Data Assimilation with 3D models of the Earth's dynamo

Alexandre Fournier[1], Julien Aubert[1], Lars Nerger[2], Sabrina M. Sanchez [1]

1: Institut de Physique du Globe de Paris, Paris, France; 2: Alfred Wegener Institute, Bremerhaven, Germany

LabEx UnivEarthS Fall School, October 2015

Messages

- The geomagnetic field is mainly created by a complex, nonlinear process taking place in the Earth's core: the **geodynamo** operates on a wide range of space and time scales
- The ever-growing and heterogeneous catalog of geomagnetic data allows us to get a more and more accurate description of this process and of its variability
- This better description is an incentive for constructing and testing physical models able to account for the record of interannual to millennial geomagnetic variations (in a data assimilation framework).
- Goals:
  - Identify the processes controlling the geomagnetic secular variation
  - Place constraints on the internal structure of the core
  - Forecast the evolution of the field and reanalyze its past variations
- This effort started about 8 years ago, and is still at the research stage.

Outline

1. **Introduction – The geomagnetic field**
2. **Application of the EnKF to a numerical model of the geodynamo**

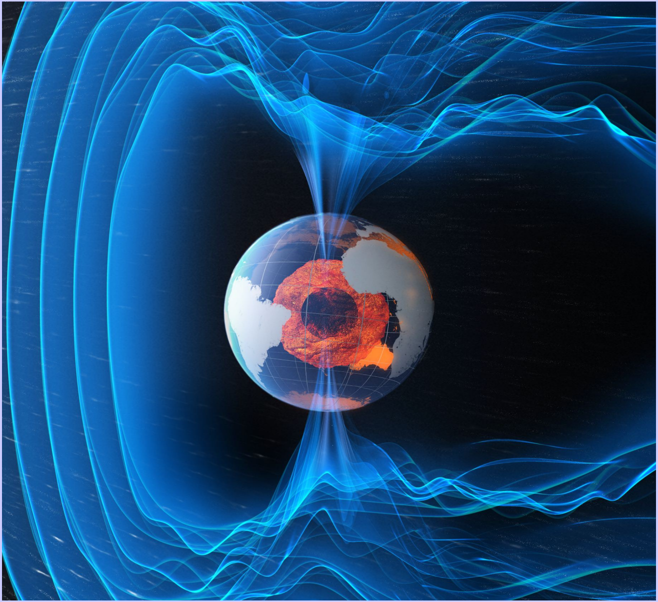
The Tesla warning

Remember

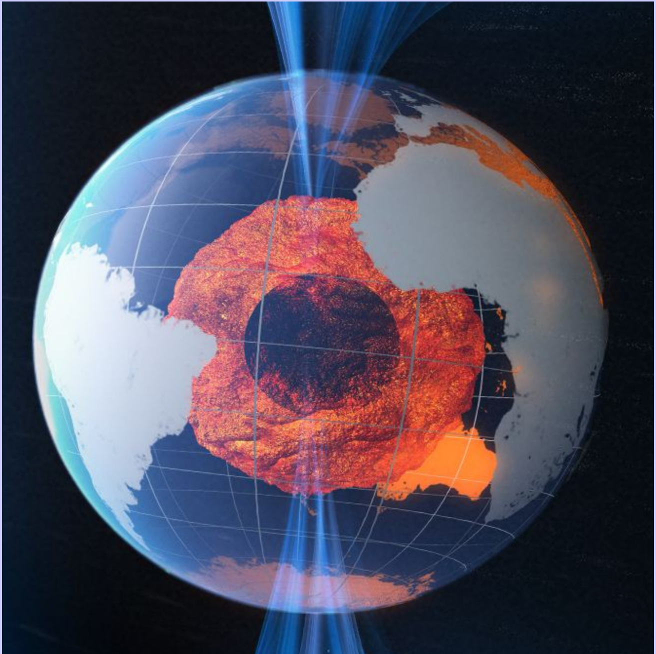
- 1 T=10 kG
- 1 mT=10 G
- 1 μT=10 mG
- 1 nT=10 μG

1. Introduction – The geomagnetic field

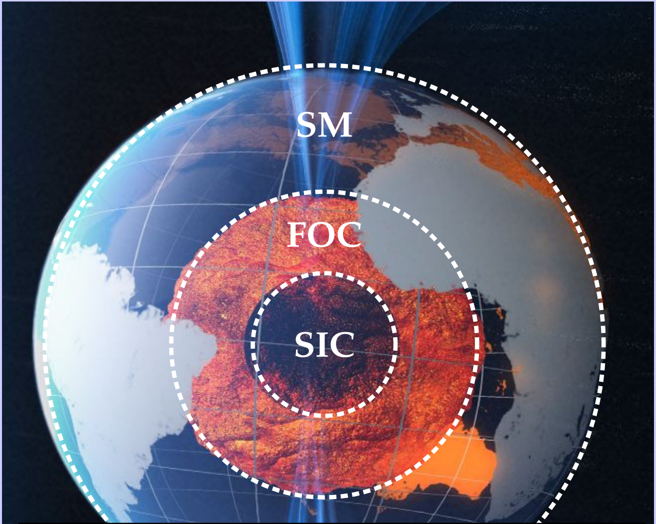
The Earth's interior and the geomagnetic field



The Earth's interior and the geomagnetic field



The Earth's interior and the geomagnetic field



SM: Solid Mantle (rocks),	0–2890 km depth
FOC: Fluid Outer Core (liquid Fe),	2890–5150 km depth
SIC: Solid Inner Core (solid Fe),	5150–6370 km depth

Sources of the geomagnetic field

The geodynamo accounts for more than 90% of the field measured at the Earth's surface.

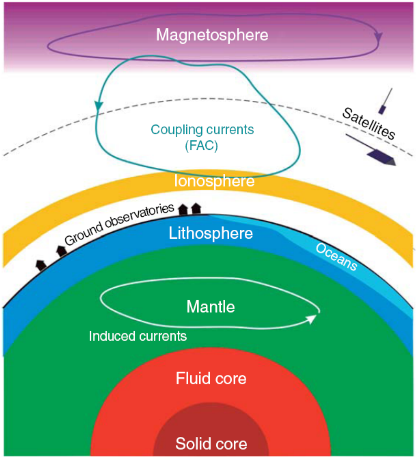
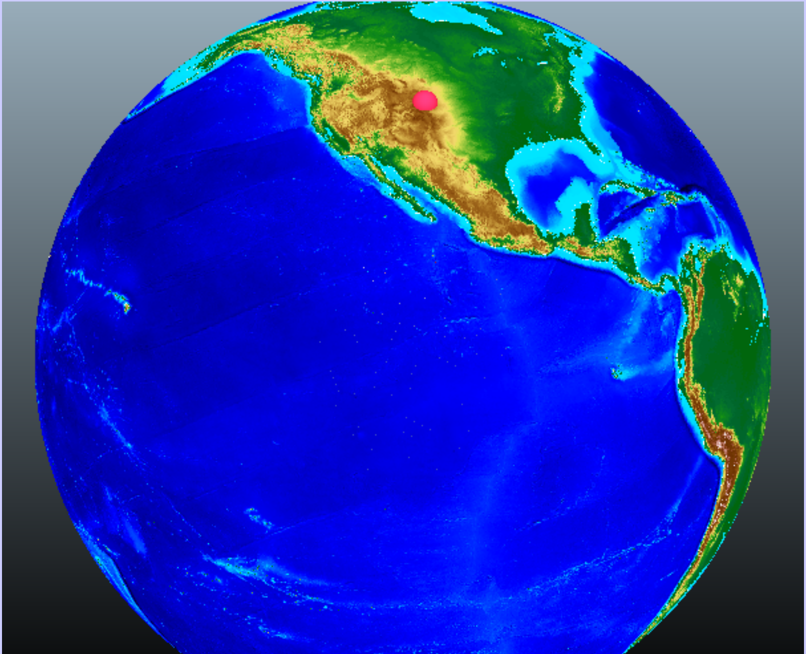


Figure 1 Sketch of the various sources contributing to the near-Earth magnetic field.

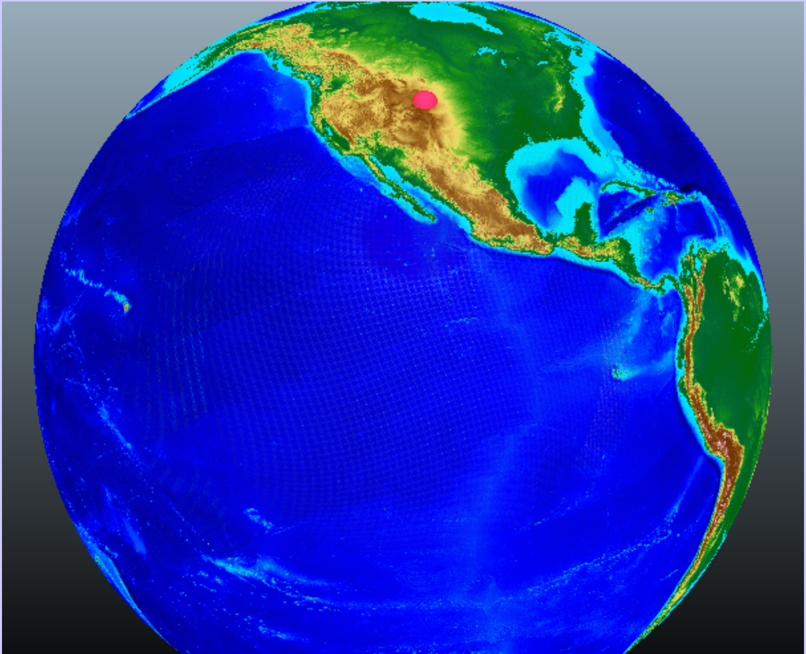
Hulot, Sabaka, Olsen, Fournier

Treatise on Geophysics, 2nd edition, 2015

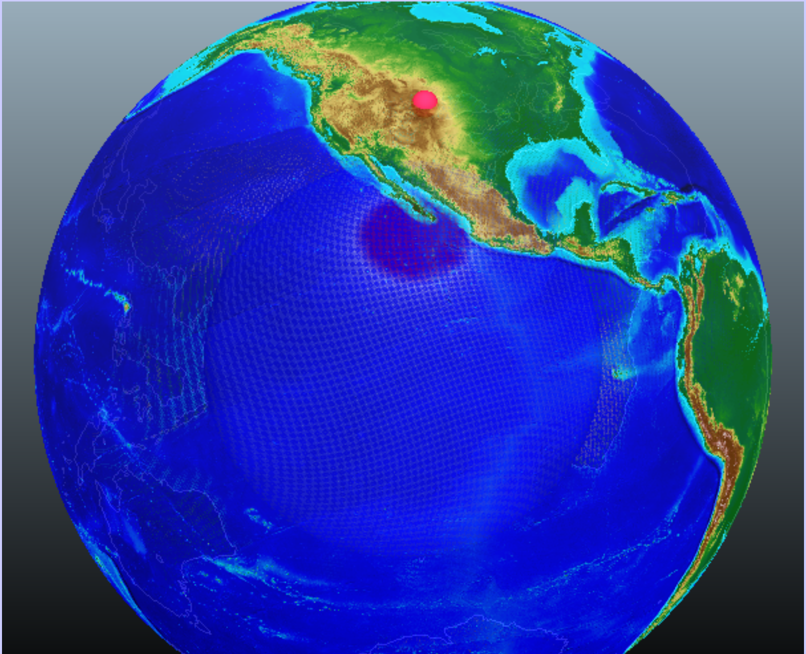
Sensitivity of a measurement to B_r at the core surface



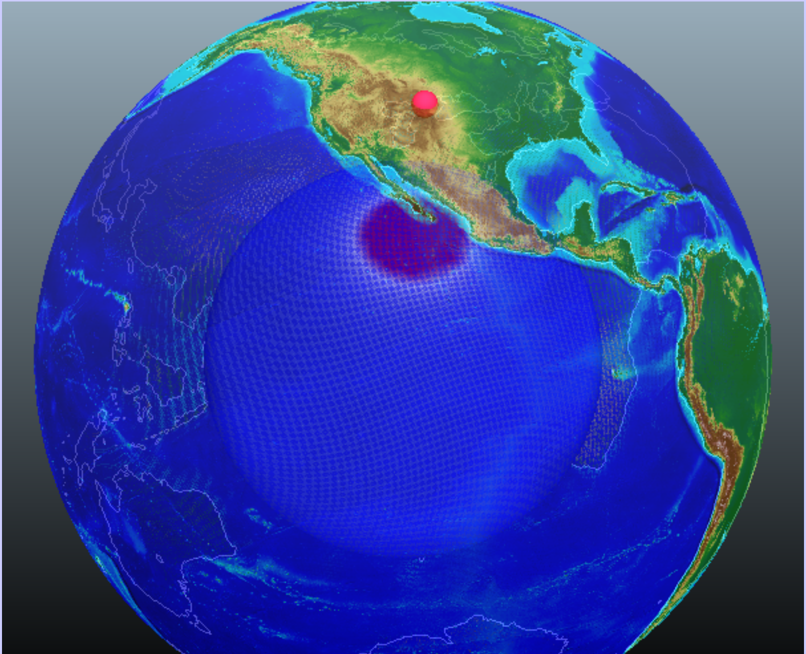
Core-Mantle boundary sensitivity



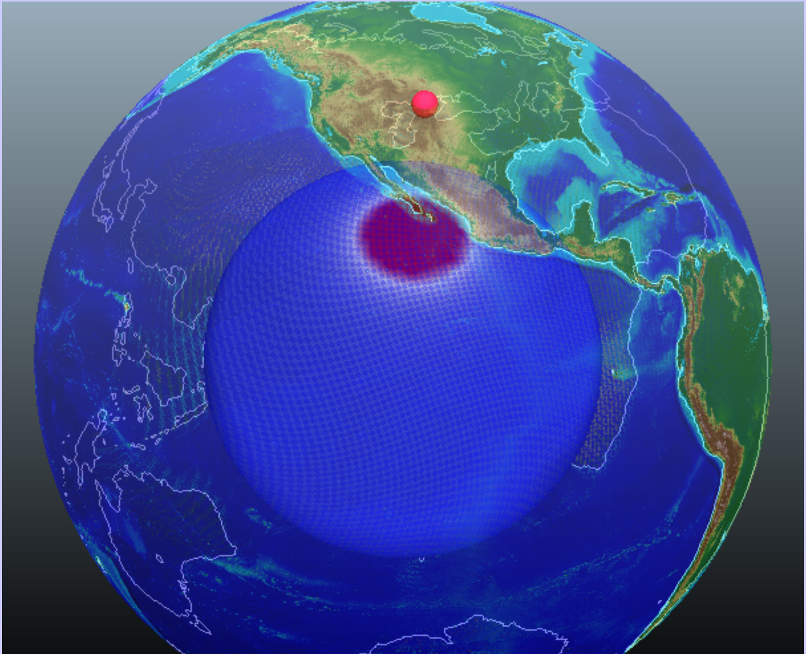
Core-Mantle boundary sensitivity



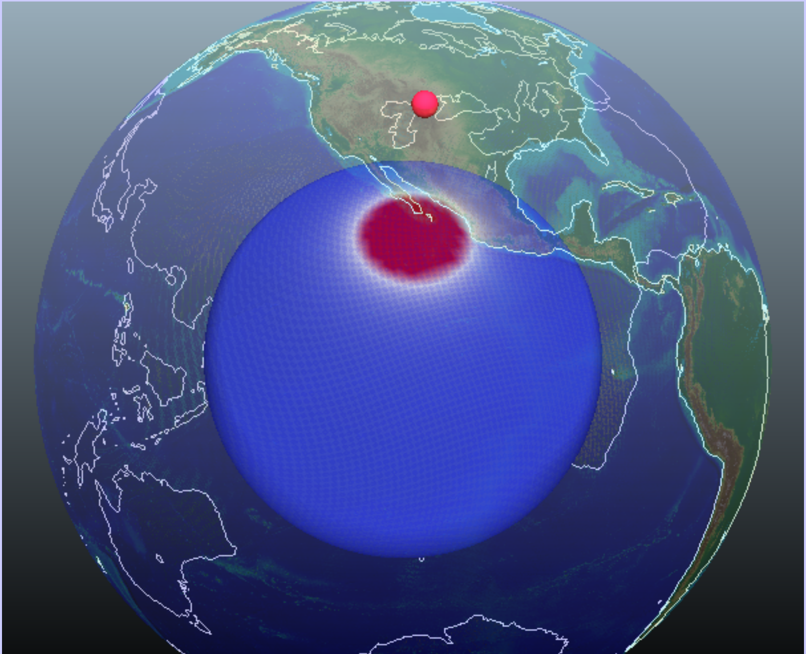
Core-Mantle boundary sensitivity



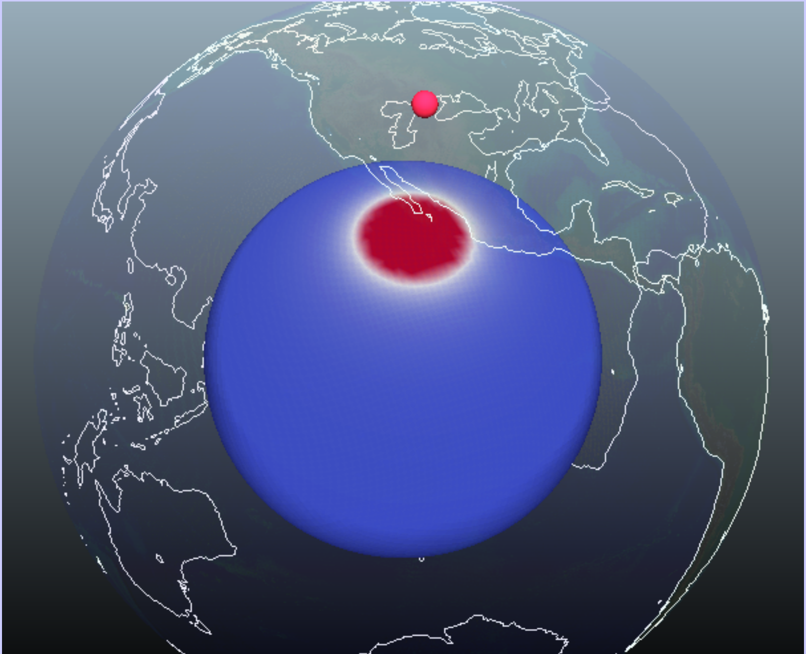
Core-Mantle boundary sensitivity



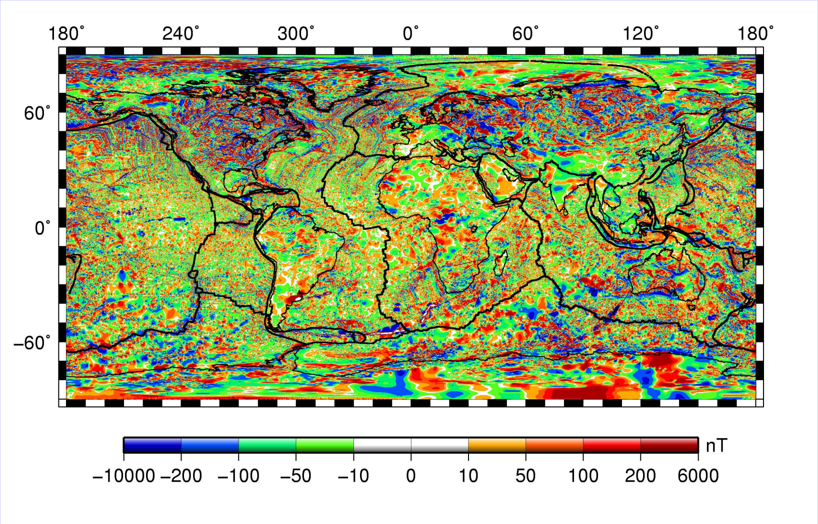
Core-Mantle boundary sensitivity



Core-Mantle boundary sensitivity



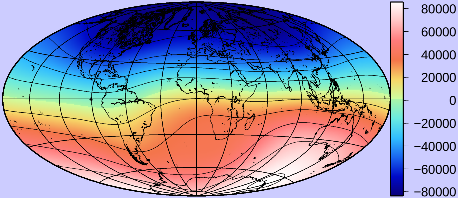
The lithosphere is magnetized



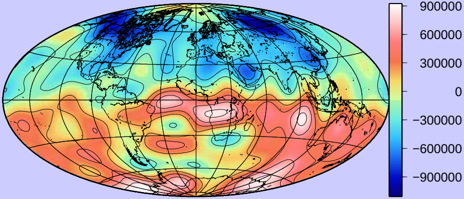
World Digital Magnetic Anomaly Map consortium

The Earth's main magnetic field

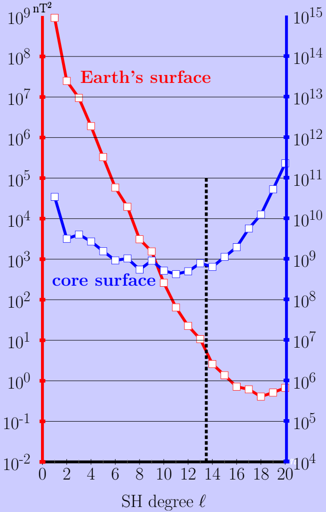
B_r (nT) at Earth's surface in 2007



B_r (nT) at the core surface in 2007



Mauersberger–Lowes spectrum



CHAOS2 model (1997-2009), Olsen et al., GJI, 2009

The main field

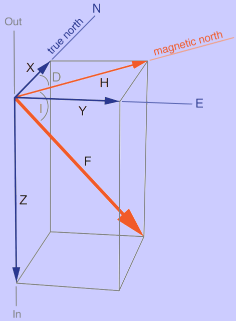
To take home

On interannual to millennial time scales, geomagnetic observations are connected with, and restricted to (by nature), the (large-scale) radial component of magnetic induction, B_r, at the core surface (the small scales are screened by the crustal field).
Large-scale: spherical harmonic degree $\ell \lesssim 13$ – lateral resolution of ~ 1500 km at the core surface. Geomagnetists are short-sighted.

The catalogs of data

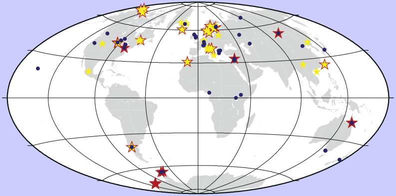
$\tau_{\text{conv}} \sim 150$ yr $\quad \tau_{\text{diff}} \sim 60,000$ yr

- Paleo-,archeomagnetism: $0 - 10(100, 1000+)$ kyr ago — D, I, F
- Mariners: $0 - 400$ yr ago — D, I
- Observatories: $0 - 150$ yr ago — X, Y, Z
- Satellites: $0 - 15$ yr ago — X, Y, Z



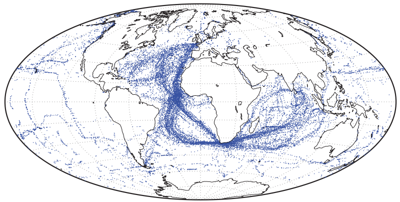
A heterogeneous record: spatial coverage (courtesy Chris Finlay)

archeo/paleo: 0 – 10+ kyr ago



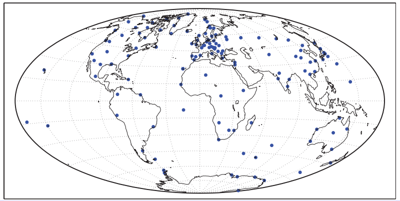
Locations of lake sediment records used to constrain the CALS10k model of **Korte et al. (EPSL, 2011)** spanning the past 10kyrs.

logbooks: 0 – 400 yr ago



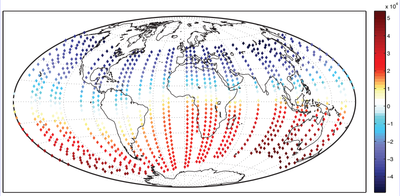
Locations of historical data (all components) between 1770 and 1790 from the **Jonkers et al. (Rev. Geophys., 2003)** database.

observatories: 0 – 150 yr ago

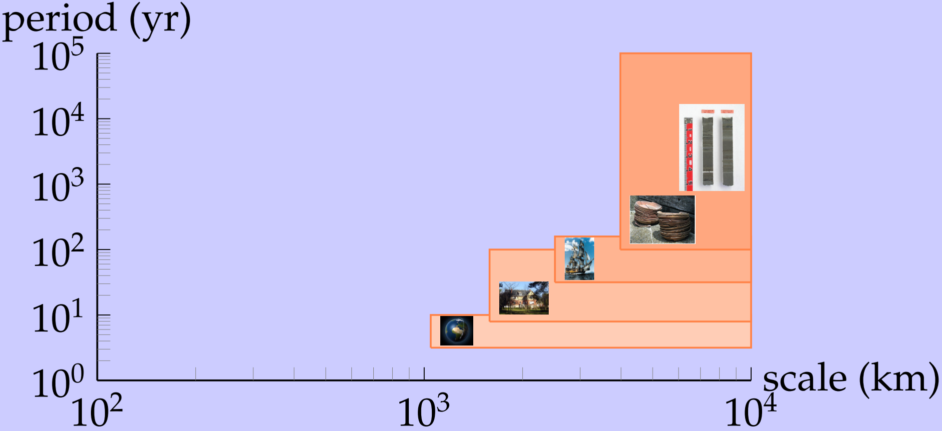


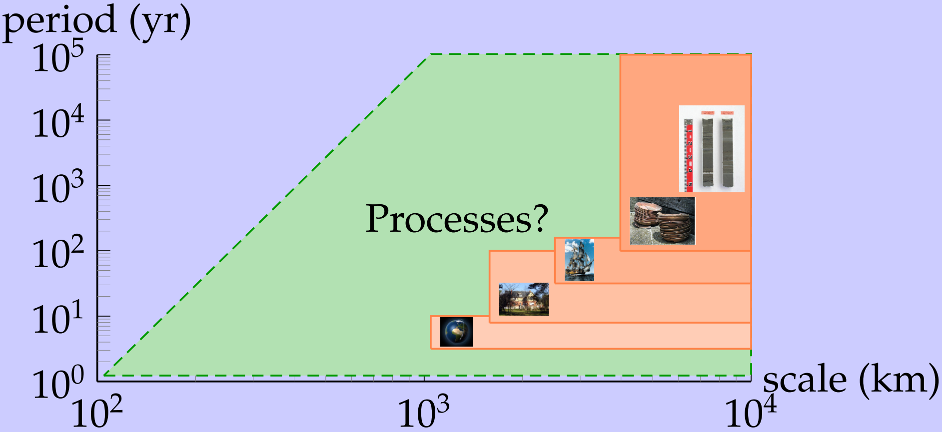
Locations of observatories used in determination of recent internal field models.

satellites: 0 – 15 yr ago



Example showing 3 days of CHAMP vector satellite data from 2009

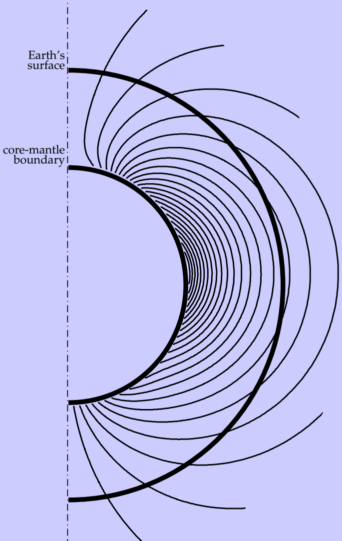

period (yr)
10^5
10^4
10^3
10^2
10^1
10^0
10^2
10^3
10^4
scale (km)


period (yr)
10^5
10^4
10^3
10^2
10^1
10^0
Processes?
scale (km)
10^2
10^3
10^4

Inference on the unseen

Another source of information: a physical model

A numerical model of core dynamics and the geodynamo.
Data assimilation: (data) && (prognostic numerical model)

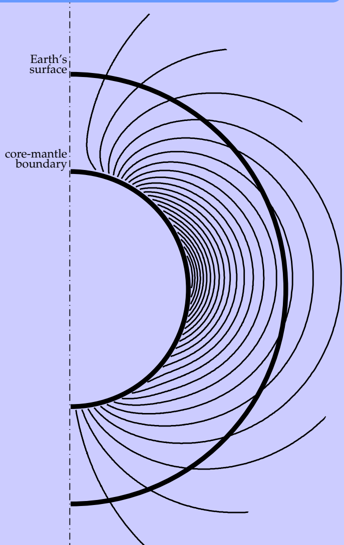


Inference on the unseen

Another source of information: a physical model

A numerical model of core dynamics and the geodynamo.
Data assimilation: (data) && (prognostic numerical model)

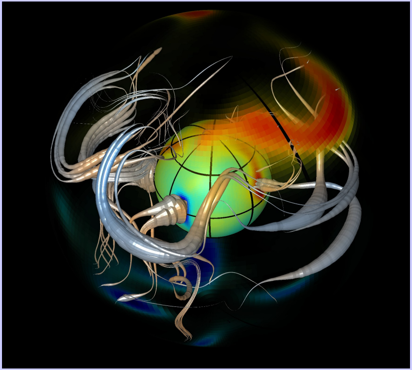
- Conservation laws (mass, momentum, energy) and Maxwell's equations (MHD approximation)
- Set of 3D non-linear coupled PDEs to solve in a spherical shell (the FOC)
- Boussinesq (Dissipation number = 0.4 in the FOC)
- Pseudo-spectral method (Glatzmaier, 1984)

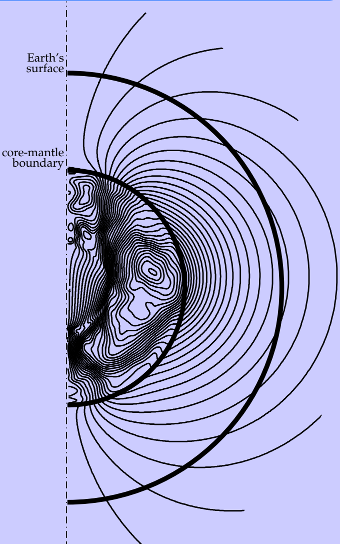


Inference on the unseen

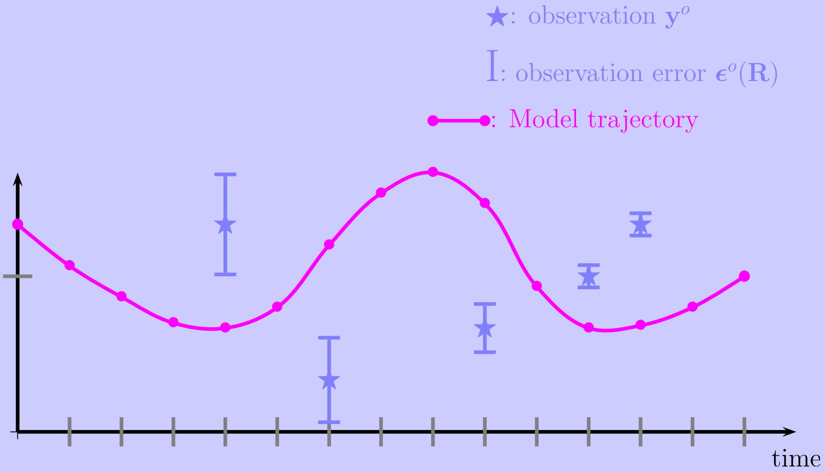
Another source of information: a physical model

A numerical model of core dynamics and the geodynamo.
Data assimilation: (data) && (prognostic numerical model)

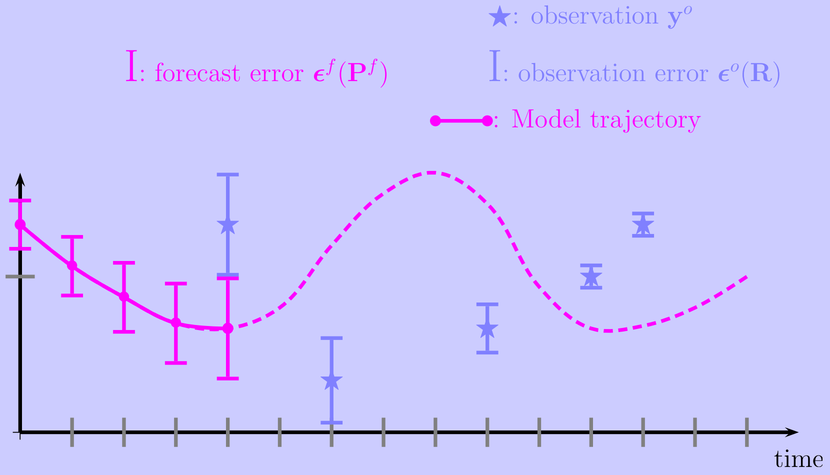




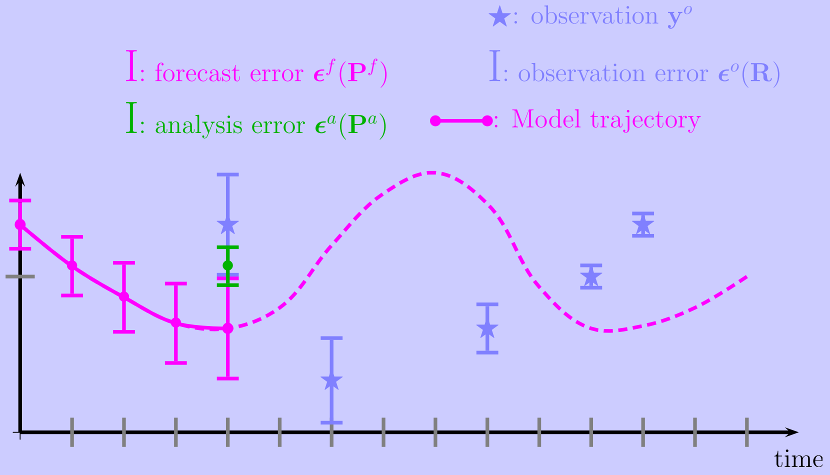
Trajectory in model space



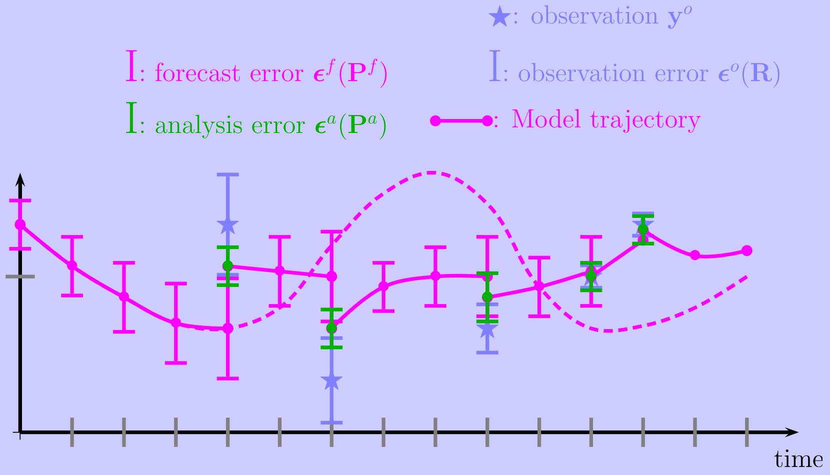
Sequential assimilation



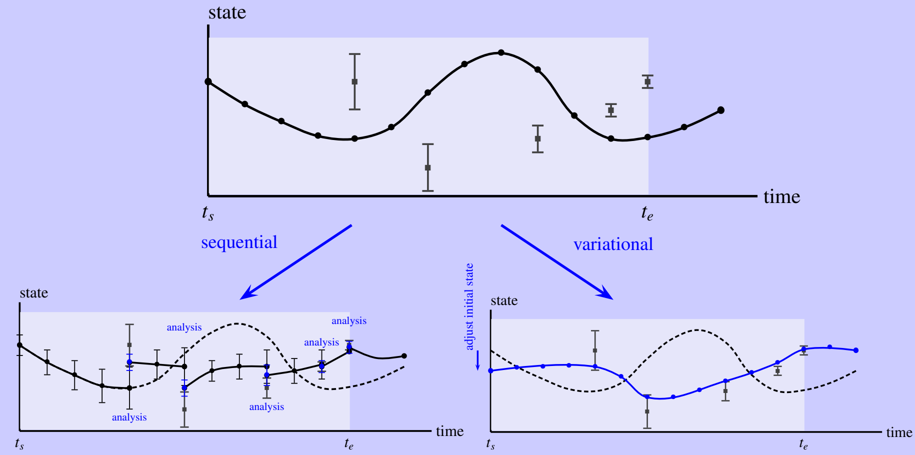
Sequential assimilation



Sequential assimilation



Sequential or variational?



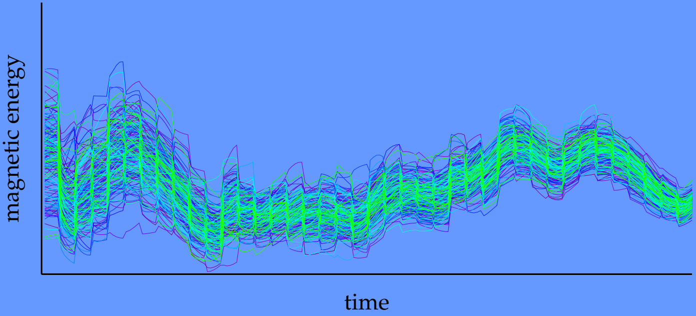
2. Application of the EnKF to a numerical model of the geodynamo

The ensemble Kalman filter: Principle

This is a sequential assimilation method, applied here to three-dimensional, convection-driven, numerical dynamo model.

Concept (Evensen, 1994)

- Have an ensemble of dynamical states evolve concurrently
- Use this ensemble to generate (on-the-fly) the statistics needed for the analysis of the stream of observations



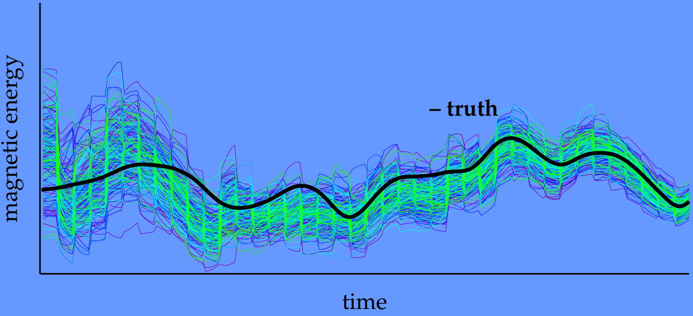
- Well-suited for moderatly nonlinear problems, and more adaptive than optimal interpolation (OI), which assumes frozen forecast error statistics (Kuang et al., 2008, 2009; Aubert & Fournier, 2011).

The ensemble Kalman filter: Principle

This is a sequential assimilation method, applied here to three-dimensional, convection-driven, numerical dynamo model.

Concept (Evensen, 1994)

- Have an ensemble of dynamical states evolve concurrently
- Use this ensemble to generate (on-the-fly) the statistics needed for the analysis of the stream of observations



- Well-suited for moderatly nonlinear problems, and more adaptive than optimal interpolation (OI), which assumes frozen forecast error statistics (Kuang et al., 2008, 2009; Aubert & Fournier, 2011).

The ensemble Kalman filter: Implementation

- The starting dynamo code:
  a modified (more modular) version of the `PARODY` code (Dormy et al., 1998; Aubert et al., 2008).
  + SHTns (Schaeffer, G^3, 2013).
- The EnKF layer:
  a suitably modified version of the Parallel Data Assimilation Framework of Nerger & Hiller (2013).


Contents lists available at SciVerse ScienceDirect
Computers & Geosciences
journal homepage: www.elsevier.com/locate/cageo

Software for ensemble-based data assimilation systems—Implementation strategies and scalability

Lars Nerger*, Wolfgang Hiller
Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, D-27570 Bremerhaven, Germany

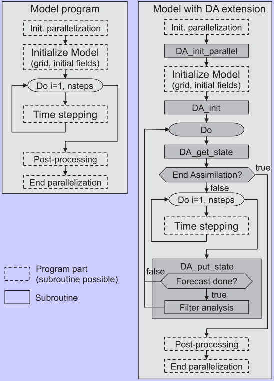


Fig. 1. Left: Flow diagram of a typical numerical model. Right: Flow diagram of the model extended to an assimilation system by calls to routines of the assimilation framework. (Based on Nerger et al., 2005b.)

Nerger & Hiller (2013)

(almost) Embarassingly parallel setup: Efficiency $>$ 99 % on 1,440 cores.

Closed-loop (twin) experiments with the EnKF

Fournier, Nerger, Aubert (G-cubed, 2013)

Outline:

We solve a time-dependent assimilation problem:

- Generate synthetic data from 1 model free run over $\sim 3,000$ yr: SH representation of maps of B_r at the core surface (**truncated at** $\ell = 13$ **or** $\ell = 5$, with noise added, **diagonal** error covariance matrix)
- Start from t_0 using a initial ensemble (N_e random samples from model free run, outside observation window) **Each member: same control parameters (same as data), different initial condition**
- Assimilate synthetic observations every 25 yr for 3,000 yr to correct the trajectory of each member of the ensemble; let go for another 500 years
- Assess quality of assimilation scheme by comparing the known 'true' dynamo state $\mathbf{x}^t$ and the estimate $\widehat{\mathbf{x}}$ (the ensemble average)

$$\widehat{\mathbf{x}} = \frac{1}{N_e} \sum_{e=1}^{N_e} \mathbf{x}_e$$

  - Retrieval of internal structure
  - Forecast quality (see what happens after the stream of observations stops)

We use a rather 'simple', low-resolution geodynamo model, $90 \times 64 \times 64$. (Size of the state vector: 10^6)

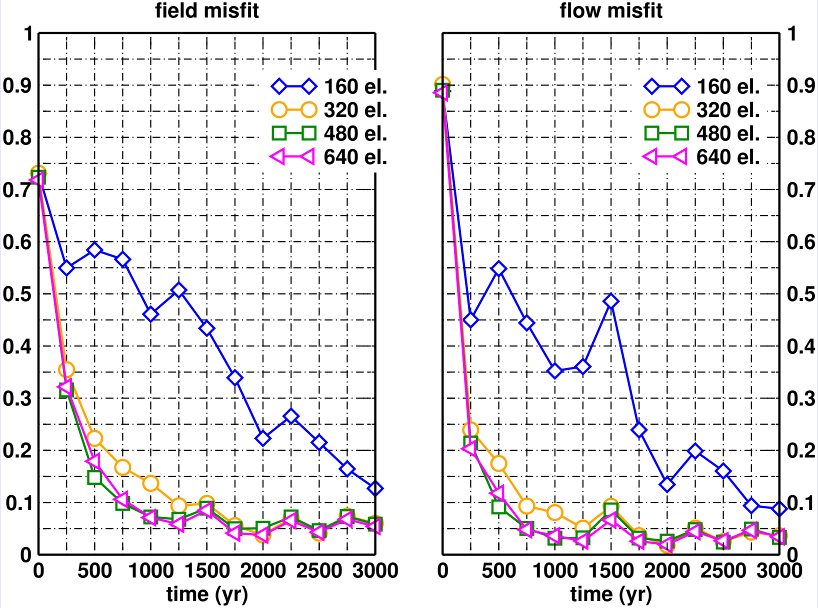
Ensemble size

Minimum ensemble size for converged statistics and proper behaviour of filter?

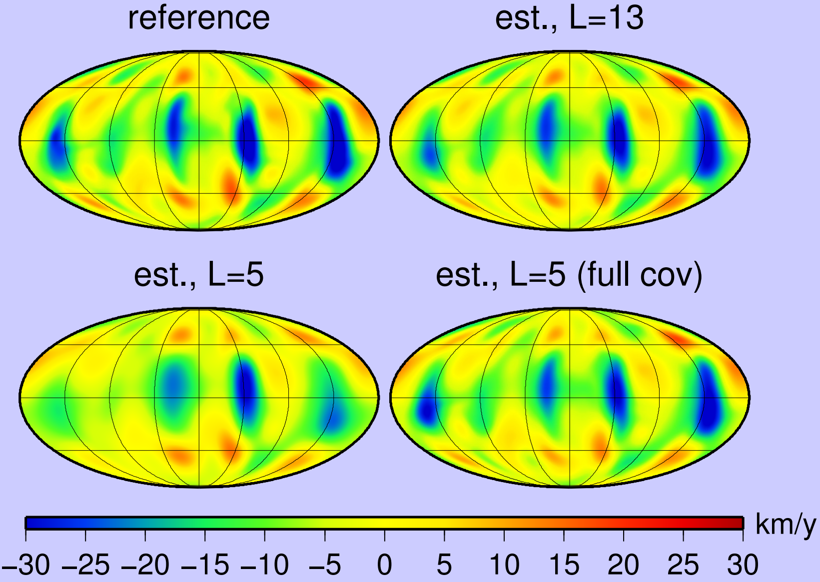
- Surface data truncation $\ell = 13$
- We measure the distance between the estimate $\widehat{\mathbf{x}}$ and the synthetic truth $\mathbf{x}^t$ with normalized quadratic misfits

$$\text{field misfit} = \frac{\int \left(\widehat{\boldsymbol{B}} - \boldsymbol{B}^t\right)^2 \mathrm{d}V}{\int (\boldsymbol{B}^t)^2 \mathrm{d}V}, \qquad \text{flow misfit} = \frac{\int \left(\widehat{\boldsymbol{u}} - \boldsymbol{u}^t\right)^2 \mathrm{d}V}{\int (\boldsymbol{u}^t)^2 \mathrm{d}V},$$

Ensemble size



Retrieval of internal structure: u_φ at mid-depth after 1000 yr



Behavior of the axial dipole

red: truth blue: EnKF estimate