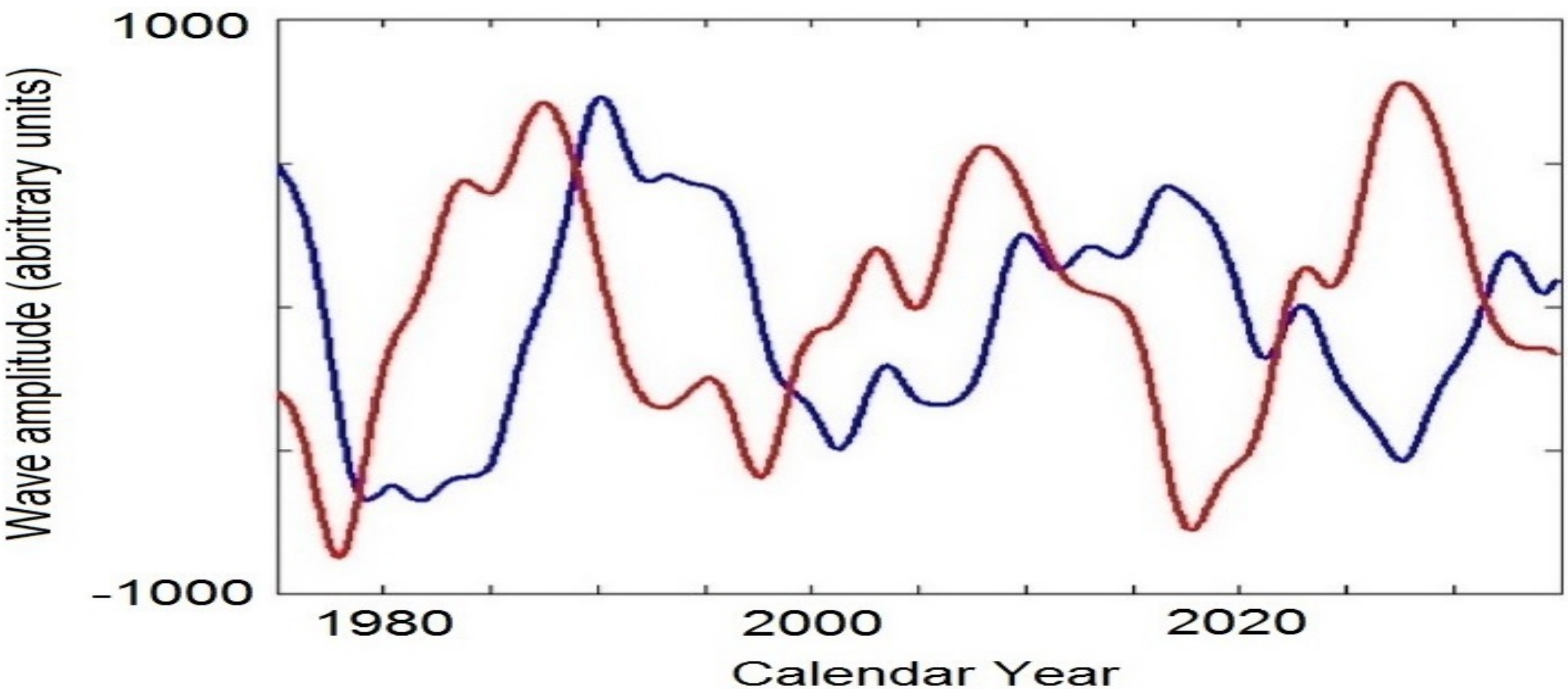


Quadrupole components of poloidal magnetic field (BMF)





Wave amplitude (abritrary units)
1200
800
400
0
-150000
-100000
-50000
0
Dynamo Number



Wave amplitude (abritrary units)
1000
-1000
1980
2000
2020
Calendar Year

Undamped Wave Equation: Solution to Initial Value Problem (2 of 3)

- Solution of a wave equation with forced oscillations

$$y(t) = \frac{F_0}{m(\omega_0^2 - \omega^2)}\left(\cos\omega t - \cos\omega_0 t\right)$$

- To simplify the solution even further, let $A = (\omega_0 + \omega)/2$ and $B = (\omega_0 - \omega)/2$. Then $A + B = \omega_0 t$ and $A - B = \omega t$. Using the trigonometric identity

$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B,$$

it follows that

$$\cos\omega t = \cos A \cos B + \sin A \sin B$$

$$\cos\omega_0 t = \cos A \cos B - \sin A \sin B$$

and hence

$$\cos\omega t - \cos\omega_0 t = 2\sin A \sin B$$

Undamped Equation: Beats (3 of 3)

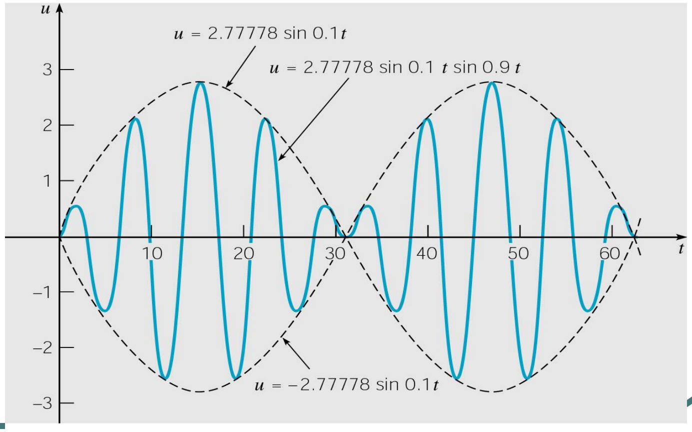
- Using the results of the previous slide, it follows that

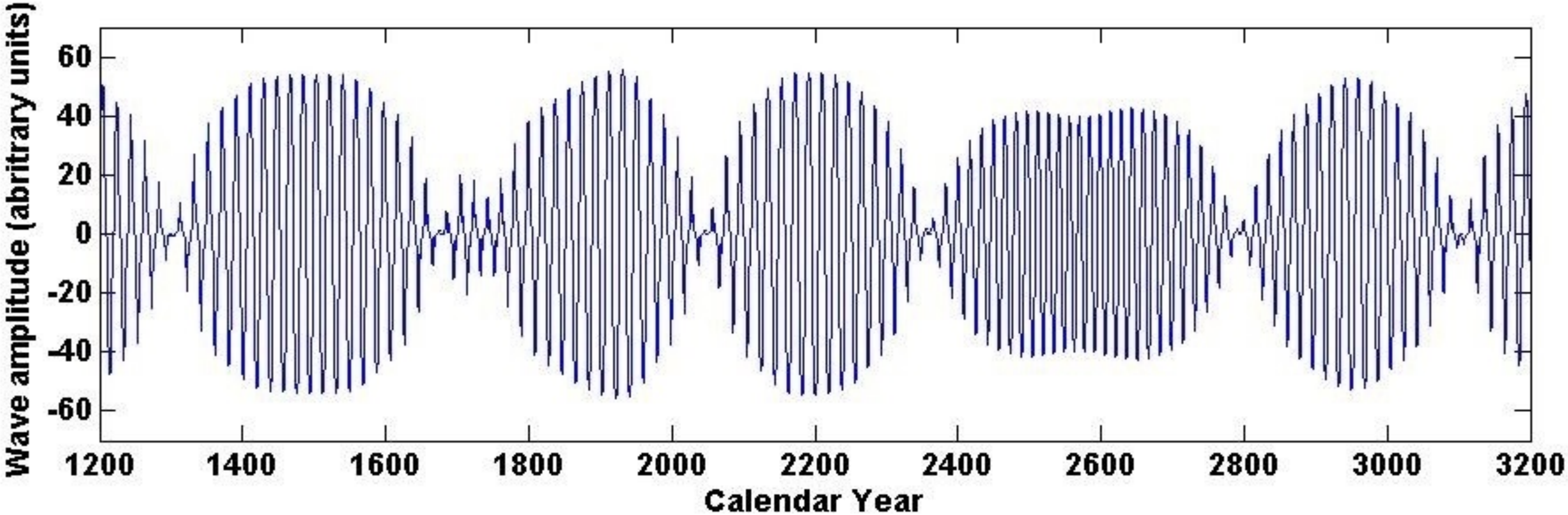
$$y(t) = \left[\frac{2F_0}{m(\omega_0^2 - \omega^2)} \sin\frac{(\omega_0 - \omega)t}{2}\right] \sin\frac{(\omega_0 + \omega)t}{2}$$

- When $|\omega_0 - \omega| \cong 0$, $\omega_0 + \omega$ is much larger than $\omega_0 - \omega$, and $\sin[(\omega_0 + \omega)t/2]$ oscillates more rapidly than $\sin[(\omega_0 - \omega)t/2]$.
- Thus motion is a rapid oscillation with frequency $(\omega_0 + \omega)/2$, but with slowly varying sinusoidal amplitude given by

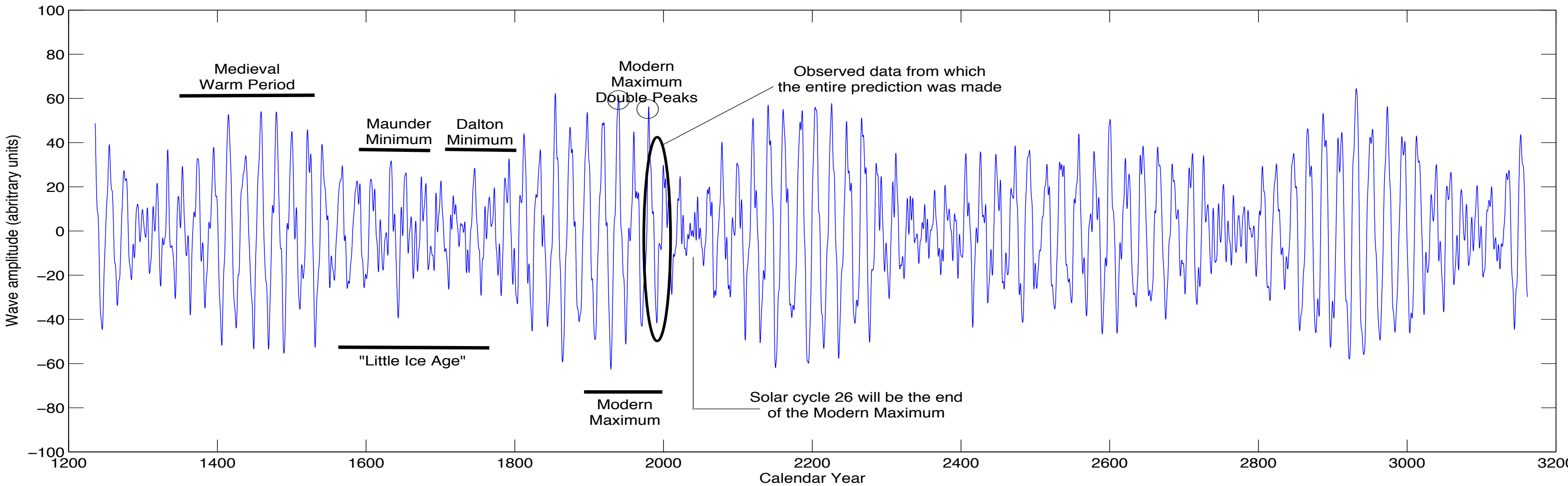
$$\frac{2F_0}{m\left|\omega_0^2 - \omega^2\right|} \left|\sin\frac{(\omega_0 - \omega)t}{2}\right|$$

- This phenomena is called a **beat**.
- Beats occur with two tuning forks of nearly equal frequency.





Wave amplitude (abritrary units)
Calendar Year



Medieval Warm Period
Maunder Minimum
Dalton Minimum
Modern Maximum Double Peaks
Observed data from which the entire prediction was made
"Little Ice Age"
Modern Maximum
Solar cycle 26 will be the end of the Modern Maximum
Wave amplitude (abritrary units)
Calendar Year

Conclusions:
EOFs components: cycles 21-23

- Principal components of SBMF are paired
- The strongest PCs cover more than 40% of variance
- These PCs are shown to reflect 2 dynamo waves travelling with increasing phase shift from one hemisphere to another
- The waves intercept with the increased turbulence one year prior and after the cycle maximum
- Cross-correlation shows a presence of quadruple sources in all the cycles and possible sextuple ones in cycle 23
- Mathematical laws derived with Hamiltonian approach (Euriqa) was used for prediction of the reduction of the solar activity in cycles 25-27 – next Maunder Minimum
- Prediction for 3000 years backwards fits the main warming and cooling periods – Sun gave us the clues!