united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

Course on "Inverse Methods in Atmospheric Science"
1 - 12 October 2001

301/1332-15

"Assimilation of Remote Sensing Data in Numerical Weather Prediction"

A. MCNALLY
European Centre for Medium-range Weather Forecasts
Reading, England

Assimilation of Remote sensing data in Numerical Weather Prediction

A. McNally

European Centre for Medium-range Weather Forecasts

Course on:

Inverse Methods in Atmospheric Science
(Trieste, October 2001)

Overview

- Key elements of an NWP system
  - Forecast model
  - observations
  - data assimilation
- Satellite data used in NWP
  - sounding data
  - surface (window) data
  - active data
- Data assimilation systems
  - optimal interpolation (retrievals)
  - Variational (3D/4D) methods (direct radiance assimilation)
- Research issues
  - background error covariances
  - systematic error
  - treatment of cloud and the surface

ECMWF:

- A European organisation with headquarters in the UK
- Established by Convention in force from November 1975
- Principal objectives:
  - development of methods for forecasting weather beyond two days ahead
  - collection and storage of appropriate meteorological data
  - daily production and distribution of forecasts to the Member States
  - provision of archival/retrieval facilities to the Member States
  - provision of computational resources to the Member States
- Staff of about 200

Member States:

Belgium
Denmark
Germany
Spain
France
Greece
Ireland
Italy
The Netherlands
Norway
Austria
Portugal
Switzerland
Finland
Sweden
Turkey
United Kingdom

Co-operation agreements with:

Croatia
Czech Republic
Iceland
Hungary
Slovenia

ECMWF activities

- Medium-range forecasts of the state of the atmosphere, land and ocean-waves to ten days ahead
  - Deterministic (single high-resolution forecast)
  - Probabilistic (ensemble of perturbed lower-resolution forecasts)
- Boundary conditions (initial conditions) for Member States' short-range regional forecasting systems
- Seasonal forecasts (including ocean circulation) to six months ahead
- Re-analyses of historical observations (for climate applications)

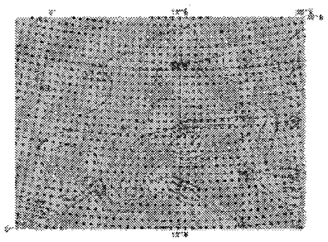
Key elements of the NWP system

- The **forecast model** time evolves fields of geophysical parameters (e.g. $T/Q/U/V/O_3$) following the laws of thermodynamics and chemistry
- The initial conditions used to start the **forecast model** are provided by the **analysis**
- The **analysis** is generated from **observations** relating to the geophysical parameters combined with *a priori* **background information** (usually a short-range forecast from the previous analysis).
- This combination process is known as **data assimilation**

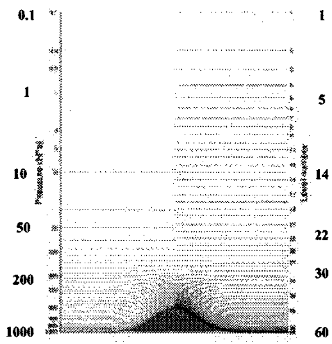
The ECMWF forecast model (1)

Global spectral primitive equation model

Horizontal resolution
T_L511 ~ 40 km

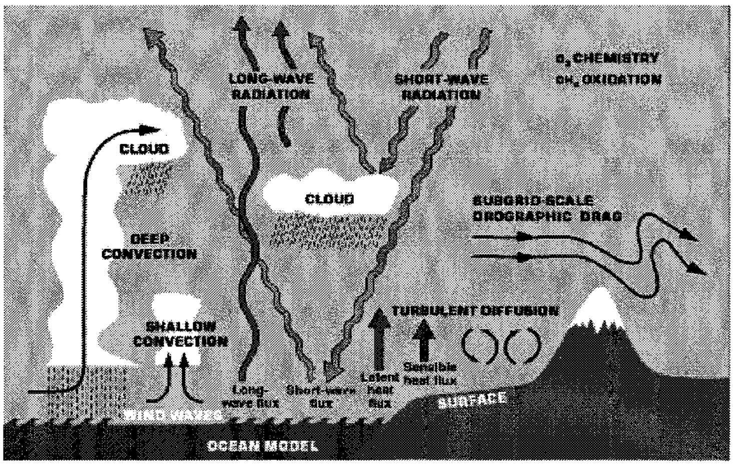


Vertical resolution
60 hybrid sigma levels

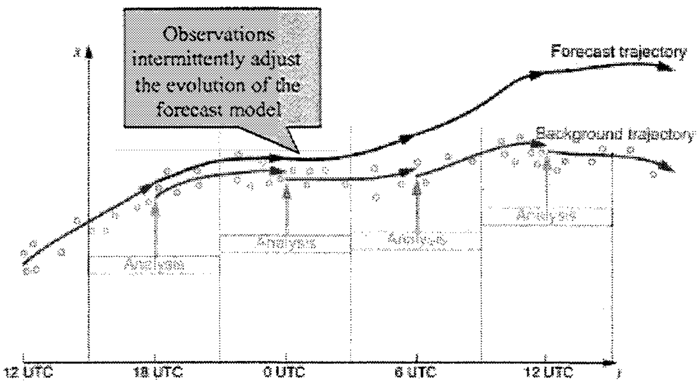


12 levels below 850 hPa

The ECMWF forecast model (2)



The Data Assimilation Process



Observations Used in NWP

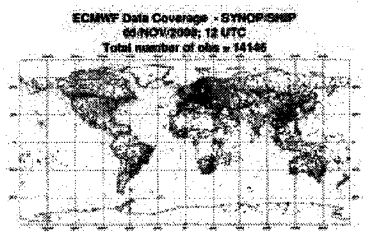
***In situ* (conventional)**

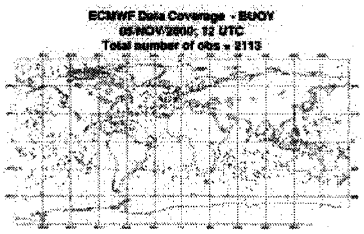
- SYNOP(surface)
  - Ps, Wind-10m, RH-2m
- AIREP
  - Wind, Temp
- DRIBU(drifting buoy)
  - Ps, Wind-10m
- TEMP(balloon))
  - Wind, Temp, Spec Humidity
- DROPSONDE
  - Wind, Temp
- PILOT/Profiler
  - Wind
- PAOB
  - Ps

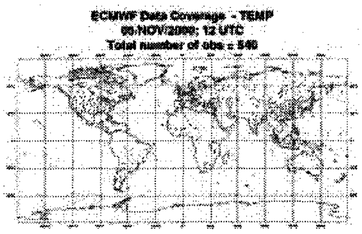
Remotely sensed (satellite)

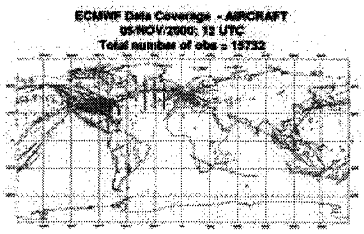
- Polar orbiting platforms
  - HIRS
  - MSU
  - AMSU-A / B
  - SSU
  - SSM/I(S)
  - QuickScat
  - ERS-scat
  - AIRS (soon)
- Geostationary platforms
  - METEOSAT (5/7)
  - GOES (E/W)
  - GMS

Coverage of *in-situ* measurements

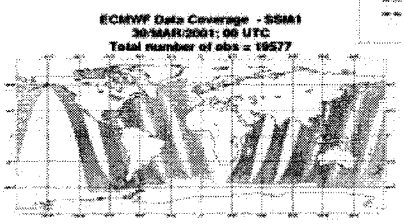


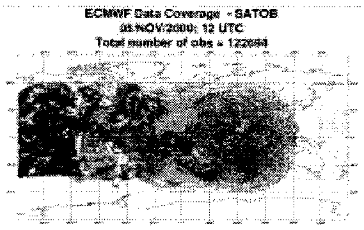


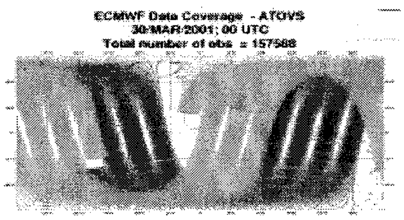


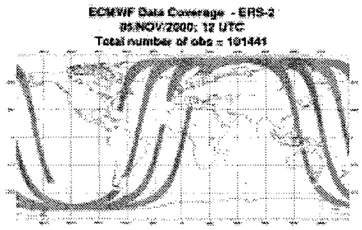


Coverage of satellite-based measurements









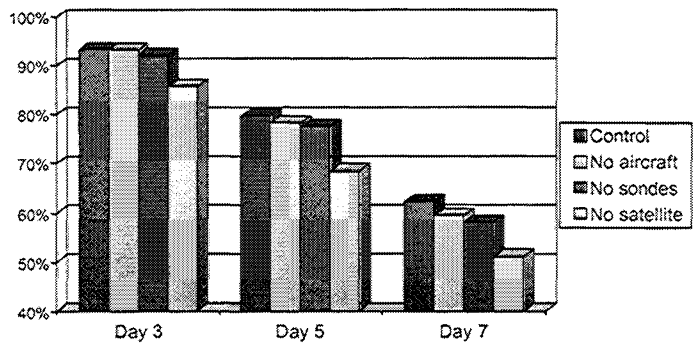
The importance of satellite data

The **limited coverage of *in-situ* observations** means that satellite data are extremely important for global numerical weather prediction, particularly in the medium-range

Improvements in the quality of satellite observations and the techniques developed to assimilate the data have resulted in **satellites now being of equal or greater importance than radiosonde observations** even in data dense regions of the Northern Hemisphere

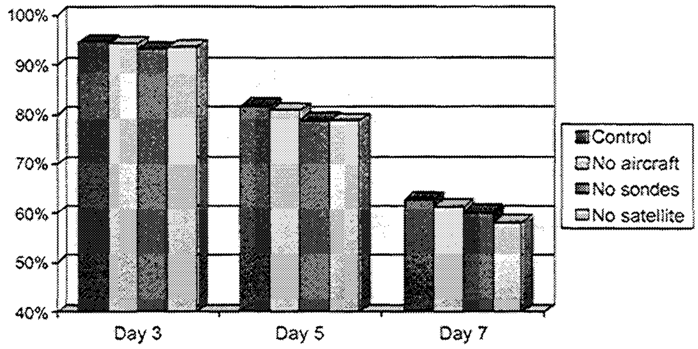
Impact of withdrawing different types of observations on forecast quality

Anomaly correlation of 500hPa height for Southern Hemisphere

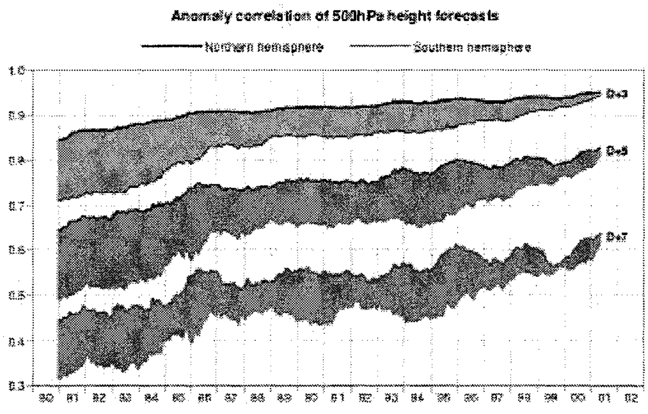


Impact of withdrawing different types of observations on forecast quality

Anomaly correlation of 500hPa height for Northern Hemisphere



Evolution of forecast skill



Break

So satellite data are very important… what do they measure

What do satellite instruments measure?

They DO NOT measure TEMPERATURE
They DO NOT measure HUMIDITY
They DO NOT measure WIND

Satellite instruments (active and passive) simply measure the radiance L that reaches the top of the atmosphere at frequency <. The measured radiance is related to geophysical atmospheric variables by the radiative transfer equation (covered in previous lectures).

$$L(\nu) = \int_0^\infty B(\nu, T(z)) \left[\frac{d\tau(\nu)}{dz} \right] dz + \text{Surface emission} + \text{Surface reflection} + \text{Surface scattering} + \text{Cloud/rain contribution}$$

FREQUENCY SELECTION

By selecting radiation at different frequencies or CHANNELS a satellite instrument can provide information on a range of geophysical variables.

In general, the channels currently used for NWP applications may be considered as one of 3 different types

- Atmospheric nadir sounding channels (passive instruments)
- Surface sensing channels (passive instruments)
- Surface sensing channels (active instruments)

In practice (and often despite their name) real satellite instruments have a combination of both atmospheric sounding and surface sensing channels

ATMOSPHERIC SOUNDING CHANNELS

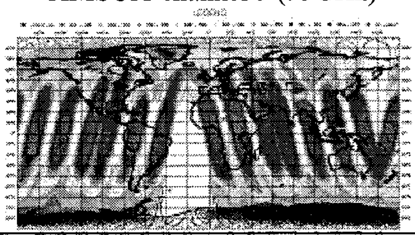
These channels are located in parts of the infra-red and microwave spectrum for which the main contribution to the measured radiance is described by:

$$L(\nu) = \int_0^\infty B(\nu, T(z)) \left[\frac{d\tau(\nu)}{dz} \right] dz$$

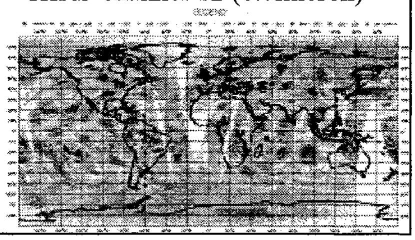
That is they avoid frequencies for which surface radiation and cloud contributions are important.

They are primarily used to obtain information about atmospheric temperature and humidity.

AMSUA-channel 5 (53GHz)



HIRS-channel 12 (6.7micron)



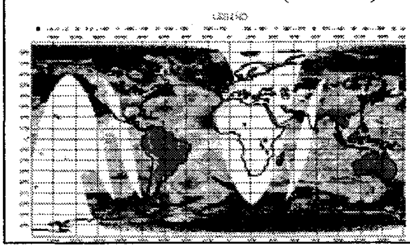
SURFACE SENSING CHANNELS (PASSIVE)

These are located in window regions of the infra-red and microwave spectrum at frequencies where there is very little interaction with the atmosphere and the main contribution to the measured radiance is:

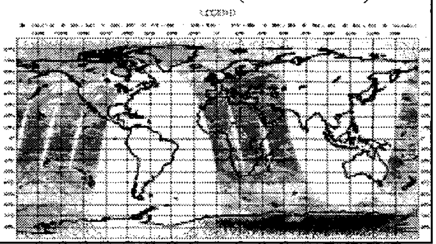
$$L(v) = \text{Surface emission } [\, T_{surf}, \varepsilon(u,v) \,]$$

These are primarily used to obtain information on the surface temperature and quantities that influence the surface emissivity such as wind (ocean) and vegetation (land). They can also be used to obtain information on clouds/rain and cloud movements (to provide wind information)

SSM/I channel 7 (89GHz)



HIRS channel 8 (11microns)



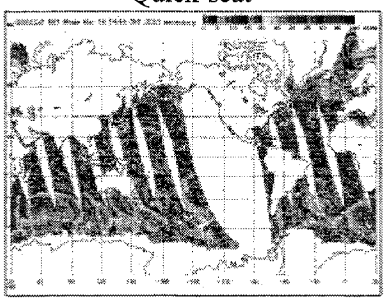
ACTIVE INSTRUMENTS

These (e.g. scatterometers) illuminate the surface in window parts of the spectrum such that

$$L(v) = \text{Surface scattering } [\, \varepsilon(u,v) \,]$$

These primarily provide information on ocean winds (via emissivity) without T_{surf} ambiguity

Quick-scat



ATMOSPHERIC TEMPERATURE SOUNDING

If radiation is selected in a sounding channel for which

$$L(v) = \int_0^{\infty} B(v, T(z)) \left[\frac{d\tau(v)}{dz} \right] dz$$

And we define a function $K(z) = \left[\frac{d\tau}{dz} \right]$

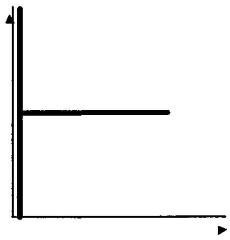
the primary absorber being a well mixed gas (e.g. oxygen or CO2) it can be seen that the measured radiance is essentially a weighted average of the atmospheric temperature profile, or

$$L(v) = \int_0^{\infty} B(v, T(z)) K(z) dz$$

The function K(z) that defines this vertical average is known as a **WEIGHTING FUNCTION**

IDEAL WEIGHTING FUNCTIONS

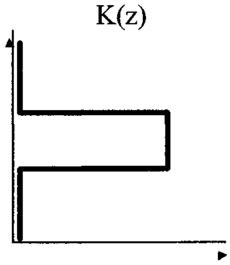
z



K(z)

If the weighting function was a delta-function, this would mean that the measured radiance is sensitive to the temperature at a single level in the atmosphere.

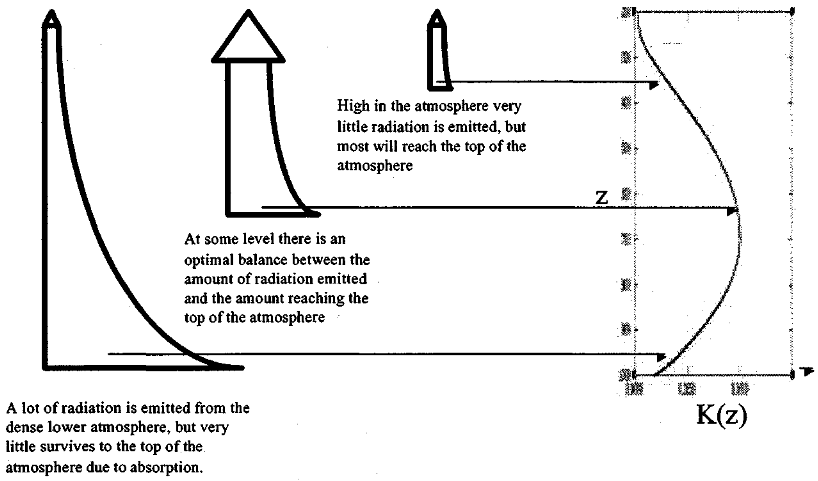
z



K(z)

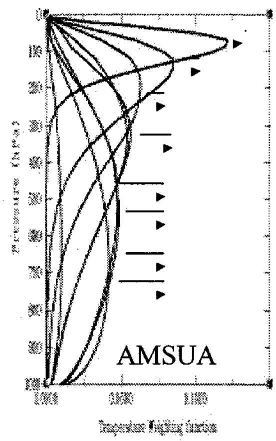
If the weighting function was a box-car function, this would mean that the measured radiance was sensitive to the mean temperature between two atmospheric levels

REAL ATMOSPHERIC WEIGHTING FUNCTIONS



REAL WEIGHTING FUNCTIONS continued...

- The altitude at which the peak of the weighting function occurs depends on the strength of absorption for a given channel
- Channels in parts of the spectrum where the absorption is strong (e.g. near the centre of CO2 or O2 lines) peak high in the atmosphere
- Channels in parts of the spectrum where the absorption is weak (e.g. in the wings of CO2 O2 lines) peak low in the atmosphere

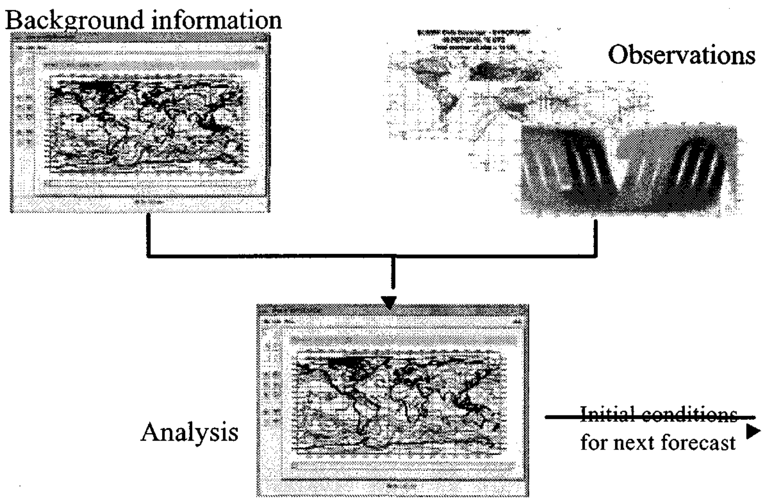


By selecting a number of channels with varying absorption strengths we sample the atmospheric temperature at different altitudes

Break

So we know what satellites measure, how do they fit in to NWP..?

The data assimilation problem (1)



The data assimilation problem (2)

The analysis is an optimal combination of *a priori* background information and new observed data.

It is optimal in that it is the Maximum Liklehood solution and respects the uncertainty in both sources of information

Using Bayes theory the analysis becomes the state of the atmosphere that minimizes a COST or PENALTY FUNCTION

It is completely analogous to the inverse problem solved for satellite retrievals.

The data assimilation problem (3) The COST function

Vector containing all observed data

Observation error covariance

$$J(x) = (x - x_b)^T \mathbf{B}^{-1}(x - x_b) + (y - \mathrm{H}[x])^T \mathbf{R}^{-1}(y - \mathrm{H}[x])$$

Multivariate 3 or 4 dimensional state of the atmosphere (background estimate shown with subscript b

Background error covariance

Operator mapping atmospheric state to observation space

The data assimilation problem (4)

In the past linear (one-step) implementations of Optimal Interpolation (OI) have been used to produce the analysis

$$x_a = \mathbf{BH}^T[\mathbf{HBH}^T + \mathbf{R}]^{-1}(y - \mathbf{H}[x_b])$$

Apart from the need to divide the globe in to small boxes (to reduce the dimensionality of the problem) another limitation of this approach was that the observations had to be linearly related to the analysis variables (T/Q/U/V)

This was fine for in-situ data (e.g. radiosondes)

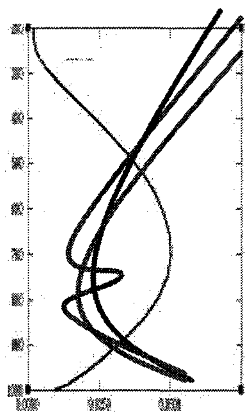
But satellite radiance data had to be converted to retrievals of (T/Q) before being supplied to the assimilation system

EXTRACTING ATMOSPHERIC TEMPERATURE FROM RADIANCE MEASUREMENTS

If we know the entire atmospheric temperature profile T(z) then we can compute (uniquely) the radiances a sounding instrument would measure using the *radiative transfer equation*. This is sometimes known as the **forward problem**

In order to extract or **retrieve** the atmospheric temperature profile from a set of measured radiances we must solve what is known as the **inverse problem**

Unfortunately with a finite number of channels and weighting functions that are generally broad, the inverse problem is **formally ill-posed (an infinite number of different temperature profiles could give the same measured radiances)**



See paper by Rodgers 1976 Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. Rev. Geophys .Space. Phys. 14, 609-624

RETRIEVAL ALGORITHMS

Three different types of retrieval have been used in NWP:

- Exact or least squares solutions to reduced inverse problems
- Regression (statistical / library search / neural net) methods
- Forecast background methods

The retrieval schemes differ in the way **prior information** is used to supplement the information of the measured radiances and solve the inverse problem !

1. Solutions to reduced inverse problems

We acknowledge that there is a limited amount of information in the measured radiances and **re-formulate** the ill-posed inverse problem in terms of a **reduced number of unknown variables** that can be solved for **uniquely.**

E.g. deep mean layer temperatures or EOF's (eigenfunctions) of the temperature profile

Unfortunately these can produce **ill-conditioned solutions** if we attempt to retain enough degrees of freedom required **for NWP** and we subjectively impose a reduced representation for which it is **difficult to quantify the accuracy** (this is very important for NWP).

2. Regression and Library search methods

Using a sample of temperature profiles matched (collocated) with a sample of radiance observations, a statistical relationship is derived that **predicts** atmospheric temperature from the measured radiance.

(e.g. NESDIS operational retrievals or the 3I approach)

These tend to be limited by the statistical characteristics of the training sample / profile library **and will not** produce **physically important** features if they are **statistically rare** in the training sample.

3. Forecast background methods

These use an explicit background or *first-guess* profile from a short range forecast and perform **optimal adjustments** using the measured radiances. The adjustments **minimize a cost function**

Forecast Background Retrievals

We formulate a 1D cost function (analogous to the 3D/4D cost function defined for the analysis

$$J(x) = (x - x_b)^T \mathbf{B}^{-1}(x - x_b) + (y - \mathrm{H}[x])^T \mathbf{R}^{-1}(y - \mathrm{H}[x])$$

1D profiles of T / Q / O3

Vector of measured radiances

Radiative transfer operator

And minimize with a single step solution (if channels / data are selected to avoid nonlinear effects

$$x_a = \mathbf{BH}^T[\mathbf{HBH}^T + \mathbf{R}]^{-1}(y - \mathbf{H}[x_b])$$

Or iterate towards a solution if H incorporates nonlinear effects (e.g.cloud)

OR ⟶

$$x_{n+1} = x_b + \mathbf{W}_n[y - H(x_n) - H(x_n)(x_b - x_n)]$$
$$\mathbf{W}_n = \mathbf{BH}(x_n)^T[\mathbf{H}(x_n)\mathbf{BH}(x_n)^T + \mathbf{R}]^{-1}$$

Forecast Background Retrievals

These have a number of advantages that make them more suitable for NWP than other methods

•The prior information (short-range forecast) is very accurate (more than statistical climatology) which improves retrieval accuracy.

•The prior information contains information about physically important features such as fronts, inversions and the tropopause.

•The error covariance of the prior information and resulting retrieval is better known (crucial for the subsequent assimilation process).

•The retrieval may be considered an intermediate step towards the direct assimilation of radiances (no external sources of prior information)

BUT the error characteristics of the retrieval may be complicated due to its correlation with the forecast background (used twice!)

Assimilation of satellite retrievals in NWP

Whatever approach is adopted to convert radiance measurements to temperature, humidity etc…The use of satellite retrievals is problematic for two main reasons:

1) They retain characteristics of the a priori information that are very difficult to remove.

2) They generally have complicated error structures that are difficult to model in the subsequent assimilation (e.g. strong correlations between levels and variables)

For these reasons the use of retrievals in global NWP has generally been superceded by the direct assimilation of radiance data.

Direct assimilation of radiances in NWP

Variational analysis methods such as 3DVAR and 4DVAR allow the direct assimilation of radiance observations (without the need for and explicit retrieval step).

This is because such methods do NOT require a linear relationship between the observed quantity (radiance) and the analysis variables (T/Q..)

The retrieval (or inversion) is essentially incorporated within the main analysis by finding the 3D or 4D state of the atmosphere that minimizes the cost function

The forecast background still provides the prior information to supplement the radiances, but the inversion is further constrained by the simultaneous assimilation of other observations.

The cost function is minimized by iteration using efficient adjoint techniques but the process is still expensive and requires super-computers

Implementation of 3DVAR

$$J(x) = (x - x_b)^T \mathbf{B}^{-1}(x - x_b) + (y - \mathrm{H}[x])^T \mathbf{R}^{-1}(y - \mathrm{H}[x]) + Jc$$

The vector x is a full global 3D vector describing the state of the atmosphere and has a dimension in excess of 10^6. In practice the analysis variables are scaled and remapped to balanced variables for which the background error covariance reduces to a computationally managable block diagonal form.

These reduced covariances are estimated offline (see later)

The incremental approach is adopted where the comparison with observations is done at full resolution, but the minimization (and gradient calculations) at a reduced resolution.

The operator H (observation operator) for in-situ data is simply a spatial interpolation, but for radiance data includes the full radiative transfer operator.

Additional constraints Jc are imposed upon the solution by the inclusion of an additional cost function term to e.g. filter gravity waves.

Implementation of 4DVAR

Instead of finding a single 3D atmospheric state that represents observations over a given time window (e.g. 6hrs), the 4DVAR searches for a time series or trajectory of atmospheric states that fits the observations at the time they were actually measured.

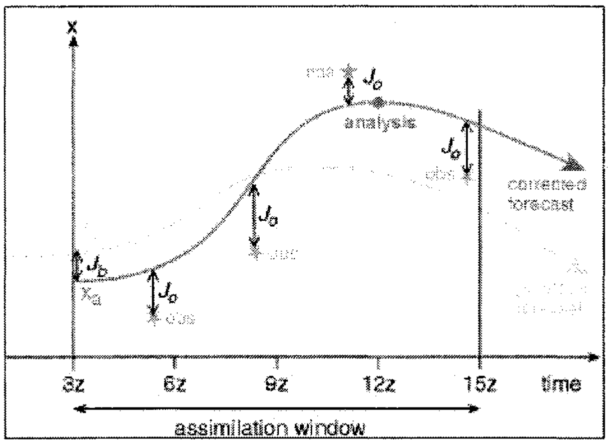
We minimize the cost function through all times slots i :

$$J(x) = (x - x_b)^T \mathbf{B}^{-1}(x - x_b) + \sum_i (y_i - \mathrm{H}[x_i])^T \mathbf{R}^{-1}(y_i - \mathrm{H}[x_i]) + Jc$$

Subject to the hard constraint that the states follow the model equations

$$\forall i, x_i = \mathbf{M}_{0 \rightarrow i}(x)$$

Schematic representation of 4DVAR



Special characteristics of 4DVAR

- Better use is made of observations far from the center of the assimilation time window (particularly important for satellite data)
- The inversion of the radiance data is constrained by the background and its covariance, but also by the constraint that radiance observations at different times force adjustments that are consistent with the forecast model physics and dynamics
- In fitting the radiances, the 4DVAR has the option of advecting warm (or moist) air and thus causes radiance data can cause wind adjustments during the assimilation

Direct assimilation of radiances

By the direct assimilation of radiances we avoid the problem of assimilating retrievals with complicated error structures.

BUT

There are still a number of significant problems that must be handled

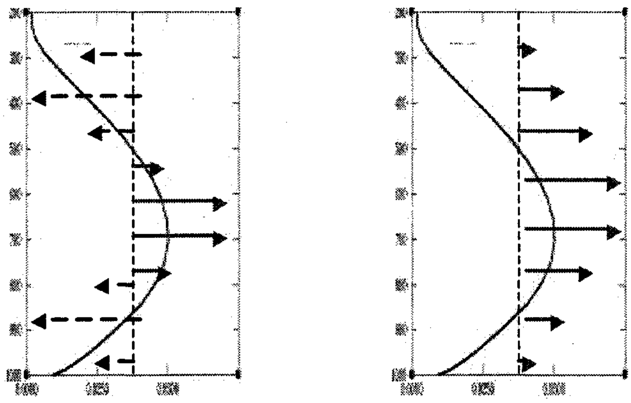
- The specification of the background error covariance
- The specification of the radiance error covariance
- Other ambiguities in the data
- Systematic radiance and RT error

Break

So much for the theory, what are the main issues …?

Specifying the background error covariance

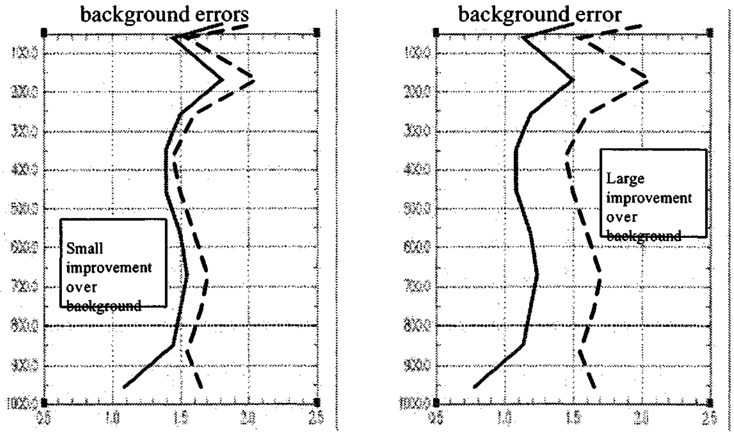
We can think of the radiance data "seeing" and correcting errors in the background state during the data assimilation process.



Difficult to correct   Easy to correct

RETRIEVAL / ANALYSIS PERFORMANCE

Sharp / anti-correlated background errors   Broad / deep correlated background error



Error standard deviation (K)   Error standard deviation (K)

ESTIMATING FORECAST ERROR CORRELATIONS

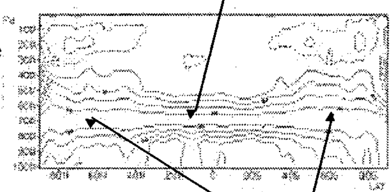
If the **background errors are mis-specified** in the retrieval / analysis this can lead to a complete mis-interpretation of the radiance information and **badly damage the analysis** (indeed producing a analysis with **larger** errors than the background state !)

Thus accurate estimation of **B** is crucial:

•comparison with radiosondes (best estimate of truth but limited coverage

•comparison of e.g. 48hr and 24hr forecasts (so called *NMC method*)

•comparison of ensembles of analyses made using perturbed observations



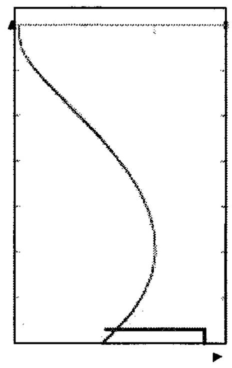
Sounding channels sensitive to the lower troposphere

By placing sounding channels in parts of the spectrum where the absorption is weak we obtain temperature (and humidity) information from the lower troposphere (low peaking weighting functions).

BUT

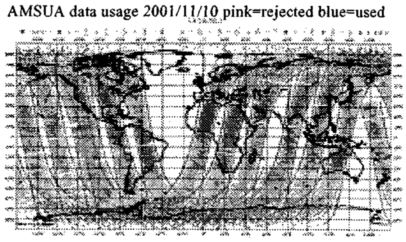
These channels (obviously) become more sensitive to surface emission and the effects of cloud and precipitation.

In some cases **surface or cloud contribution can dominate the atmospheric signal** and it is difficult to use the data safely for temperature / humidity sounding.



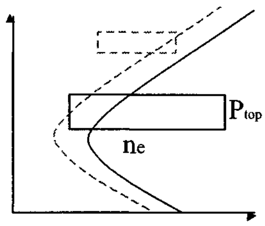
OPTIONS FOR USING LOWER TROPOSPHERIC SOUNDING CHANNELS

AMSUA data usage 2001/11/10 pink=rejected blue=used

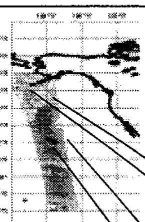


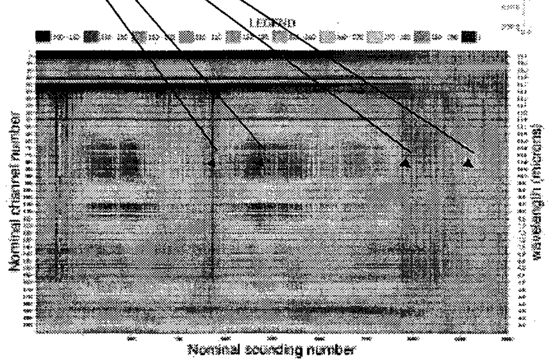
• Screen the data carefully and only use situations for which the surface and cloud radiance contributions can be computed very accurately *a priori* (e.g. cloud free situations over sea). But meteorologically important areas are often cloudy!

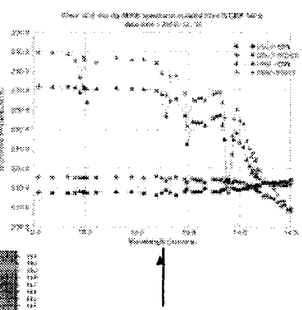
•Simultaneously estimate atmospheric temperature, surface temperature / emissivity and cloud parameters within the analysis or retrieval process (need very good background statistics !) Can be dangerous.



Clear and Cloudy AIRS spectra







Characteristic spectral patterns

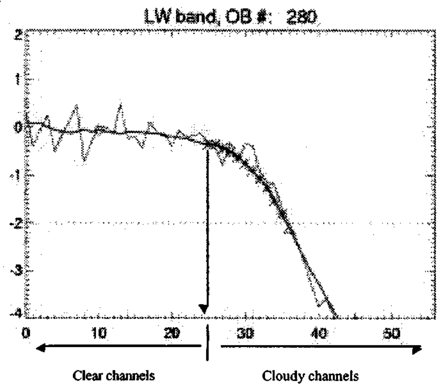
What do we know about the cloud signal ?

• Over warm surfaces (non-frozen) it is always negative

•In band split / ranked channels it increases monotonically negative

•We can identify an "obviously" contaminated channel and step backwards with a digital filter to locate the first channel with discernable cloud contamination

•All channels ranked as higher peaking can safely be assimilated as clear



SYSTEMATIC ERRORS

Systematic error must be removed before the assimilation otherwise biases will propogate in to the analysis.

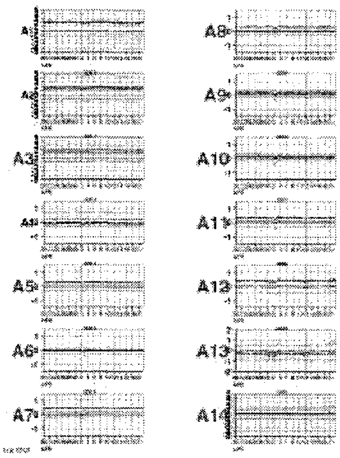
Sources of systematic error in radiance assimilation include

• instrument error (calibration)

•radiative transfer error (spectroscopy or RT model)

•cloud/rain screening errors

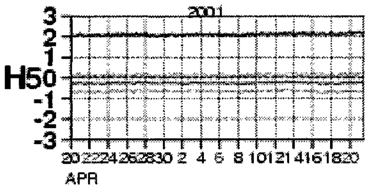
Mean corrected and uncorrected (obs-fg) radiance departure



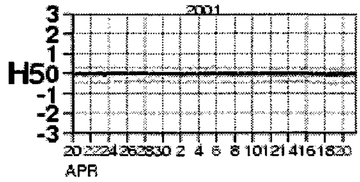
AMSUA for May 20001

DIAGNOSING SYSTEMATIC ERRORS

Systematic errors in observations are usually identified by monitoring against the forecast background (or analysis) in the vicinity of constraining radiosonde data. **How do we know the source of the bias** ?



HIRS channel 5 (peaking around 600hPa on NOAA-14 satellite has +2.0K radiance bias against model

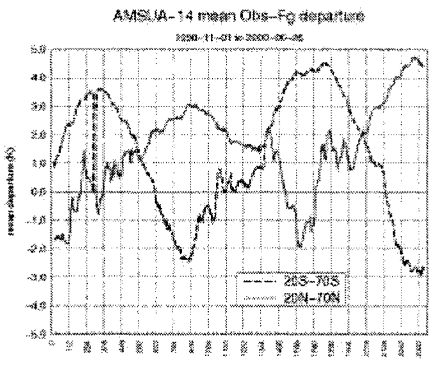


HIRS channel 5 (peaking around 600hPa on NOAA-16 satellite has no radiance bias against model.

DIAGNOSING SYSTEMATIC ERRORS

What if the model is wrong ?

This time series shows an apparent systematic error in AMSU channel 14 (peaking at 1hPa). By checking against other research data (HALOE and LIDAR data) the bias was confirmed as a model bias and the channel is now assimilated with no bias correction



Wind adjustments with radiance data

Radiances can influence the model wind field during the data assimilation process in a number of ways:

•Directly through the use of frequent cloud imagery

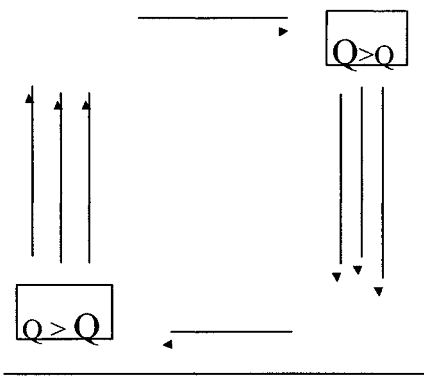
•Directly via surface emissivity (mostly microwave)

•Indirectly through model physics (humidity)

•Indirectly through passive tracing(humidity and ozone)

We must ensure that the adjustments from different data types are consistent within the system (satellite vs *in-situ*)

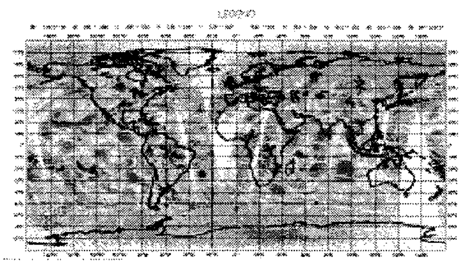
Indirect forcing of the wind field through model physics



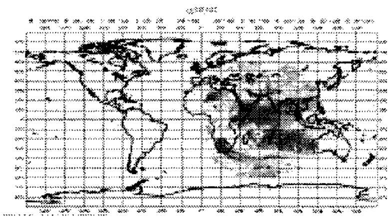
By adding humidity to the lower troposphere or removing moisture from the upper troposphere the satellite humidity information can cause large scale wind adjustments !

Indirect forcing of the wind field by passive tracing

By observing humidity or ozone signals in the radiance data, the 4DVAR can advect these fields to fit the radiances causing wind adjustments.



This is particularly true with high temporal density radiance from GEO satellites



Review of key concepts (1)

- Satellite data are extremely important in NWP, even in areas with a dense network of in-situ observations
- Data assimilation combines observations and a priori information in an optimal way and is analogous to the retrieval inverse problem
- Modern data assimilation systems have largely moved to variational approaches and use radiance observations directly (not retrievals)

Review of key concepts (2)

- The limited vertical resolution of satellite radiances makes the specification of background error covariances crucial

- Systematic errors can be very harmful, particularly in 4D systems where they have a multivariate (wind) impact on the analysis

- Dealing with cloud and surface emission remains one of the most difficult areas of research.

Suggested reading

Andersson E., Thepaut J.N., Eyre J.R., McNally A.P., Kelly G., Courtier P., and PailleauxJ., 1994: Use of cloud cleared radiances in three/four dimensional variational data assimilation. Q.J.R. Meteorol. Soc., 120, pp. 627-653.

Daley, R. 1991 Atmospheric Data Analysis. Cambridge University Press

Eyre J.R., Kelly G., McNally A.P., Andersson E., and Persson A., Assimilation of TOVS radiances through one-dimensional variational analysis. Q.J.R. Meteorol. Soc. (1993), 119,pp.1427-1463

Le Dimet F., and Talagrand O., 1986: Variational algorithms for analysis and assimilation of meteorological observations. Tellus, 37A, 97-110.

Lorenc A.C., 1981: A global three-dimensional multivariate statistical interpolation scheme. Mon. Wea. Rev., 109, 701-721.

McNally A.P. 2000: Estimates of short-range forecast temperature error correlations and theimplications for radiance data assimilation. Q.J.R. Meteorol. Soc. 126 pp 361-373.

Parrish, D.F., Derber, J.C. 1992 The National Meteorological Center's Spectral Statistical Interpolation Analysis system. Mon. Weather Rev. 120, 1747-1763

Rabier F., A. McNally, E. Andersson, P. Courtier, P. Unden, J. Eyre, A. Hollingsworth andF. Bouttier 1998 . The ECMWF implementation of three-dimensional variational assimilation (3D-Var) II Structure functions. Q.J.R. Meteorol. Soc. 124 part B, pp 1809..

Rodgers C.D., 1976: Retrieval of Atmospheric Temperature and Composition From Remote measurements of Thermal Radiation. Rev. Geophys. and Space Phys., 14, 609-624.