Non Linear and Non Stationary Forecast Errors:

Time to revise the forecast strategy?

Magdalena A. Balmaseda
Frederic Vitart

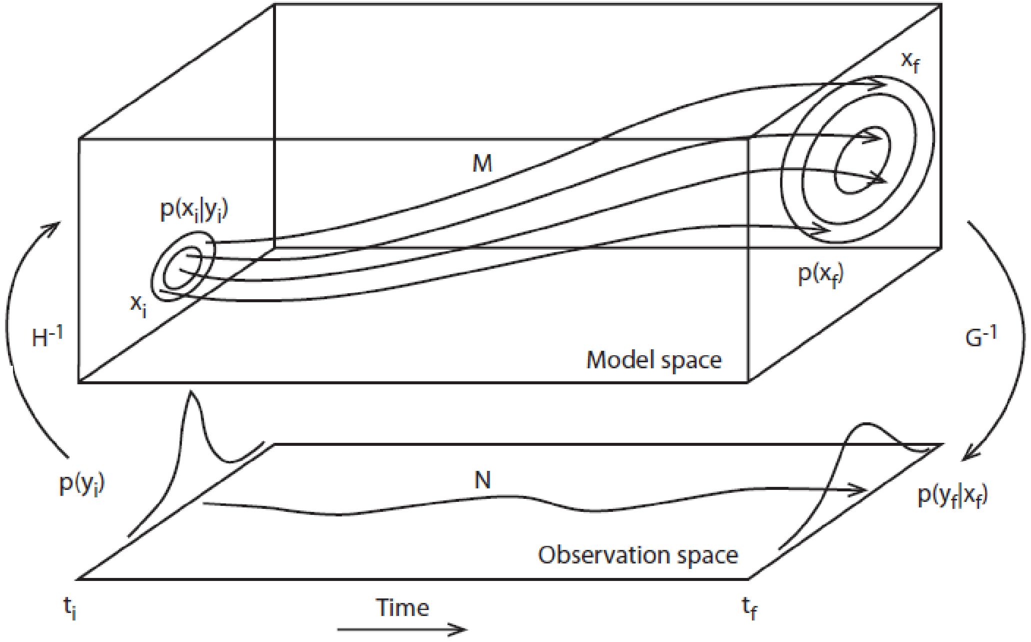
ECMWF, Shinfield Park, RG2 9AX, Reading, UK

End to end Initialized Probabilistic Forecasting System

2) **Propagating information, uncertainty and errors into the future**:
Forecast model

1) Initialization
Data Assimilation

$$p(x_i|y_i) = \frac{p(y_i|x_i)p(x_i)}{p(y_i)}$$



3) Calibration
Forecast Assimilation

$$p(y_f|x_f) = \frac{p(x_f|y_f)p(y_f)}{p(x_f)}$$

Stephenson et al 2005

$$J_{x|y} = (x - x_b)^T B^{-1}(x - x_b) + (y - Hx)^T R^{-1}(y - Hx).$$

Significant progress in representation of initial and model uncertainty
but
The treatment of model error is inconsistent along these 3 stages

ECMWF

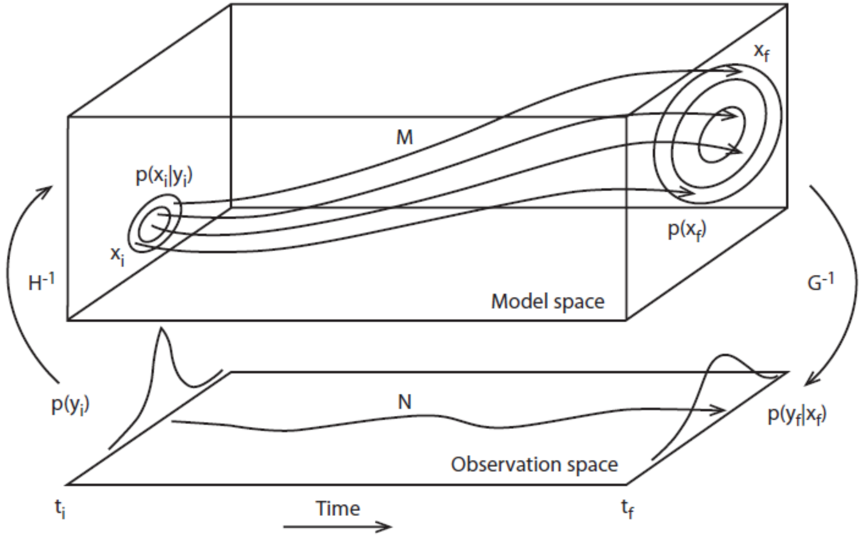
Inconsistent treatment of model errors

2) **Propagating information, uncertainty and error into the future**:
Forecast model

- ✓ Stochastic parameterizations for sub-grid processes.
- X Other missing processes and earth system components not represented
- X Model bias is not targeted

1) Initialization
Data Assimilation

- ✓ Initial uncertainty considered.
- ✓ Model uncertainty starts being considered.
- ✓ Observation uncertainty considered
- ✓ Observation bias considered
- X Model bias ignored



Model: $x = x + x + \varepsilon \downarrow x$

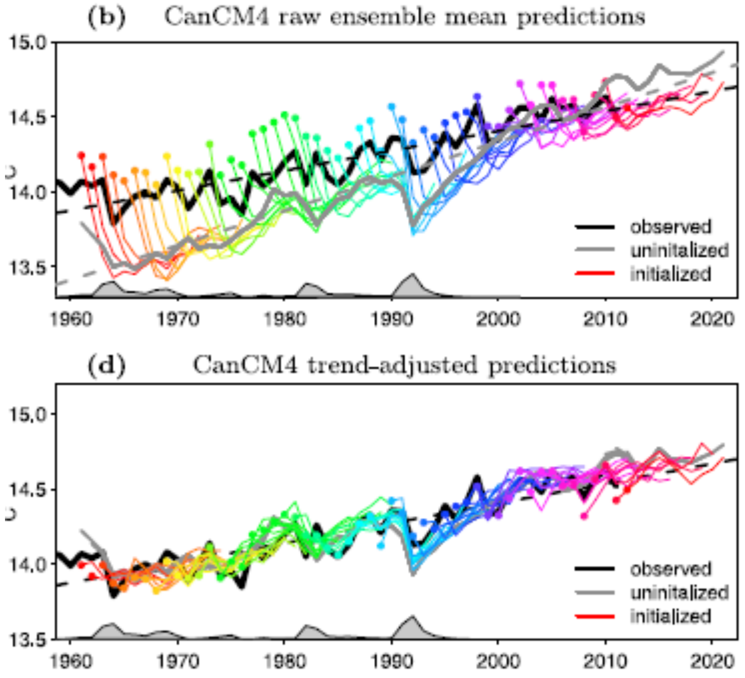
Observations: $y = y + y + \varepsilon \downarrow y$

3) Calibration
Forecast Assimilation

- **Model Bias accounted for: removed a posterirori.** Stockdale et al 1997
- Model uncertainty considered (ensemble)
- Observation error neglected *
- Residuals are non stationary, non gaussian.
  - Limitation to forecast skill
  - calibration is more difficult

Calibration

$$x = y + \mathbf{K}(x - x) + \mathbf{F}\varepsilon_x + \mathbf{T}(t) + \mathbf{G}(y_0)$$



From Kharin eta la 2012

Bias correction ($x \neq y$)

K: linear transformation of anomalies

F: Adjustment of ensemble spread

T: detrending

G: other flow dependent corrections

Stephenson et al 2005
Kharin et al 2012
Fukar et al 2014

- Need very long reforecast records for robust estimation of many parameters
  Expensive, not records long enough are available
- **Error in mean state errors degrades variability and forecast skill**

Mean state error influencing model fidelity and predictability

Correcting model biases leads to better representation of variability (or model fidelity) :

(several papers: D'Andrea and Vautard 2000, Balmaseda et al 2010, Scaife 2011, ….)

Correcting bias in tropical SST improves seasonal forecast skill of ENSO, tropical cyclones…

Magnusson et al 2012, Vecchi et al 2014:

Correcting biases in atmosphere improves seasonal atmospheric predictability:

Kharin and Scinocca 2012

Correcting North Atlantic SST bias improves subseasonal skill over North Atlantic and Europe

Vitart 2018

Impact on atmospheric mean errors in variability and predictability

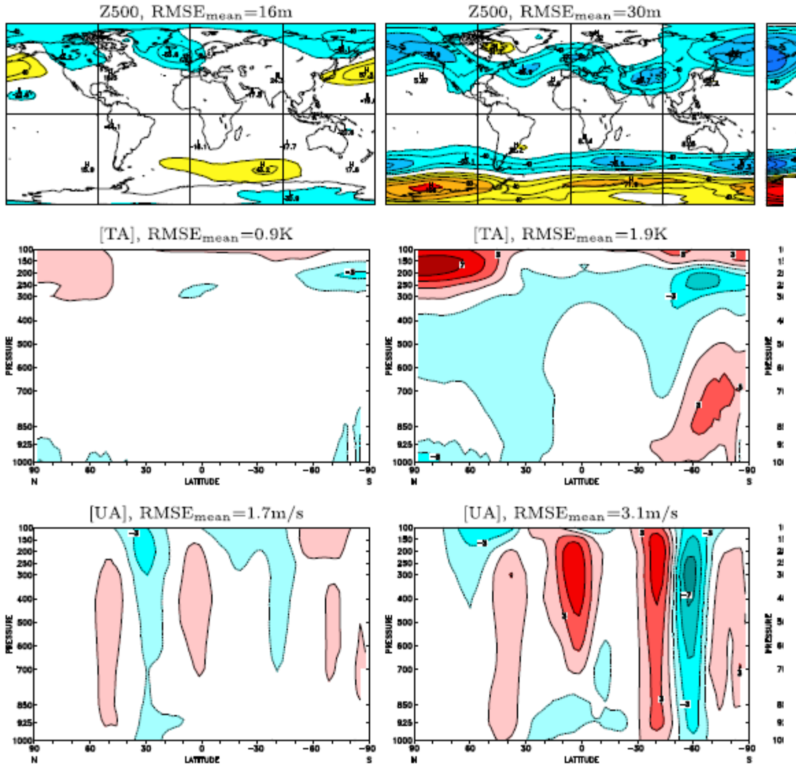
Nudging Method in AMIP runs

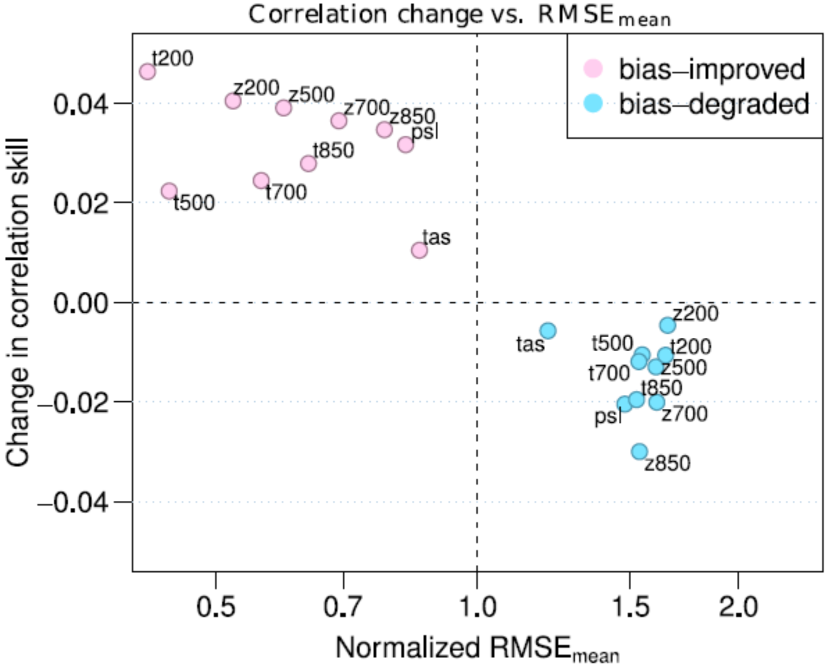
1 $$\frac{\partial X}{\partial t} = F(X) - \frac{1}{\tau}(X - X_R)$$

2 $$G = -\frac{1}{\tau}\overline{(X - X_R)}^{AC}.$$

3 $$\frac{\partial X}{\partial t} = F(X) + G$$

BIAS corrected CNTL BIAS degraded





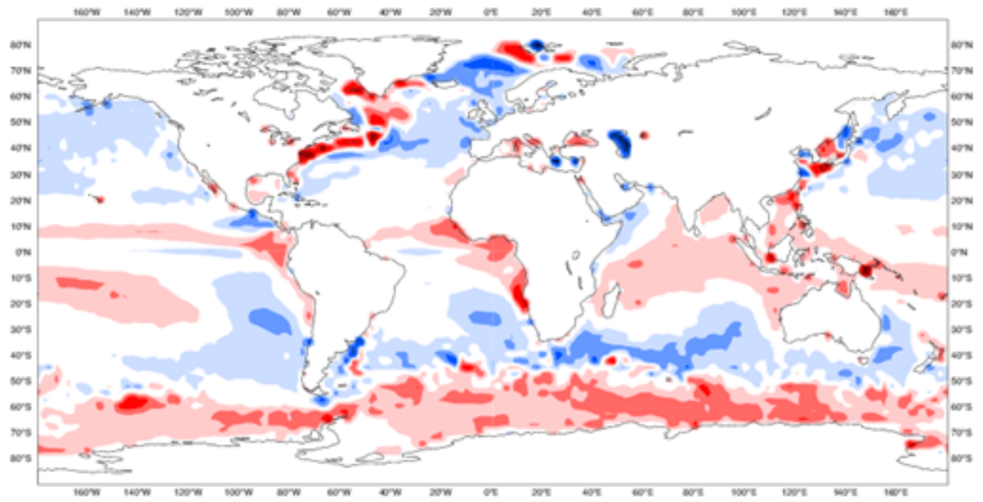
From Kharin and Scinocca 2012

Non linear interactions: North Atlantic SST errors impact subseasonal forecast skill

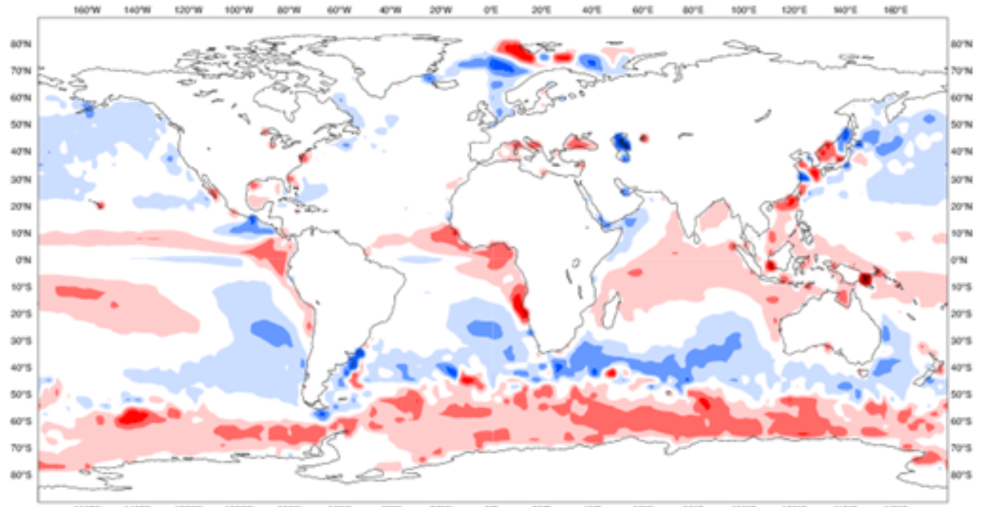
SST Biases Week 4 (day 26-32)

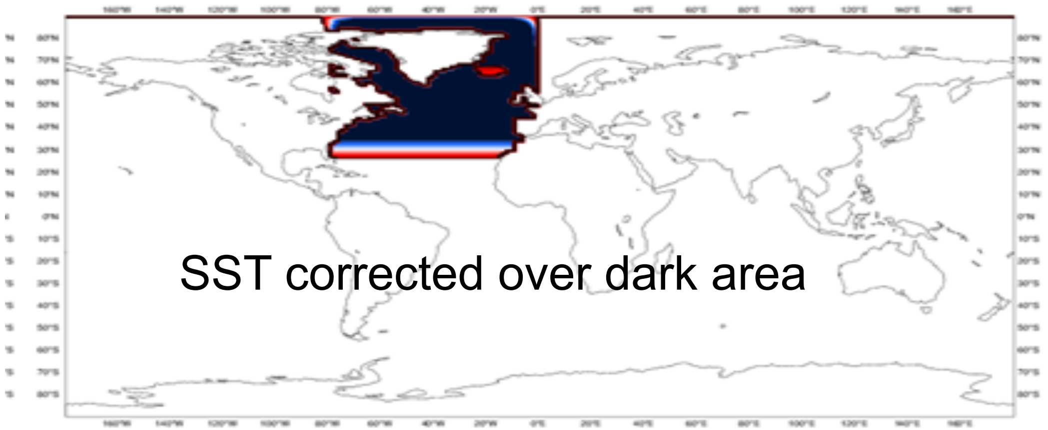
DJF (162 start dates)

CONTROL



BC AT



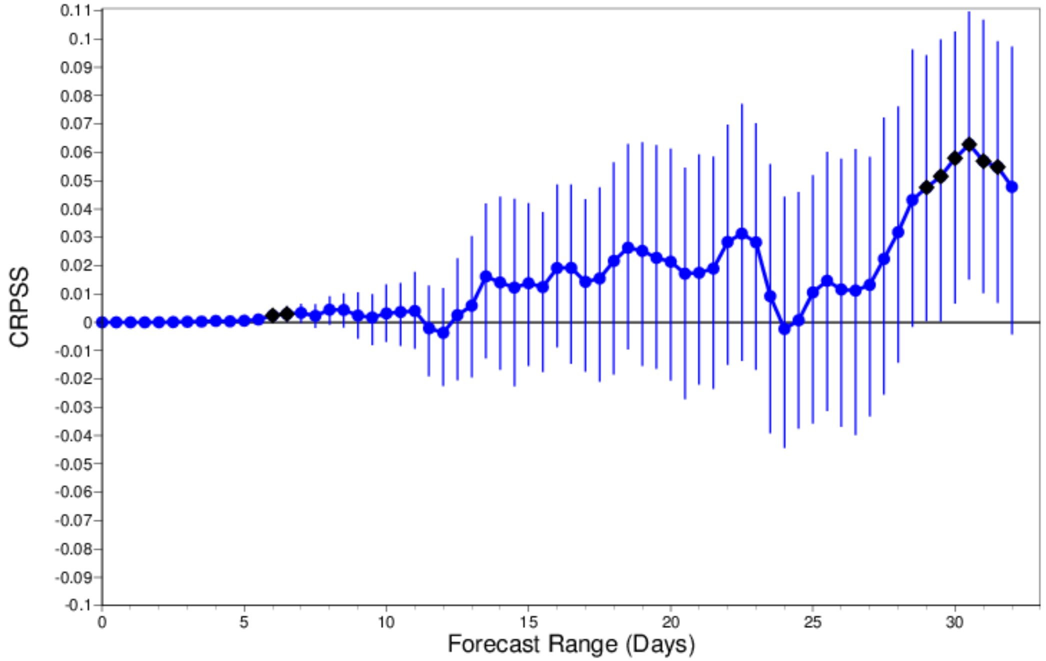


SST corrected over dark area

From Vitart et al 2018

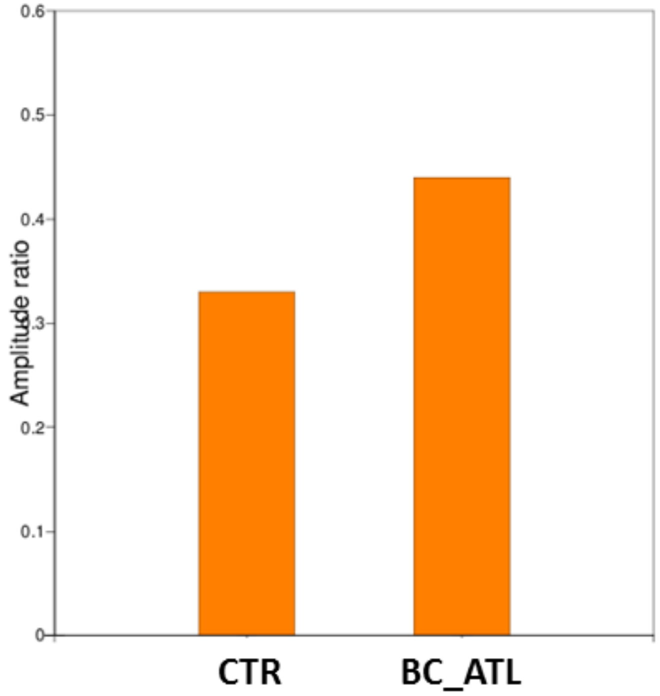
Non linear interactions: North Atlantic SST errors impact subseasonal forecast skill

NAO CRPSS differences
+ve is improvement



From Vitart et al 2018

3rd pentad after MJO Phase 7



Main impact is on MJO/ NAO –ve teleconnections

Treatment of model error during the forecast

A. Stochastic parameterizations of sub-grid scale processes

- SPPT, SPP, SKEB, intrinsic stochastic parameterizations. See Berner et al 2017 for a review.
- They increase the ensemble spread (Leutbecher et al 2018). Important for tropical convection and ENSO (Weisheimer et al 2014).
- They do not tackle model bias explicitly, but change model climate (Christensen et al 2017, Berner et al 2018)
- Choice of parameters: tuned to calibrate ensemble spread or first principles
- **No optimal control with observational constraint on**

C. Model error estimation based on observational "optimal" control : data assimilation to estimate model error (or approximations)

- D'Andrea and Vautard 2000
- Piccolo and Cullen 2016

$$\mathbf{x}_i = M_i(\mathbf{x}_{i-1}) + \eta_i$$

$$J(\eta) = \frac{1}{2}\sum_{i=1}^{n} \eta_i^T \mathbf{Q}^{-1} \eta_i + \frac{1}{2}\sum_{i=1}^{n} (H_i(\mathbf{x}_i) - \mathbf{y}_i)^T \mathbf{R}^{-1} (H_i(\mathbf{x}_i) - \mathbf{y}_i)$$

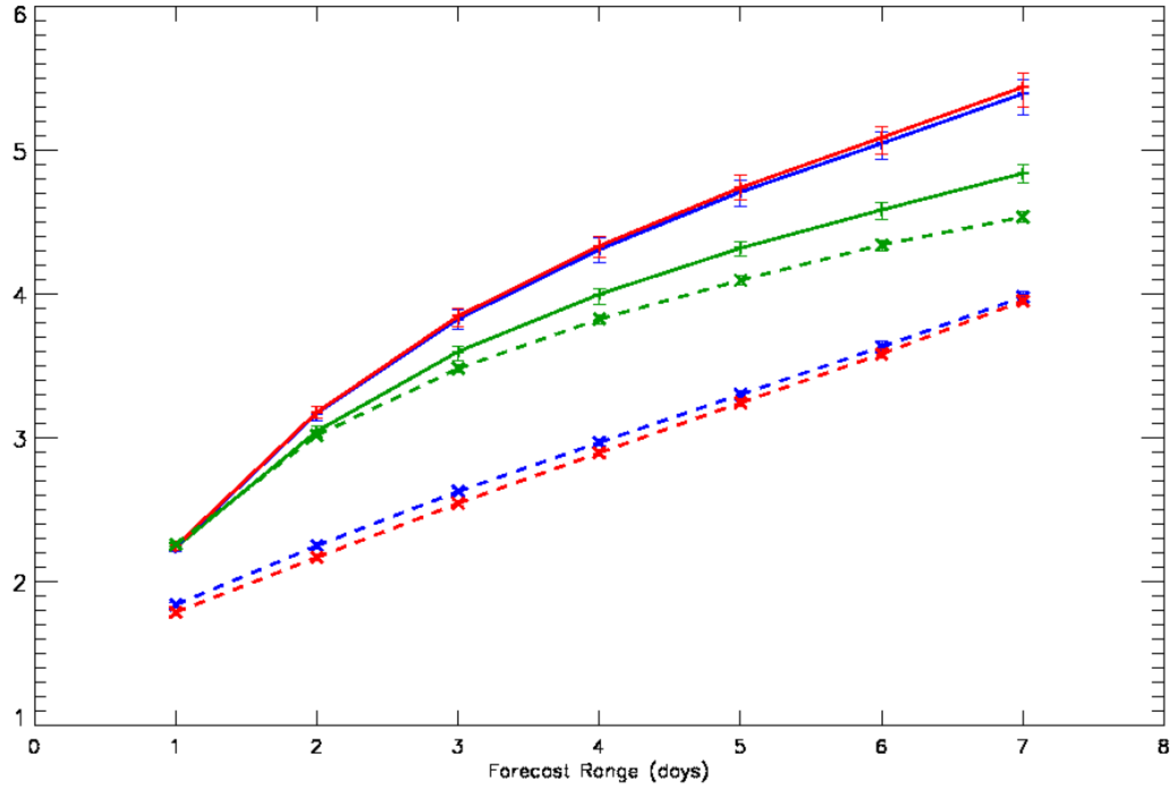
-Proxi: nugdging terms as in Kharin and Scinocca 2012…

Comparison of model error approaches: medium range

Perturbed Parameter SPPT Analysis increments

MOGREPS
Verification against analysis
250 hPa winds (m/s)
Tropics

Solid: RMSE
Dash: spread



Credit: Chiara Piccolo; see also Piccolo and Cullen 2016

Assimilation increments sample mean error. Likely reason for improved peformance
Caveat: model bias is currently ignored in atmospheric data assimilation systems.

Model bias treated explicitly in ocean data assimilation

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{b}^f + \mathbf{K}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)]$$

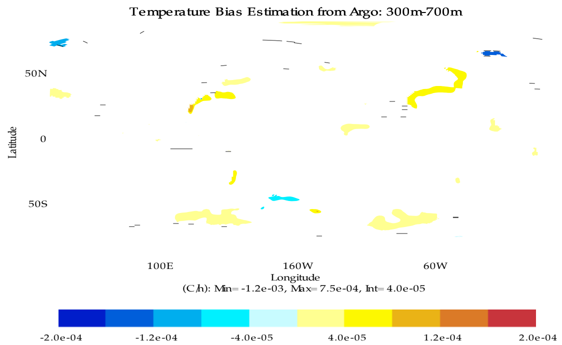
$$\mathbf{b}^a = \mathbf{b}^f + \mathbf{L}[\mathbf{y} - \mathbf{H}(\mathbf{x}^f + \mathbf{b}^f)]$$

Dee and DaSilva 1998
Dee and Todling 2001
Tremolet 2007

Example: scheme in ECMWF ORAS

$$\mathbf{b}^f{}_k = \bar{\mathbf{b}}_k + \mathbf{b}'^f_k$$

Offline climatological estimate



Adaptive term; AR1 process, constrained by observations

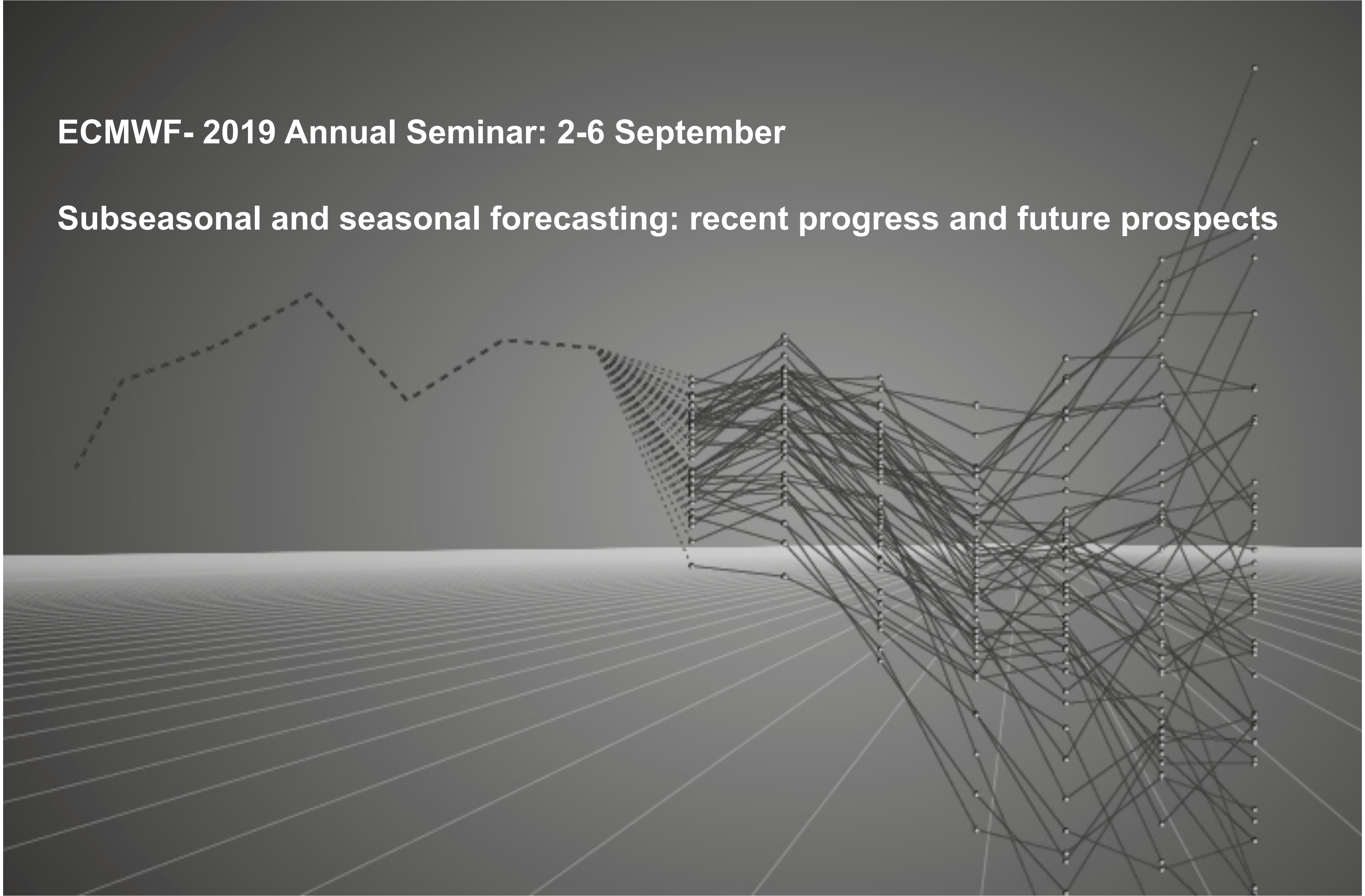
$$\mathbf{b}'^f_k = \mathbf{A}\,\mathbf{b}'^f_{k-1} + \boldsymbol{\varepsilon}_k$$

$$\mathbf{b}'^a_k = \mathbf{A}\,\mathbf{b}'^f_{k-1} + \mathbf{L}\,[\mathbf{y}-\mathbf{H}(\mathbf{x}+\mathbf{b}'^f_k)]$$

Balmaseda et al 2007

Summary

- It is possible to produce **more skilful predictions at extended and seasonal range by correcting model bias** during forecast phase
- It is possible **to design a consistent framework for treatment of model bias**:
  - estimation of model bias during data assimilation phase using observational constrain.
  - bias estimate applied during forecast phase. Complementary to stochastic physics
  - This should produce improved forecast, easier to calibrate .
- The nudging residuals provide a starting point for experimentation
- **Machine learning can be applied to model error terms** (assimilation increments, nudging terms)
- **Future forecasting systems:**
  - **Combination dynamical models for the signals and empirical stochastic models for the errors.**
  - **Consistent treatment of model error in the 3 stages of probabilistic forecast.**


ECMWF- 2019 Annual Seminar: 2-6 September
Subseasonal and seasonal forecasting: recent progress and future prospects