

# Datenassimilation

## ein Überblick

Timur Olzhabaev

Arbeitsbereich Wissenschaftliches Rechnen  
Fachbereich Informatik  
Fakultät für Mathematik, Informatik und Naturwissenschaften  
Universität Hamburg

21. Januar 2013

# Übersicht

Problemstellung

Definitionen

Theorie

Anwendung

Aussicht

# Problemstellung



Wir möchten eine Vorhersage treffen...

Wir möchten eine Vorhersage treffen...

...also modellieren wir das System...

Wir möchten eine Vorhersage treffen...

...also modellieren wir das System...

...und implementieren eine Simulation.

Wir möchten eine Vorhersage treffen...

...also modellieren wir das System...

...und implementieren eine Simulation.

Was ist der Startzustand?



Wir möchten eine Vorhersage treffen...

...also modellieren wir das System...

...und implementieren eine Simulation.

Was ist der Startzustand?

⇒ wir müssen Beobachtungen hinzuziehen!

Beobachtungen sind...

Beobachtungen sind...

...teuer,

Beobachtungen sind...

...teuer,

...fehlerhaft,

Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Modelle sind...

Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Modelle sind...

...teuer,

Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Modelle sind...

...teuer,

...fehlerhaft,



Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Modelle sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Modelle sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Wie füllen wir die Datenlücken?

Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Modelle sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Wie füllen wir die Datenlücken?

Wie korrigieren wir das Modell?

Beobachtungen sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Modelle sind...

...teuer,

...fehlerhaft,

... und unvollständig.

Wie füllen wir die Datenlücken?

Wie korrigieren wir das Modell?

⇒ Datenassimilation!

# Definitionen

*“Data assimilation is the process by which observations are incorporated into a computer model of a real system. Applications of data assimilation arise in many fields of geosciences, perhaps most importantly in weather forecasting and hydrology. Data assimilation proceeds by **analysis cycles**. In each analysis cycle, observations of the current (and possibly past) state of a system are combined with the results from a numerical weather prediction mode (**the forecast**) to produce an **analysis**, which is considered as ‘the best’ estimate of the current state of the system. This is called the **analysis step**. Essentially, the analysis step tries to balance the uncertainty in the data and in the forecast. The model is then advanced in time and its result becomes the forecast in the next analysis circle.”*

[http://en.wikipedia.org/wiki/Data\\_assimilation](http://en.wikipedia.org/wiki/Data_assimilation)

*“Data assimilation is the process by which **observations are incorporated into a computer model of a real system**. Applications of data assimilation arise in many fields of geosciences, perhaps most importantly in weather forecasting and hydrology. Data assimilation proceeds by **analysis cycles**. In each analysis cycle, observations of the current (and possibly past) state of a system are combined with the results from a numerical weather prediction mode (**the forecast**) to produce an **analysis**, which is considered as ‘the best’ estimate of the current state of the system. This is called the **analysis step**. Essentially, the analysis step tries to balance the uncertainty in the data and in the forecast. The model is then advanced in time and its result becomes the forecast in the next analysis circle.”*

[http://en.wikipedia.org/wiki/Data\\_assimilation](http://en.wikipedia.org/wiki/Data_assimilation)

*“Data assimilation is the process by which **observations are incorporated into a computer model of a real system**. Applications of data assimilation arise in many fields of geosciences, perhaps most importantly in weather forecasting and hydrology. Data assimilation proceeds by **analysis cycles**. In each analysis cycle, **observations of the current (and possibly past) state of a system are combined with the results from a numerical weather prediction mode (the forecast) to produce an analysis, which is considered as ‘the best’ estimate of the current state of the system**. This is called the **analysis step**. Essentially, the analysis step tries to balance the uncertainty in the data and in the forecast. The model is then advanced in time and its result becomes the forecast in the next analysis circle.”*

[http://en.wikipedia.org/wiki/Data\\_assimilation](http://en.wikipedia.org/wiki/Data_assimilation)



*“The data assimilation methodology adds value to the observations by filling in the observational gaps and to the model by constraining it with observations. In this way, the data assimilation allows one to ‘make sense’ of the observations.”*

Data Assimilation - Making Sense of Observations, W. Lahoz, B. Khattatov, R. Menard

*“The data assimilation methodology adds value to the observations by filling in the observational gaps and to the model by constraining it with observations. In this way, the data assimilation allows one to ‘make sense’ of the observations.”*

Data Assimilation - Making Sense of Observations, W. Lahoz, B. Khattatov, R. Menard

Datenassimilationsverfahren kombinieren Modell und Beobachtungen um ein möglichst genaues Abbild der Realität zu erhalten.

# Theorie: Grundgerüst

# Grundbegriffe

Systemzustand

# Grundbegriffe

## Systemzustand

$\mathbf{x}$  *state vector*

$\mathbf{x}^t$  *true state*

$\mathbf{x}^b$  *background state*

$\mathbf{x}^a$  *analysed state*

# Grundbegriffe

## Systemzustand

$\mathbf{x}$  *state vector*

$\mathbf{x}^t$  *true state*

$\mathbf{x}^b$  *background state*

$\mathbf{x}^a$  *analysed state*

## Beobachtungen

# Grundbegriffe

## Systemzustand

$\mathbf{x}$  *state vector*

$\mathbf{x}^t$  *true state*

$\mathbf{x}^b$  *background state*

$\mathbf{x}^a$  *analysed state*

## Beobachtungen

$\mathbf{y}$  *observational space*

$\mathbf{y}^o$  *observation vector*

$\epsilon^o = \epsilon^f + \epsilon^u$  *observation error*



# Beobachtungsgleichung

$$\mathbf{y} = h(\mathbf{x})$$

# Beobachtungsgleichung

$$\mathbf{y} = h(\mathbf{x})$$

$$\mathbf{y}^o = h(\mathbf{x}) + \epsilon^\mu$$

# Beobachtungsgleichung

$$\mathbf{y} = h(\mathbf{x})$$

$$\mathbf{y}^o = h(\mathbf{x}) + \epsilon^\mu$$

$$\mathbf{y}^o = \mathcal{H}(\mathbf{x}^t) + \epsilon^o$$

# Beobachtungsgleichung

$$\mathbf{y} = h(\mathbf{x})$$

$$\mathbf{y}^o = h(\mathbf{x}) + \epsilon^\mu$$

$$\mathbf{y}^o = \mathcal{H}(\mathbf{x}^t) + \epsilon^o$$

$$\mathbf{y}^o = \mathcal{H}(\mathbf{x}^b) + \epsilon^o + \epsilon^b$$

# Modellgleichung

# Modellgleichung

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k)$$

# Modellgleichung

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k)$$

$$\mathbf{x}_{k+1}^t = \mathcal{M}_{k,k+1}(\mathbf{x}_k^t) + \eta_{k,k+1}$$

# Modellgleichung

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k)$$

$$\mathbf{x}_{k+1}^t = \mathcal{M}_{k,k+1}(\mathbf{x}_k^t) + \eta_{k,k+1}$$

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$



# Modellgleichung

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k)$$

$$\mathbf{x}_{k+1}^t = \mathcal{M}_{k,k+1}(\mathbf{x}_k^t) + \eta_{k,k+1}$$

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$

Zu bestimmen: Analyse  $\mathbf{x}^a$

# Modellgleichung

$$\mathbf{x}_{k+1} = g(\mathbf{x}_k)$$

$$\mathbf{x}_{k+1}^t = \mathcal{M}_{k,k+1}(\mathbf{x}_k^t) + \eta_{k,k+1}$$

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$

Zu bestimmen: Analyse  $\mathbf{x}^a$

Allgemeine Annahme:  $\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$

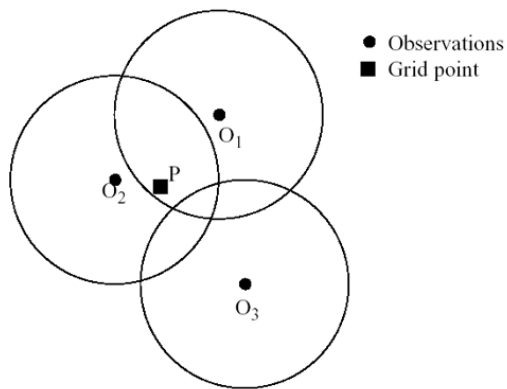
# Theorie: Verfahren

# Verfahren

- Sukzessive Korrektur
- Nudging
- Optimale Interpolation
- Variationsverfahren
- Kalman Filter

# Cressman's objective analysis

# Cressman's objective analysis



Observations  $O_1$  and  $O_2$  influence grid point  $P$ ,  $O_3$  does not.

Quelle: National Centre for Earth Observation (UK) <http://www.nceo.ac.uk/>

# Cressman's objective analysis

# Cressman's objective analysis

$$\mathbf{x}_j^a = \mathbf{x}_j^b + \frac{\sum_{i=1}^S w(i,j)(\mathbf{y}_i - \mathbf{x}_i^b)}{\sum_{i=1}^S w(i,j)}$$



# Cressman's objective analysis

$$\mathbf{x}_j^a = \mathbf{x}_j^b + \frac{\sum_{i=1}^S w(i,j)(\mathbf{y}_i - \mathbf{x}_i^b)}{\sum_{i=1}^S w(i,j)}$$

$$w(i,j) = \begin{cases} \frac{R^2 - r(i,j)^2}{R^2 + r(i,j)^2} & \text{wenn } r(i,j) \leq R \\ 0 & \text{wenn } r > R \end{cases}$$

# Nudging

# Nudging

$$\frac{\delta \mathbf{x}}{\delta t} = \mathcal{M}(\mathbf{x}) + \alpha(\mathbf{y} - \mathbf{x})$$

# Optimale Interpolation

# Optimale Interpolation

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

# Optimale Interpolation

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

Annahme:  $\mathcal{H}$  ist linear  $\Rightarrow \mathcal{H} = \mathbf{H}$

# Optimale Interpolation

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

Annahme:  $\mathcal{H}$  ist linear  $\Rightarrow \mathcal{H} = \mathbf{H}$

Zu berechnen:  $\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$

# Optimale Interpolation

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

Annahme:  $\mathcal{H}$  ist linear  $\Rightarrow \mathcal{H} = \mathbf{H}$

Zu berechnen:  $\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$

**B**: Kovarianzmatrix des Backgrounds (hier konstant!)

**R**: Kovarianzmatrix der Beobachtungen



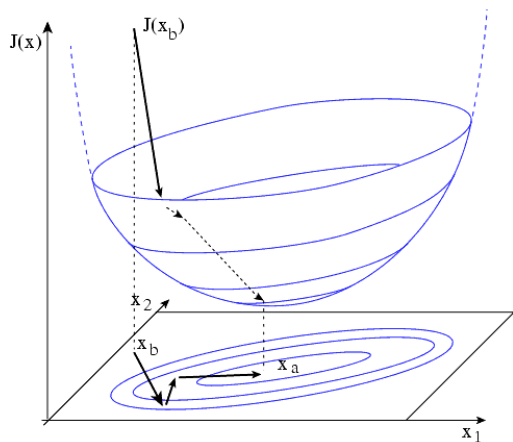
# 3D-Var

$$\mathbf{x}^a = \arg \min \mathcal{J}$$

$$\mathcal{J}(\mathbf{x}) = (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + (\mathbf{y}^o - \mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}))$$

$$\nabla \mathcal{J} = 2\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - 2\mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}))$$

# 3D-Var

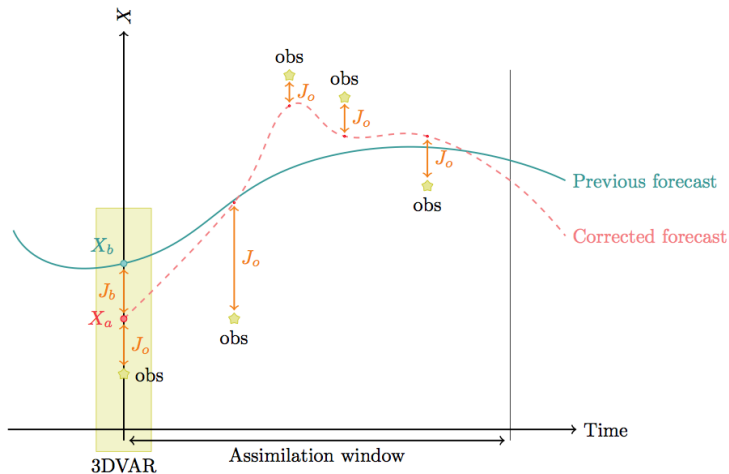


Quelle: European Centre for Medium-Range Weather Forecasts <http://www.ecmwf.int/>

# 4D-Var

$$\mathcal{J}(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \sum_{i=0}^n (\mathbf{y}_i^o - \mathbf{H}_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{y}_i^o - \mathbf{H}_i(\mathbf{x}_i))$$

## 4D-Var



Quelle: Introduction to Data Assimilation (2011), E. Blayo, E. Cosme, M. Nodet, A. Vidard

# Extended Kalman Filter

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) + \eta_{k,k+1}$$

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

---

$$\mathbf{x}_0^a = \mathbf{x}^b, \mathbf{P}_0^a = \mathbf{B}$$

$$\mathbf{x}_{k+1}^f = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a)$$

$$\mathbf{P}_{k+1}^f = \mathbf{M}\mathbf{P}_k^a\mathbf{M}^T + \mathbf{Q}$$

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^T(\mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R})^{-1}$$

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^f$$

# Ensemble Kalman Filter

Erweiterung des EKF mit mehreren simultan berechneten Assimilationspfaden

# Anwendung

# Anwendungsgebiete



# Anwendungsgebiete

- Meteorologie
  - Numerische Wettervorhersage

# Anwendungsgebiete

- Meteorologie
  - Numerische Wettervorhersage
- Meereskunde

# Anwendungsgebiete

- Meteorologie
  - Numerische Wettervorhersage
- Meereskunde
- Mathematische Geophysik

# Anwendungsgebiete

- Meteorologie
  - Numerische Wettervorhersage
- Meereskunde
- Mathematische Geophysik
- Wirtschaft (Preisvorhersage)

# Anwendungsgebiete

- Meteorologie
  - Numerische Wettervorhersage
- Meereskunde
- Mathematische Geophysik
- Wirtschaft (Preisvorhersage)
- Bioinformatik

# Numerische Wettervorhersagen

# Numerische Wettervorhersagen

...sind rechnergestützte Wettervorhersagen

# Numerische Wettervorhersagen

...sind rechnergestützte Wettervorhersagen

...basierend auf mathematischen Modellen der Erde



# Numerische Wettervorhersagen

...sind rechnergestützte Wettervorhersagen

...basierend auf mathematischen Modellen der Erde

...und aktuellen Beobachtungen.

# Modelle

# Modelle

...basieren auf Gleichungen,

# Modelle

...basieren auf Gleichungen,

...die analytisch nicht zu lösen sind,

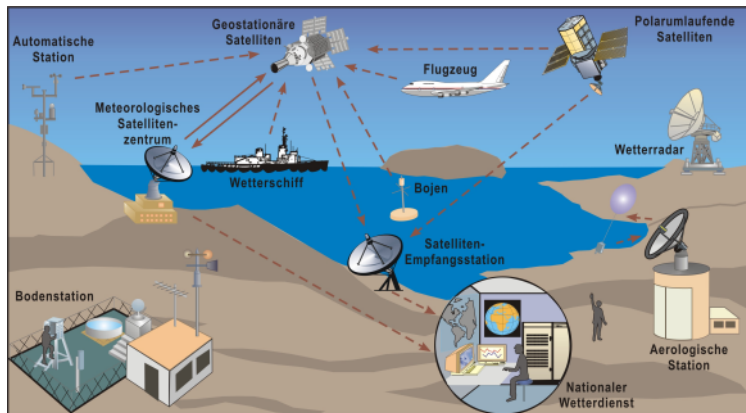
# Modelle

...basieren auf Gleichungen,

...die analytisch nicht zu lösen sind,

...und daher numerisch approximativ gelöst werden.

# Beobachtungen

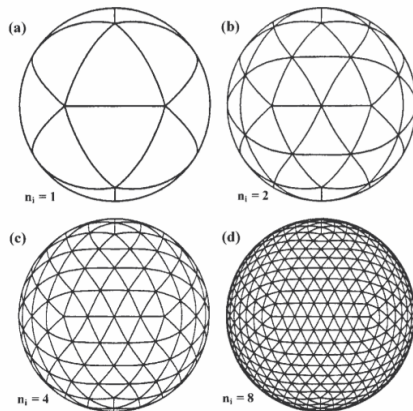


Quelle: Deutscher Wetterdienst

[www.dwd.de](http://www.dwd.de)

# Globalmodell Europa (GME)

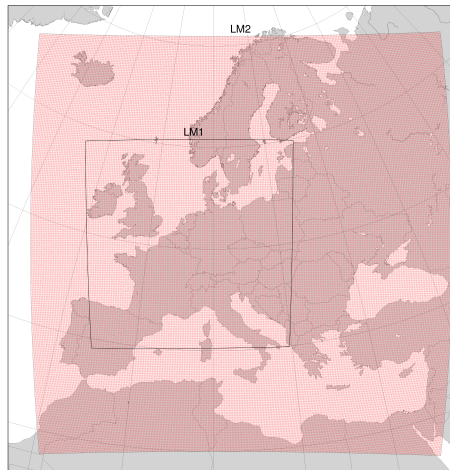
- seit 1999 im Einsatz
- Assimilation durch Optimale Interpolation
- Dreiecksgitter (20 km Seitenlänge)
- $1474550 + 12 = 1474562$  Gitterpunkte
- 3 Stunden Kurzfristvorhersage für Assimilation
- 4 Modellrechnungen am Tag



Quelle: Das Global-Modell GME, D. Majewski, B. Ritter

# COSMO-EU

- seit 1999 im Einsatz
- Assimilation durch Nudging
- $665 \cdot 657 = 436905$   
Gitterpunkte
- 40 Atmosphäreschichten (bis 24 km)
- ⇒  $436905 \cdot 40 = 17.5$  Millionen  
Gitterpunkte
- 3 Stunden Zeitintervall seit 2005
- Parameter pro Schicht:  
Luftdruck, Temperatur und  
Luftfeuchtigkeit



Quelle: Atmacs - Atmospheric attraction computation service  
<http://atmacs.bkg.bund.de/>  
[www.dwd.de](http://www.dwd.de)



# Aussicht

# Aussicht

Heute: **OI** und **3D-Var**

# Aussicht

Heute: **OI** und **3D-Var**

Morgen: **KF** und **4D-Var**

# Aussicht

Heute: **OI** und **3D-Var**

Morgen: **KF** und **4D-Var**

Bessere Daten = größte Qualitätsverbesserung

# Mars

## Atmosphäre

- Kohlenstoffdioxid: 95,32%
- Stickstoff: 2,7%
- Argon: 1,6%
- Sauerstoff: 0,13%
- Andere: < 0,025 %

Druck:  $6 \cdot 10^{-3}$  bar (0,63% der Erdatmosphäre)

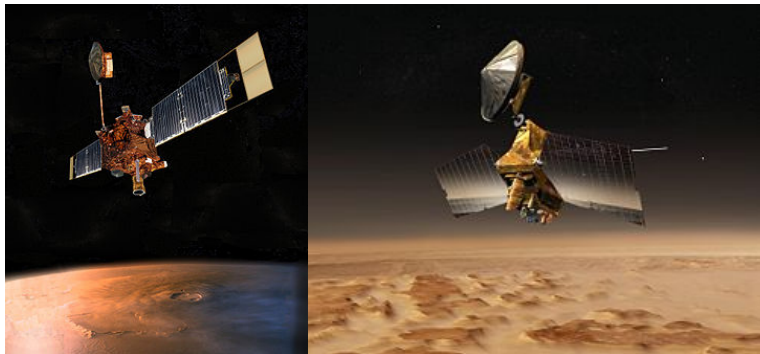
## Temperaturen

- Min: -133 °C
- Mittel: -55 °C
- Max: +27 °C



Quelle: [http://en.wikipedia.org/wiki/Atmosphere\\_of\\_Mars](http://en.wikipedia.org/wiki/Atmosphere_of_Mars)

# Mars Global Surveyor & Mars Reconnaissance Orbiter



Quelle links: [http://en.wikipedia.org/wiki/Mars\\_Global\\_Surveyor](http://en.wikipedia.org/wiki/Mars_Global_Surveyor)

Quelle rechts: [http://en.wikipedia.org/wiki/Mars\\_Reconnaissance\\_Orbiter](http://en.wikipedia.org/wiki/Mars_Reconnaissance_Orbiter)

# Mars Wettervorhersage

- Vorhersage zu Wissenschaftlichen Zwecken
- Daten: Temperatur, Staubbichte
- Assimilation durch sukzessive Korrektur und EnKF

# Zusammenfassung



# Zusammenfassung

Beobachtungen + Modell = gute Schätzung des Zustandes

# Zusammenfassung

Beobachtungen + Modell = gute Schätzung des Zustandes  
Datenassimilation bietet 'theoretisches Toolset'

# Zusammenfassung

Beobachtungen + Modell = gute Schätzung des Zustandes

Datenassimilation bietet 'theoretisches Toolset'

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k (\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

# Zusammenfassung

Beobachtungen + Modell = gute Schätzung des Zustandes

Datenassimilation bietet 'theoretisches Toolset'

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

Hauptanwendung in Numerischer Wettervorhersage

# Zusammenfassung

Beobachtungen + Modell = gute Schätzung des Zustandes

Datenassimilation bietet 'theoretisches Toolset'

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

Hauptanwendung in Numerischer Wettervorhersage

Im Einsatz: **OI**, **Nudging** und **3D-Var**

# Zusammenfassung

Beobachtungen + Modell = gute Schätzung des Zustandes

Datenassimilation bietet 'theoretisches Toolset'

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^b))$$

Hauptanwendung in Numerischer Wettervorhersage

Im Einsatz: **OI**, **Nudging** und **3D-Var**

Kommend: **4D-Var** und **Kalman-Filter**

# Quellen

Data Assimilation - Making Sense of Observations - W. Lahoz, B. Khattatov, R. Menard

Introduction to Data Assimilation - E. Blayo, E. Cosme, M. Nodet, A. Vidard

Data Assimilation Concepts and Methods - F. Bouttier, P. Courtier

Data Assimilation tutorial on the Kalman filter - A. Fowler

[http://www.atmos.millersville.edu/~lead/Obs\\_Data\\_Assimilation.html](http://www.atmos.millersville.edu/~lead/Obs_Data_Assimilation.html)

<http://www.dwd.de/>

<http://atmacs.bkg.bund.de/>

[http://en.wikipedia.org/wiki/Atmosphere\\_of\\_Mars](http://en.wikipedia.org/wiki/Atmosphere_of_Mars)