CERA-20C Uncertainty Estimation

-

Per Dahlgren

CERA-20C Uncertainty Estimation

Per Dahlgren

Model space diagnostics

Compare with proxy truth in model space

Observation space diagnostics

CERA-20C produced with a 10-member Ensemble Data Assimilation (EDA) system

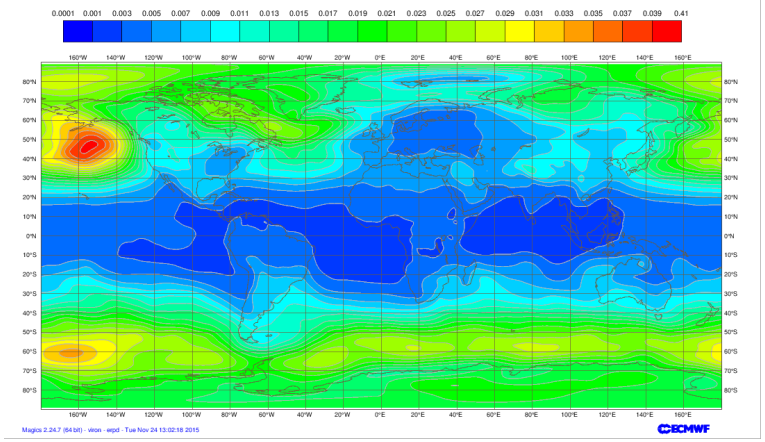
EDA is used to provide a flow dependent background error in 4D-Var

The ensemble information is also an estimate of the uncertainty in the analysed atmospheric state

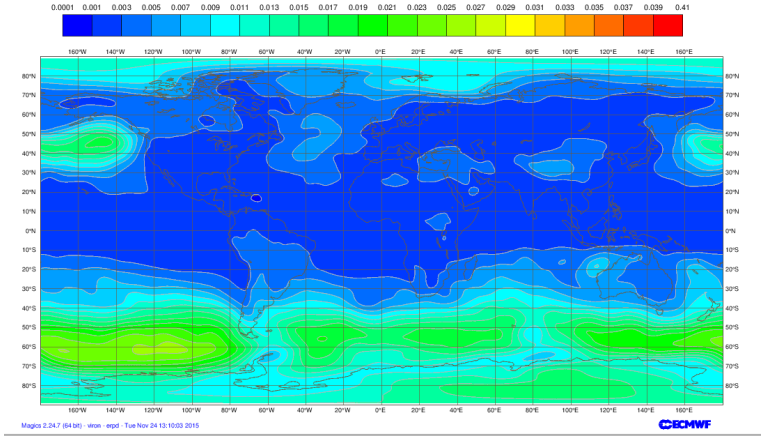
Model space diagnostics

Ensemble spread on LnPs, monthy average

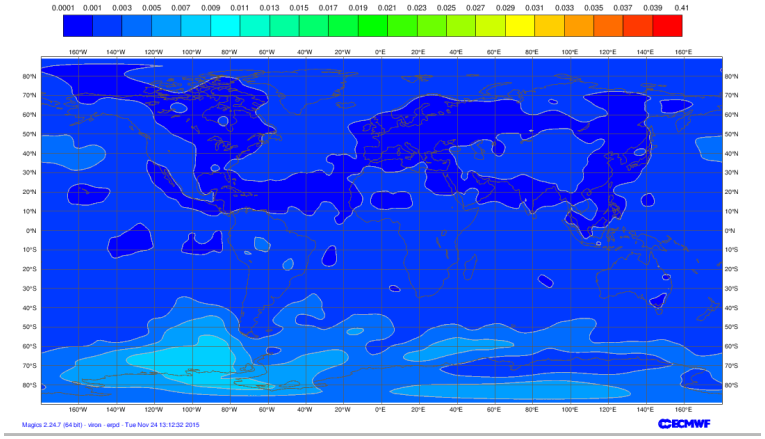
190002 Avg. ens. spread. scaled standard deviation LNPS
EXP 2366 CERA-20C



194802 Avg. ens. spread. scaled standard deviation LNPS
EXP 2372 CERA-20C



200402 Avg. ens. spread. scaled standard deviation LNPS
EXP 2379 CERA-20C



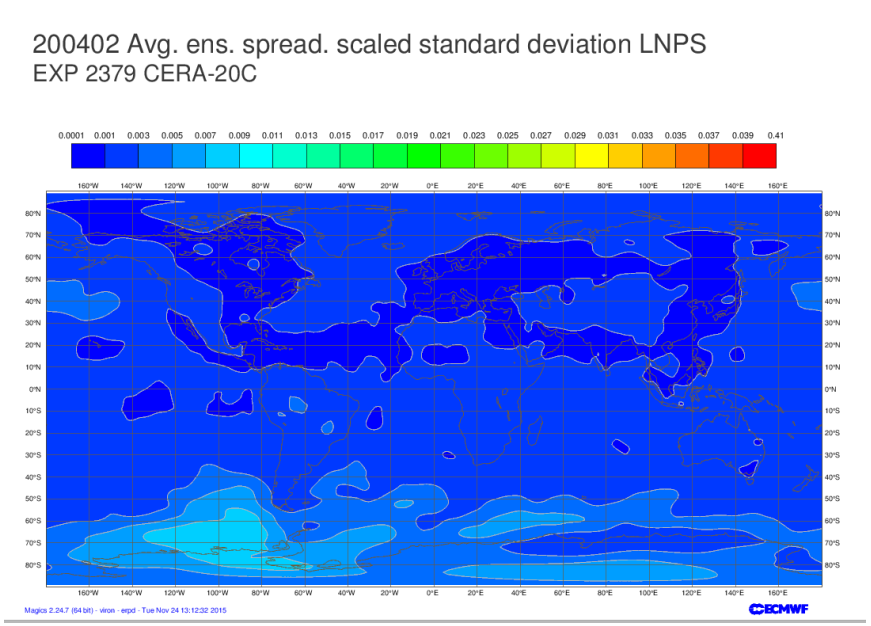
Geographical distribution looks reasonable

Spread decreases during the century =>
EDA responds to observing system changes

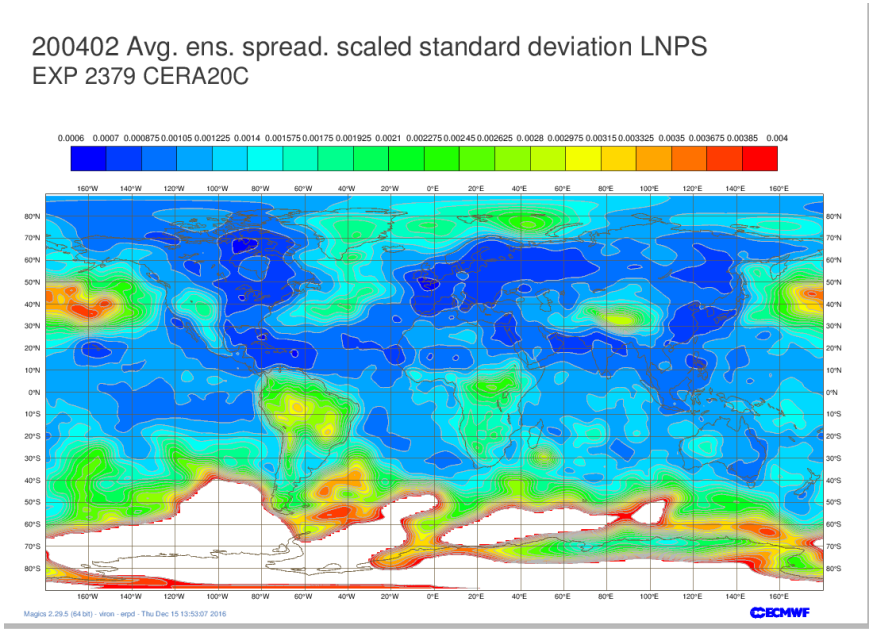
In 2004, largest spread in southern hemisphere extratropics

Model space diagnostics

Ensemble spread on LnPs, monthly average



Feb 2004 again



Feb 2004, different colour shading

High spread in NH extratropics over sea (baroclinic instability)

Low spread over Europe, well observed

High spread over mountains

Model space diagnostics

Analysis ensemble spread, times-series

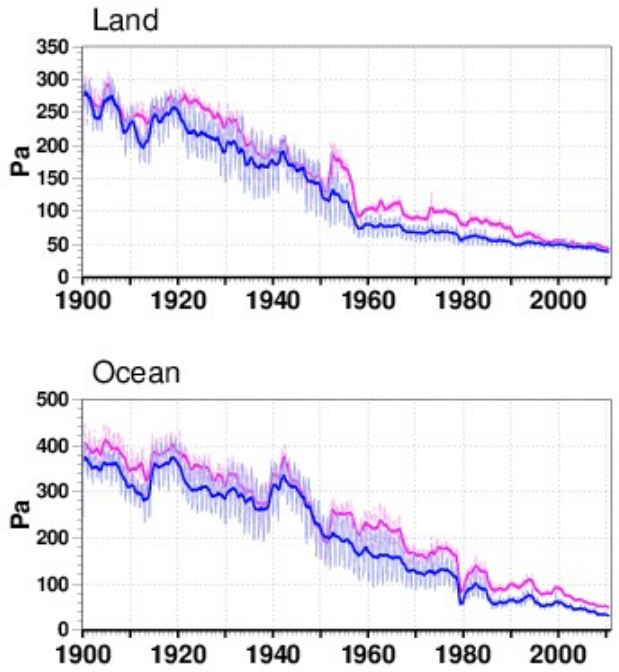
Yearly and monthly (thin lines) average

Blue=CERA20C
Pink=20CR

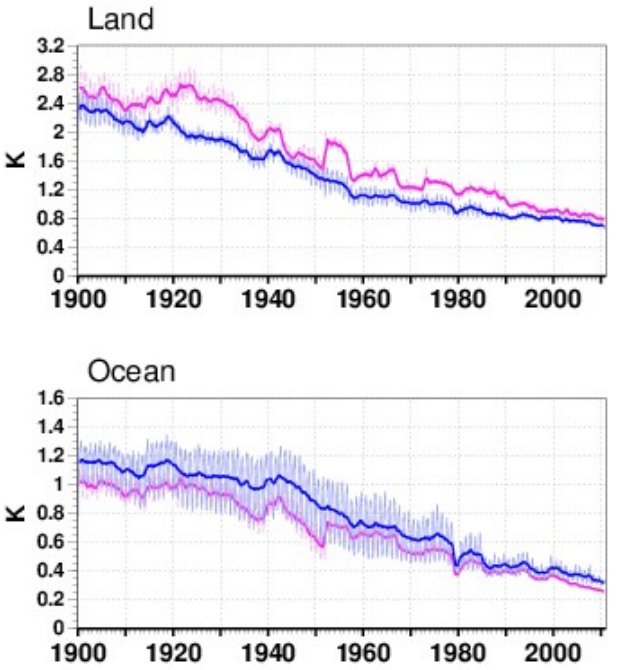
Spread decreases with time

20CR generally bigger spread estimate: larger ensemble size

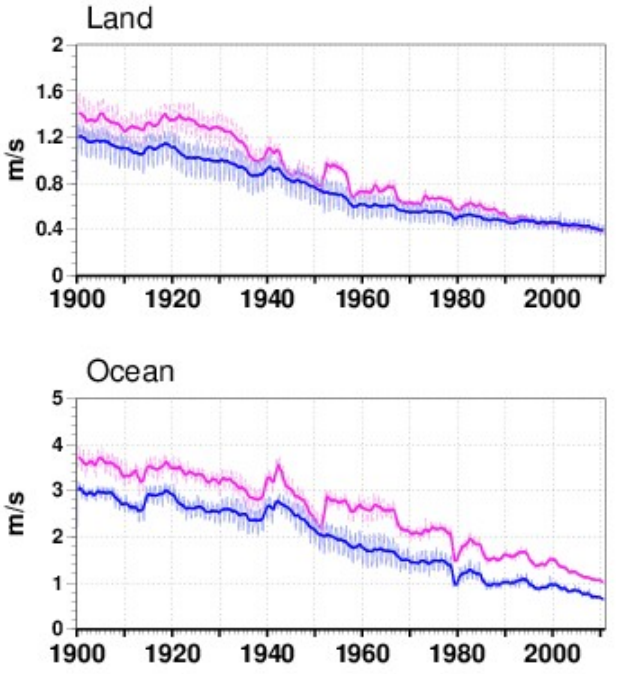
MSLP



T2m

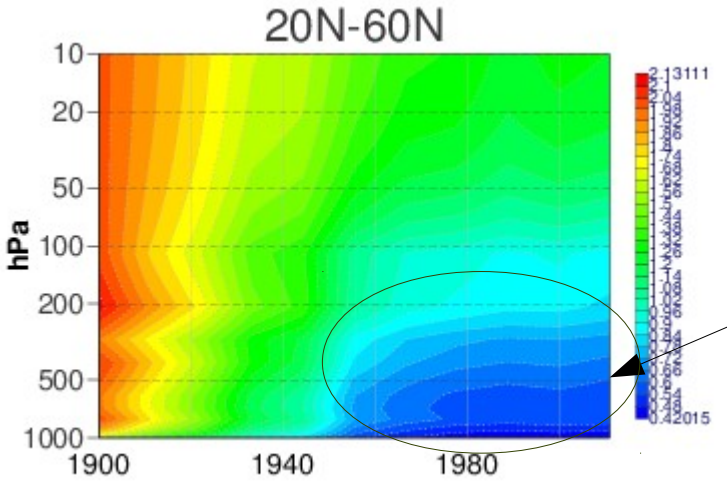


U10m



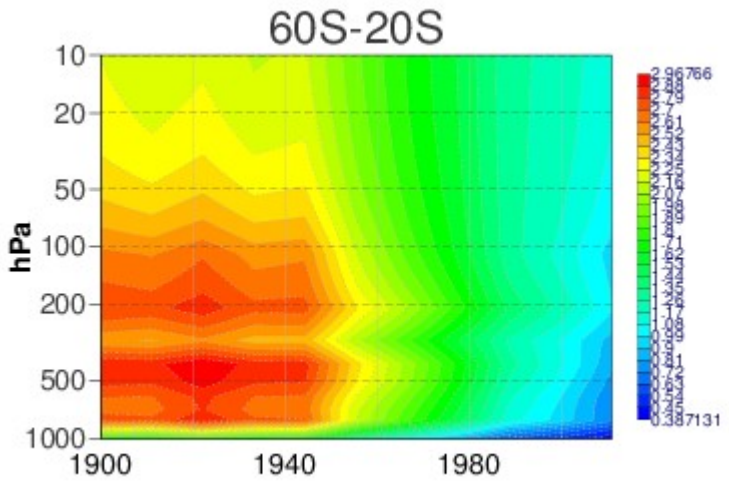
Model space diagnostics

Temperature analysis ensemble spread, times-series



Northern hemisphere extratropics
Large increase in number of observations

Surface pressure observations have significant impact on the atmospheric state throughout the troposphere => B matrix structure functions



Southern hemisphere extratropics

Compare with proxy truth in model space

Ideal case:
Ensemble spread should describe the error
of the ensemble mean

The CERA20C 10 member ensemble probably
underdispersive => gives too low error estimate

Compare with proxy truth in model space

Ideal case:
Ensemble spread should describe the error
of the ensemble mean

The CERA20C 10 member ensemble probably underdispersive => gives too low error estimate

ES=ensemble spread
Xm=ensemble mean
Xt=true state

$$ES=RMS(Xm-Xt)$$

True state not known

Use ERA-Interim as proxy truth in recent years:
* Assimilates the full observing system
* Higher horizontal resolution T255 (CERA20C T159)

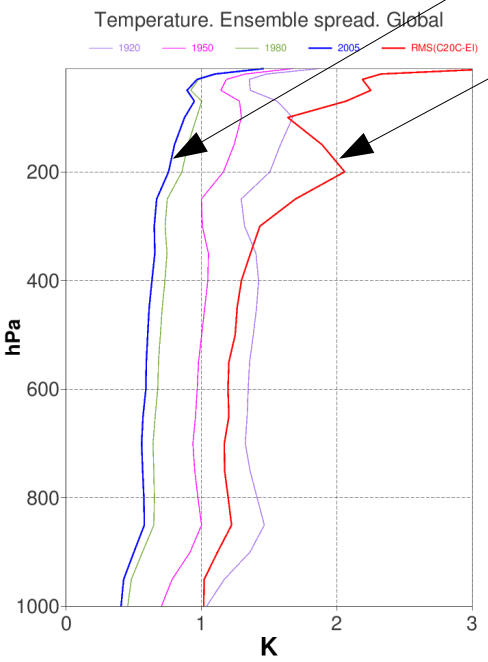
=> ERA-Interim should lie closer to the true state than CERA20C

Compare with proxy truth in model space

ES=CERA20C ensemble spread
Xm=CERA20C ensemble mean
Xt=ERA-Interim

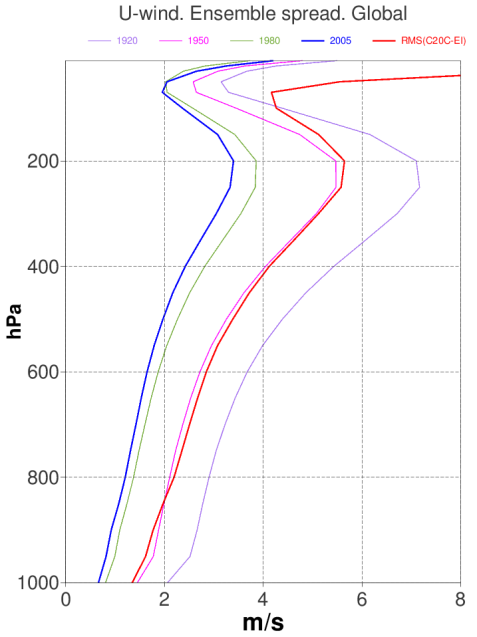
Compare thick blue line with thick red line

ES=RMS(Xm-Xt)



CERA20C error estimate at least a factor of two too small

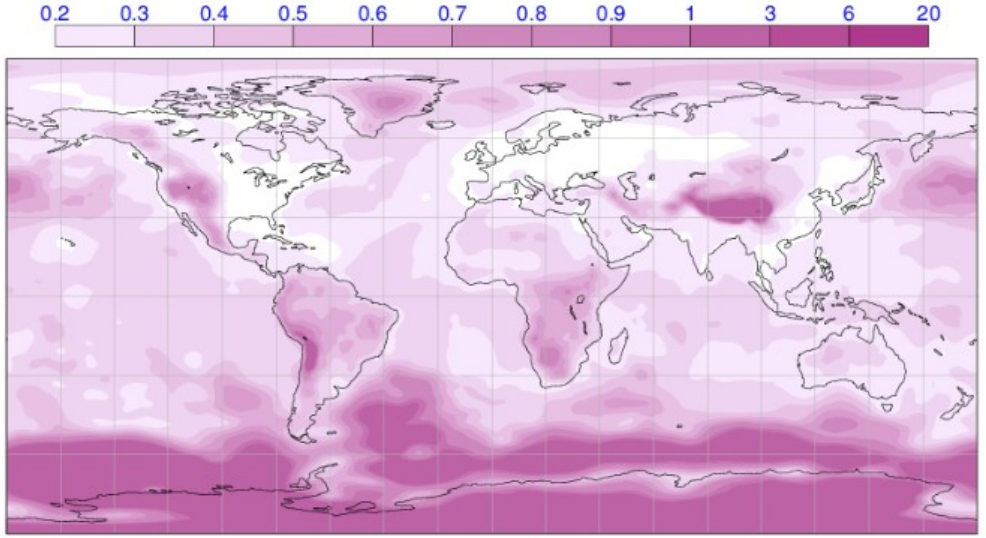
But, this is a global average



Compare with proxy truth in model space

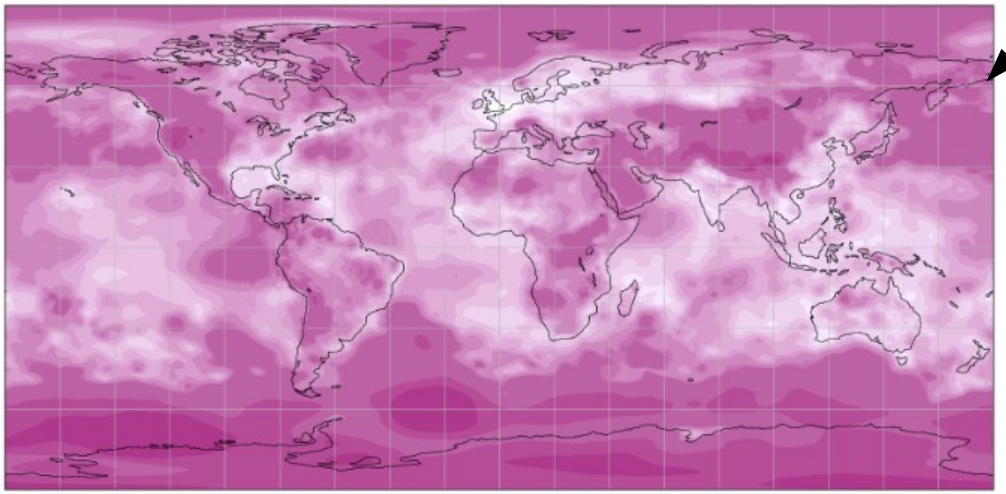
ES=CERA20C ensemble spread
Xm=CERA20C ensemble mean
Xt=ERA-Interim

DJF 2010



ES=RMS(Xm-Xt)

RMSE (CERA-20C mean – "truth")



Horizontal structures very similar

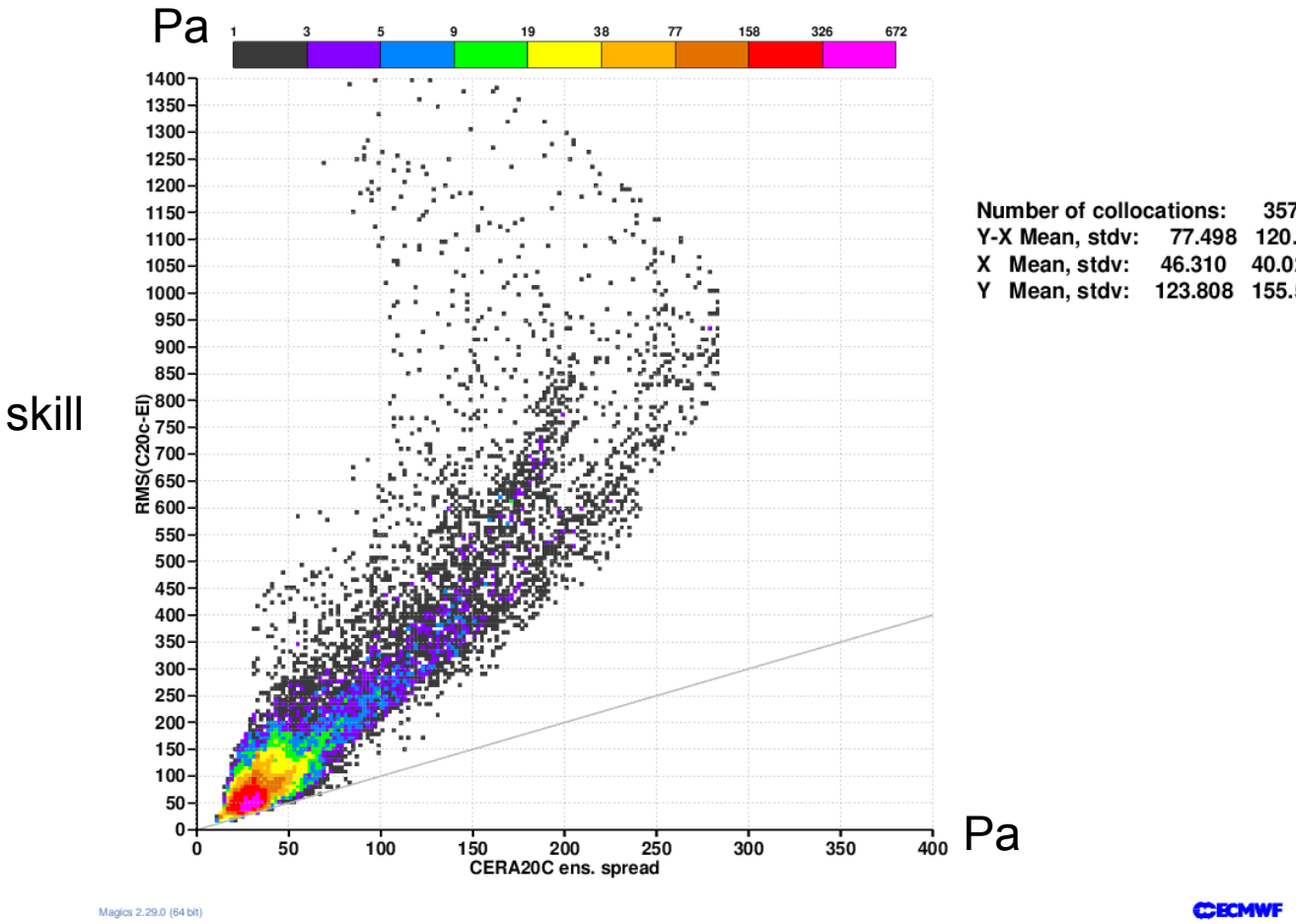
CERA20C EDA correctly captures where the uncertainties are

Ensemble spread too small

From P. Laloyaux

Compare with proxy truth in model space

MSLP 1 year statistics, 2005



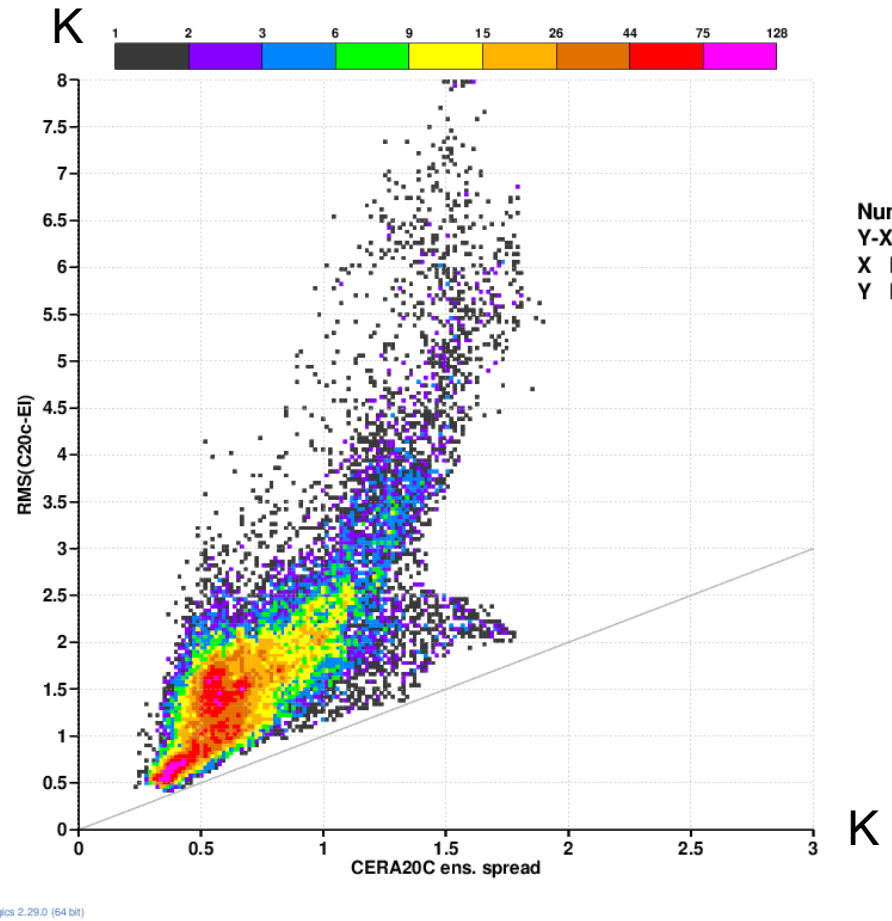
skill

spread

Compare with proxy truth in model space

Temperature 850hPa

1 year statistics, 2005



Number of collocations: 35718
Y-X Mean, stdv: 0.936 0.700
X Mean, stdv: 0.698 0.276
Y Mean, stdv: 1.634 0.906

skill

K

spread

Observation space diagnostics

Spread-skill relation in observation space

o=observation
b=first guess, NWP model interpolated to obs location

σ_o=Observation error
σ_b=Background error

Ideally, the following relation should be fulfilled

$$RMS(o-b) = \sqrt{\sigma_o^2 + \sigma_b^2}$$

Skill/error

Spread

Observation space diagnostics

Spread-skill relation in observation space

o=observation
b=first guess, NWP model interpolated to obs location

σ_o=Observation error
σ_b=Background error

Ideally, the following relation should be fulfilled

Skill/error Spread

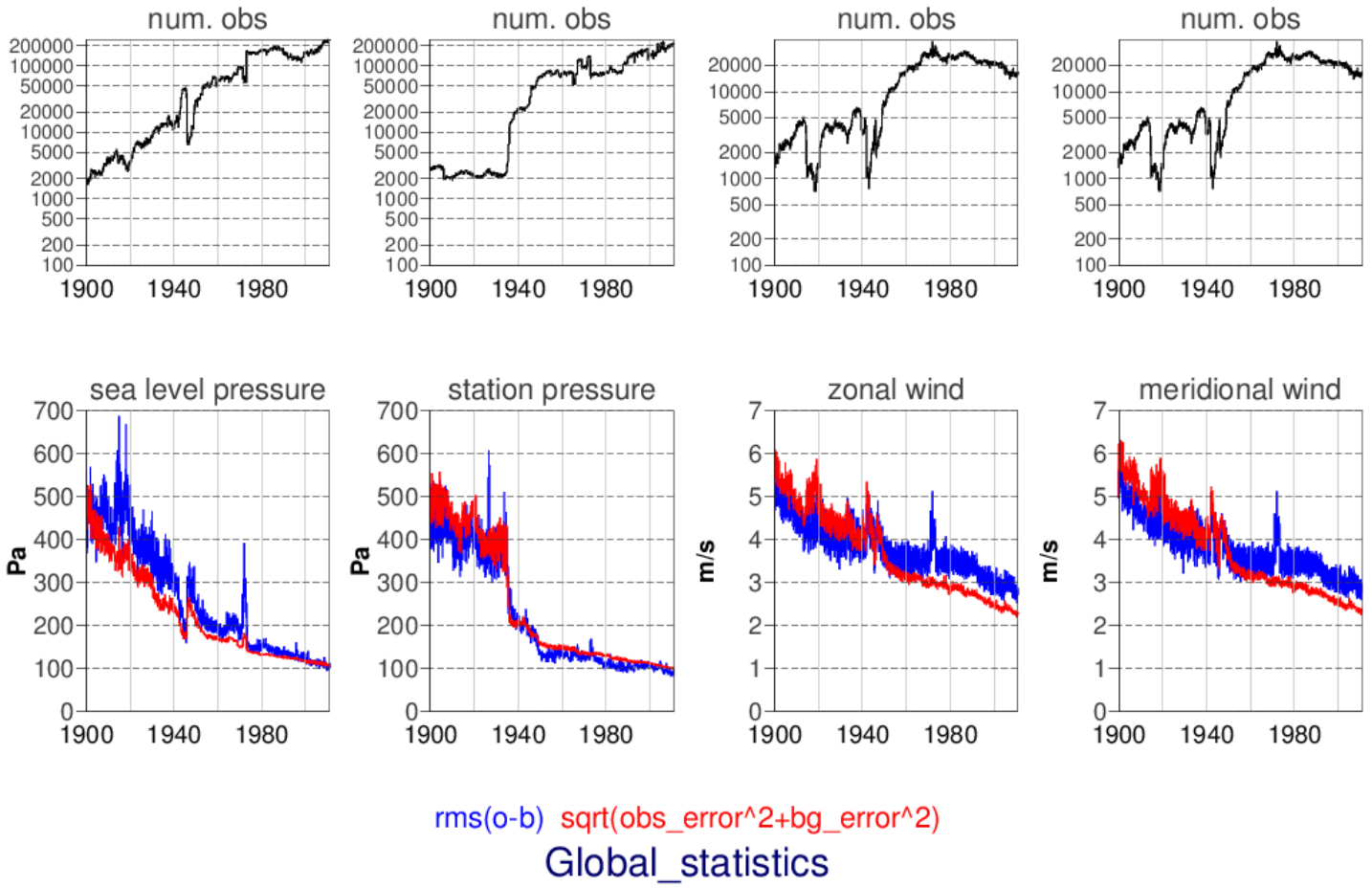
$$RMS(o-b) = \sqrt{\sigma_o^2 + \sigma_b^2}$$

fg_depar obs_error eda_spread

Available in ODB feedback from IFS assilmilation

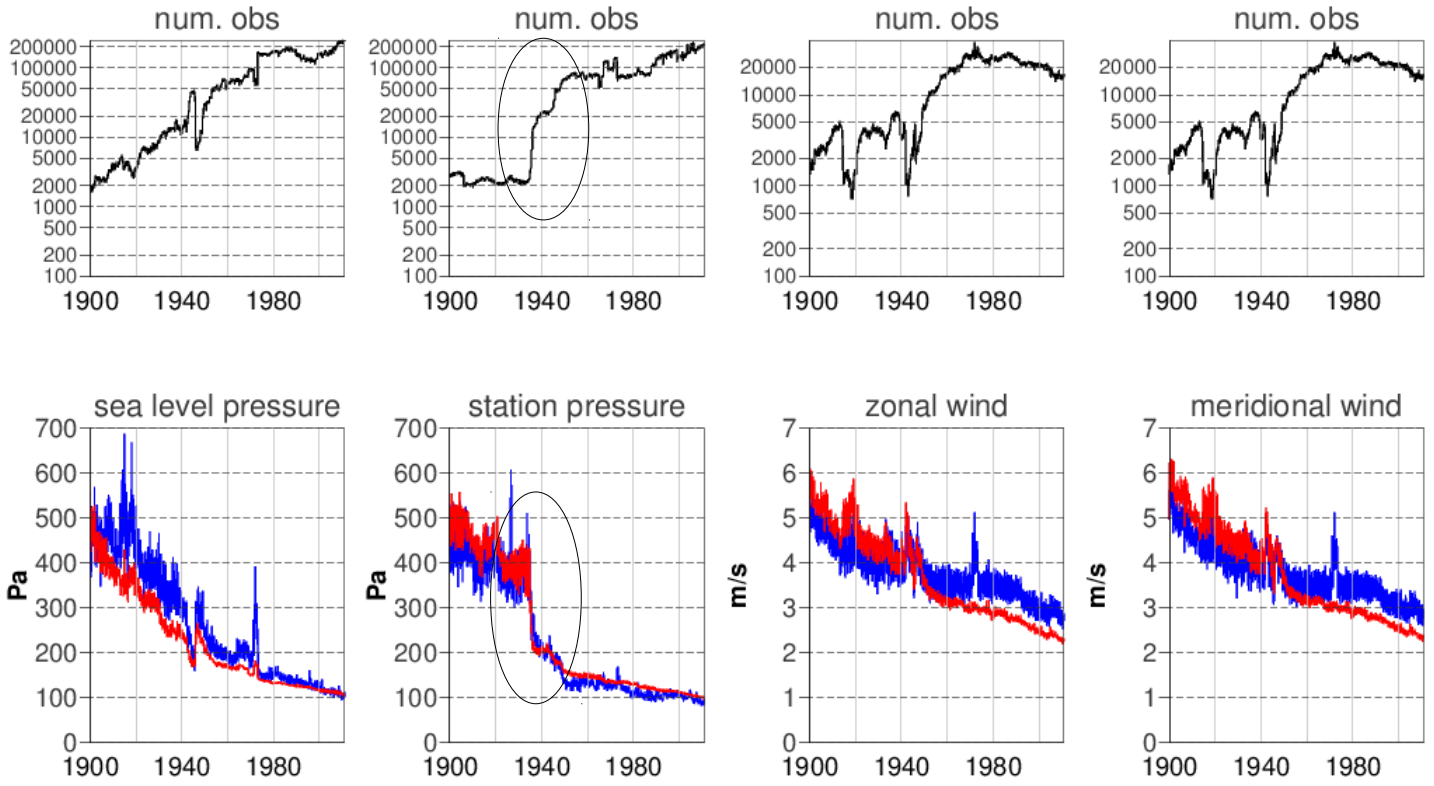
Observation space diagnostics

Spread-skill relation in observation space



Observation space diagnostics

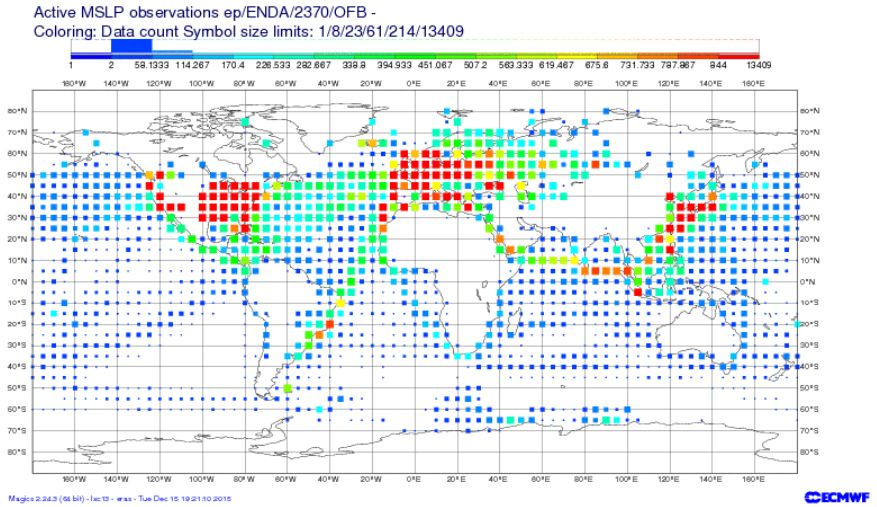
Spread-skill relation in observation space



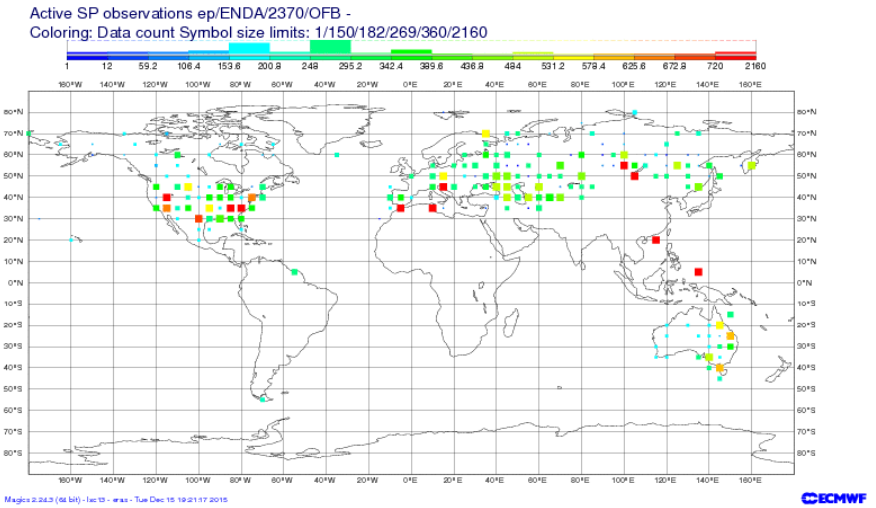
Observation space diagnostics

Number of assimilated observations in 1933 vs 1948

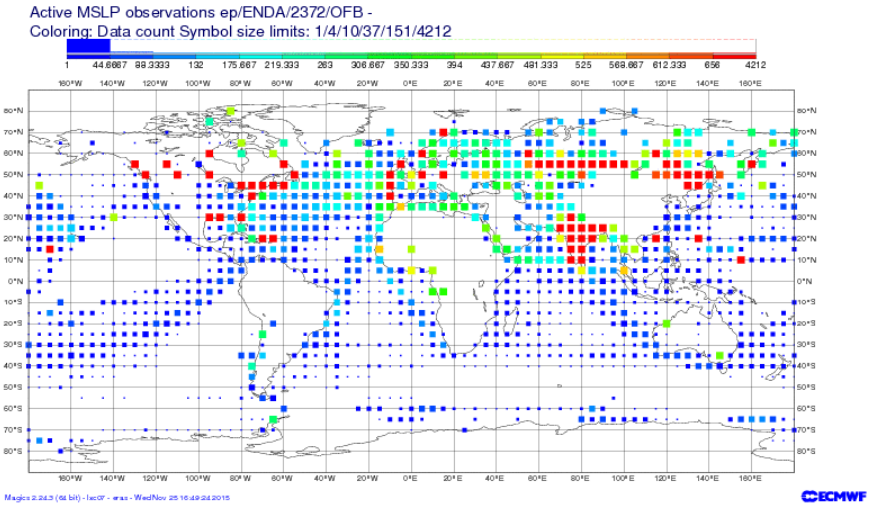
1933 MSLP



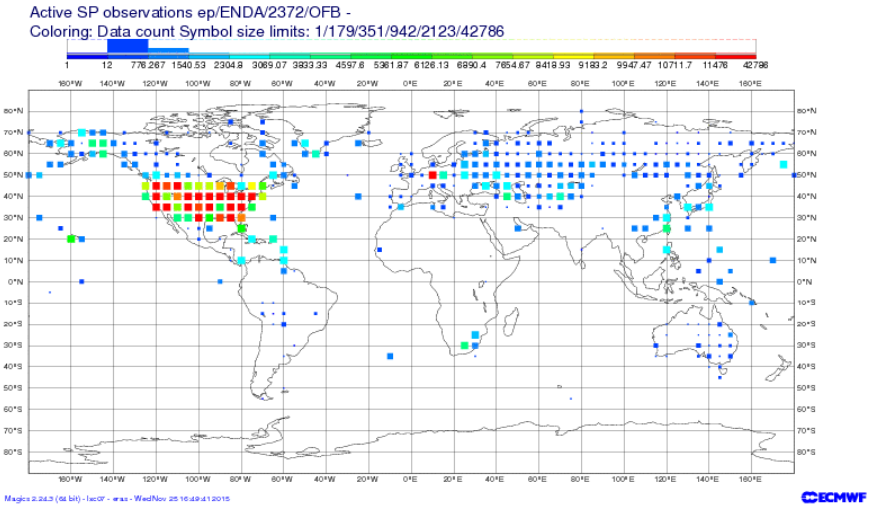
1933 Ps



1948 MSLP



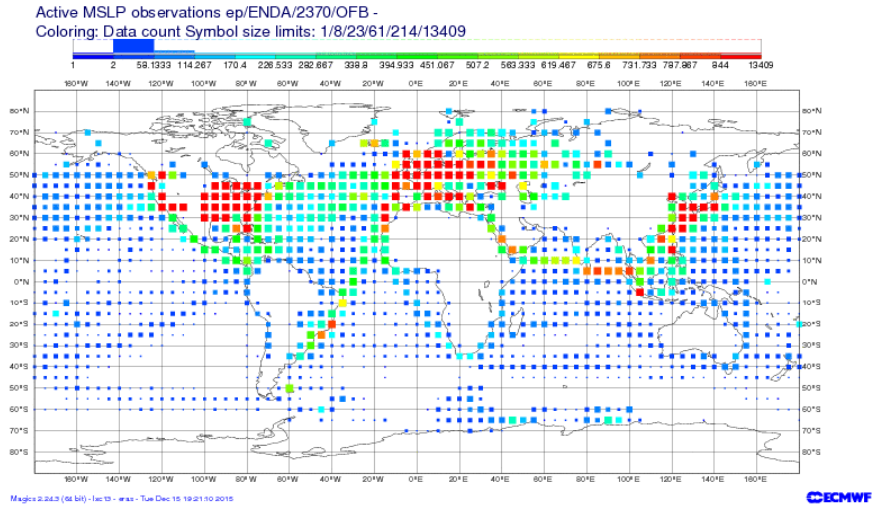
1948 Ps



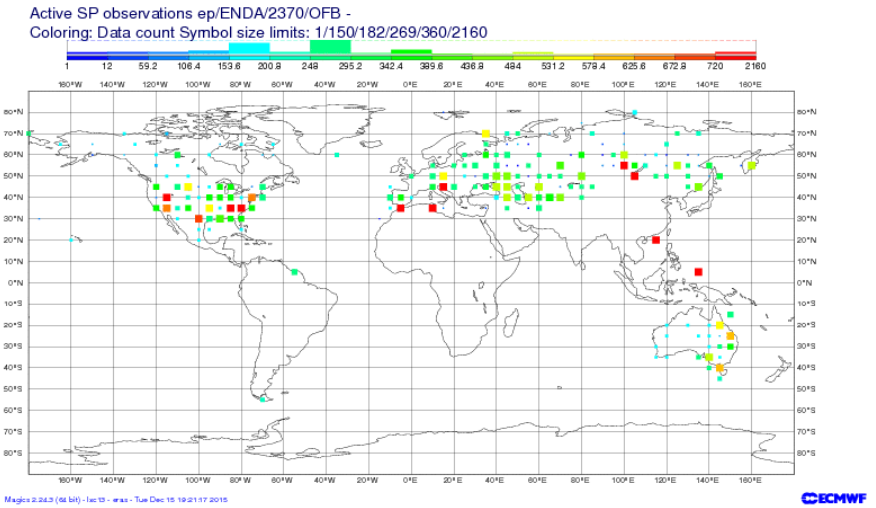
Observation space diagnostics

Number of assimilated observations in 1933 vs 1948

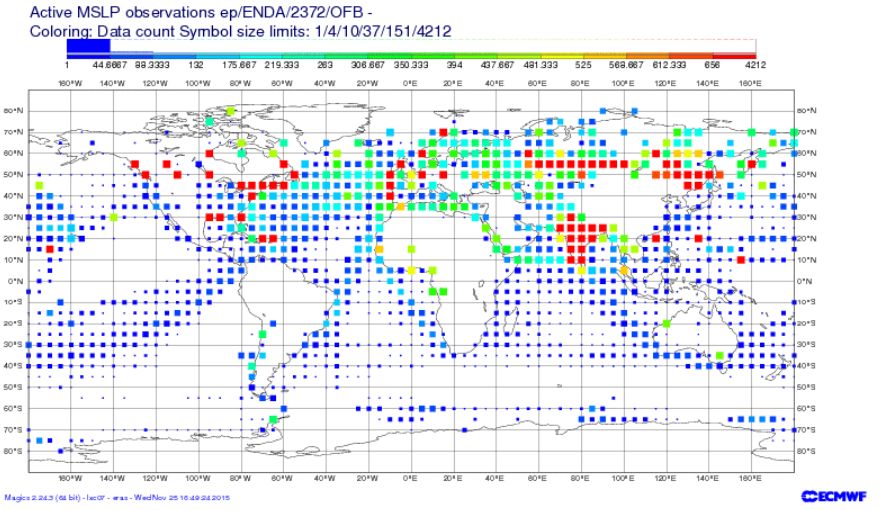
1933 MSLP



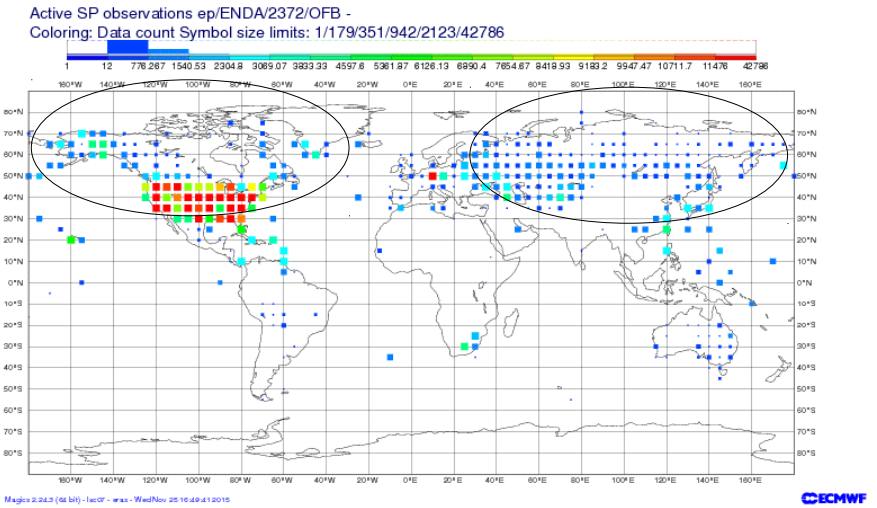
1933 Ps



1948 MSLP



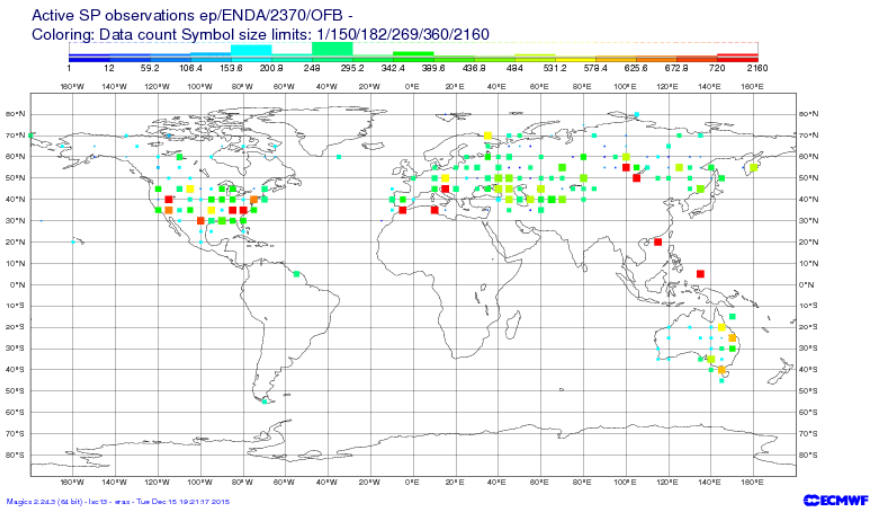
1948 Ps



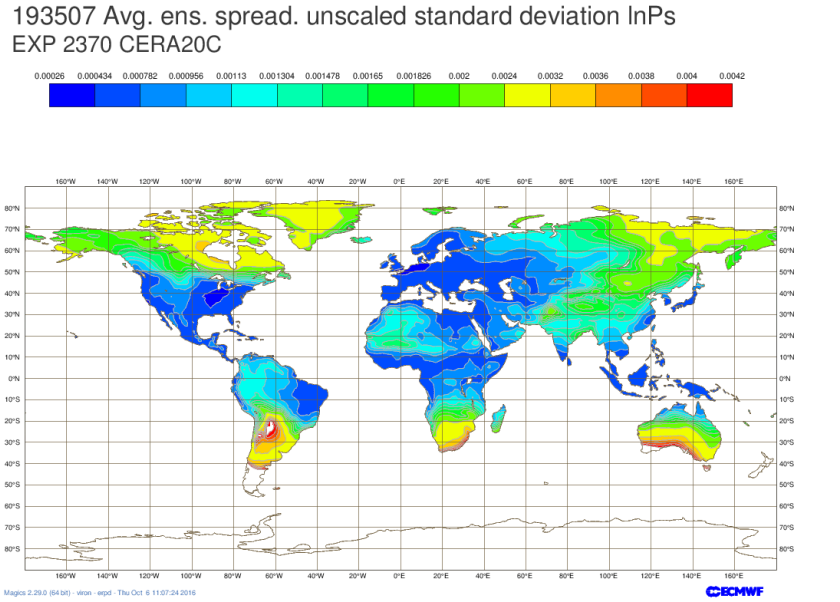
Observation space diagnostics

Number of assimilated observations in 1933 vs 1948

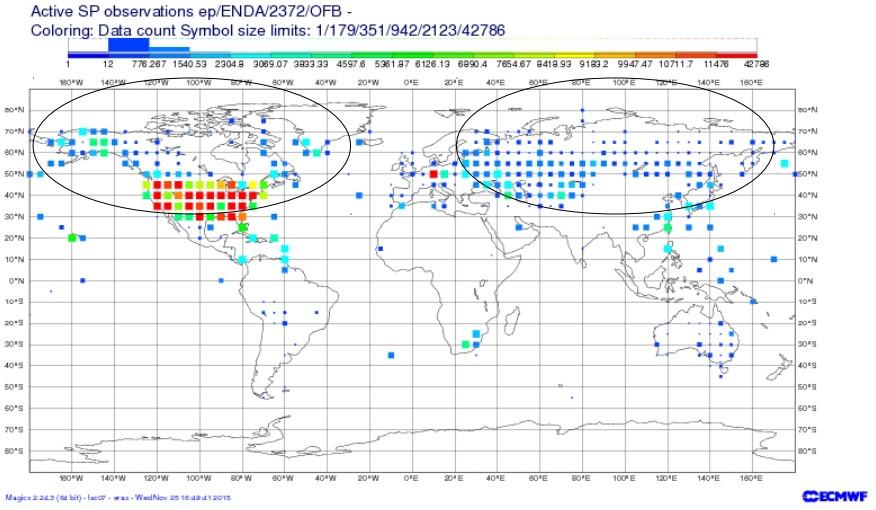
1933 Ps



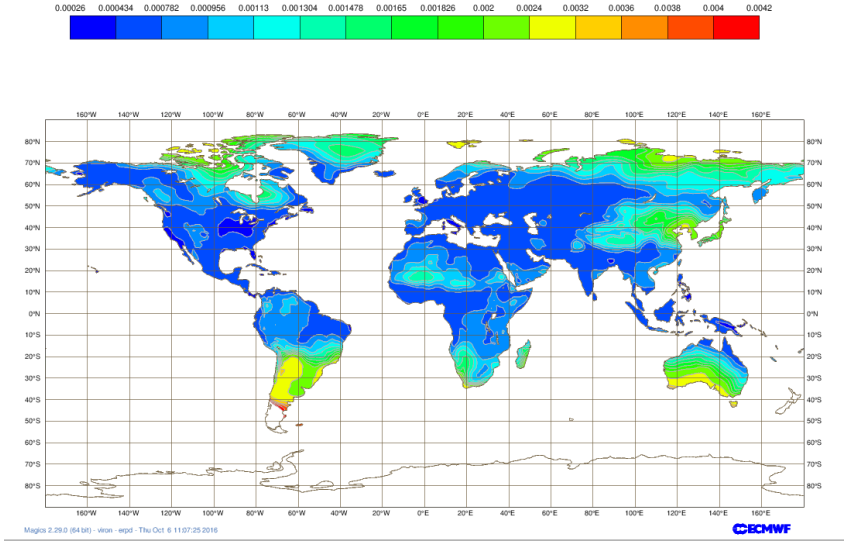
1935 ens spread



1948 Ps

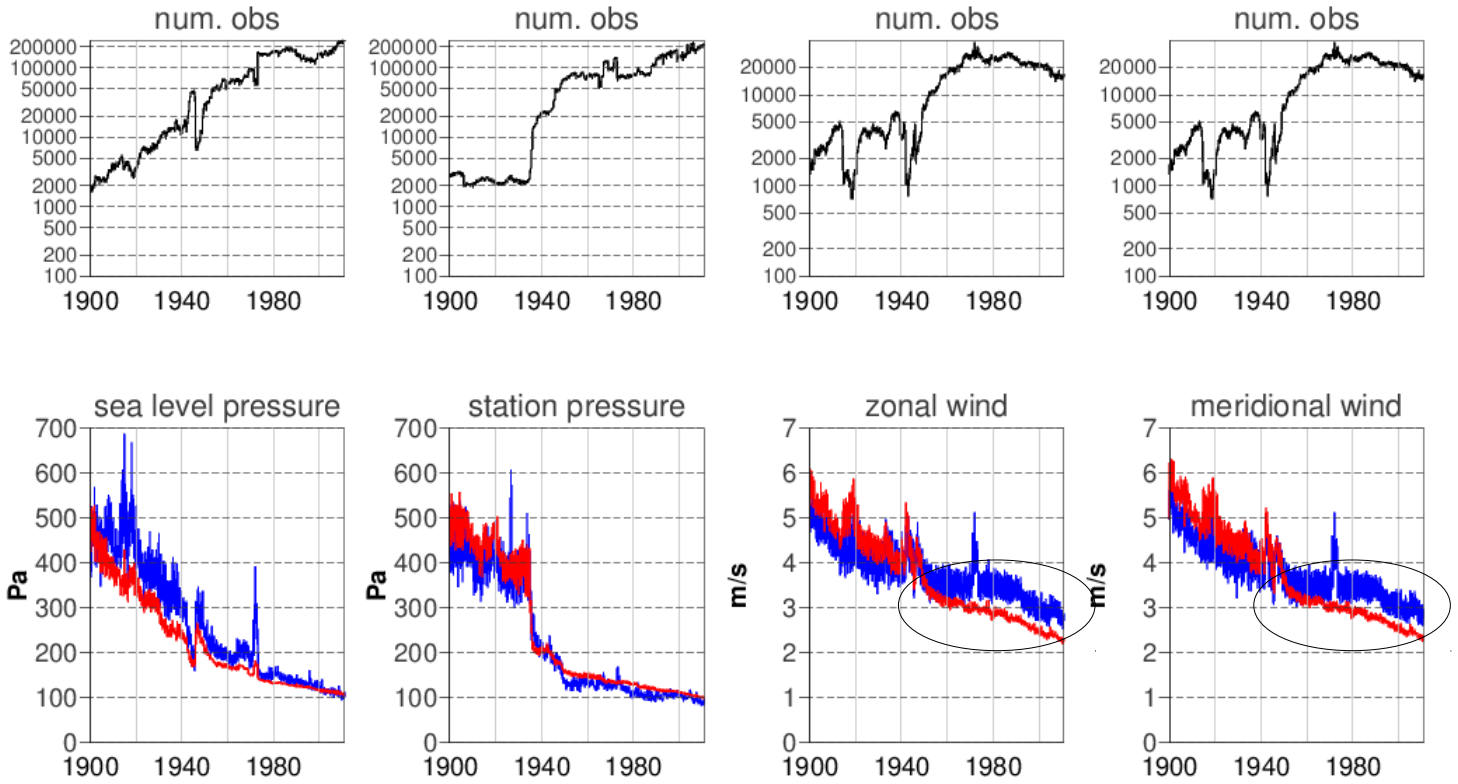


1945 ens spread



Observation space diagnostics

Spread-skill relation in observation space

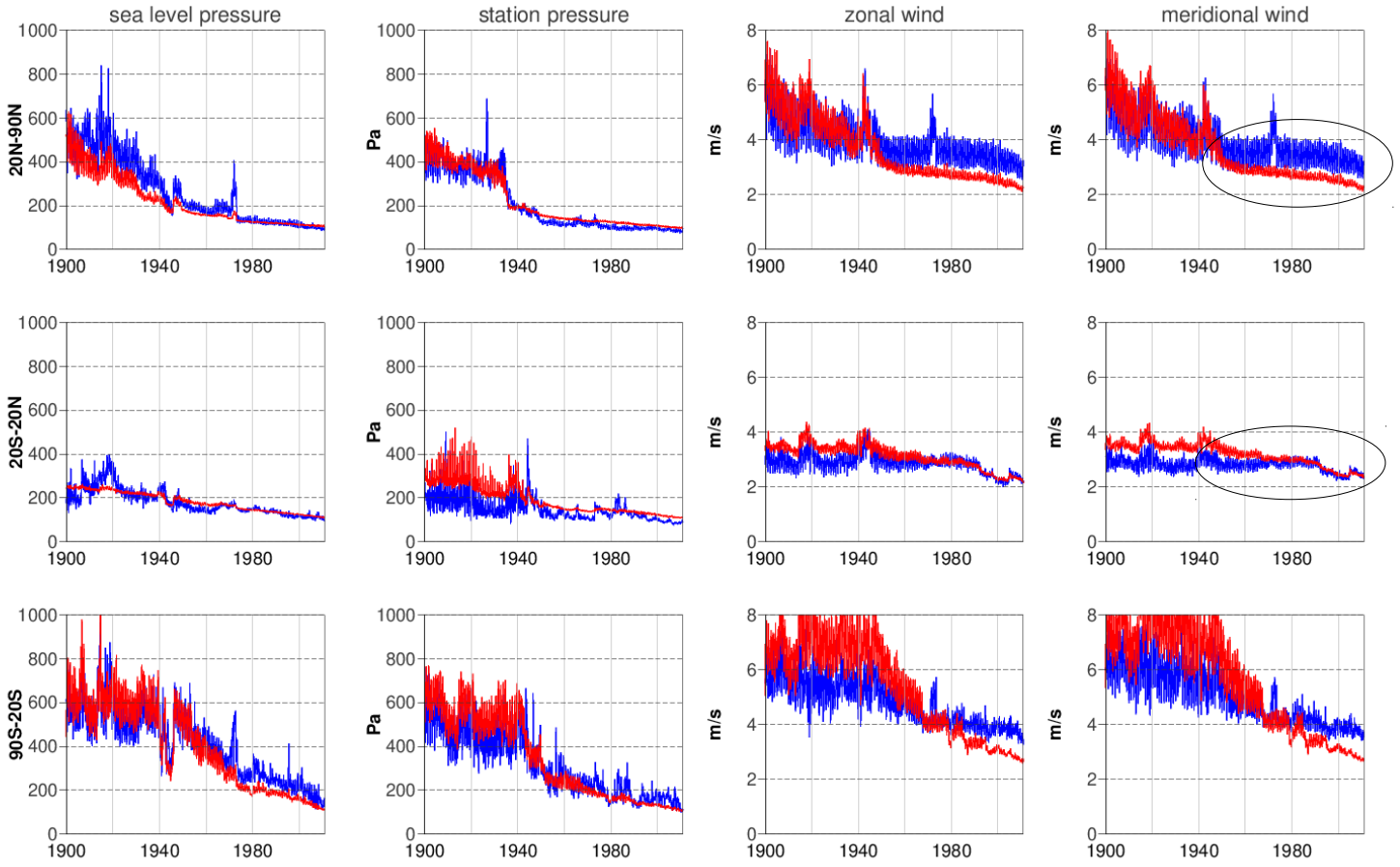


rms(o-b) sqrt(obs_error^2+bg_error^2)

Global_statistics

Observation space diagnostics

Spread-skill relation in observation space



rms(o-b) sqrt(obs_error^2+bg_error^2)

CERA-20C Uncertainty Estimation

Summary

CERA-20C ensemble:

- captures where the uncertainties are
- responds well to changes in the observing system
- is underdispersive

Seems to be a mismatch between observation and background error settings for wind