École Doctorale des Sciences de l'Environnement d'Île-de-France

Année Universitaire 2017-2018

Modélisation Numérique de l'Écoulement Atmosphérique et Assimilation de Données

Olivier Talagrand

Cours 2

26 Avril 2018

- Prévision Météorologique Numérique. Performances (d'après CEPMMT)

- Le système d'observation météorologique

- Assimilation. Les bases de l'estimation statistique

Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT, Reading, GB)

(European Centre for Medium-range Weather Forecasts, ECMWF)

Depuis mars 2016 :

Troncature triangulaire TCO1279 / O1280 (résolution horizontale ≈ 9 kilomètres)

137 niveaux dans la direction verticale (0 - 80 km)

Discrétisation en éléments finis dans la direction verticale

Dimension du vecteur d'état correspondant > 10^9

Pas de discrétisation temporelle (schéma semi-Lagrangien semi-implicite): 450 secondes

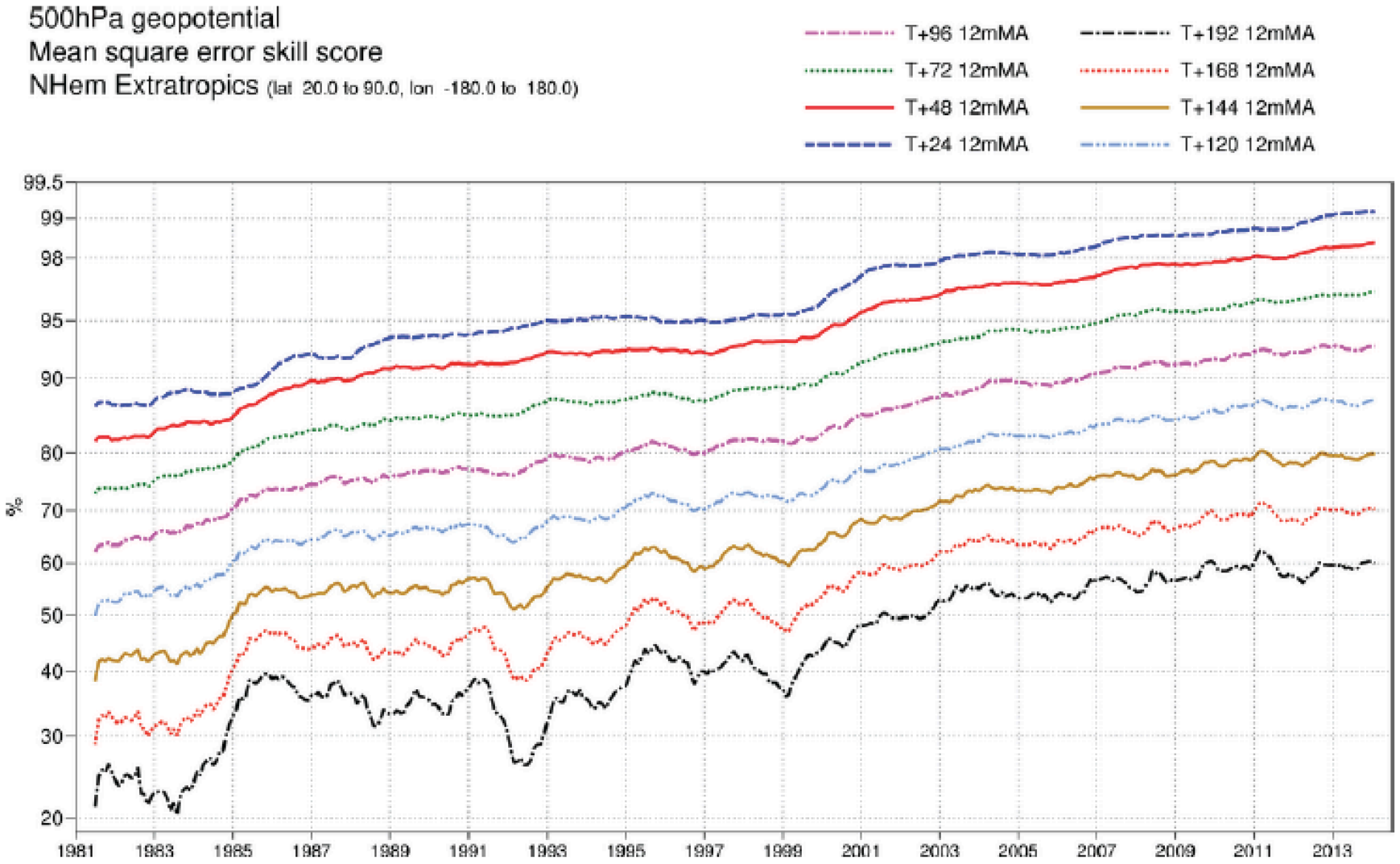


Figure 3: 500 hPa geopotential height mean square error skill score for Europe (top) and the northern hemisphere extratropics (bottom), showing 12-month moving averages for forecast ranges from 24 to 192 hours. The last point on each curve is for the 12-month period August 2013–July 2014.

Persistence = 0 ; climatology = 50 at long range

Results extracted from :

Haiden *et al.*, 2017, *Evaluation of ECMWF forecasts, including 2016-2017 upgrades*, Technical Memorandum 817, ECMWF, Reading, UK.

Available at the address

https://www.ecmwf.int/sites/default/files/elibrary/2017/17913-evaluation-ecmwf-forecasts-includuing-2016-2017-upgrades.pdf

(see also site of ECMWF)

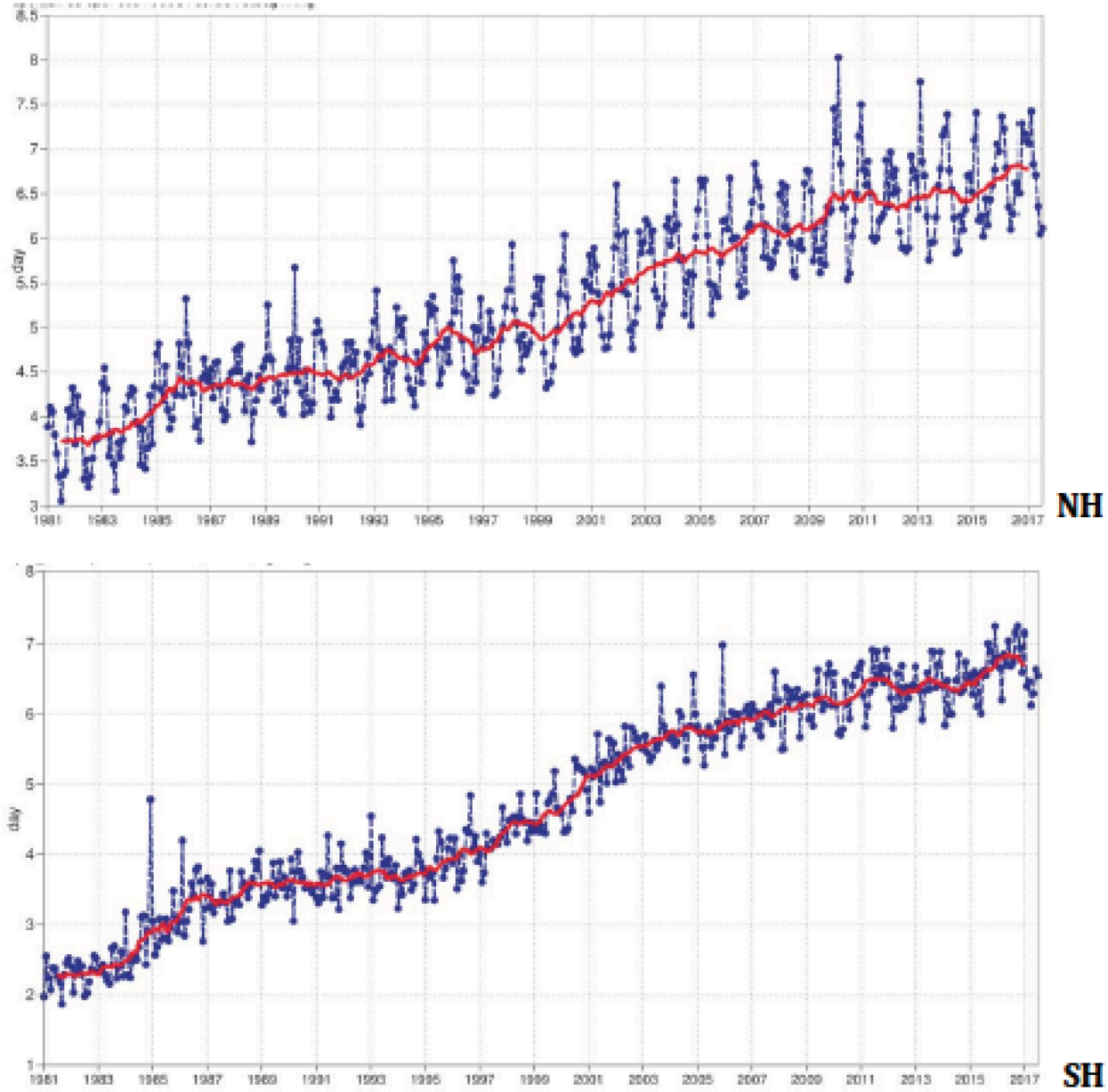


Figure 5: Primary headline score for the high-resolution forecasts. Evolution with time of the 500 hPa geopotential height forecast performance – each point on the curves is the forecast range at which the monthly mean (blue lines) or 12-month mean centred on that month (red line) of the forecast anomaly correlation (ACC) with the verifying analysis falls below 80% for Europe (top), northern hemisphere extratropics (centre) and southern hemisphere extratropics (bottom).

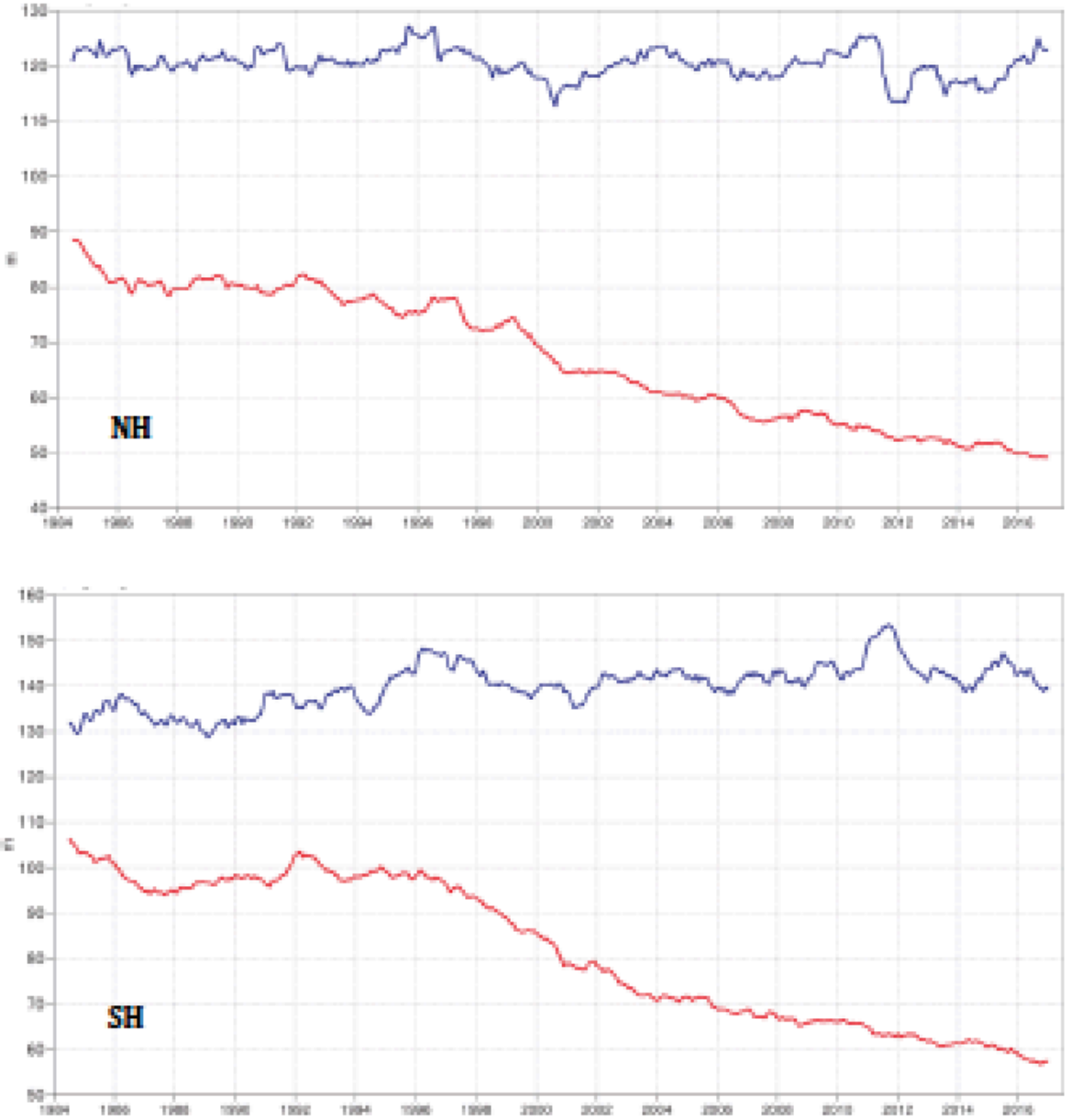


Figure 6: Root mean square (RMS) error of forecasts of 500 hPa geopotential height (m) at day 6 (red), verified against analysis. For comparison, a reference forecast made by persisting the analysis over 6 days is shown (blue). Plotted values are 12-month moving averages; the last point on the curves is for the 12-month period August 2016–July 2017. Results are shown for the northern extra-tropics (top), and the southern extra-tropics (bottom).

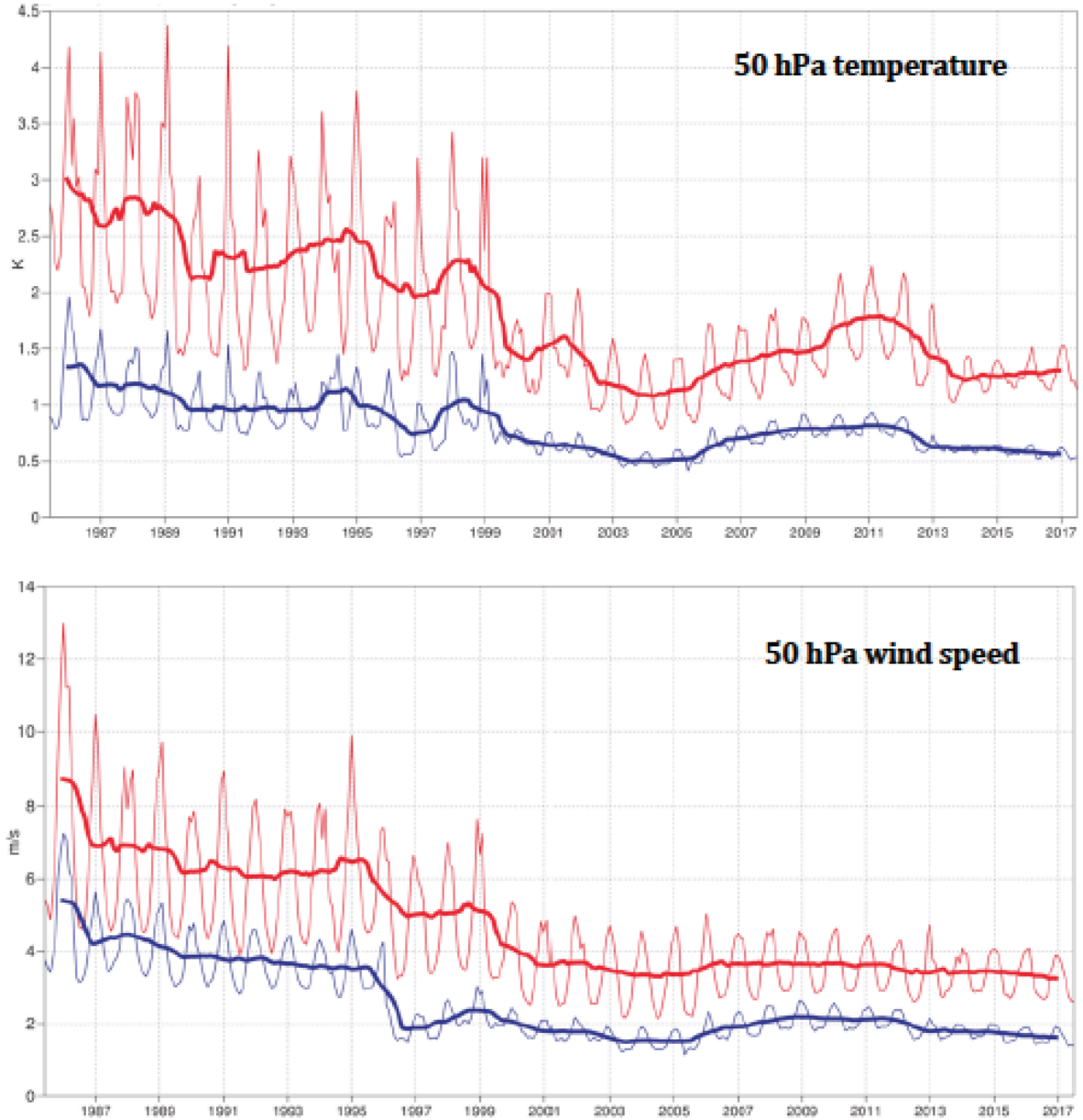


Figure 8: Model scores for temperature (top) and wind (bottom) in the northern extratropical stratosphere. Curves show the monthly average RMS temperature and vector wind error at 50 hPa for one-day (blue) and five-day (red) forecasts, verified against analysis. 12-month moving average scores are also shown (in bold).

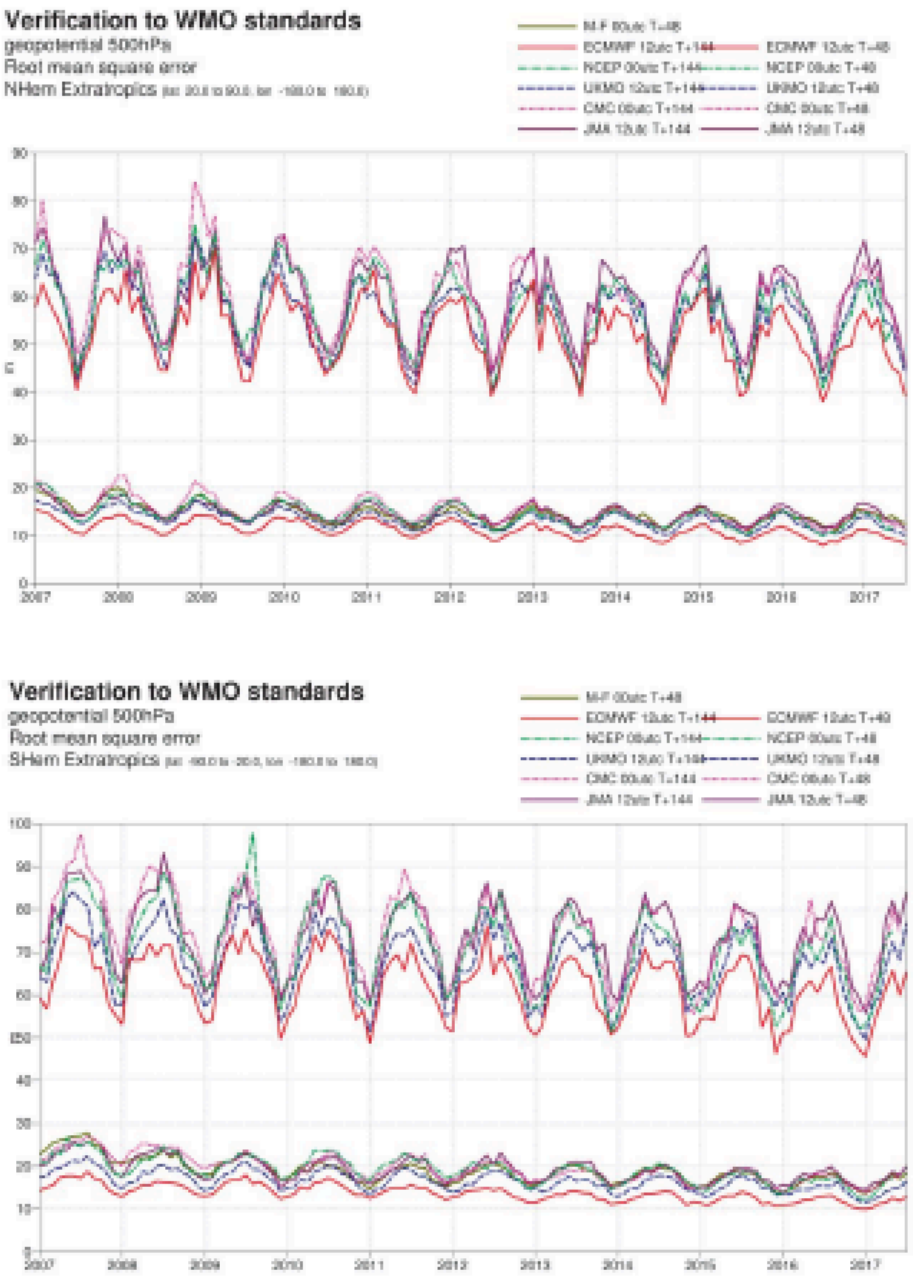
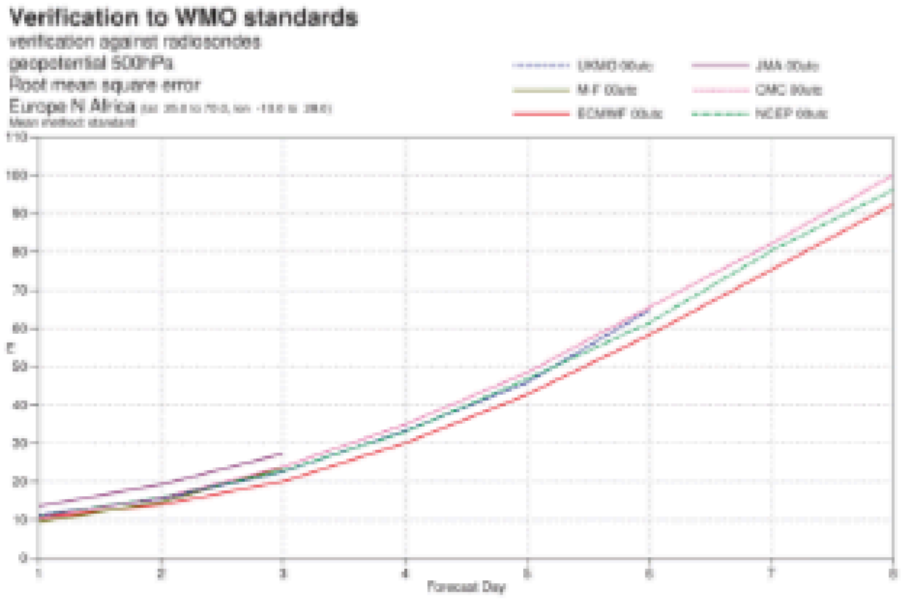


Figure 15: WMO-exchanged scores from global forecast centres. RMS error of 500 hPa geopotential height over northern (top) and southern (bottom) extratropics. In each panel the upper curves show the six-day forecast error and the lower curves show the two-day forecast error. Each model is verified against its own analysis. JMA = Japan Meteorological Agency, CMC = Canadian Meteorological Centre, UKMO = the UK Met Office, NCEP = U.S. National Centers for Environmental Prediction, M-F = Météo France.



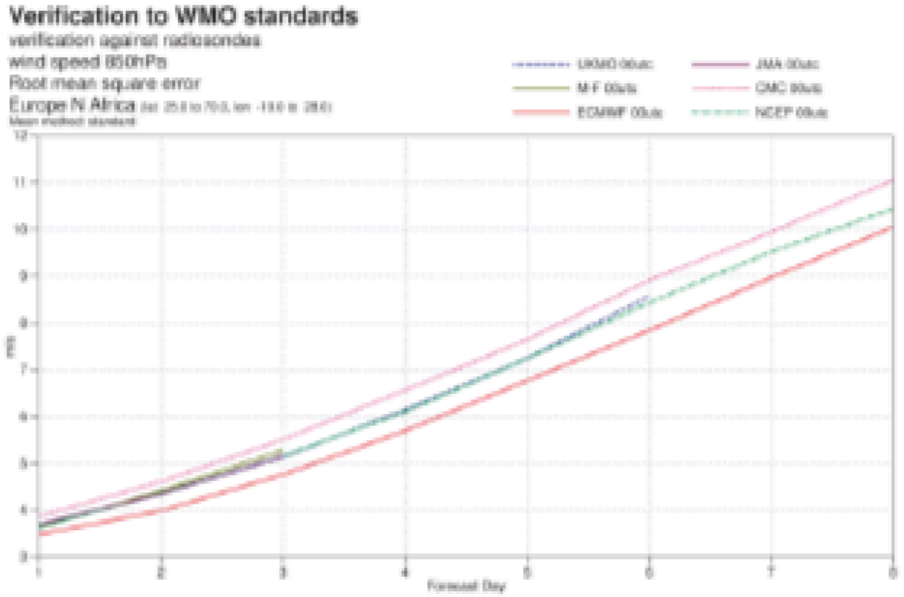


Figure 16: WMO-exchanged scores for verification against radiosondes: 500 hPa height (top) and 850 hPa wind (bottom) RMS error over Europe (annual mean August 2016–July 2017).

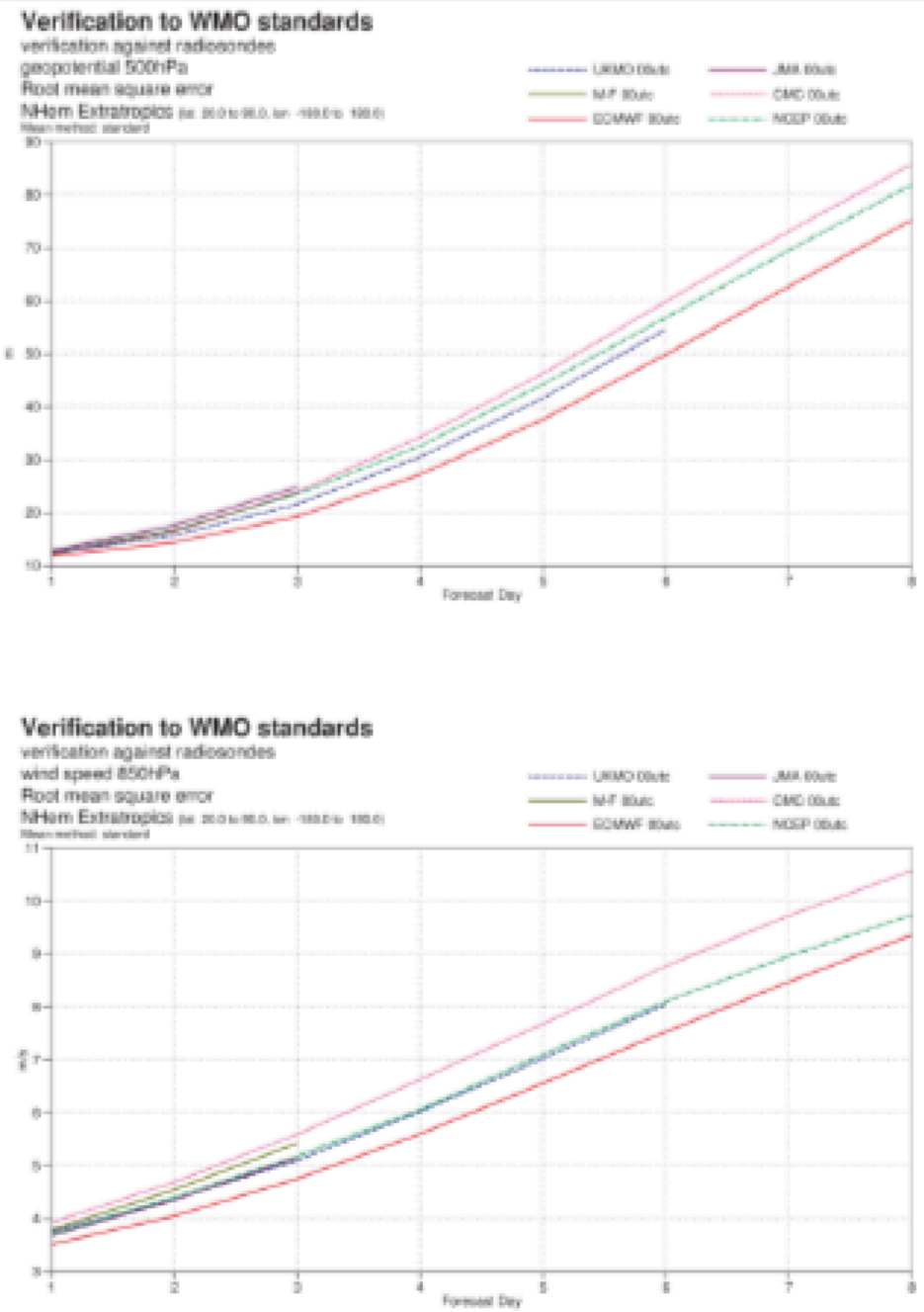


Figure 17: As Figure 16 for the northern hemisphere extratropics.

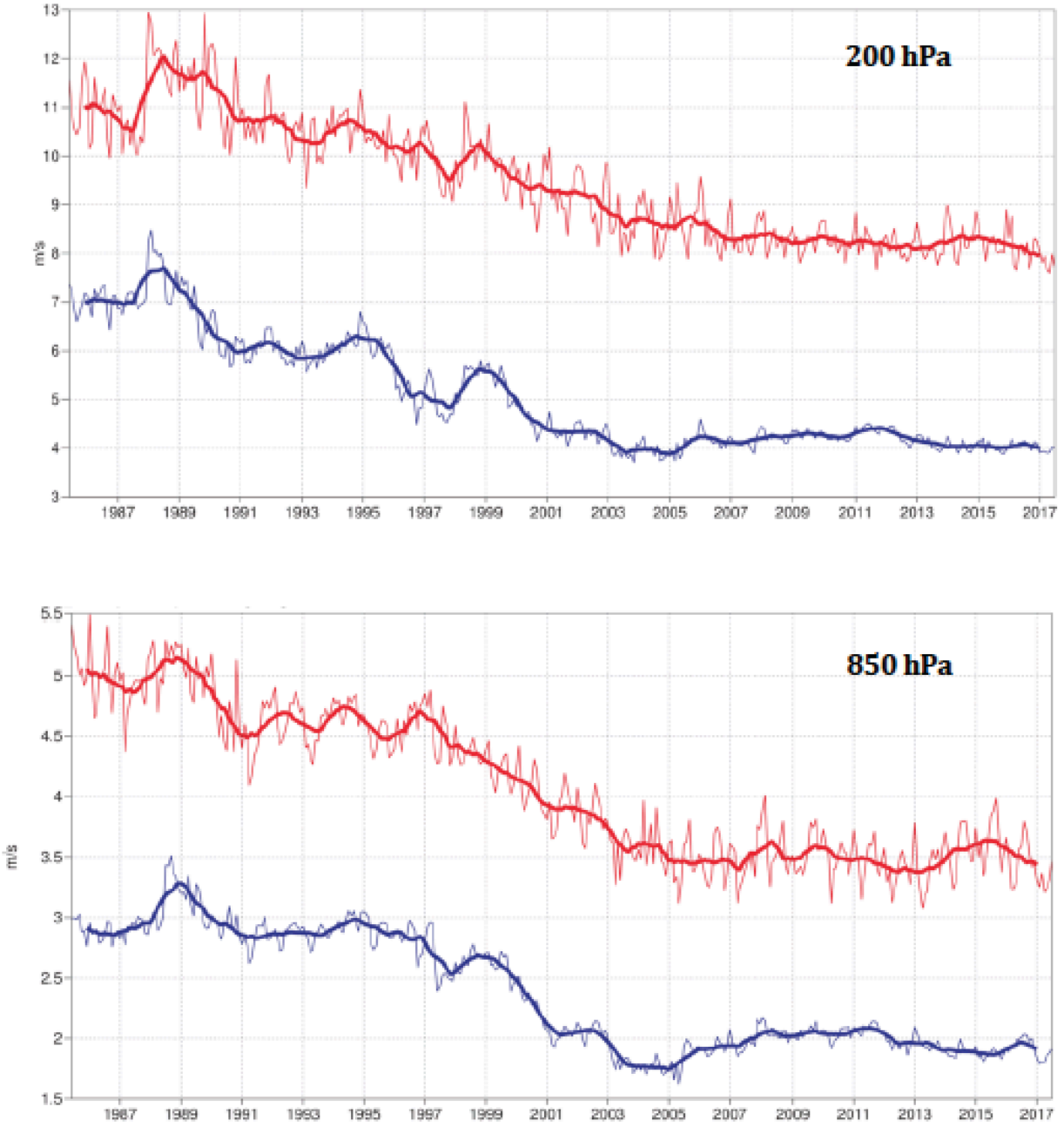


Figure 14: Forecast performance in the tropics. Curves show the monthly average RMS vector wind errors at 200 hPa (top) and 850 hPa (bottom) for one-day (blue) and five-day (red) forecasts, verified against analysis. 12-month moving average scores are also shown (in bold).

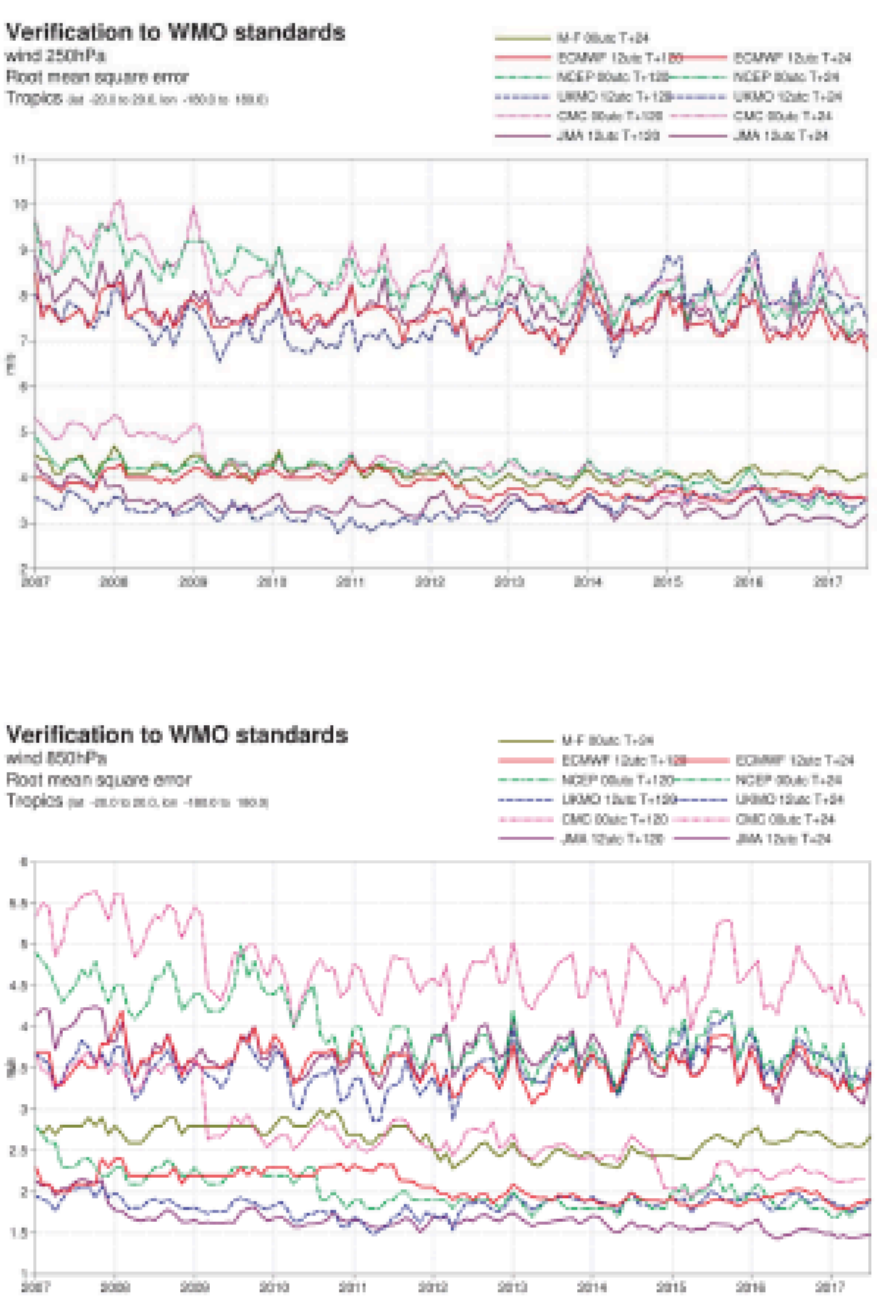


Figure 18: WMO-exchanged scores from global forecast centres. RMS vector wind error over tropics at 250 hPa (top) and 850 hPa (bottom). In each panel the upper curves show the five-day forecast error and the lower curves show the one-day forecast error. Each model is verified against its own analysis.

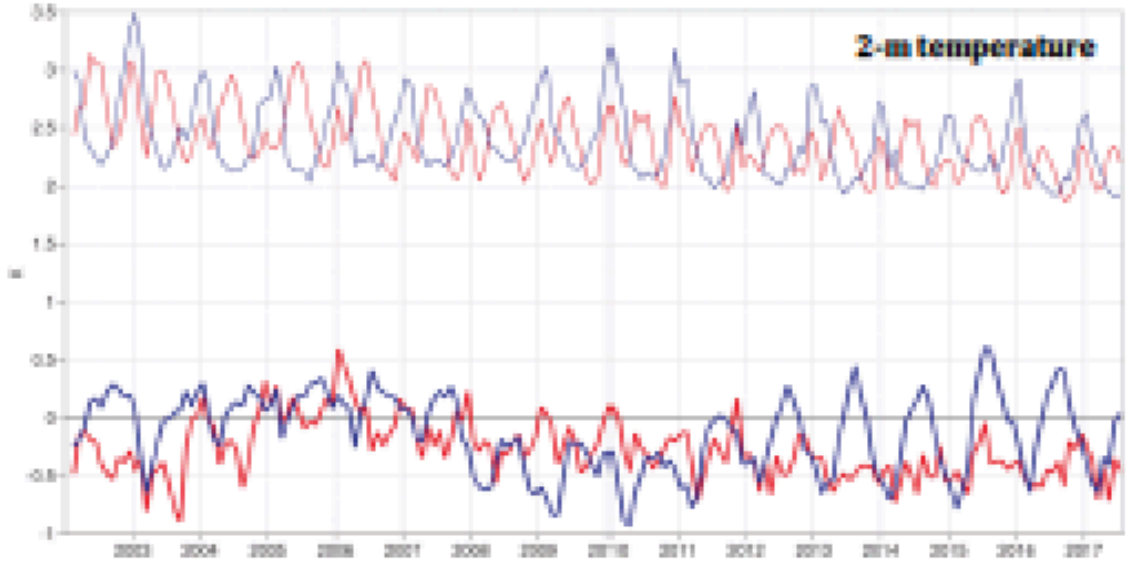


Figure 22: Verification of 2 m temperature forecasts against European SYNOP data on the GTS for 60-hour (night-time) and 72-hour (daytime) forecasts. Lower pair of curves shows bias, upper curvesare standard deviation of error.

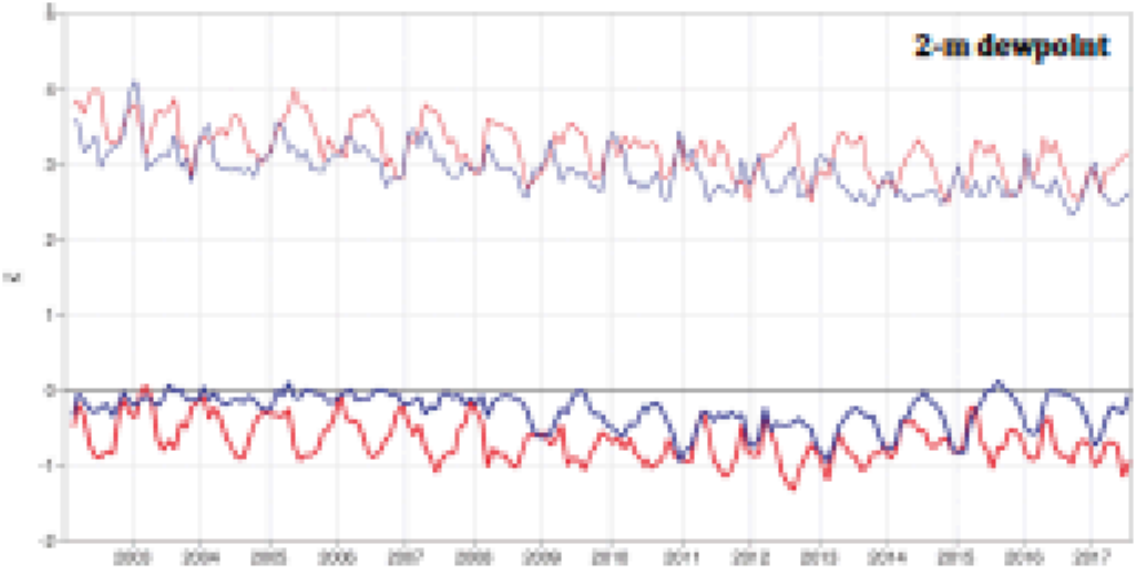


Figure 23: Verification of 2 m dew point forecasts against European SYNOP data on the Global Telecommunication System (GTS) for 60-hour (night-time) and 72-hour (daytime) forecasts. Lower pair of curves shows bias, upper curves show standard deviation of error.

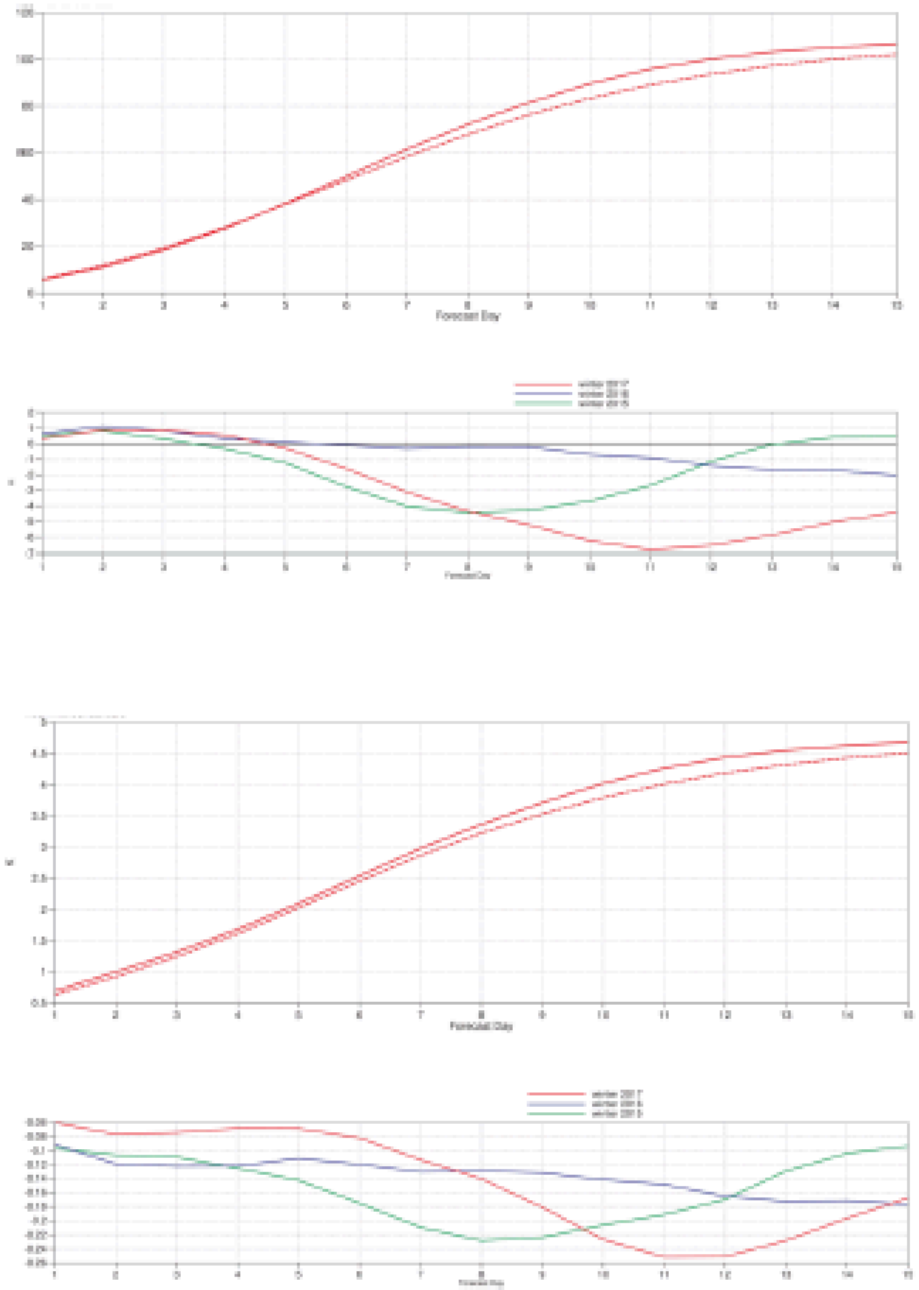


Figure 10: Ensemble spread (standard deviation, dashed lines) and RMS error of ensemble-mean (solid lines) for winter 2016–2017 (upper figure in each panel), and differences of ensemble spread and RMS error of ensemble mean for last three winter seasons (lower figure in each panel, negative values indicate spread is too small); verification is against analysis, plots are for 500 hPa geopotential (top) and 850 hPa temperature (bottom) over the extratropical northern hemisphere for forecast days 1 to 15.

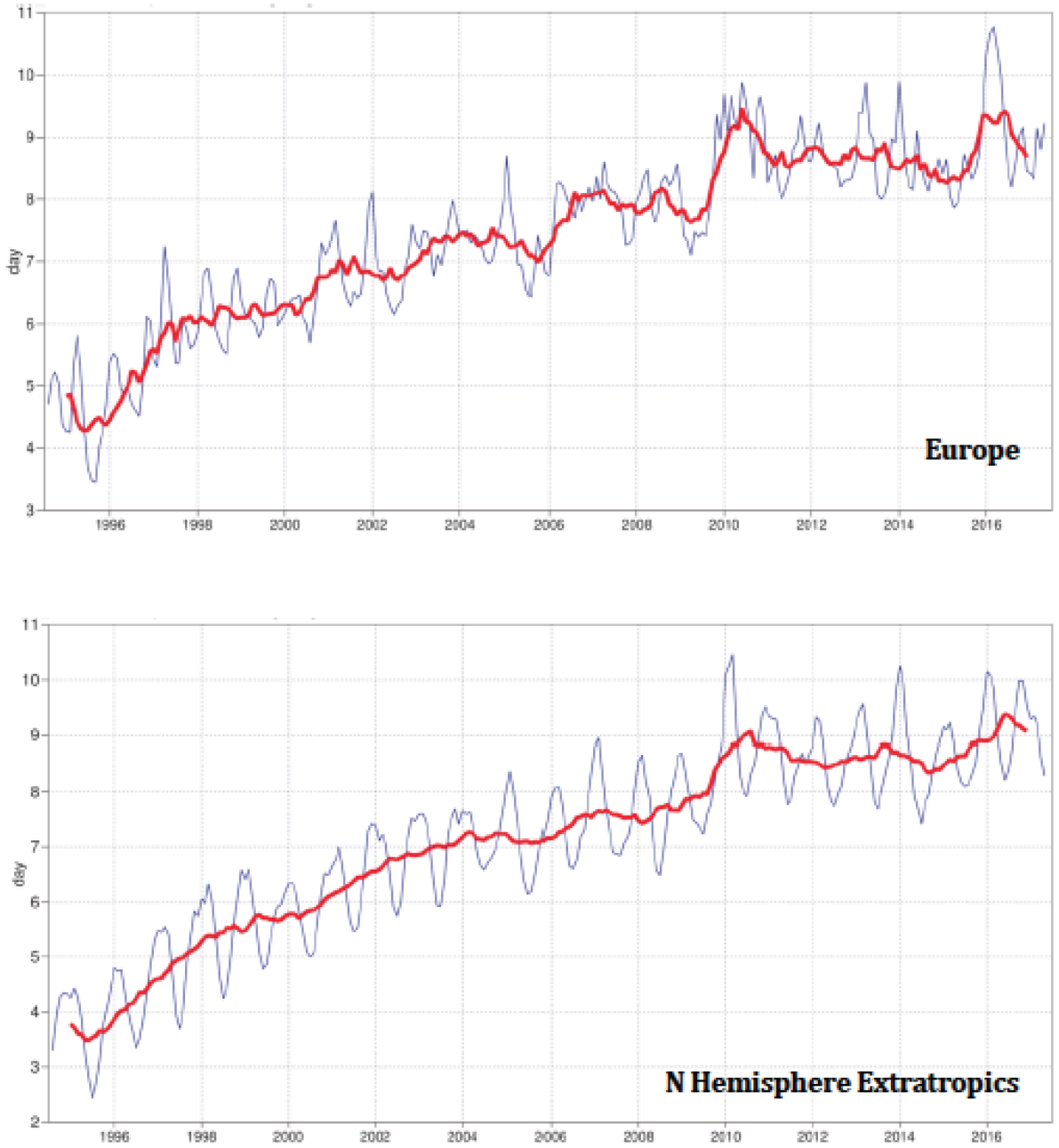


Figure 9: Primary headline score for the ensemble probabilistic forecasts. Evolution with time of 850 hPa temperature ensemble forecast performance, verified against analysis. Each point on the curves is the forecast range at which the 3-month mean (blue lines) or 12-month mean centred on that month (red line) of the continuous ranked probability skill score (CPRSS) falls below 25% for Europe (top), northern hemisphere extratropics (bottom).

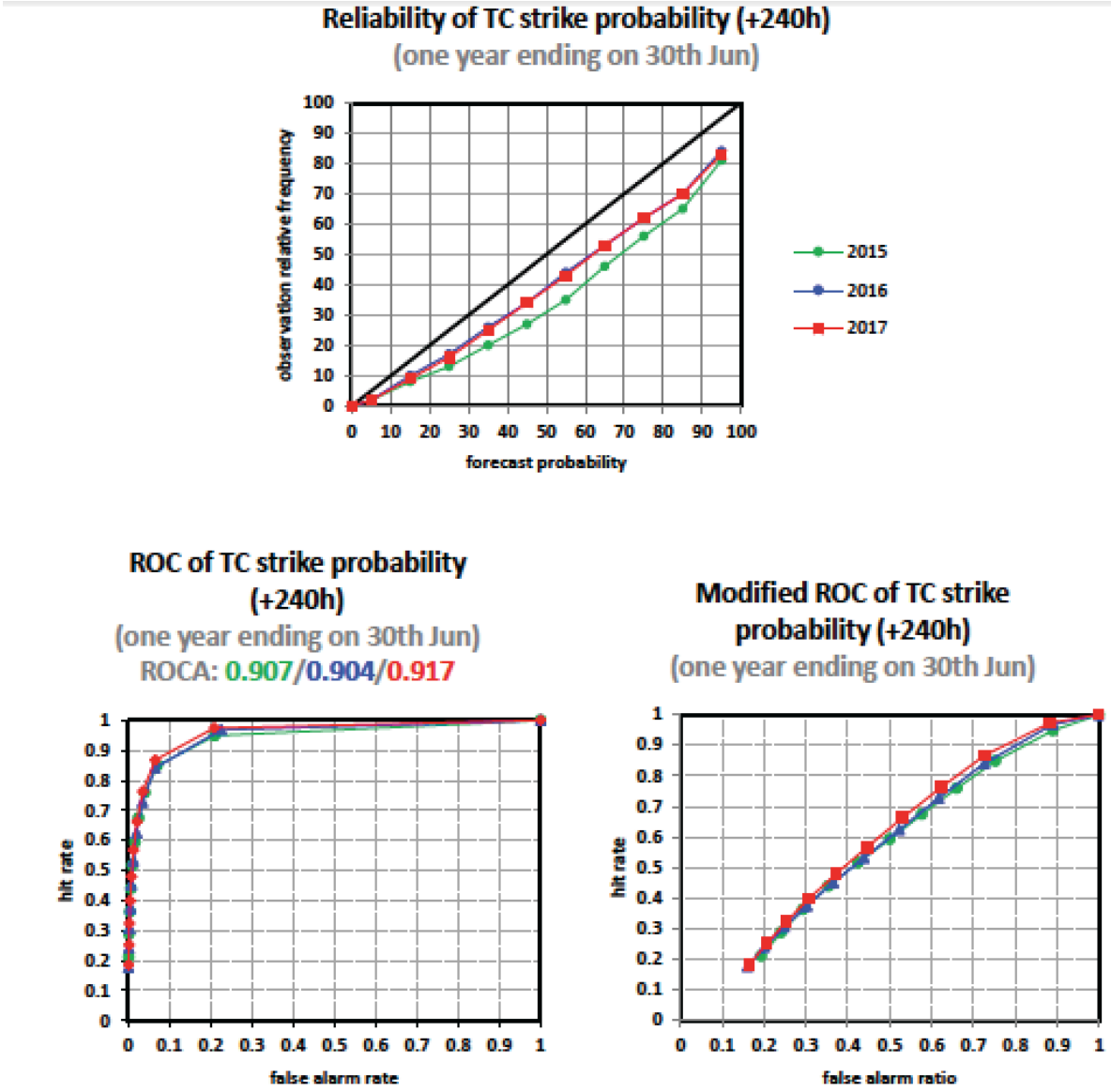
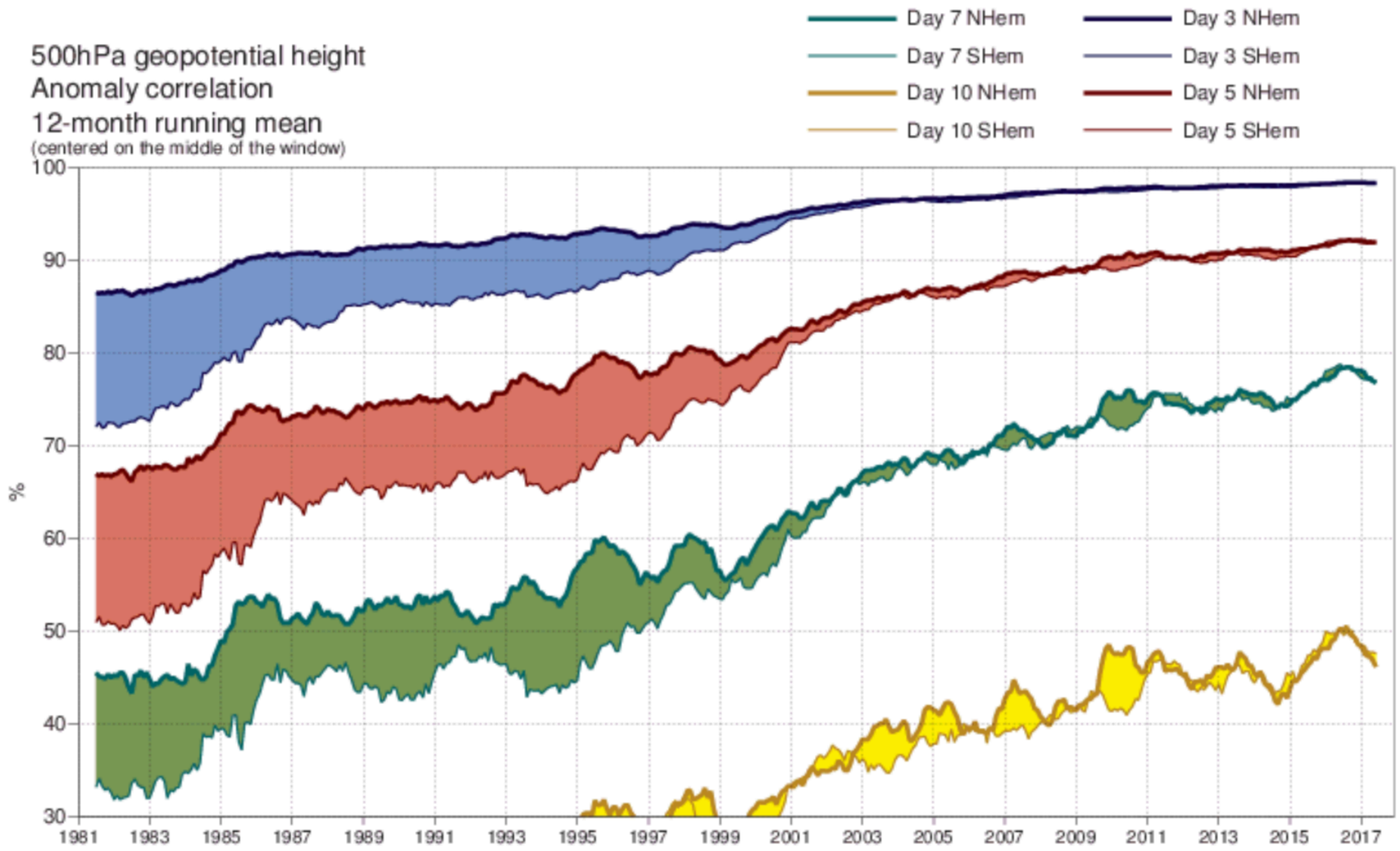


Figure 33: Probabilistic verification of ensemble tropical cyclone forecasts at day 10 for three 12-month periods: July 2014–June 2015 (green), July 2015–June 2016 (blue) and July 2016–June 2017 (red). Upper panel shows reliability diagram (the closer to the diagonal, the better). The lower panel shows (left) the standard ROC diagram and (right) a modified ROC diagram, where the false alarm ratio is used instead of the false alarm rate. For both ROC and modified ROC, the closer the curve is to the upper-left corner, the better, indicating a greater proportion of hits, and fewer false alarms.



Anomaly correlation of ECMWF 500hPa height forecasts

ECMWF

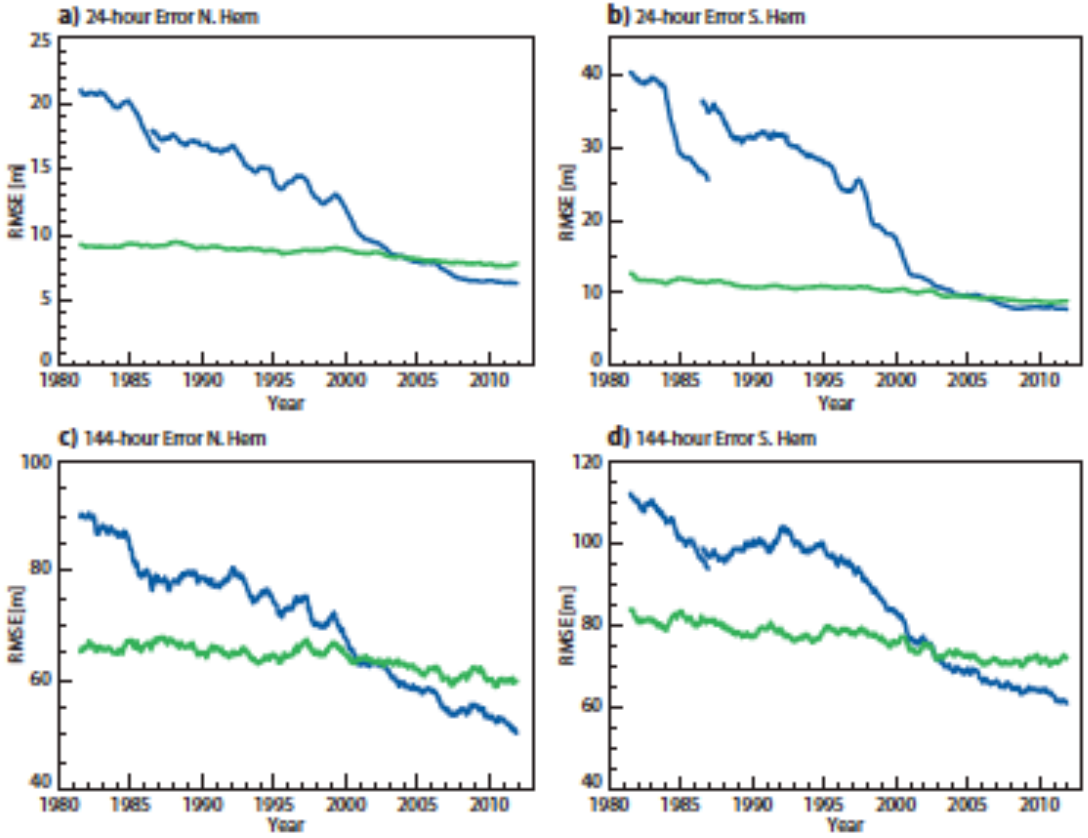


FIG. 3. Evolution of forecast errors from 1981 to 2012 for N.Hem (a and c) and S.Hem (b and d). Operational forecasts (blue) and ERA Interim (green). Note that before 1986 the operational analysis is used to verify the operational forecasts, after 1986 ERA Interim is used for the verification (with an overlap of 6 months present).

Magnusson and Källén, 2013, *Mon. Wea. Rev.*, **141**, 3142–3153

Problèmes restants

- Cycle de l'eau (évaporation, condensation, influence sur le rayonnement absorbé ou émis par l'atmosphère)

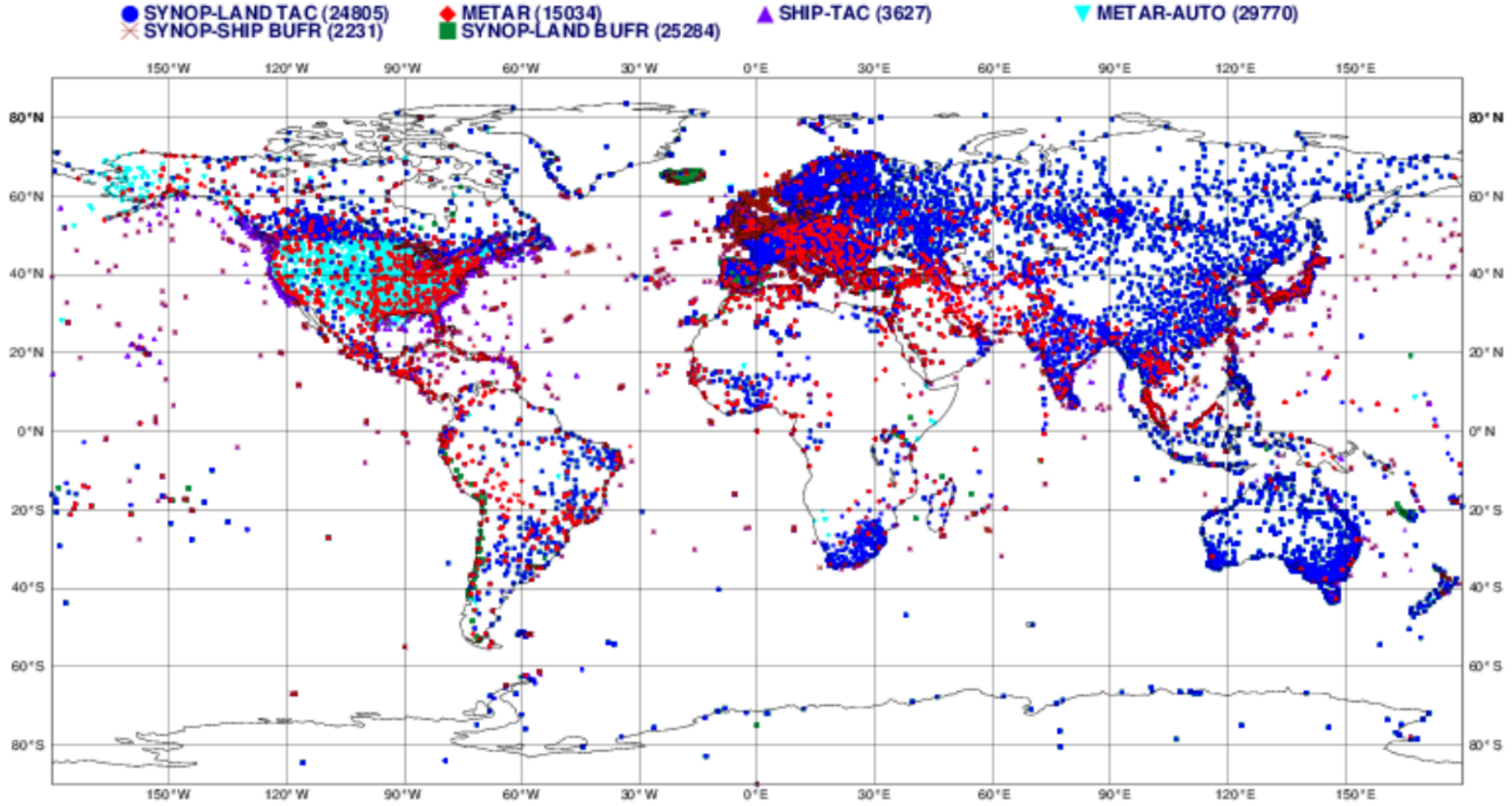
- Échanges avec l'océan ou la surface continentale (chaleur, eau, quantité de mouvement, …)

- …

ECMWF data coverage (all observations) - SYNOP-SHIP-METAR

25/04/2018 00

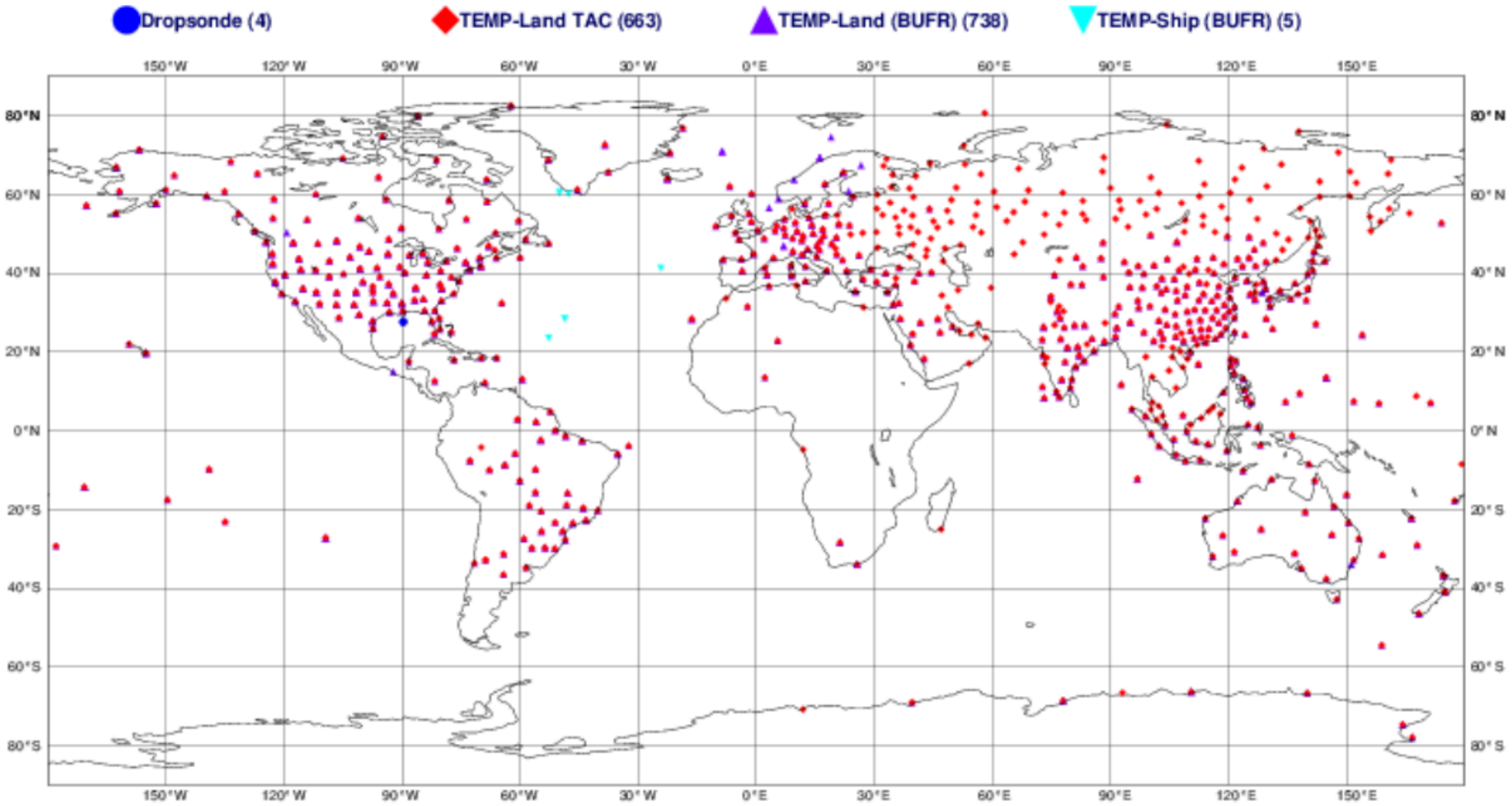
Total number of obs = 100751



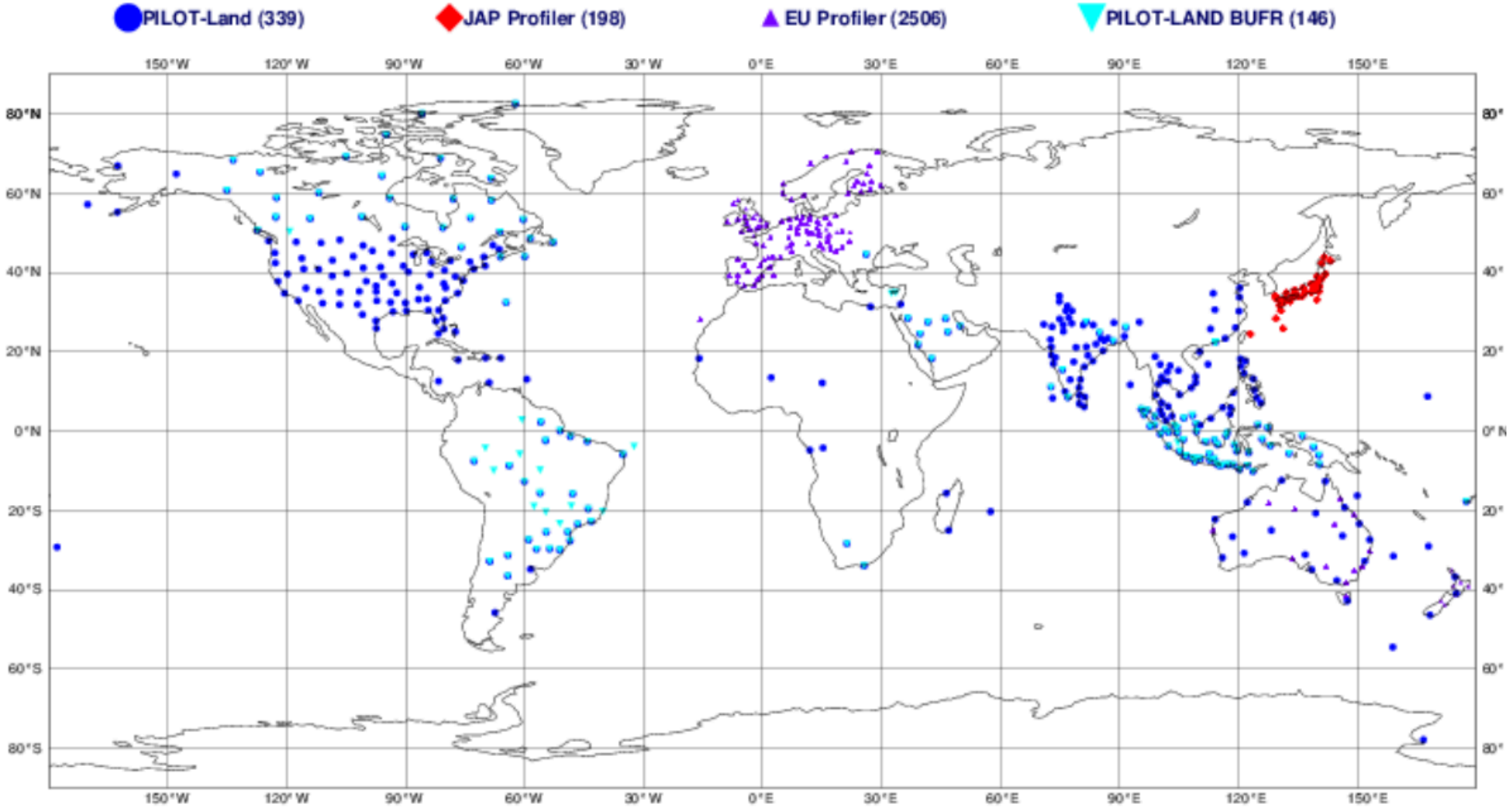
ECMWF data coverage (all observations) - RADIOSONDE

25/04/2018 00

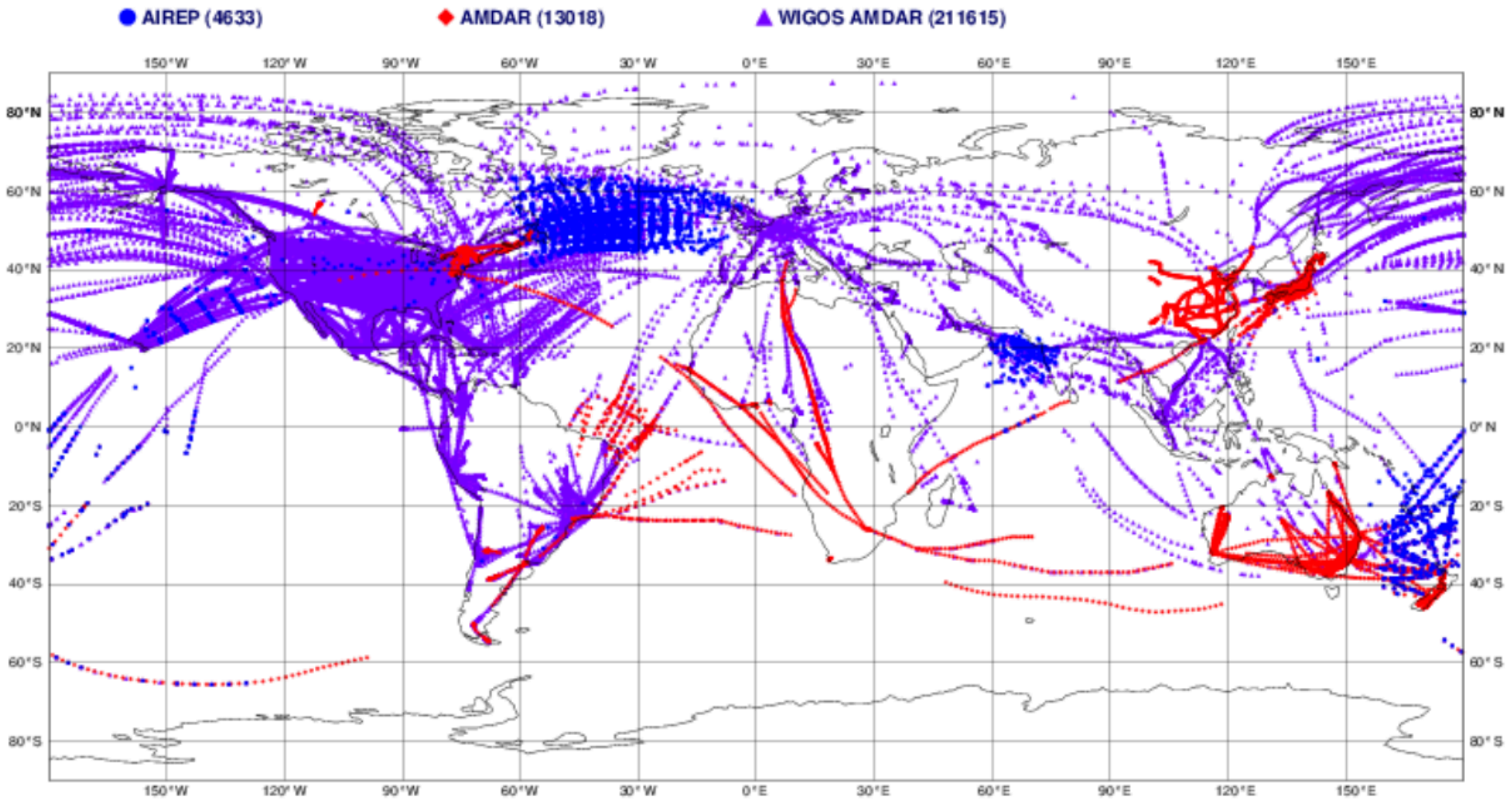
Total number of obs = 1410



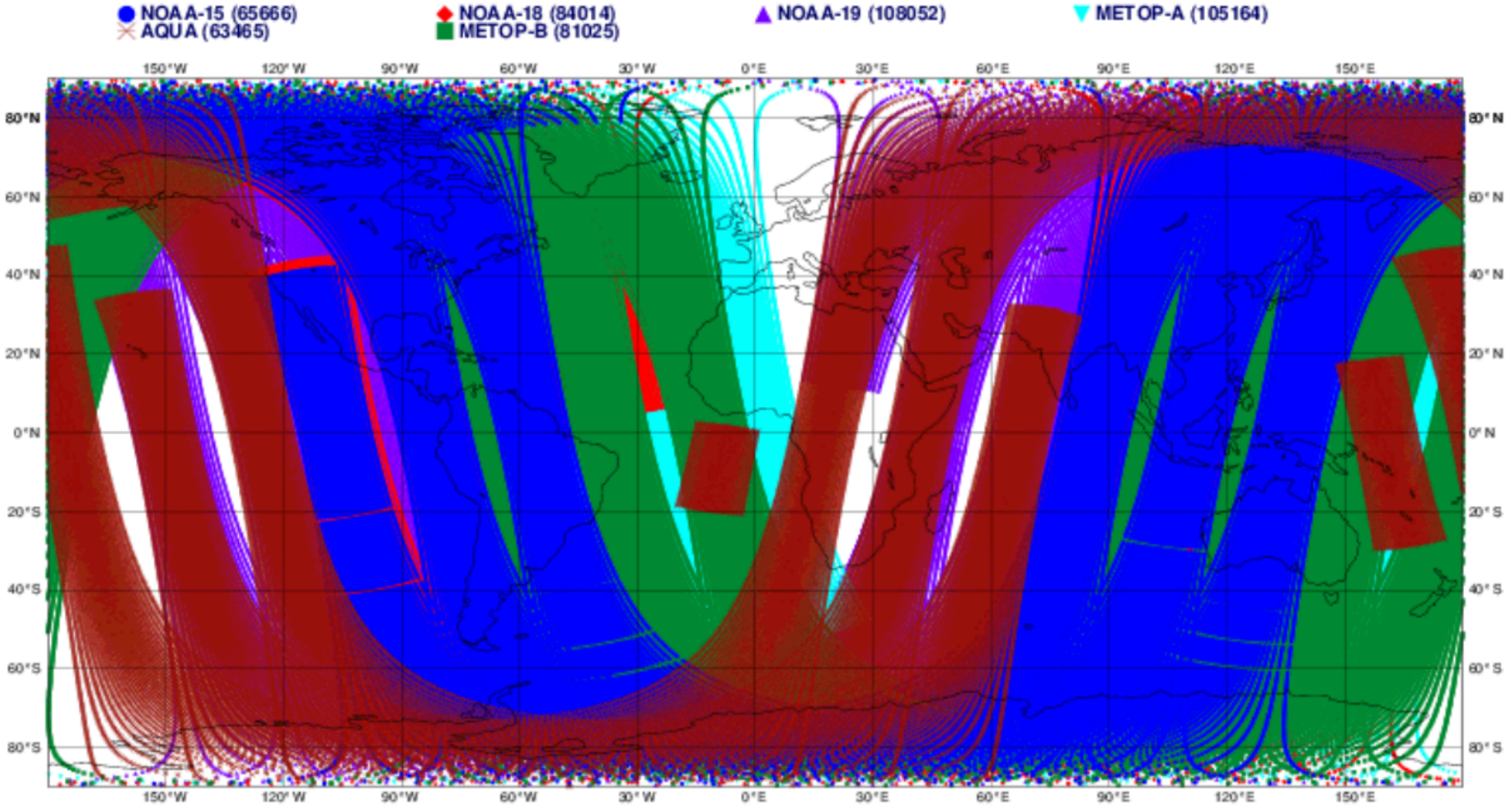
ECMWF data coverage (all observations) - PILOT
25/04/2018 00
Total number of obs = 3189



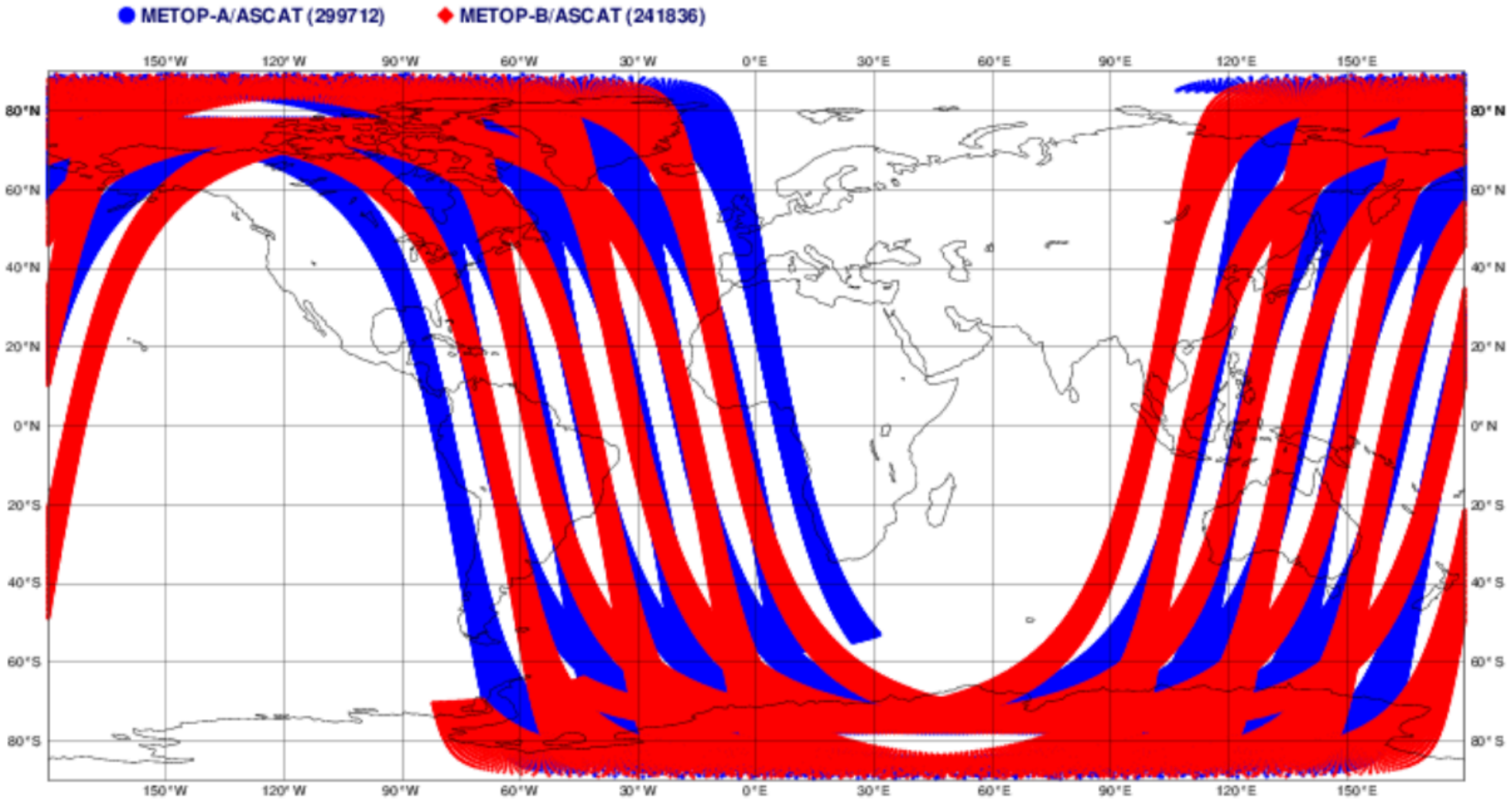
ECMWF data coverage (all observations) - AIRCRAFT
25/04/2018 00
Total number of obs = 229266



ECMWF data coverage (all observations) - AMSUA
25/04/2018 00
Total number of obs = 507386



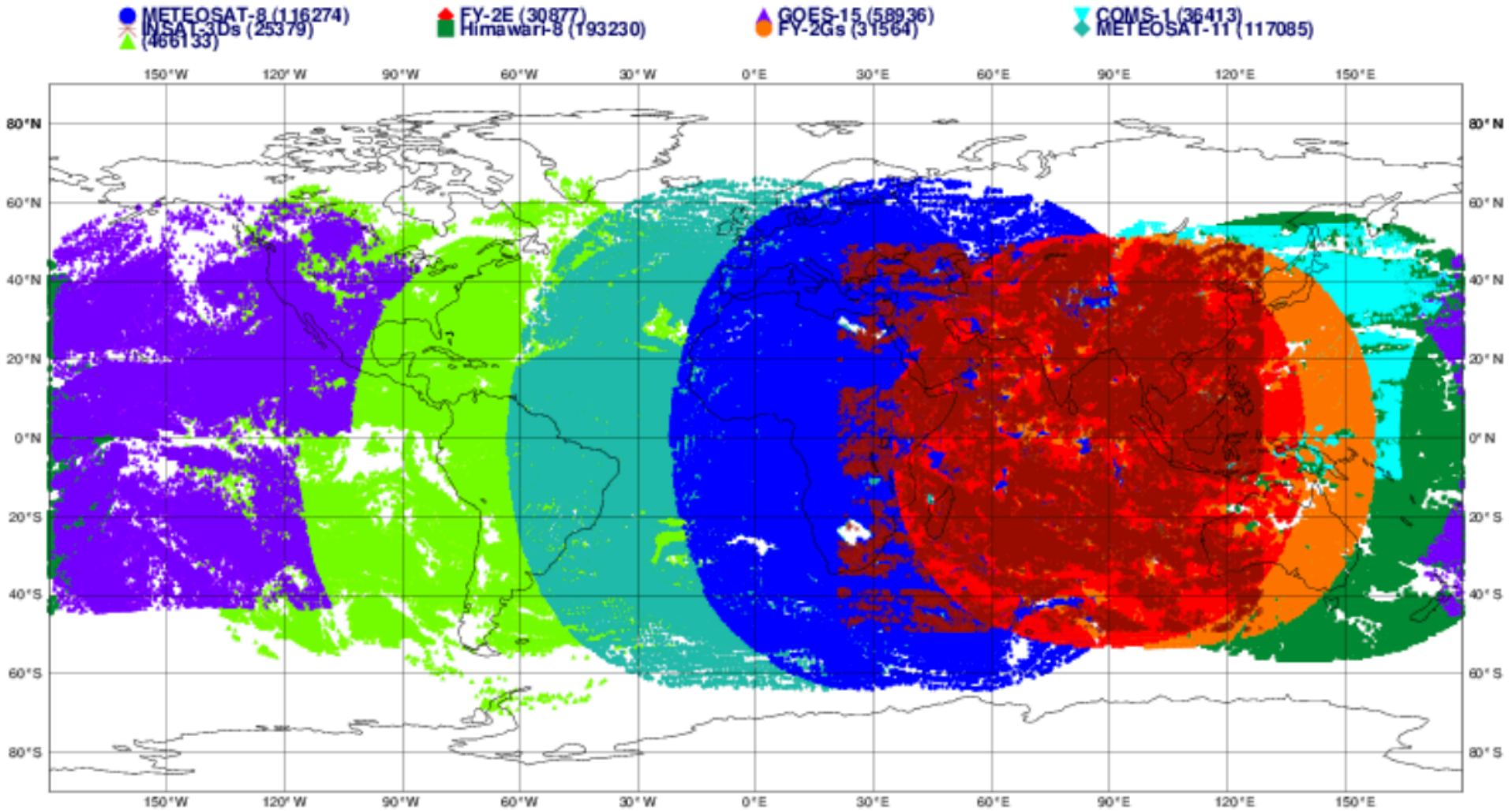
ECMWF data coverage (all observations) - SCATTEROMETER
25/04/2018 00
Total number of obs = 541548



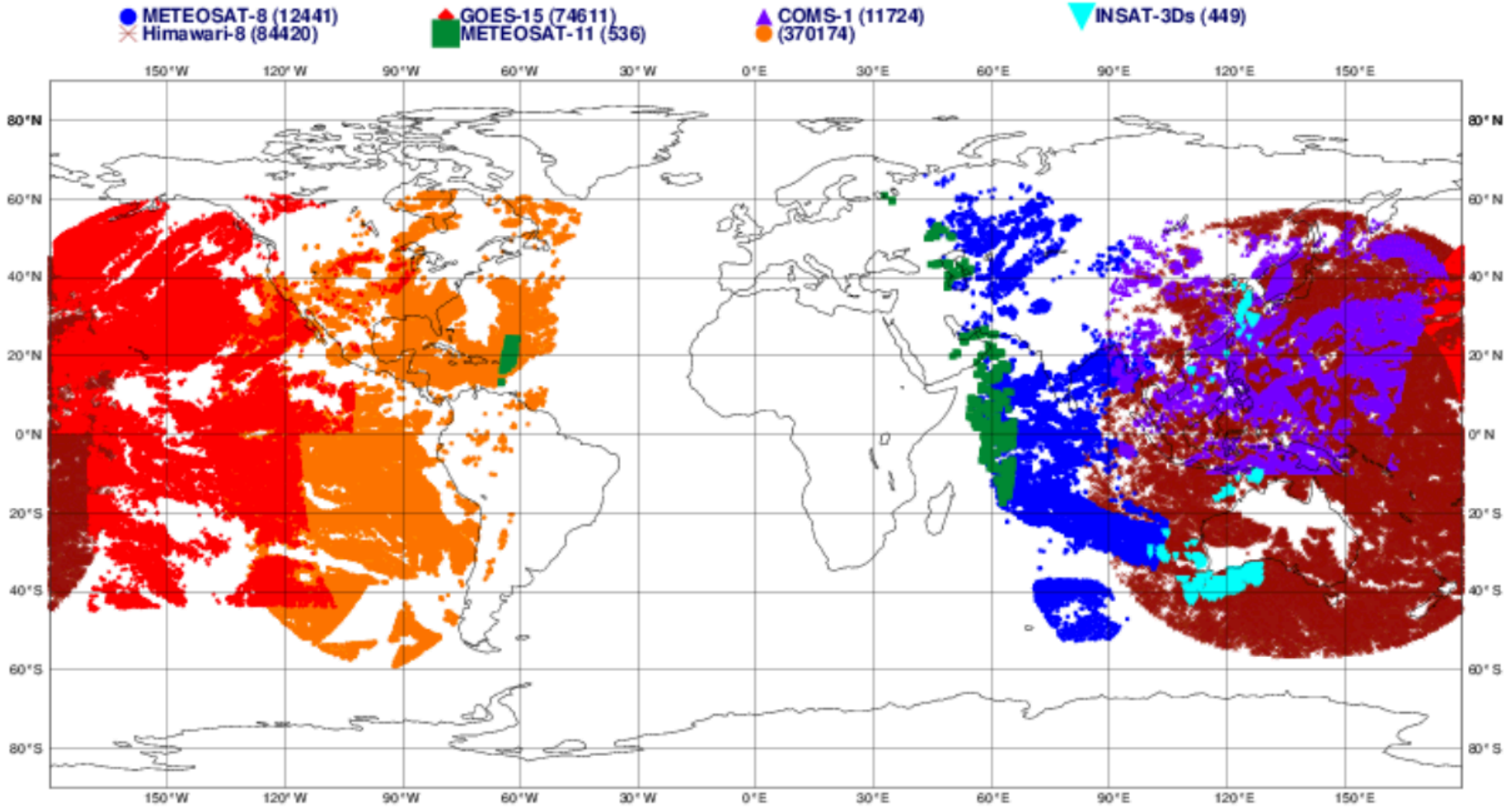
ECMWF data coverage (all observations) - AMV WV

25/04/2018 00

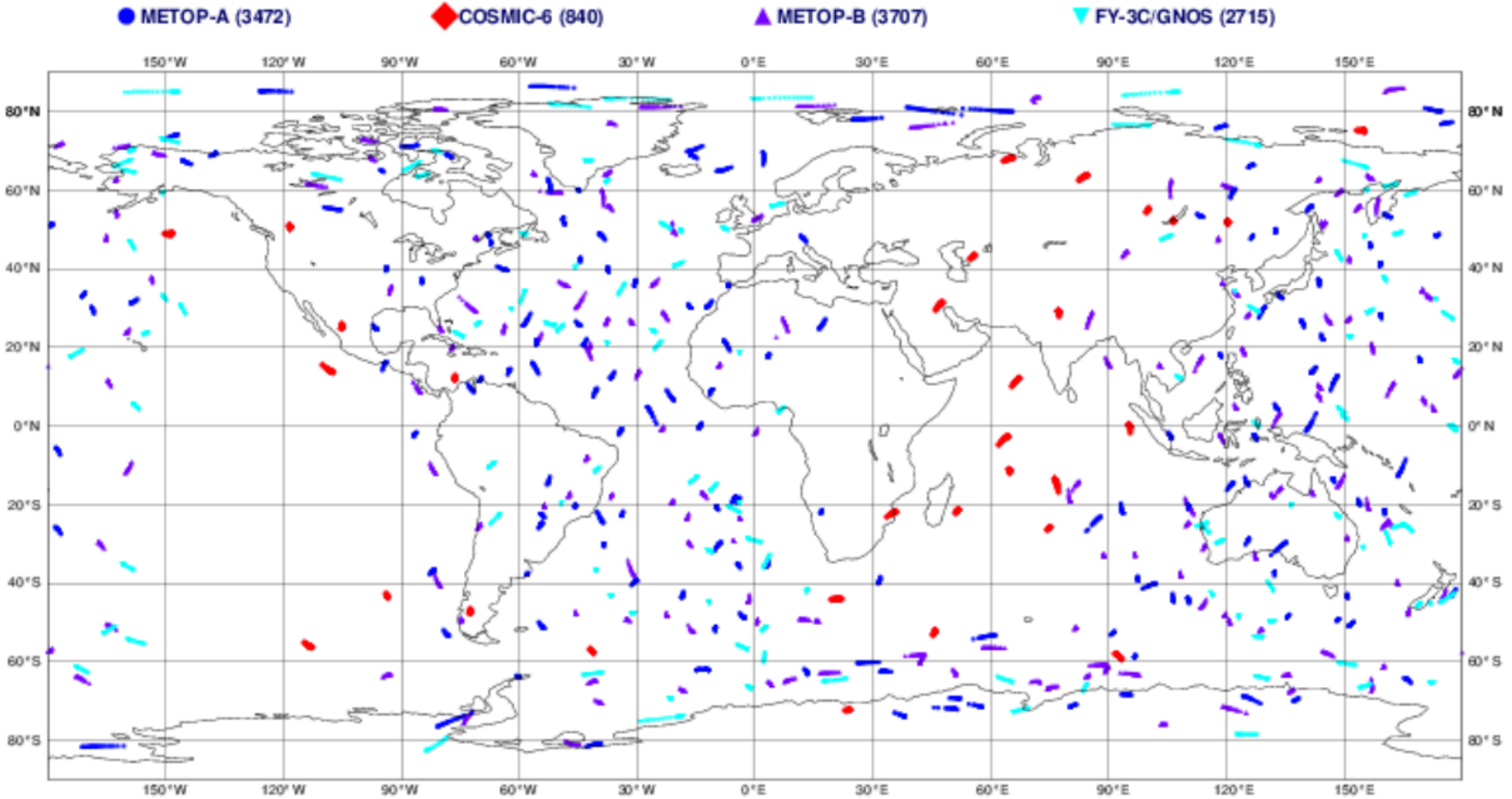
Total number of obs = 1075891



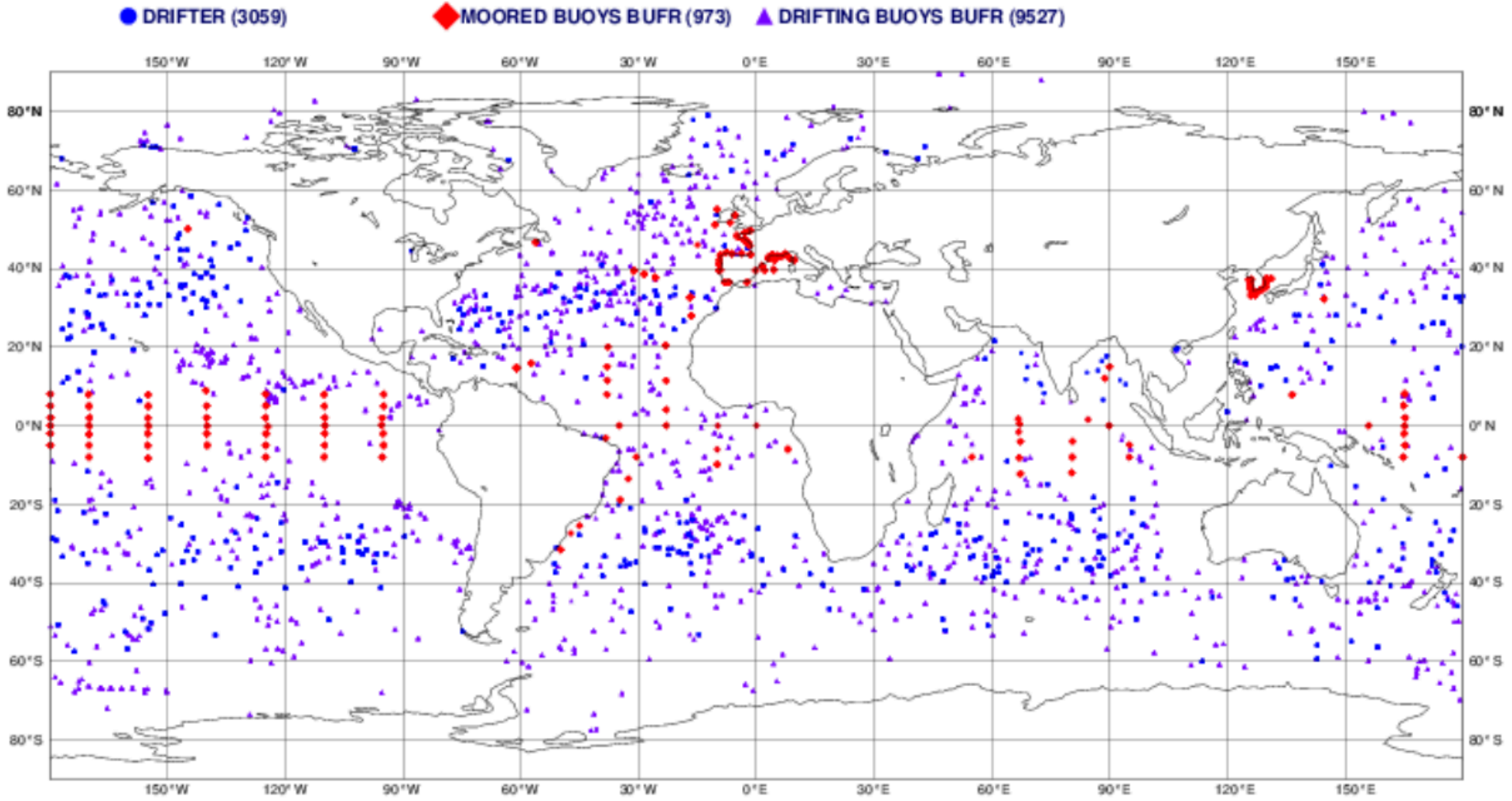
ECMWF data coverage (all observations) - AMV VIS
25/04/2018 00
Total number of obs = 554355



ECMWF data coverage (all observations) - GPSRO
25/04/2018 00
Total number of obs = 10734



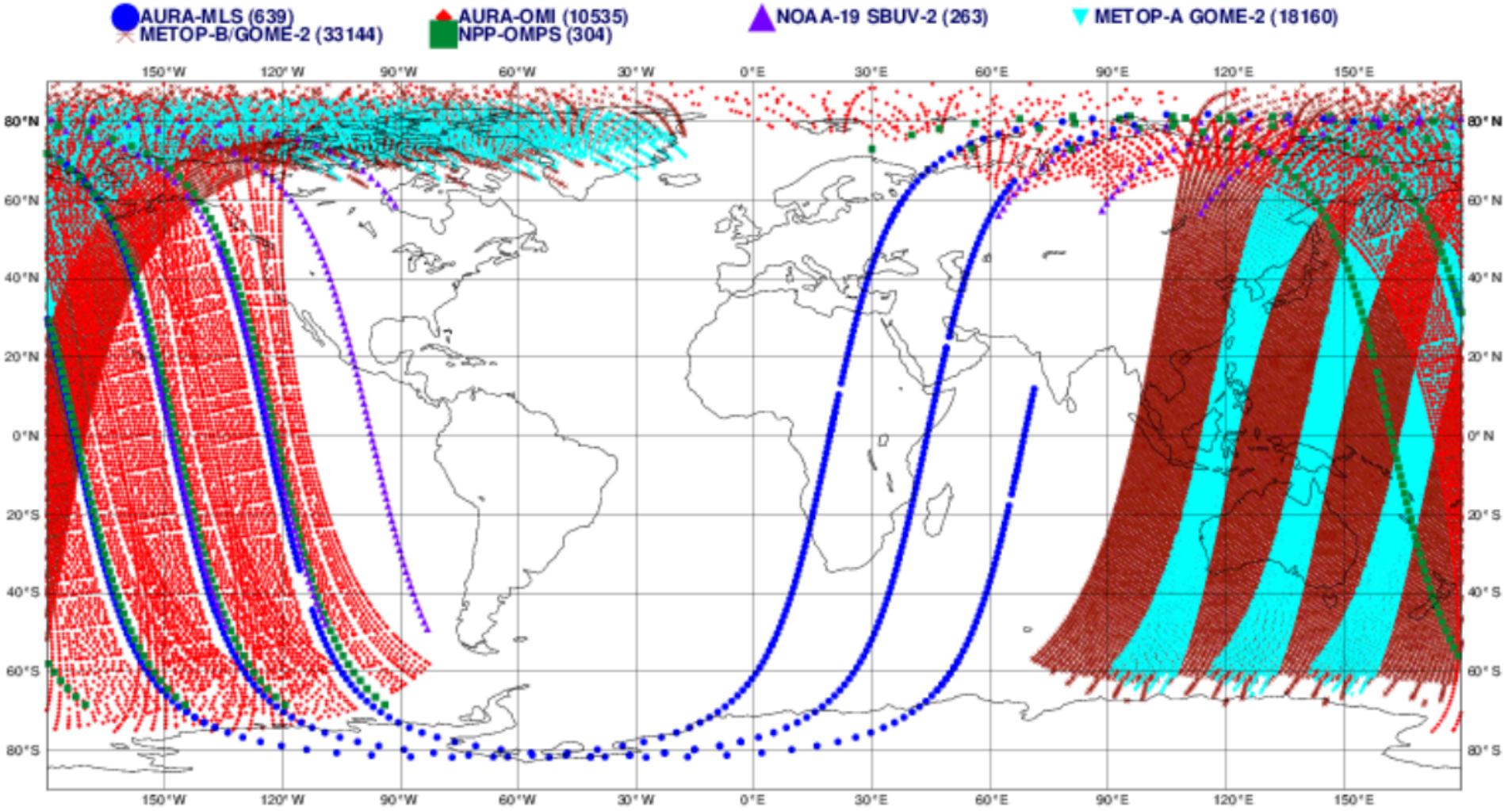
ECMWF data coverage (all observations) - BUOY
25/04/2018 00
Total number of obs = 13559



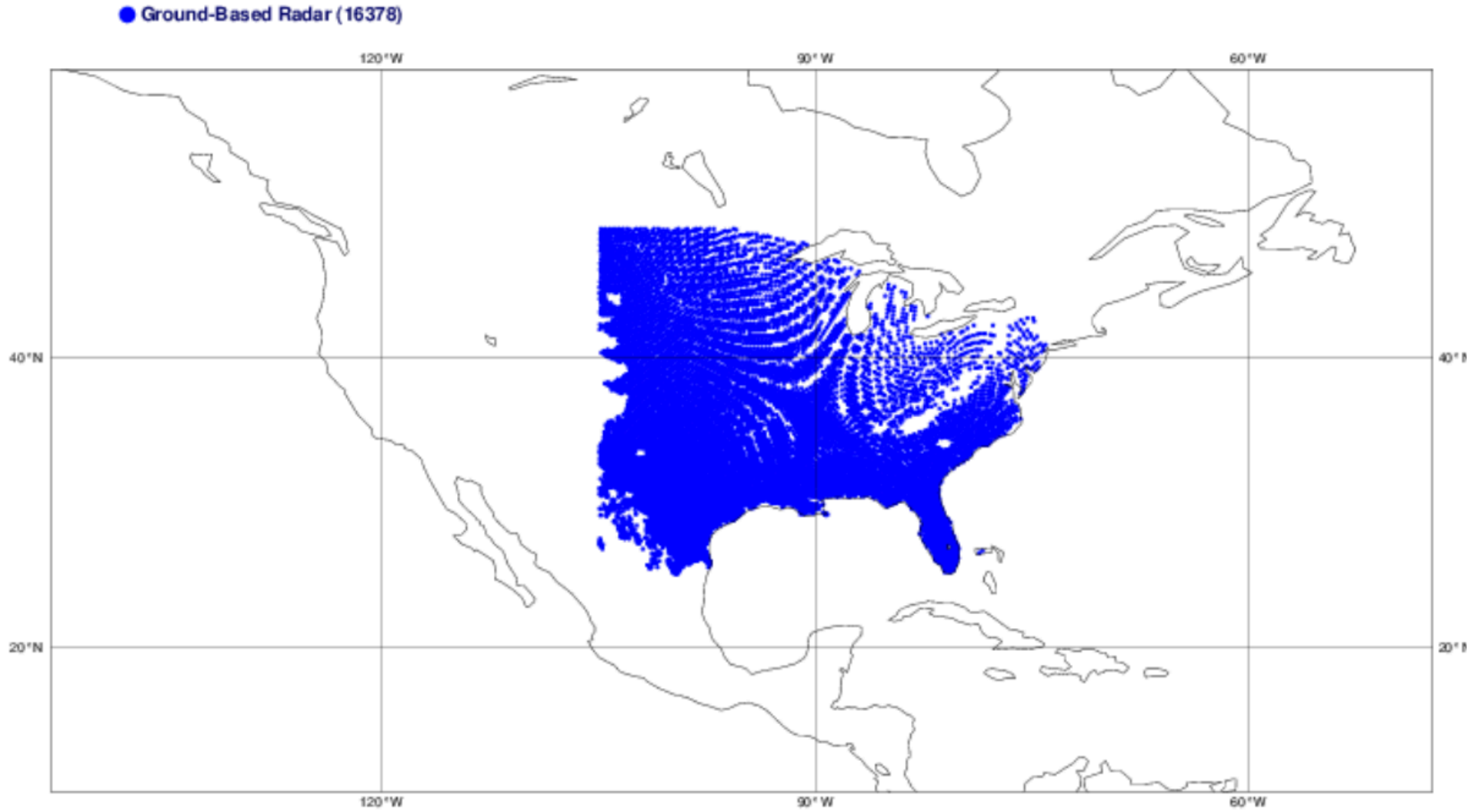
ECMWF data coverage (all observations) - RESAT

25/04/2018 00

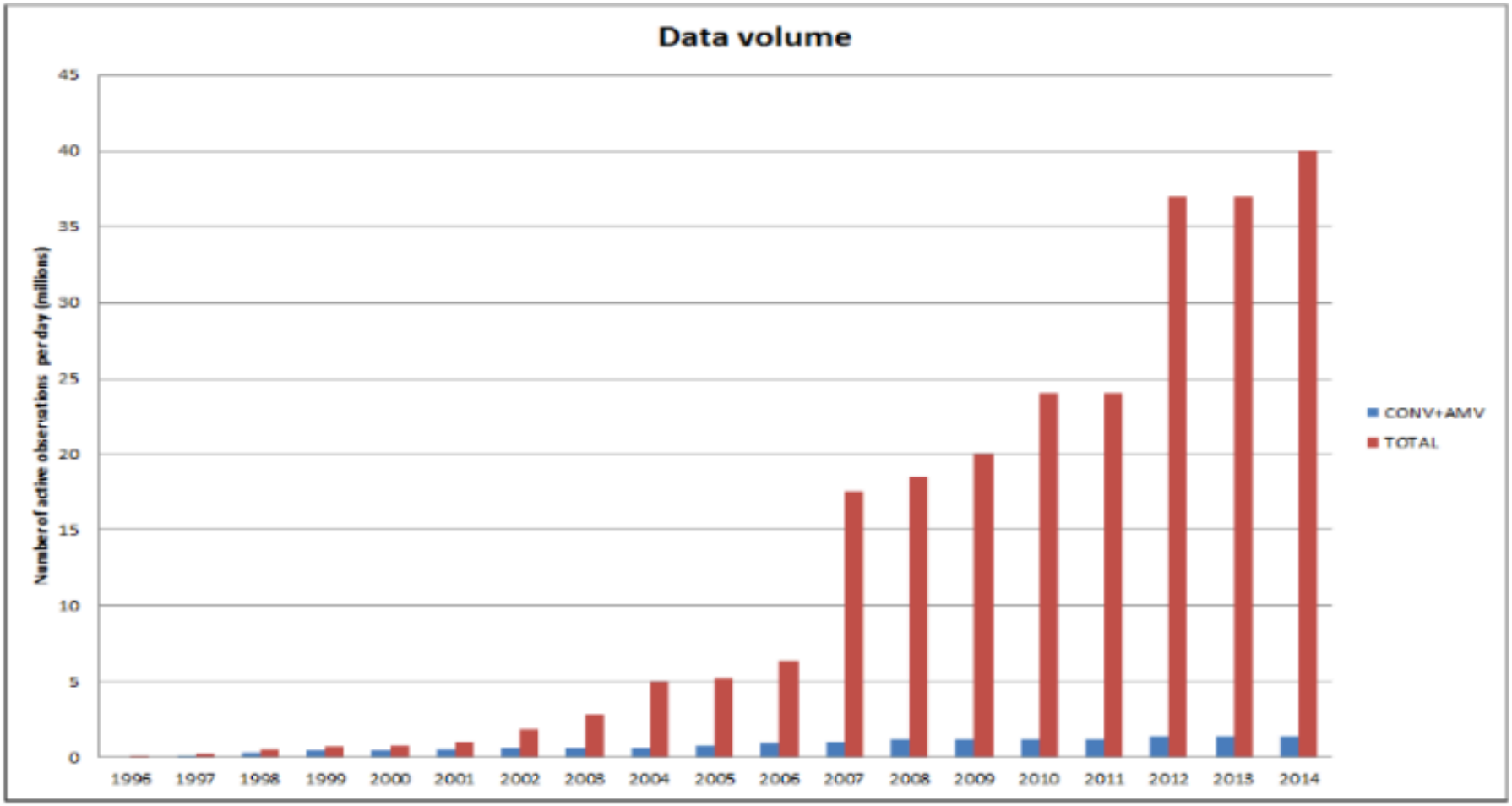
Total number of obs = 63045



ECMWF data coverage (all observations) - NEXRAD PRECIPITATION
25/04/2018 00
Total number of obs = 16378



ECMWF


Data volume
Number of active observations per day (millions)
45
40
35
30
25
20
15
10
5
0
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
CONV+AMV
TOTAL

- *Synoptic* observations (ground observations, radiosonde observations), performed simultaneously, by international agreement, in all meteorological stations around the world (00:00, 06:00, 12:00, 18:00 UTC)

- *Asynoptic* observations (satellites, aircraft), performed more or less continuously in time.

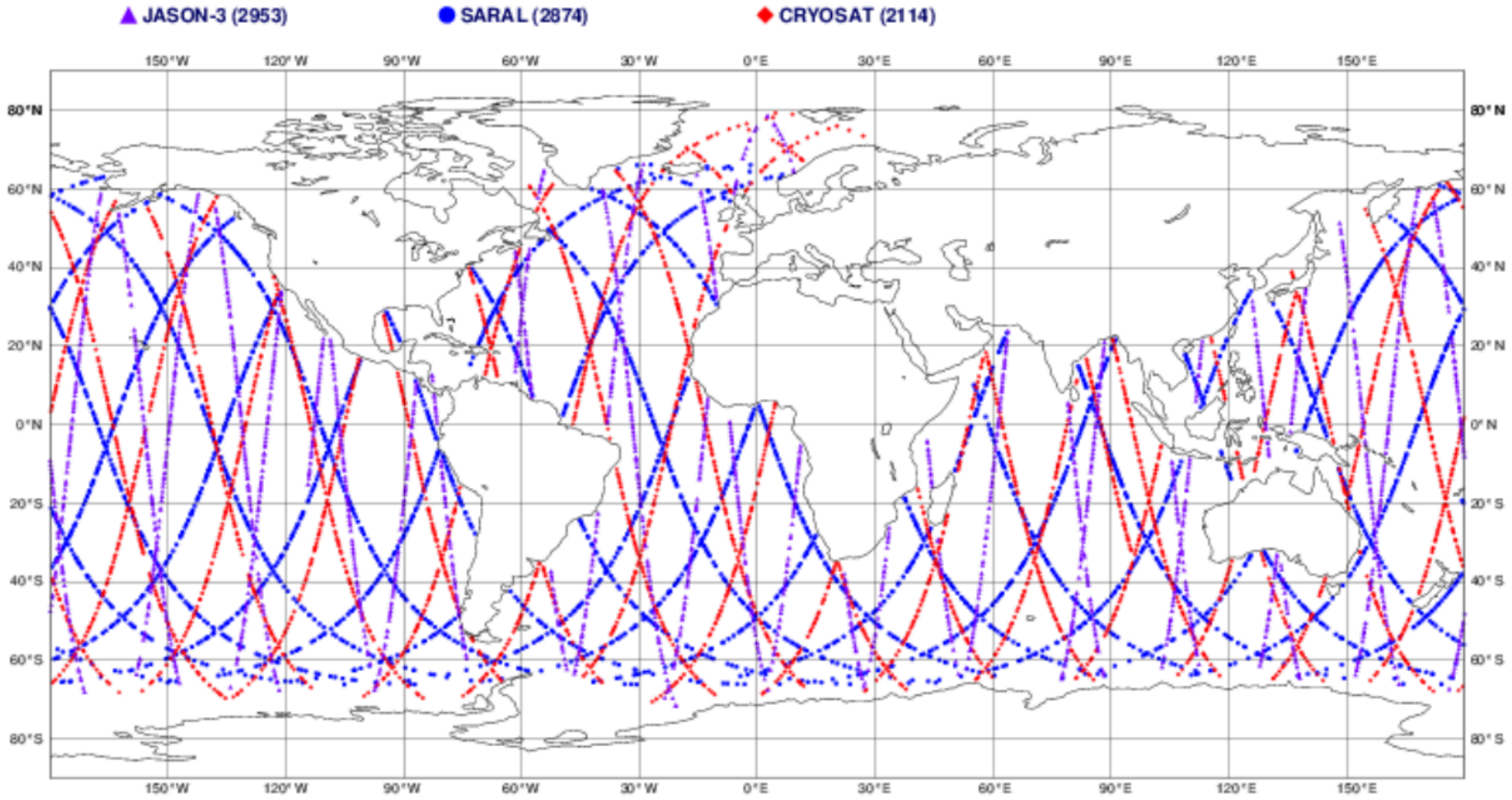
- *Direct* observations (temperature, pressure, horizontal components of the wind, moisture), which are local and bear on the variables used for for describing the flow in numerical models.

- *Indirect* observations (radiometric observations, …), which bear on some more or less complex combination (most often, a one-dimensional spatial integral) of variables used for for describing the flow

$$y = H(\boldsymbol{x})$$

H : *observation operator (*for instance, radiative transfer equation)

ECMWF data coverage (all observations) - SEA LEVEL ANOMALY
20180422 00
Total number of obs = 7941



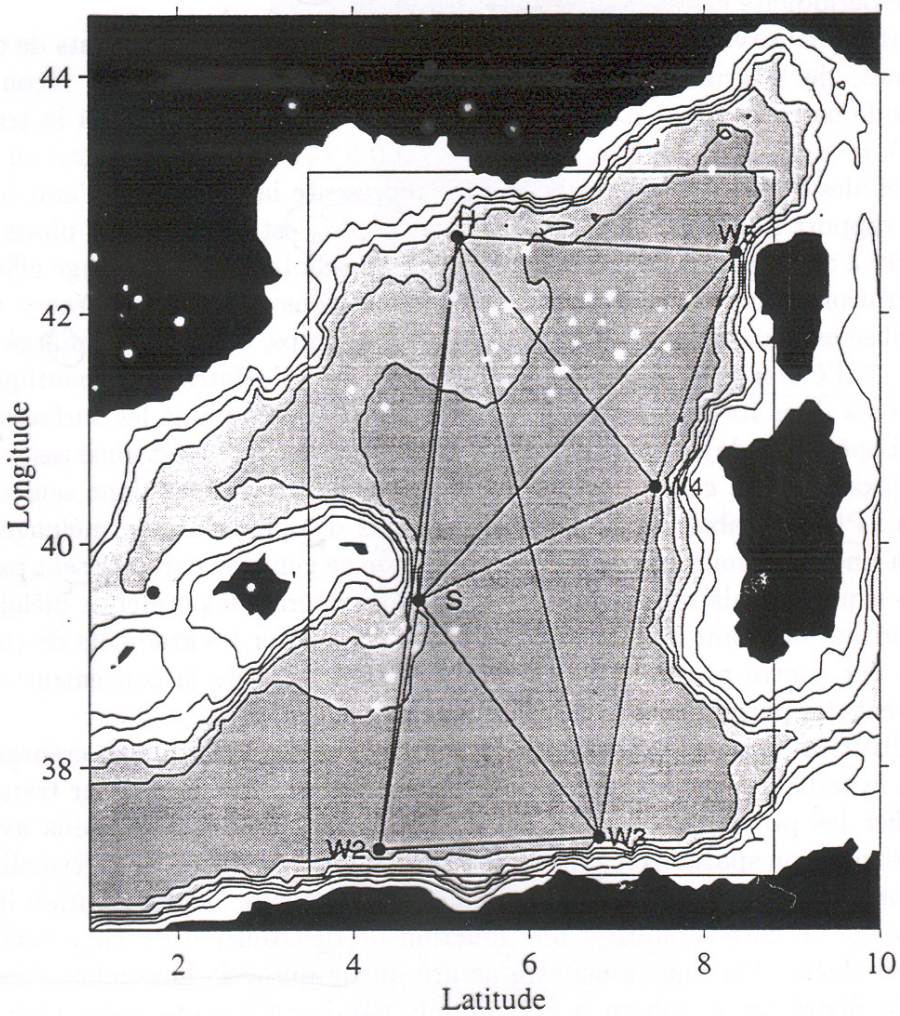


FIG. 1 – *Bassin méditerranéen occidental: réseau d'observation tomographique de l'expérience Thétis 2 et limites du domaine spatial utilisé pour les expériences numériques d'assimilation.*

E. Rémy, Doctoral Dissertation, 1999

Purpose of assimilation : reconstruct as accurately as possible the state of the atmospheric or oceanic flow, using all available appropriate information. The latter essentially consists of

- The observations proper, which vary in nature, resolution and accuracy, and are distributed more or less regularly in space and time.

- The physical laws governing the evolution of the flow, available in practice in the form of a discretized, and necessarily approximate, numerical model.

- 'Asymptotic' properties of the flow, such as, *e*. *g*., geostrophic balance of middle latitudes. Although they basically are necessary consequences of the physical laws which govern the flow, these properties can usefully be explicitly introduced in the assimilation process.

Assimilation is one of many '*inverse problems*' encountered in many fields of science and technology

- solid Earth geophysics
- plasma physics
- 'nondestructive' probing
- navigation (spacecraft, aircraft, ….)
- …

Solution most often (if not always) based on Bayesian, or probabilistic, estimation. 'Equations' are fundamentally the same.

Difficulties specific to assimilation of meteorological observations :

- Very large numerical dimensions ($n \approx 10^6$-10^9 parameters to be estimated, $p \approx 4$-5.10^7 observations per 24-hour period). Difficulty aggravated in Numerical Weather Prediction by the need for the forecast to be ready in time.

- Non-trivial, actually chaotic, underlying dynamics

Both observations and ‘model’ are affected with some uncertainty ⇒ uncertainty on the estimate.

For some reason, uncertainty is conveniently described by probability distributions (don’t know too well why, but it works; see, *e.g*. Jaynes, 2007, *Probability Theory: The Logic of Science,* Cambridge University Press).

Assimilation is a problem in bayesian estimation.

Determine the conditional probability distribution for the state of the system, knowing everything we know (see Tarantola, A., 2005*, Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM).

Coût des différentes composantes de la chaîne de prévision opérationnelle du CEPMMT (septembre 2015, J.-N. Thépaut) :

4DVAR: 9.5%

HRES FC: 4.5%

EDA: 30%

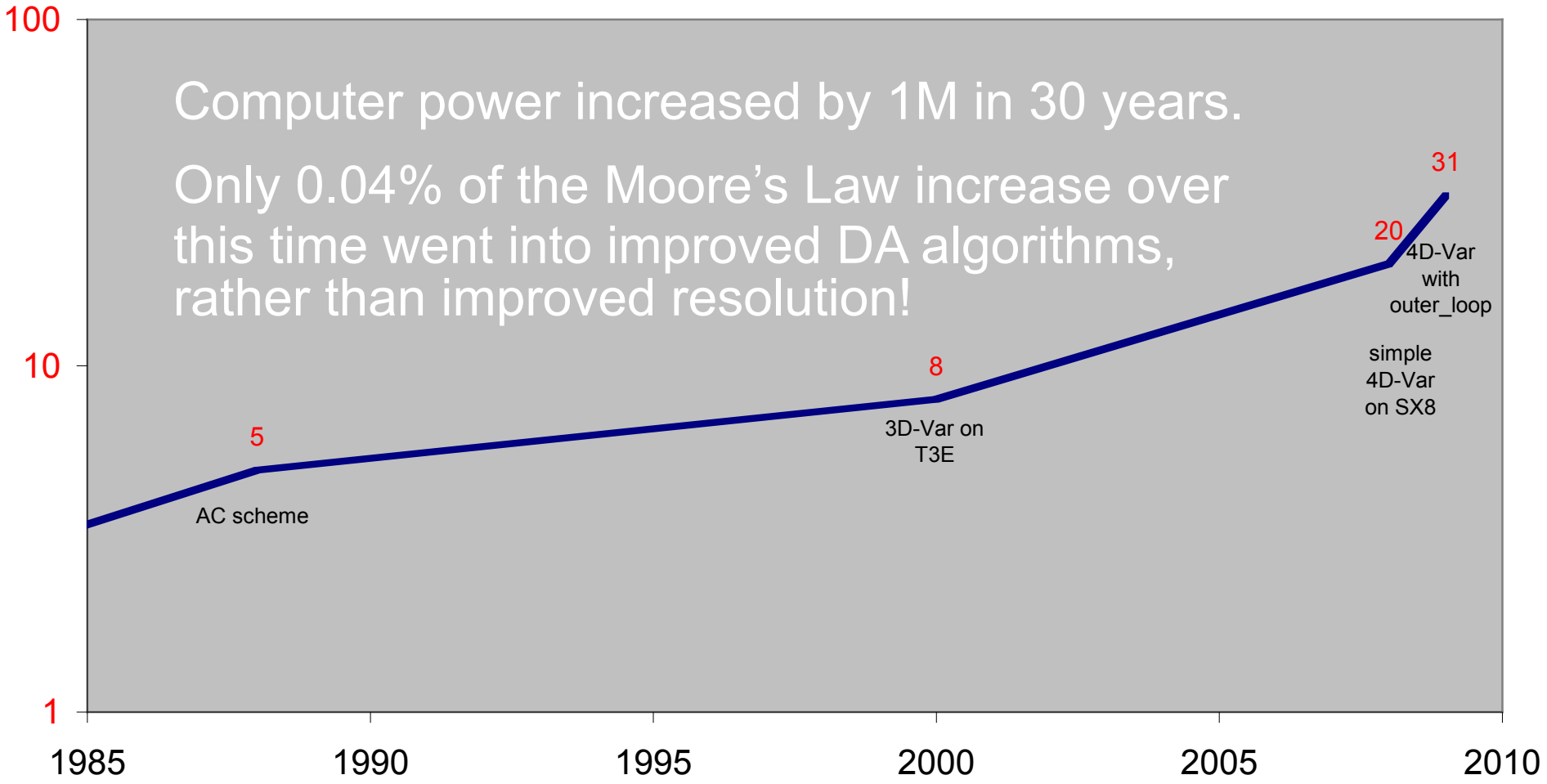
ENS: 22%

ENS: hindcasts 14%

Other: 20% of which BC AN: 3.5% BC FC: 4% BC ENS: 9.5%

L'EDA fournit à la fois les variances d'erreur d'ébauche du 4D-Var, et les perturbations initiales (en complément des vecteurs singuliers) de l'EPS.

ratio of supercomputer costs:
1 day's assimilation / 1 day forecast



Courtesy A. Lorenc

Bayesian Estimation

Determine conditional probability distribution of the state of the system, given the probability distribution of the uncertainty on the data

$z_1 = x + \zeta_1$ $\qquad$ $\zeta_1 = \mathcal{N}[0, s_1]$

density function $p_1(\zeta) \propto \exp[- (\zeta^2)/2s_1]$

$z_2 = x + \zeta_2$ $\qquad$ $\zeta_2 = \mathcal{N}[0, s_2]$

density function $p_2(\zeta) \propto \exp[- (\zeta^2)/2s_2]$

- ζ_1 and ζ_2 mutually independent

What is the conditional probability $P(x = \xi \mid z_1, z_2)$ that x be equal to some value ξ ?

$z_1 = x + \zeta_1$ density function $p_1(\zeta) \propto \exp[- (\zeta^2)/2s_1]$

$z_2 = x + \zeta_2$ density function $p_2(\zeta) \propto \exp[- (\zeta^2)/2s_2]$

ζ_1 and ζ_2 mutually independent

$x = \xi \;\Leftrightarrow\; \zeta_1 = z_1 - \xi$ and $\zeta_2 = z_2 - \xi$

- $P(x = \xi \mid z_1, z_2) \propto p_1(z_1 - \xi)\, p_2(z_2 - \xi)$

$\propto \exp[- (\xi - x^a)^2/2p^a]$

where $1/p^a = 1/s_1 + 1/s_2$, $x^a = p^a\,(z_1/s_1 + z_2/s_2)$

Conditional probability distribution of x, given z_1 and z_2 : $\mathcal{N}[x^a, p^a]$

$p^a < (s_1, s_2)$ independent of z_1 and z_2

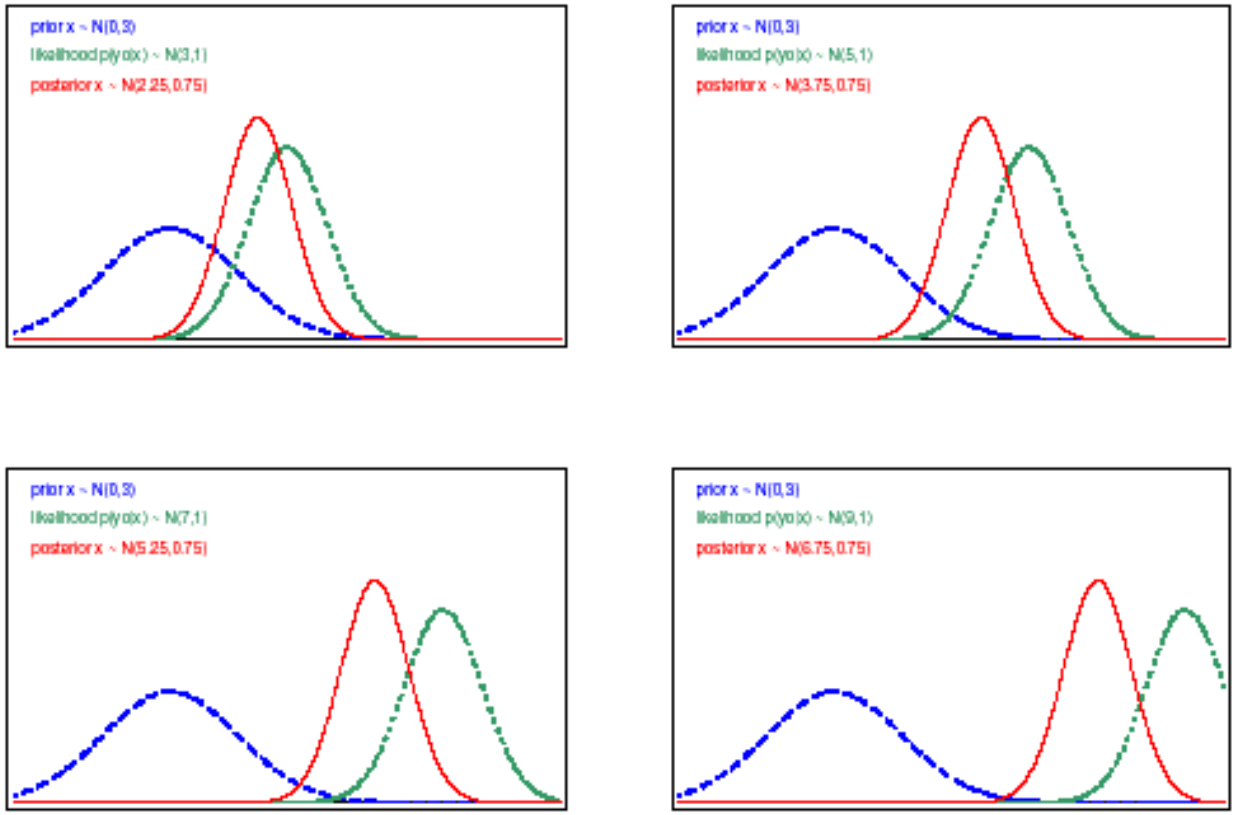


Fig. 1.1: Prior pdf $p(x)$ (dashed line), posterior pdf $p(x|y^o)$ (solid line), and Gaussian likelihood of observation $p(y^o|x)$ (dotted line), plotted against x for various values of y^o. (Adapted from Lorenc and Hammon 1988.)

$z_1 = x + \zeta_1$
$z_2 = x + \zeta_2$

Same as before, but ζ_1 and ζ_2 are now distributed according to exponential law with parameter a, *i. e.*

$$p(\zeta) \propto \exp[-|\zeta|/a] \quad ; \quad \mathrm{Var}(\zeta) = 2a^2$$

Conditional probability density function is now uniform over interval $[z_1, z_2]$, exponential with parameter $a/2$ outside that interval

$E(x \mid z_1, z_2) = (z_1+z_2)/2$

$\mathrm{Var}(x \mid z_1, z_2) = a^2 (2\delta^3/3 + \delta^2 + \delta + 1/2) / (1 + 2\delta)$, with $\delta = |z_1 - z_2|/(2a)$
Increases from $a^2/2$ to ∞ as δ increases from 0 to ∞. Can be larger than variance $2a^2$ of original errors (probability 0.08)

~~(Entropy $-\int p \ln p$ always decreases in bayesian estimation)~~

Cours à venir

~~Jeudi 19 avril~~

~~Jeudi 26 avril~~

Jeudi 3 mai

Lundi 14 mai

Jeudi 17 mai

Jeudi 24 mai

Jeudi 7 juin

Jeudi 14 juin

De 10h00 à 12h30, Salle E314, 3ième étage, Département de Géosciences, École Normale Supérieure, 24, rue Lhomond, Paris 5