Regional Cooperation for
Limited Area Modeling in Central Europe

LACE
nwp central europe

Minimization, B matrix computation from practical point of view

Antonín Bučánek
ALADIN/HIRLAM common DA training week, 10-15 February 2019

DHMZ SHMU ARSO METEO Slovenia OMSZ ZAMG ROMANIA ANM

Minimization

- Model configuration 131 (NCONF)
- Reads the observation minus first-guess departures from the CCMA database and runs the inner loop of the minimization.
- The algorithm consists actually of two iterative processes : the iterative quasi-Newton M1QN3 and iterative recomputation of gradients by the simulator.
- Thus, two maximum number of iterations have to be defined in the namelist:
  NITER is the max number of M1QN3 iterations
  NSIMU is the max number of simulations of the gradient

Minimization – Inputs

- ALADIN executable - MASTERODB
- First guess – FA file
  - `cp guess ICMRF${CNMEXP}0000`
  - `ln -sf ICMRF${CNMEXP}0000 ICMSH${CNMEXP}INIT`
  - `ln -sf ICMRF${CNMEXP}0000 ICMSH${CNMEXP}IMIN`
- Compressed observation database ODB – CCMA
- Namelists (setting code related parameters)
  - fort.4
  - In case of assim IASI – IASI_CLDDET.NL, iasichannels
- Background error covariance matrix
  - Statistical balance operators -- stabal96.bal
  - Auto-covariances -- stabal96.cv
- Specific constants
  - For conv observation only "amv_p_and_tracking_error" is necessary
  - rtcoef*, rmtberr, … (AROME: MCICA, RADSRTM)
- Variational bias correction - VARBC.cycle

Minimization – ODB Environment

```
# ODB env for e131, e002
(path to CCMA)
export ODB_CMA=CCMA
export IOASSIGN=$TMPDIR/IOASSIGN
export ODB_SRCPATH_CCMA=$TMPDIR/CCMA
export ODB_DATAPATH_CCMA=$TMPDIR/CCMA
```

```
# general ODB settings
export EC_PROFILE_HEAP=0
export TO_ODB_ECMWF=0
export TO_ODB_SWAPOUT=0
export ODB_DEBUG=0
export ODB_CTX_DEBUG=0
export ODB_REPRODUCIBLE_SEQNO=2
export ODB_STATIC_LINKING=1
export ODB_IO_METHOD=1
export ODB_ANALYSIS_DATE=${YYYY}${MM}${DD}
export ODB_ANALYSIS_TIME=${NT}0000
export TIME_INIT_YYYYMMDD=${YYYY}${MM}${DD}
export TIME_INIT_HHMMSS=${NT}0000
```

Minimization – Namelist (fort.4)

- NITER, NSIMU – number of iteration, simulation
- NGRATS, NFRGRA – write-out of gradient
- RCVGE – stop criterion for the gradient norm reduction in the minimizer
- **REDNMC** – multiplying factor for all σ_b
- **SIGMAO_COEF** – normalization coefficient of σ_o
- LSPFCE (T/F) – horizontally constant σ_b / map of σ_b
- **LSPRT=.F.** – use T not Tv in 3D-Var, Tv part bugged!
  AROME cy40 only LSPRT=T is working.

- NOUTPUT=2 – NODE file for all CPUs (debugging)

Minimization – Namelist (fort.4)

- **NOTVAR** – switch on/off the use of the obs type and variable in Jo (see **YOMCOSJO** for the ordered list of variables) (**Fischer, 2007**)
- Two dimensional – **NOTVAR (param, obstype)**
- **Obstype**: Synop=1, Airep=2, Satob=3, Dribu=4, Temp=5, Pilot=6, Rad=7, PAOB=8, Scatt=9, Limb sounde=10, Radar=13
- **Param**: wind=1, wind10m=2, rh=5, rh2m=6, temperature=7, geopotential=8, T2m=11, spec.humidity=19, radinace=13,
- NOTVAR(param,obstype)=-1/0; **0 use**, **-1 not use**
- TEMP example:

```
NOTVAR(1, 5)= 0,0,-1,-1,-1,-1,0,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,0,-1,-1,-1,-1,-1,-1,
```

Minimization – AROME specific

- ALARO x AROME differs in NAMGFL, NAMPHY*, NAMFA due to physics only!
- AROME uses grid-point humidity – **LREADGPTRAJ=.T.**
- AROME is used with surfex soil model
- The 3dvar is **not** prepared to work with surfex!
- The old isba surface scheme is used to compute screen level parameters in minimization but the first guess is not containing all fields needed by isba
- Usually program **addsurf** is used to add missing fields to first guess
- Where to get missing fields? (climfile, ARPEGE analysis, surfex)

Minimization – Outputs

- Listing -- NODE.001_01
- Jobout
- Analysis -- MX${CNMEXP}999+0000
- New VARBC.cycle

Minimization – Missing fields

- Prognostic fields like hydrometeors, NH parameters are not analyzed and they are missing in analysis.
- The usual way is to just cycle them. So they are added from guess to analysis before integration of forecast.
- Two options:
  1) Either before minimization copy guess to resulting analysis,
  2) Or after minimization add missing prognostic fields from guess to analysis by program addsurf or blend

Minimization – check execution (1) NODE.001_01

- completion of the execution successful : **grep ’END CNT0’**
- completion of the model setup successful : grep ’END OF SETUP’. Before this line, you can find all the printouts from your model setup. After this line, the printouts of the execution of any configuration do start.
- For screening and minimization, the NOTVAR tables are repeated not far below, which can be useful.
- the evolution of the cost functions, and of the gradient:
  grep ’GREPCOST’ -- lists the evolution of Jb and Jo by iteration
  grep ’GREPGRAD’ -- lists the evolution of the norm of the total gradient
- Search NODE.001_01 for **Diagnostic JO-table** where you can find used observations statistics for first and last iteration of minimization

Minimization – check execution (2)

```
Diagnostic JO-table (JOT) MINIMISATION JOB T0215 NCONF=   131 NSIM4D=   999 NUPTRA=     0
================================================================================

       Obstype     1 === SYNOP, Land stations and ships
       --------------------------------------------------

       Obstype     2 === AIREP, Aircraft data
       --------------------------------------------------
            Codetype   141 === AIREP Aircraft Report
               Variable      DataCount          Jo_Costfunction        JO/n       ObsErr      BgErr
                  U             30           2.472950019077            0.08    0.832E+01   0.301E+01
                  T             15          0.1392491304865            0.01    0.302E+01   0.106E+01
            Codetype   144 === AMDAR Aircraft Report
               Variable      DataCount          Jo_Costfunction        JO/n       ObsErr      BgErr
                  U            592           32.62280142782            0.06    0.706E+01   0.238E+01
                  T            300           9.044523595306            0.03    0.318E+01   0.104E+01
                            ----------   ---------------------------   --------
      ObsType  2 Total:        937           44.27952417269            0.05

       Obstype     3 === SATOB, Atmospheric motion winds
       --------------------------------------------------

 ---------------------------------------------------------------------------
            Jo Global :       8120           2472.982639491            0.30
 ===========================================================================

 End of JO-table (JOT)
```

Minimization – check execution (3)

- ODB ECMA/CCMA
- ECMA after screening contains reasons why particular obs (datum) is passive, rejected or blacklisted
- CCMA contains only active reports
- http://www.umr-cnrm.fr/aladin/meshtml/DOC_odb/odb.php
- https://www.ecmwf.int/sites/default/files/elibrary/2016/16646-part-i-observations.pdf
- obsvalue, varno, an_depar, fg_depar, lat,lon, statid
- datum_status

Table 2.28 *Datum status.*

Bit Position	No. of Bits	Value – Description
0	1	1 – Report Active
1	1	1 – Passive Report
2	1	1 – Rejected Report
3	1	1 – Blacklisted Report

Minimization – check execution (4)

- datum_event1

- odbsql -q “SELECT obsvalue, varno, an_depar, fg_depar, lat,lon, statid, datum_status, datum_event1 FROM hdr, body” -f odb

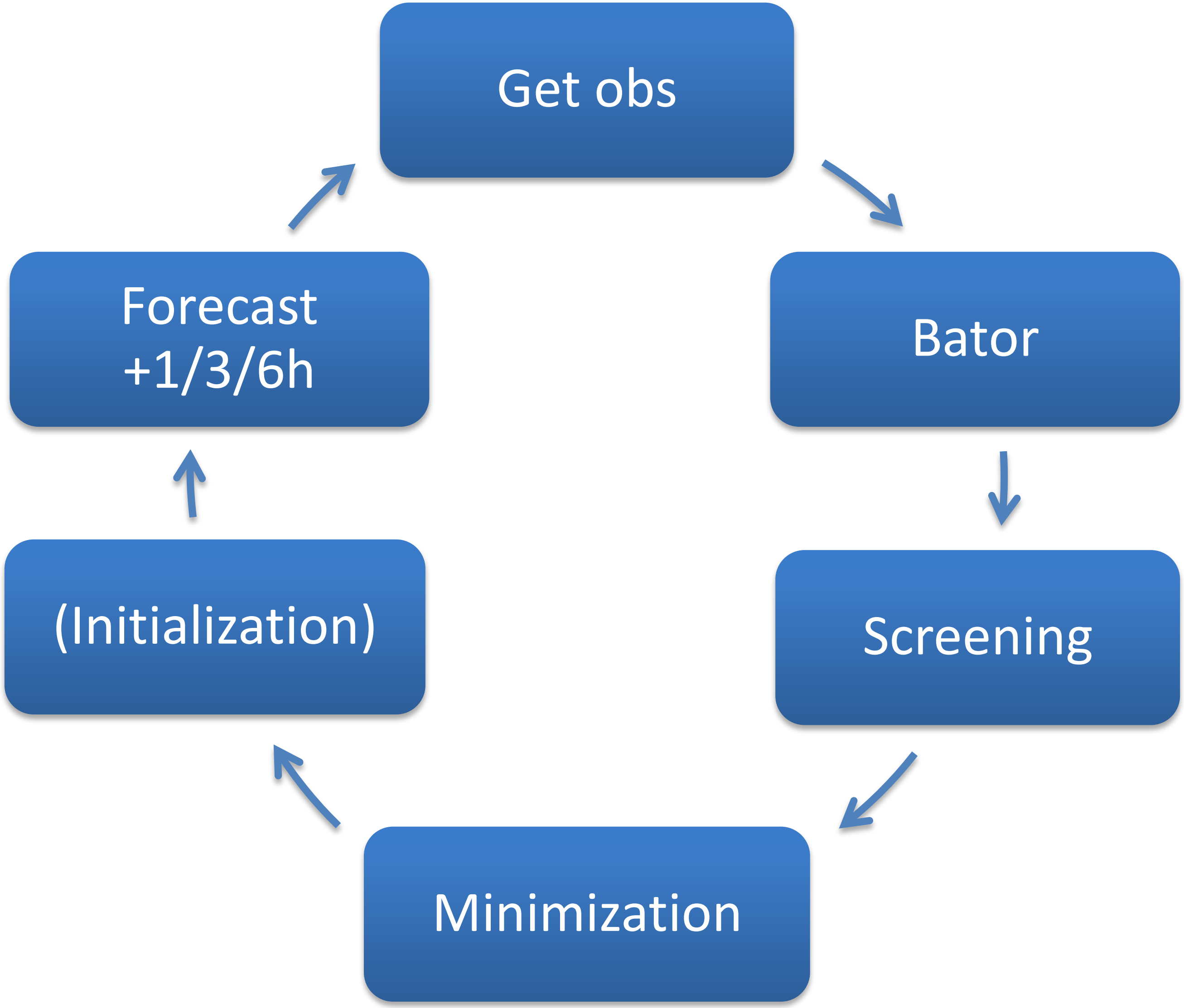
Table 2.29 *Global datum events.*

Bit Position	No. of Bits	Value – Description
0	1	1 – Missing Vertical Coordinate
1	1	1 – Missing Observed Value
2	1	1 – Missing Background (First Guess) Value
3	1	1 – Rejected due to RDB Flag
4	1	1 – Activated due to RDB Flag
5	1	1 – Activated by Whitelist
6	1	1 – Bad Reporting Practice
7	1	1 – Vertical Position out of Range
8	1	1 – Reference Level Position out of Range
9	1	1 – Too Big First Guess Departure
10	1	1 – Too Big Departure in Assimilation
11	1	1 – Too Big Observation Error
12	1	1 – Redundant Datum
13	1	1 – Redundant Level
14	1	1 – Report Over Land
15	1	1 – Report Over Sea
16	1	1 – Not Analysis Variable
17	1	1 – Duplicate Datum/Level
18	1	1 – Too Many Surface Data
19	1	1 – Multi Level Check
20	1	1 – Level Selection
21	1	1 – Vertical Consistency Check
22	1	1 – Vertical Coordinate Changed from Z to P
23	1	1 – Datum Rejected via Namelist
24	1	1 – Combined Flagging
25	1	1 – Datum Rejected due to Rejected Report
26	1	1 – Variational QC Performed
27	1	1 – Observation Error Increased
28	1	1 – Cloud Contamination
29	1	1 – Rain Contamination
30	1	1 – Aerosol Contamination
31	1	1 – Missing or Not Sensible Emissivity Values

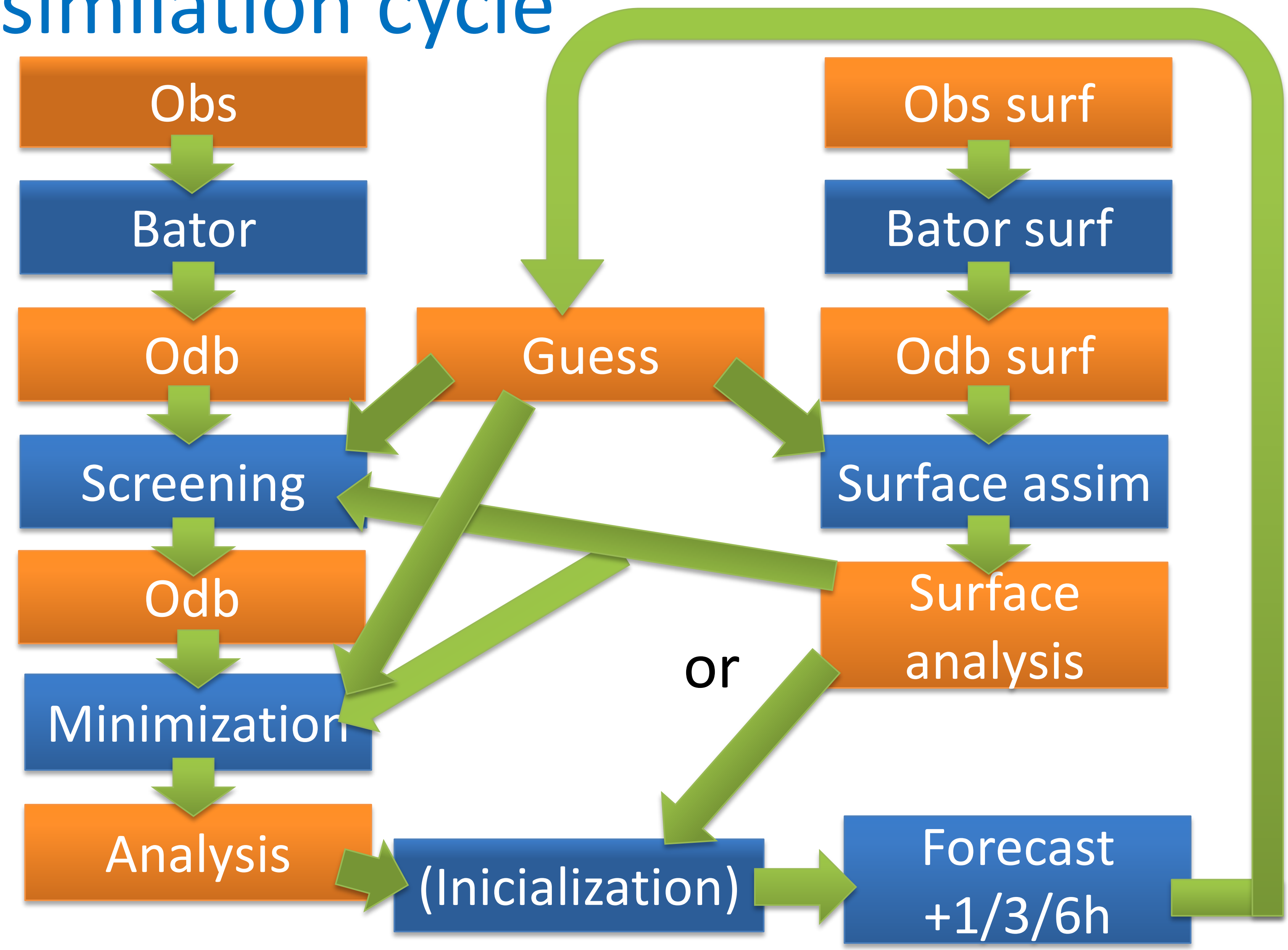
Minimization – problems, crashes

- Observation not assimilated - check NOTVAR, find the reason in odb ECMA, CCMA
- Crash in readobs.F90
  - Do you use MASTERODB, is odb linked correctly?
- Crash in mkglobstab_model.F90
  - `OBS <NO> AT <lat> <lon> NOT FOUND`
  - `MKGLOBSTAB_MODEL: IWRONG1 NE 0`
  - This happens when observation out of lam domain in odb – check bator setting of lamflag (BATOR_LAMFLAG=1)
- Wrong Bmatrix
  - Bad vertical structure
    `CALL ABOR1(' Vertical balance interpol not implemented')`
  - Low res B in high res run
    `CALL ABOR1(' Stat balance is only up to NSMAX')`
  - High res B in low res run – **no crash just warring !!!**
    `'Truncating balance'`
- Other crashes, look in 3dvar documentation Fischer (2007), https://hirlam.org/trac/wiki, http://www.rclace.eu/forum/

Assimilation cycle



Assimilation cycle



B matrix computation

B matrix computation

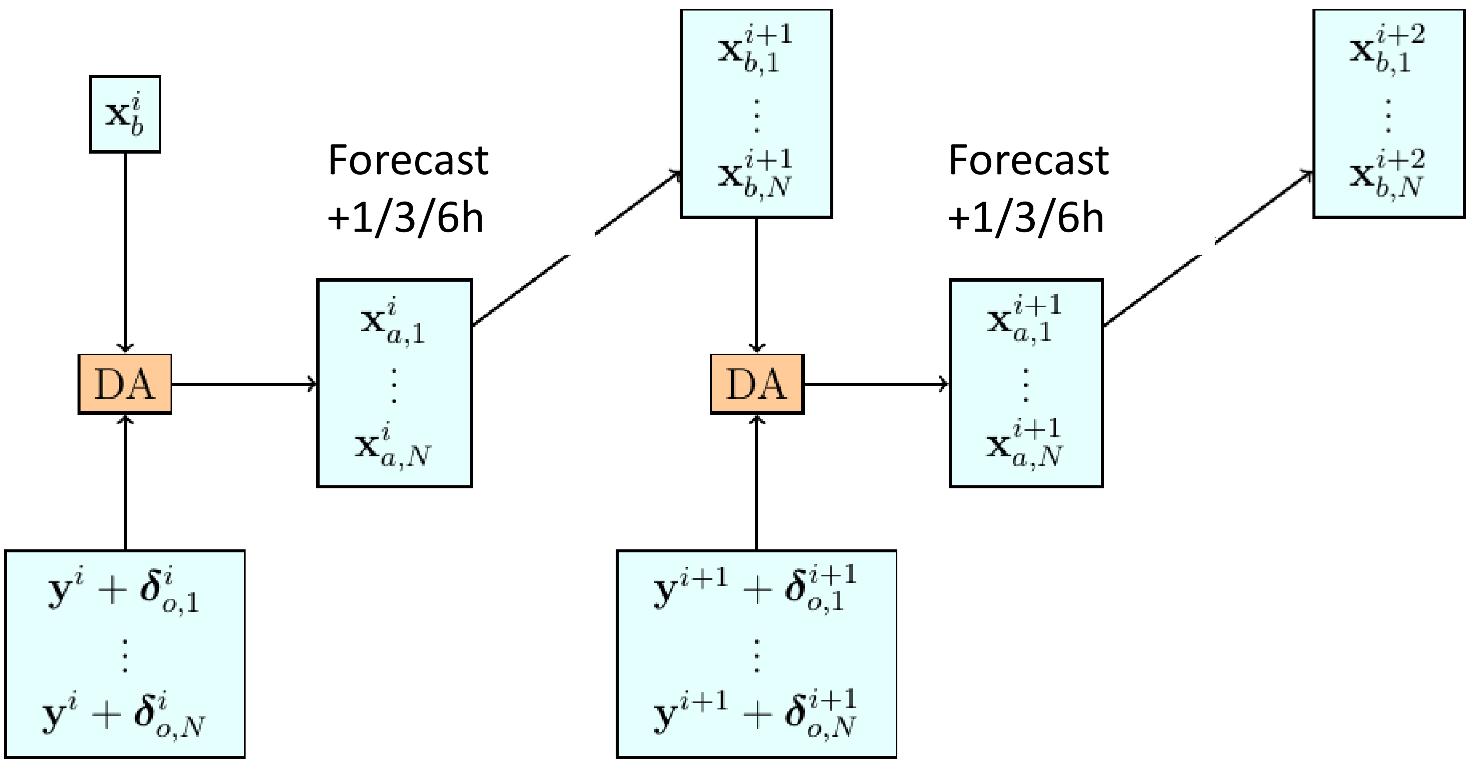
- The role of background error covariances is to scale, spatially filter and propagate the observed information away from observation point.
- Due to balance properties analysis corrections are propagated from one variable to the other
- The bg error statistics depend on:
  - model and its resolution (Global, lam)
  - geographical area (midlatidude, tropical)
  - weather regime, time of day, season
  - density of the observation network. (Dense x sparse obs)

B matrix computation

- Above mentioned + size of Full B => need for simplification
  1. Climatological B (longer period)
  2. Homogeneity and isotropy – in spectral space it means that waves with different wave numbers are uncorrelated
- Methods for computation of B matrix
  – Hollingsworth and Lonnberg 1986 - based on innovation vectors
  – NMC (National Meteorological Center, Parrish and Derber, 1992), - differences between two differently long forecast valid at the same time (deterministic)
  – Ensemble base techniques (EDA – ensemble data assimilation)
    - Spinup B – Dynamical adaptation of global ensemble
    - Full ensemble B – LAM assimilation ensemble
    - They tend to simulate the growth of errors during assimilation cycle
    - Belo-Pereira and Berre, 2006

B matrix – EDA

- Ensemble of assimilation cycles
- observations are perturbed N(0, σ_o)
- first guesses could be perturbed



B matrix – EDA

- Every member have the same set of observations but with different perturbation.
- Initial perturbation are forgotten after several cycles of EDA (El Ouaraini a Berre, 2011)
- Analysis error

$$\mathbf{e}_a^{i+1} = \mathbf{e}_b{}^{i+1} + \mathbf{K}\left(\mathbf{e}_o^{i+1} - \mathbf{H}\mathbf{e}_b^{i+1}\right)$$

- Evolution of ensemble members differences

$$\boldsymbol{\varepsilon}_a^{i+1} = \boldsymbol{\varepsilon}_b{}^{i+1} + \mathbf{K}\left(\boldsymbol{\varepsilon}_o^{i+1} - \mathbf{H}\boldsymbol{\varepsilon}_b^{i+1}\right)$$

- True B matrix

$$\mathbf{B}_* = \overline{\mathbf{e}_b\mathbf{e}_b^T}$$

Berre et al. (2006)

- Ensemble based B matrix

$$\mathbf{B}_\varepsilon = \overline{\boldsymbol{\varepsilon}_b\boldsymbol{\varepsilon}_b^T}$$

$$\mathbf{B}_\varepsilon = 2\mathbf{B}_*$$

B matrix computation practical

- Preparation of EDA
  - Perturbation of upper-air observations could be done in screening (NAMSCC – **LPERTURB=.T.**, **NAEMEMBER=xxMEMBxx**, NAENSEMBLE=1) or offline before or after screening
  - Perturbations of SYNOP and DRIBU could be done in bator (LPERTOBS=.T.) or offline
  - Perturbation of Sea Surface Temperature
  - Inflation of FG perturbations (Raynaud et al., 2012)

B matrix computation practical 2

- **Femars** – Computing ensemble members differences
  - configuration e001
  - necessary to set **LFEMARSD=.T.** and **LSPRT=.F.** in the namelist
  - suppress in-line fullpos and computation of fluxes (NFPOS=0, NAMXFU - LXFU=.F.)
  - **1cpu only**
  - To compute difference between two forecasts valid at the same time from two members you have to rename the one member to ICMSH${CNMEXP}FGIN and second member to ICMSH${CNMEXP}ANIN, CNMEXP is namelist parameter
  - Output file is a grib file with name gribdiff

B matrix computation practical 3

- Festat – B matrix computation from gribdiffs
  - compile by gmkpack, cycle 43 needs adaptation - http://www.rclace.eu/forum/viewtopic.php?f=30&t=62&p=2096#p2096
  - Option "-no-wrap-margin" for intel compiler
  - Inputs: gribdiff files called **ensdiff${ncase}**,
    special namelist "**fort.4**"
  - Outputs: B matrix files $name.cv, $name.bal, $name.cvt – where $name is specified in the namelist, some diagnostics
- Fediacov – Produces set of diagnostics
  - **1cpu only**
  - inputs: **$name.cv** or **$name.cvt**, namelist "**fort.4**"

B matrix computation practical 4

- Period
  - Number of gribdiffs must be larger than the number of model vertical levels to have positive definite B.
- REDNMC – scaling of B
  - theoretical value of REDNMC is .7 since REDNMC^2 is applied to B matrix
- SIGMAO_COEF – scaling of “R”
  - It is vector of obs error scalings per obstype
  - Should be set in Bator, screening, minim
  - It is not applied to all obs parameters.

Thank you for your attention

References

- Fischer, C., 2007. The variational computation inside ARPEGE/ALADIN: cycle CY32. *Technical report*. Available at: www.umr-cnrm.fr/gmapdoc//IMG/ps/main_var.ps
- Kertész, S., 2007. Overview of the observation usage in the ALADIN variational data assimilation system. *HMS Budapest Hungary, RC LACE report*. Available at: http://netfam.fmi.fi/HMS07/obsdoc.pdf
- Berre, L., Ştefănescu, S.E. & Pereira, M.B., 2006. The representation of the analysis effect in three error simulation techniques. *Tellus A*, 58(2), pp.196–209.
- Hollingsworth, A. & Lönnberg, P., 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus A: Dynamic Meteorology and Oceanography*, 38(2), pp.111–136.
- Parrish, D.F. & Derber, J.C., 1992. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review*, 120(8), pp.1747–1763
- El Ouaraini, R. & Berre, L., 2011. Sensitivity of ensemble-based variances to initial background perturbations. *Journal of Geophysical Research*, 116(D15), p.D15106
- Raynaud, L., Berre, L. & Desroziers, G., 2012. Accounting for model error in the Météo-France ensemble data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 138(662), pp.249–262.