Representation of model uncertainties for ensemble forecasts

ECMWF Ensemble Prediction Section

Martin Leutbecher, Simon Lang, **Sarah-Jane Lock**,

Pirkka Ollinaho, Antje Weisheimer

sarah-jane.lock@ecmwf.int

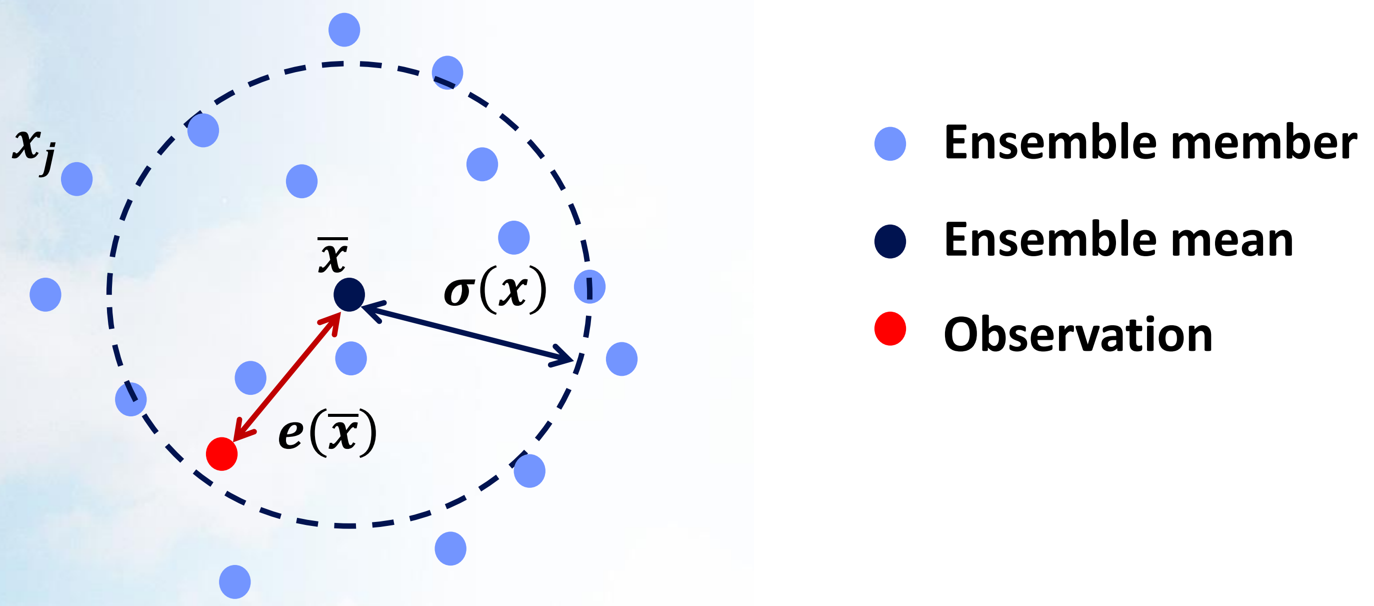
Representation of model uncertainties

- Why represent model uncertainty in an ensemble forecast?
- What are the sources of model uncertainty?
- How do we represent model uncertainty?
  - 2 stochastic physics schemes in the IFS
- Impact of stochastic physics schemes in the IFS:
  - Medium-range ensemble (ENS)
  - Seasonal forecast (S4)

sarah-jane.lock@ecmwf.int

Forecast uncertainty via ensemble reliability

- In a reliable ensemble, **ensemble spread** is a predictor of **ensemble error**



i.e. averaged over many ensemble forecasts,

$$e(\bar{x}) \approx \sigma(x)$$

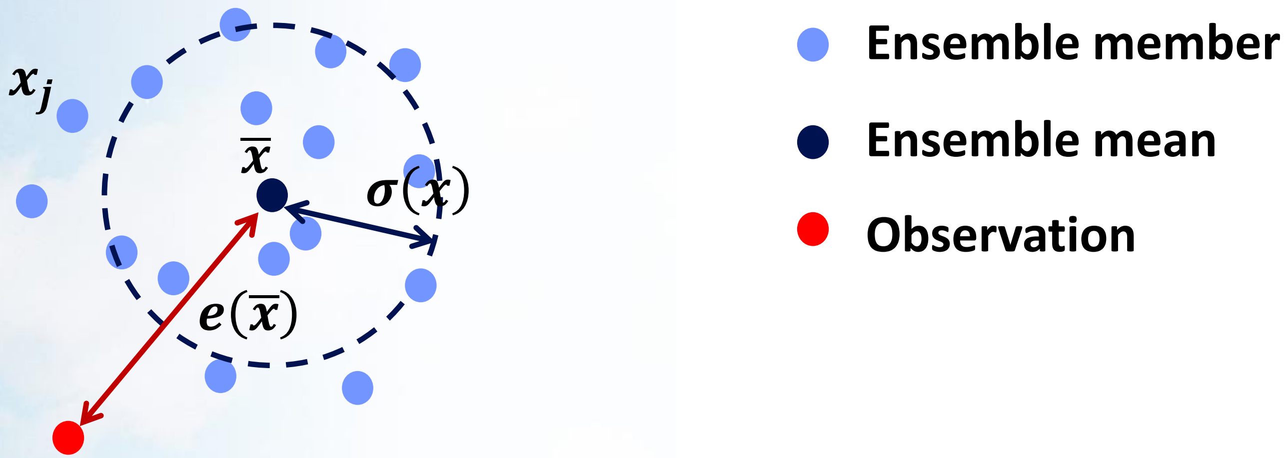
*Different ensemble members arise due to perturbations to **initial conditions** and to **model integrations***

sarah-jane.lock@ecmwf.int

Forecast uncertainty via ensemble reliability

- In an under-dispersive ensemble,

$$e(\bar{x}) \gg \sigma(x)$$

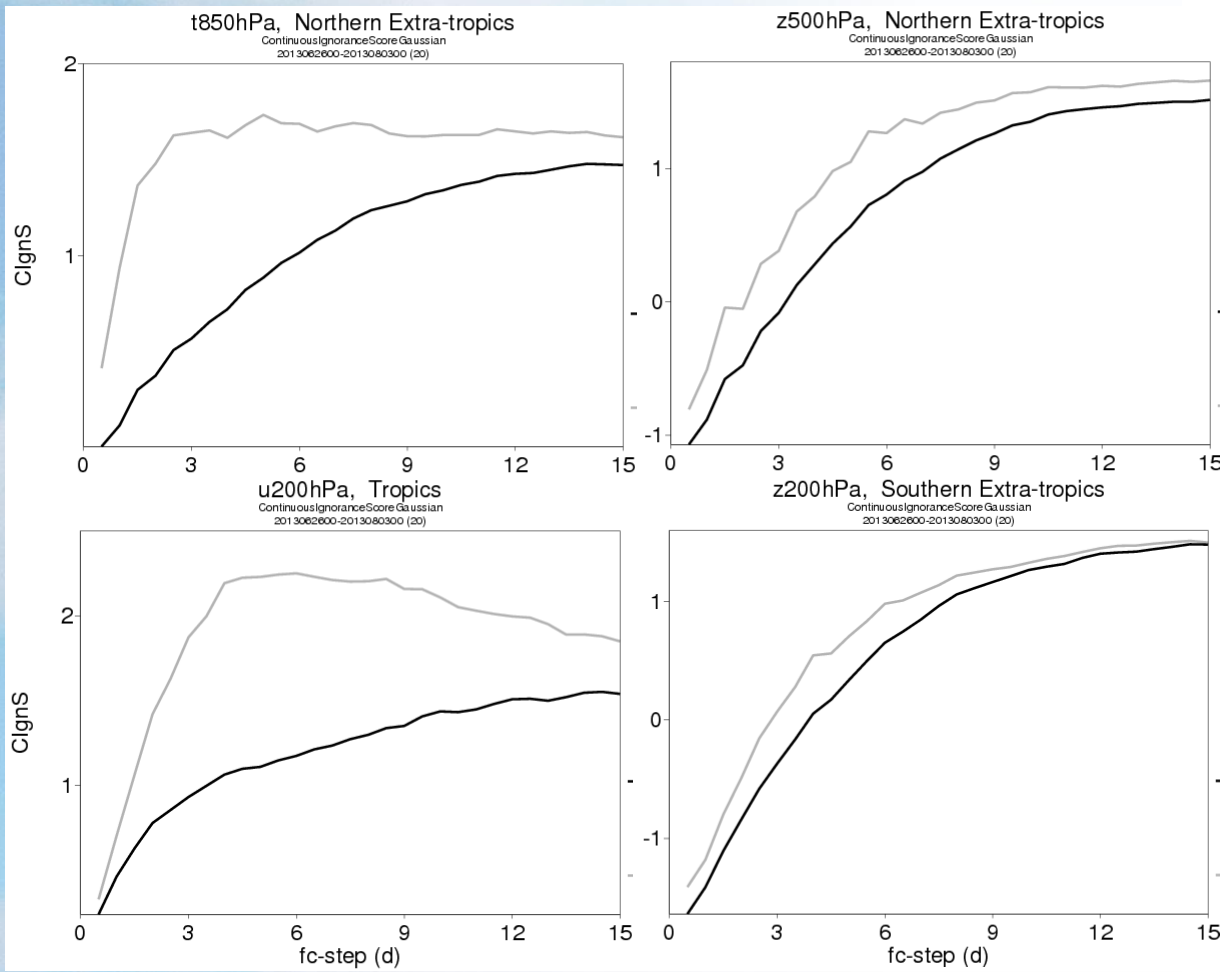


and ensemble spread does not provide a good estimate of error.

What happens when the ensemble forecast includes no representation of model uncertainty?

sarah-jane.lock@ecmwf.int

What happens with no accounting for model uncertainty?



Ensemble skill score

(Continuous Ignorance Score)

forecast times up to day 15

Key:

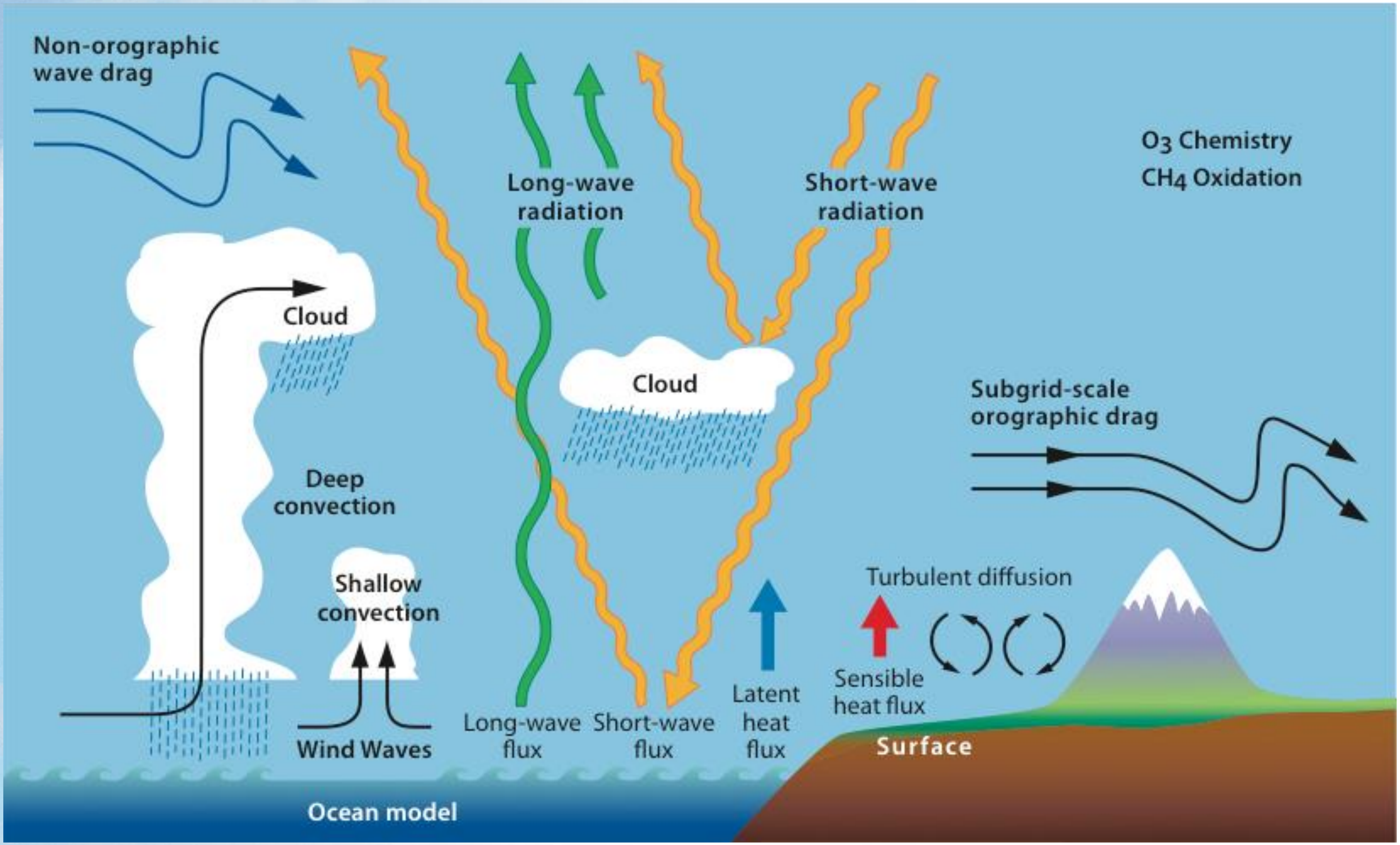
Initial perturbations ONLY

Initial + model uncertainty perturbations

sarah-jane.lock@ecmwf.int

Model uncertainty: where does it come from?

- Processes parametrised in the model:



sarah-jane.lock@ecmwf.int

Model uncertainty: where does it come from?

- Any other sources: processes missing from the underlying model?

- Atmosphere exhibits upscale propagation of kinetic energy (KE)
  - Occurs at ALL scales: no concept of “resolved” and “unresolved” scales
  - How can the model represent upscale KE transfer from unresolved to resolved scales?

sarah-jane.lock@ecmwf.int

Model uncertainty: how to simulate it?

- What do the model errors look like?
- What is the relative size of model error from different sources?

- How can we represent model errors?

- Multi-model ensembles
- Multi-physics ensembles
- Perturbed parameter ensembles
- “Stochastic parametrisations”

ECMWF
sarah-jane.lock@ecmwf.int

Stochastic physics schemes in IFS

- IFS ensemble forecasts (ENS and S4) include 2 model uncertainty schemes:
  - Stochastically perturbed physics tendencies (SPPT) scheme
  - Stochastic kinetic energy backscatter (SKEB) scheme

- SPPT scheme: simulates uncertainty due to sub-grid parametrisations
- SKEB scheme: parametrises a missing and uncertain process

sarah-jane.lock@ecmwf.int

SPPT scheme

- Initially implemented in IFS, 1998 (**Buizza et al., 1999**); revised in 2009:
- Simulates model uncertainty due to physical parameterisations by
  - taking the net parameterized physics tendency:

$$\boldsymbol{X} = [X_U, X_V, X_T, X_Q]$$

coming from $\left\{ \begin{array}{l} \textit{radiation} \\ \textit{gravity wave drag} \\ \textit{vertical mixing} \\ \textit{convection} \\ \textit{cloud physics} \end{array} \right\}$ schemes

  - and perturbing with multiplicative noise $r \in [-1, +1]$ as:

$$\boldsymbol{X}' = (1 + \mu r)\boldsymbol{X}$$

where $\mu \in [0,1]$ tapers the perturbations to zero near the surface & in the stratosphere.

Shutts et al. (2011, ECMWF Newsletter); Palmer et al., (2009, ECMWF Tech. Memo.)

sarah-jane.lock@ecmwf.int

SPPT pattern

- 2D random pattern in spectral space:
  - First-order auto-regressive [AR(1)] process for evolving spectral coefficients $\hat{r}$

$$\hat{r}(t + \Delta t) = \phi \hat{r}(t) + \rho \eta(t)$$

where $\phi = \exp(-\Delta t / \tau)$ controls the correlation over timestep Δt;

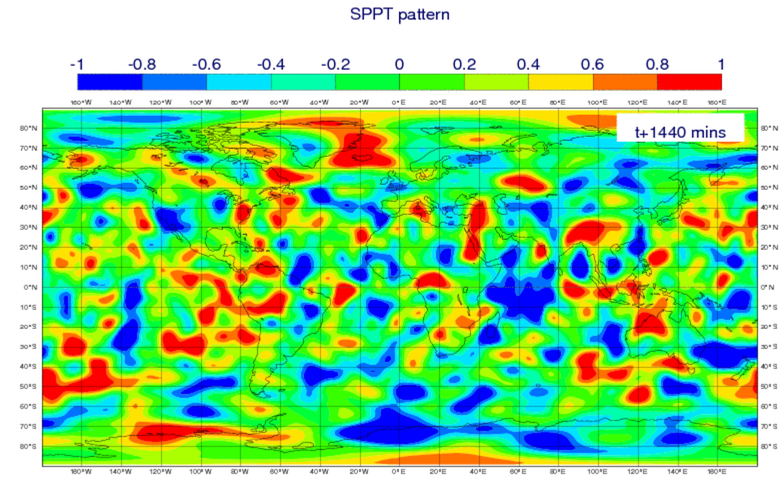
and spatial correlations (Gaussian) for each wavenumber define ρ for random numbers, η

- Resulting pattern in grid-point space r:
  - clipped such that $r \in [-1, +1]$
  - applied at all model levels to preserve vertical structures**
  - ***Except*: tapered to zero at model top/bottom, avoiding:
    - instabilities due to perturbations in the boundary layer;
    - perturbations in the stratosphere due to well-constrained clear-skies radiation

sarah-jane.lock@ecmwf.int

SPPT pattern

- 2D random pattern of spectral coefficients, r:
  - Time-correlations: AR(1)
  - Space-correlations: Gaussian
  - Clipped such that $r \in [-1, +1]$
- Applied at all model levels to preserve vertical structures**
- ***Except*: tapered to zero at model top/bottom

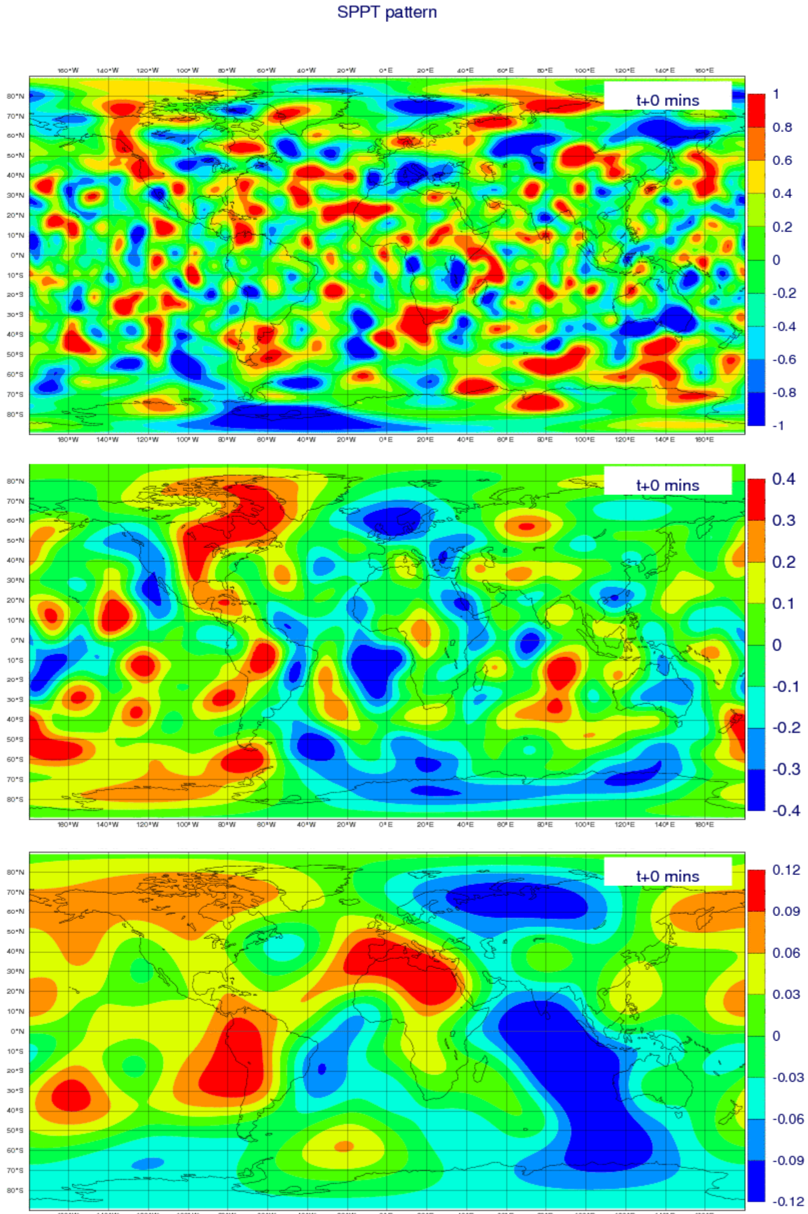


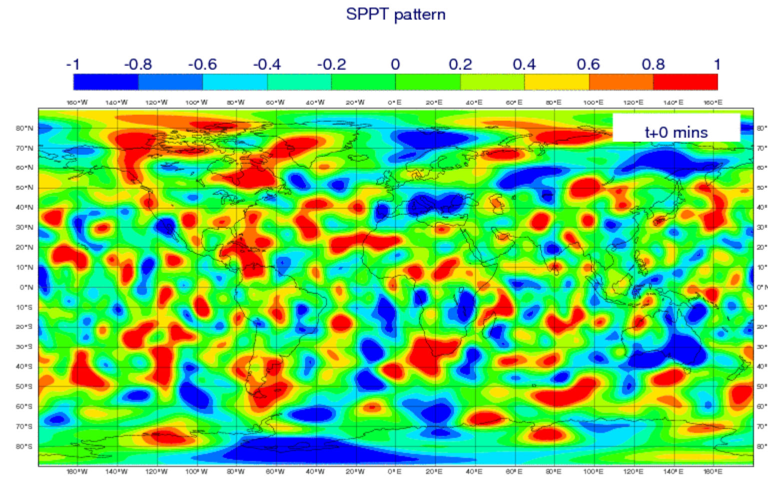
3 correlation scales:

i)	6 hours,	500 km,	$\sigma = 0.52$
ii)	3 days,	1 000 km,	$\sigma = 0.18$
iii)	30 days,	2 000 km,	$\sigma = 0.06$

sarah-jane.lock@ecmwf.int

SPPT pattern





3 correlation scales:

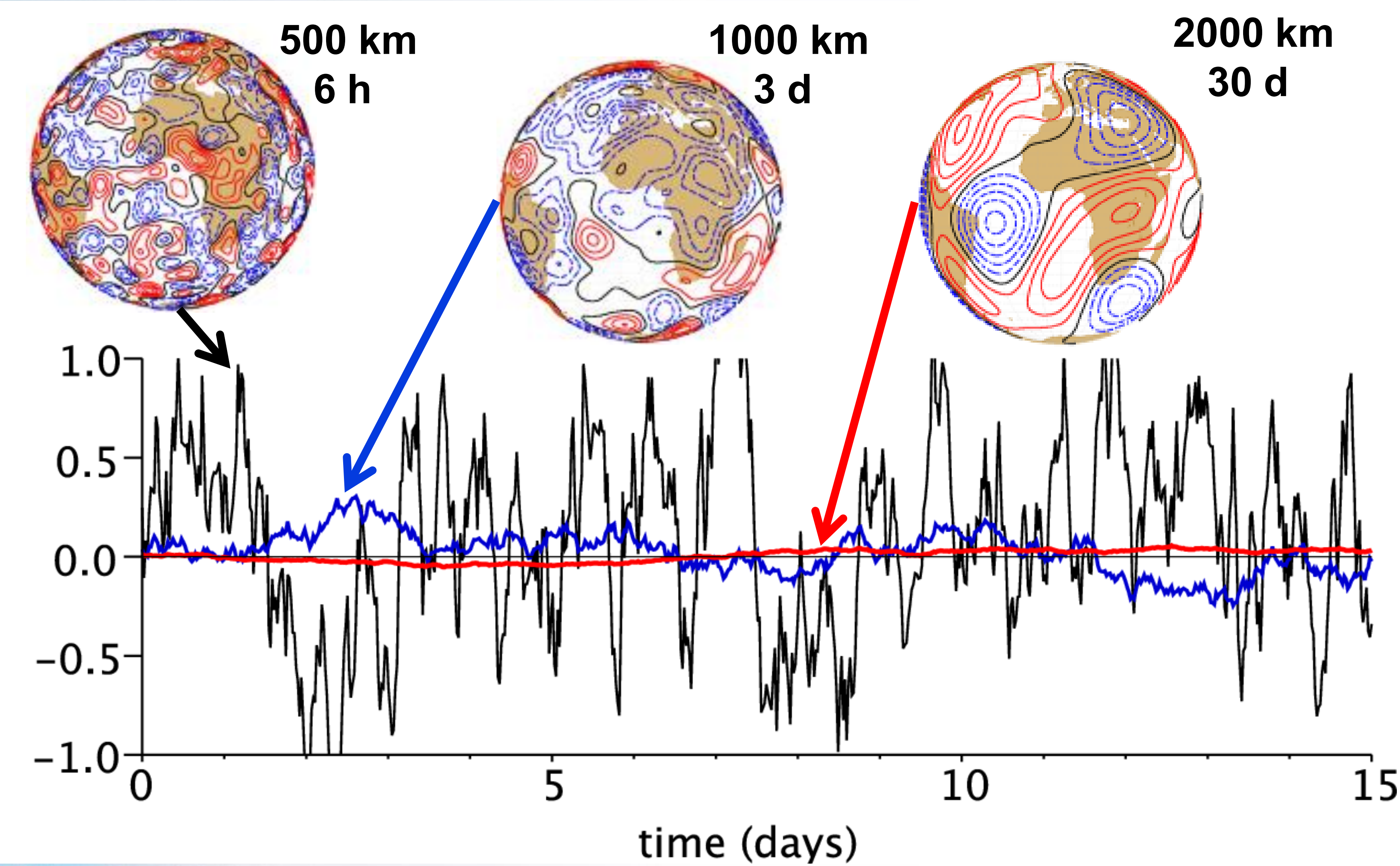
i) 6 hours, 500 km, $\sigma = 0.52$

ii) 3 days, 1 000 km, $\sigma = 0.18$

iii) 30 days, 2 000 km, $\sigma = 0.06$

sarah-jane.lock@ecmwf.int

SPPT pattern



sarah-jane.lock@ecmwf.int

SKEB scheme

- Introduced into IFS, 2010:
- Attempting to simulate a process otherwise absent from the model –

upscale transfer of energy from sub-grid scales

- Represents backscatter of Kinetic Energy (KE) by adding perturbations to U and V via a forcing term to the streamfunction:

$$F_{\varphi} = \left(\frac{b_R D_{\text{tot}}}{B_{\text{tot}}}\right)^{1/2} F^*$$

where F^* is a 3D random pattern field,

B_{tot} is the mean KE input by F^* alone,

D_{tot} is an estimate of the total dissipation rate due to the model,

b_{R} is the backscatter ratio – a scaling factor.

Shutts et al. (2011, ECMWF Newsletter); Palmer et al., (2009, ECMWF Tech. Memo.); Shutts (2005, QJRMS); Berner et al. (2009, JAS)

sarah-jane.lock@ecmwf.int

SKEB pattern

$$F_{\varphi} = \left(\frac{b_R D_{\mathrm{tot}}}{B_{\mathrm{tot}}}\right)^{1/2} F^*$$

- 3D random pattern field F^*:

– First-order auto-regressive [AR(1)] process for evolving F^*

$$F^*(t + \Delta t) = \phi F^*(t) + \rho\eta(t)$$

where $\phi = \exp(-\Delta t/\tau)$ controls the correlation over timestep Δt;

and spatial correlations (power law) for wavenumbers define ρ for random numbers, η

– vertical space-(de)correlations: random phase shift of η between levels

sarah-jane.lock@ecmwf.int

SKEB perturbations

$$F_{\varphi} = \left(\frac{b_R D_{\text{tot}}}{B_{\text{tot}}} \right)^{1/2} F^*$$

- D_{tot} is an estimate of sub-grid scale production of KE, and includes:
  - D_{num} = numerical dissipation from
    - explicit horizontal diffusion (bi-harmonic, ∇^2); and
    - estimate due to semi-Lagrangian interpolation error
  - D_{con} = estimated KE generated by updraughts and detrainment within sub-grid deep convection

sarah-jane.lock@ecmwf.int

How are the perturbation patterns determined?

- Characteristics of model errors cannot be determined from observations:
  - uncertain processes are small-scale (space and time)
  - lack of observational coverage

- Can attempt to use models: **coarse-graining** studies (e.g. Shutts and Palmer, 2007)
  - take high-resolution model simulations as "truth"
  - coarse-grain: average high-res model fields and tendencies (or streamfunction) to a grid-resolution typical of the forecast model
  - compare the contribution of "sub-grid" scales in the coarse-grained simulation with parametrisations in the forecast model
  - coarse-graining studies have been used to justify and inform scales in SPPT and SKEB perturbation patterns

sarah-jane.lock@ecmwf.int

IFS ensembles: ENS and System 4 (S4)

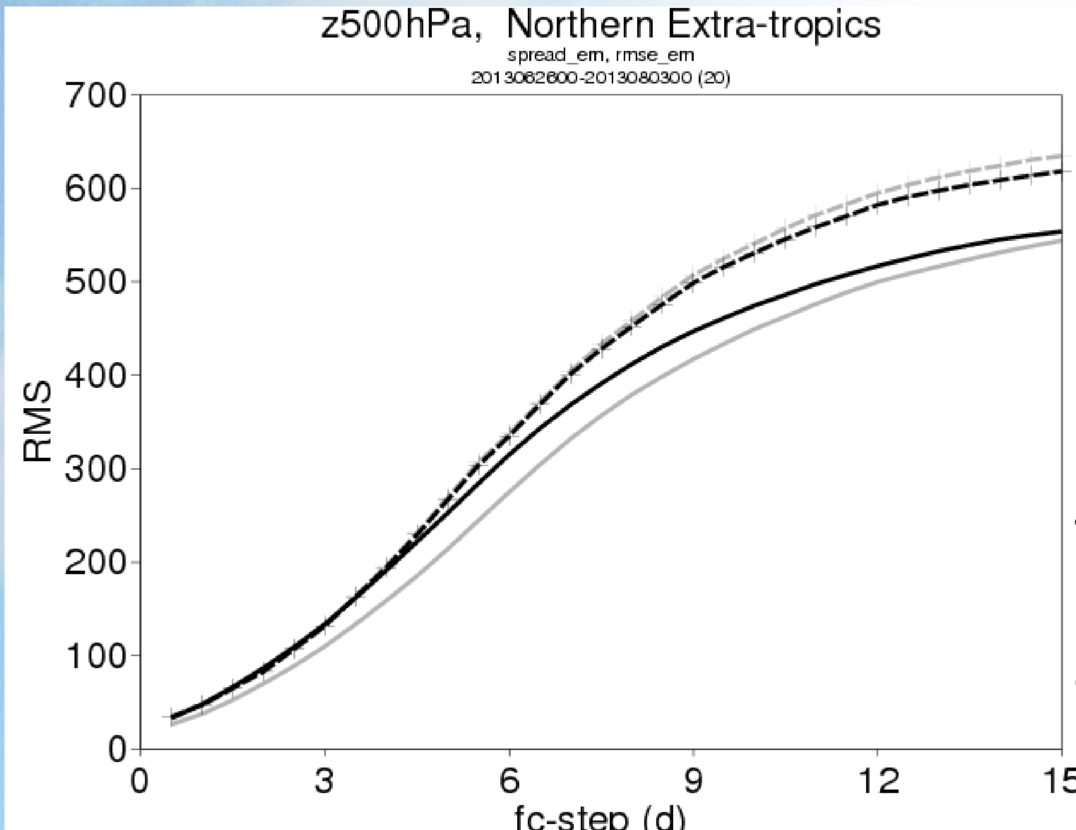
- **ENS** = ensemble prediction system for
  - medium-range forecasts (up to 15 days) and
  - monthly forecasts (up to 46 days)
- **S4** = seasonal forecasting system
  - up to 7 months

- Both systems represent model uncertainty with SPPT and SKEB

- **ENS**:
  - 1 control forecast + 50 perturbed members
  - T639 (~32 km) resolution to day 10; T319 (~65 km) days 10-15
  - 91 vertical levels, up to 0.01hPa

sarah-jane.lock@ecmwf.int

Impact of SPPT and SKEB in ENS



Key:

[dashed] = Ensemble mean RMS error

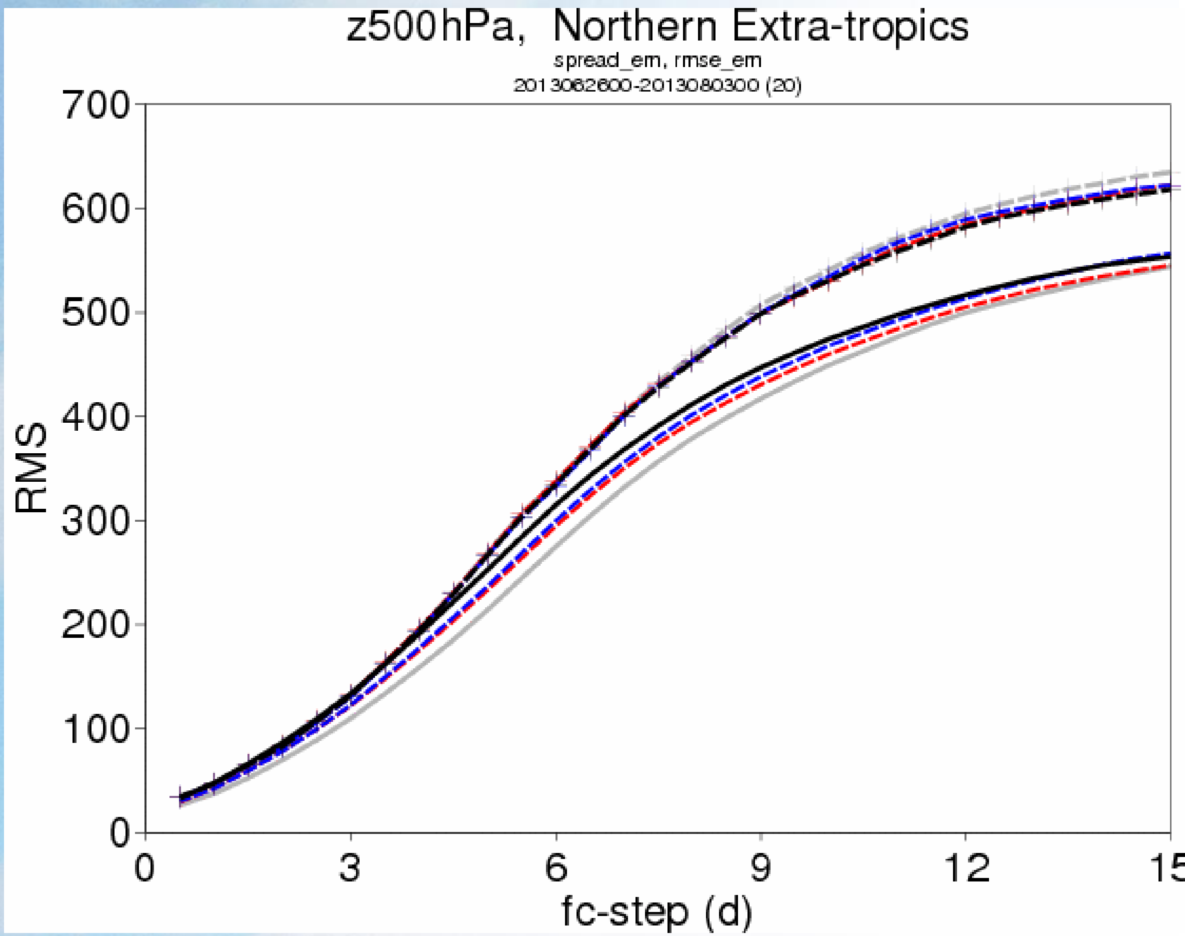
[solid] = Ensemble standard deviation

Initial & SPPT+SKEB perturbations

Initial perturbations ONLY

sarah-jane.lock@ecmwf.int

Impact of SPPT and SKEB in ENS



Key:

[-+-] = Ensemble mean RMS error

[---] = Ensemble standard deviation

Initial & SPPT+SKEB perts

Initial + SPPT perts

Initial + SKEB perts

Initial perturbations ONLY

sarah-jane.lock@ecmwf.int

Impact of SPPT and SKEB in ENS

- Adding SPPT + SKEB perturbations:
  - increases ensemble "spread" (= ensemble standard deviation), i.e. ensemble members describe greater region of the parameter space
  - some reduced ensemble mean errors
- In the extra-tropics:
  - SPPT and SKEB each have a similar impact, i.e. perturbations are successfully adopted and evolved by the model
  - *Experiments: perturbations in days 0-5 contribute most effect*
- In the tropics:
  - SPPT has a much greater impact (in terms of both spread and error) than SKEB, i.e. SPPT perturbations more able to excite modes that the model can evolve
  - *Experiments: effect of perturbations rapidly lost at all times*

sarah-jane.lock@ecmwf.int

Impact of SPPT and SKEB in S4

- System 4 (S4), November 2011: introduction of revised SPPT and SKEB
- Operational configuration:
  - T255 (~80 km), 91 vertical levels (up to 0.01 hPa)
  - Coupled ocean model: NEMOv3.0, 1 degree (~110 km), 42 vertical levels
  - 51 members
  - Initialised on 1st of each month
  - Forecast lead times: to 7 months

- Recent work with S4 to assess impact of stochastic schemes
- For longer time-scales, consider impact in terms of:
  - Noise-induced drift, i.e. change in model mean
  - Noise-activated regime transition, e.g. Pacific-N. American region regimes

sarah-jane.lock@ecmwf.int

Impact of SPPT and SKEB in S4

- Recent work with S4 to assess impact of stochastic schemes:
  - Hindcast period: 1981-2010
  - Start dates: May, Aug & Nov
  - Ensemble size: 51
  - Forecasts to lead times: 4-7 months

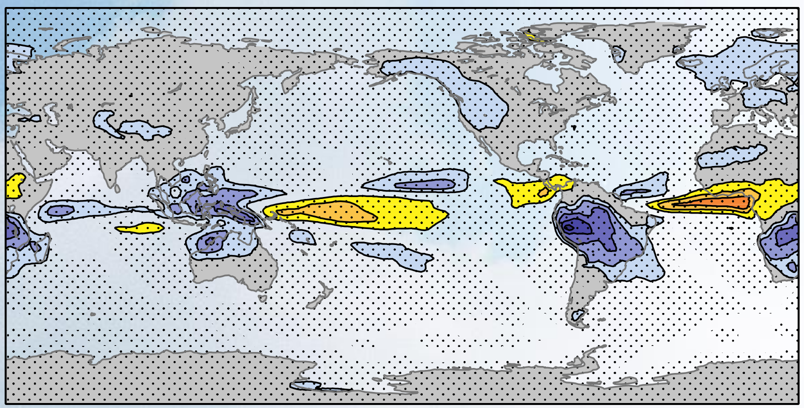
- Considers impact of SPPT + SKEB on:
  - Systematic errors
  - Madden-Julian Oscillation (MJO) statistics
  - Circulation regimes over the Pacific-North American region

sarah-jane.lock@ecmwf.int

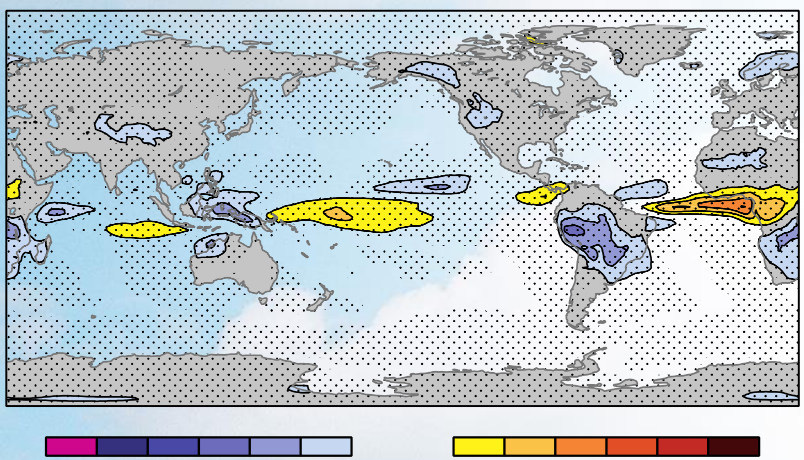
Impact of SPPT and SKEB in S4: systematic errors

Outgoing Longwave Radiation **(DJF 1981-2010)**

stochphysOFF **– ERA-I**



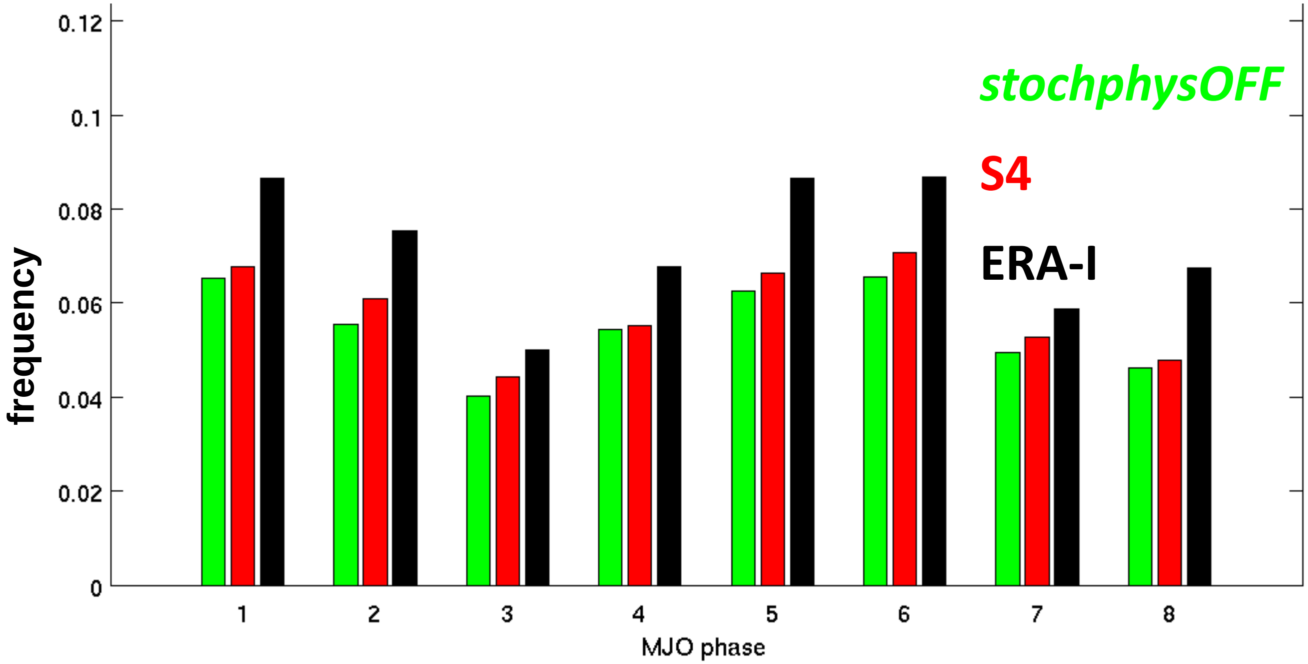
S4 – ERA-I



- SPPT+SKEB: reduction of overly active tropical convection
- Similar reductions in excessive:
  - Total cloud cover
  - Total precip
  - Zonal winds (850 hPa)
- SPPT is responsible for most of the difference; SKEB has little impact

sarah-jane.lock@ecmwf.int

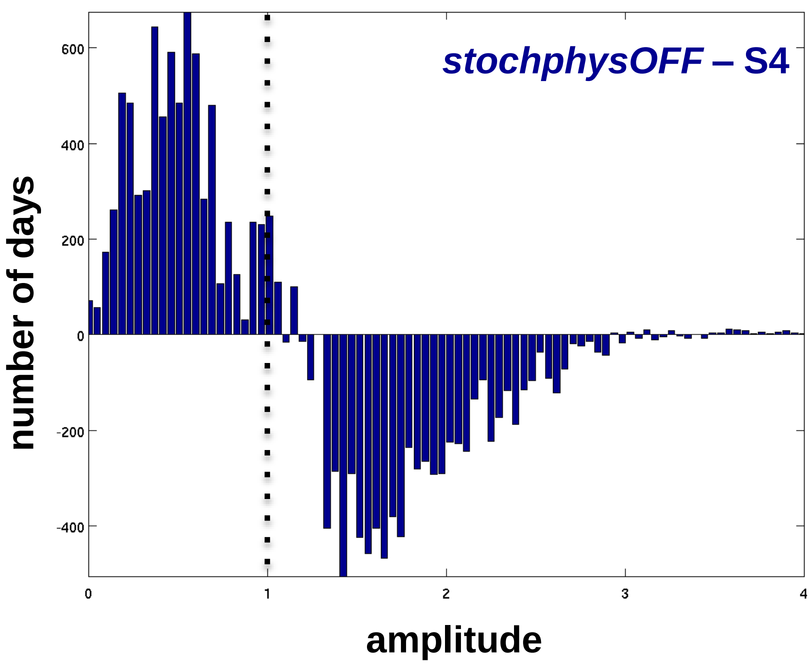
Impact of SPPT and SKEB in S4: MJO

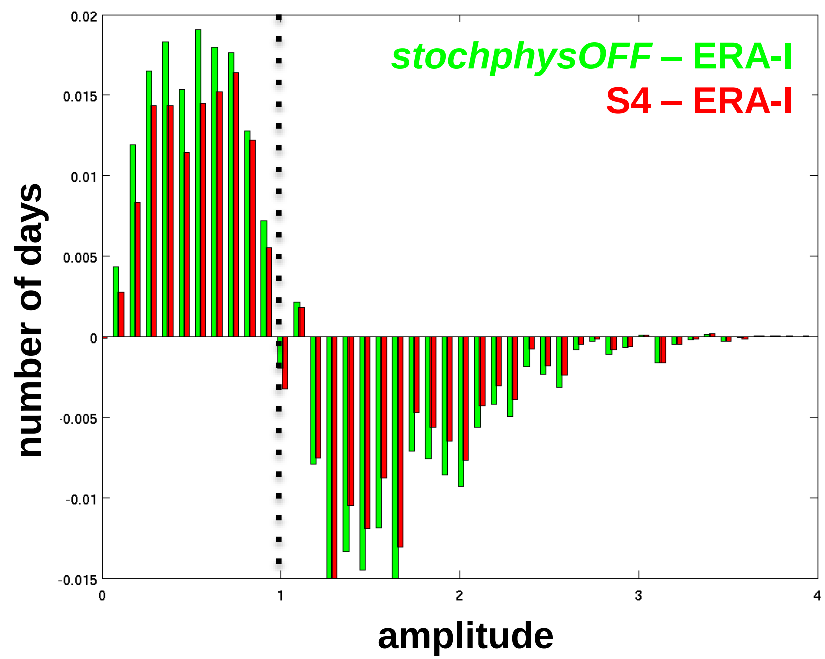


Wheeler and Hendon Index: projection of daily data on 2 dominant combined EOFs of OLR, u200 and u850 over 15° N-15° S

SPPT+SKEB:

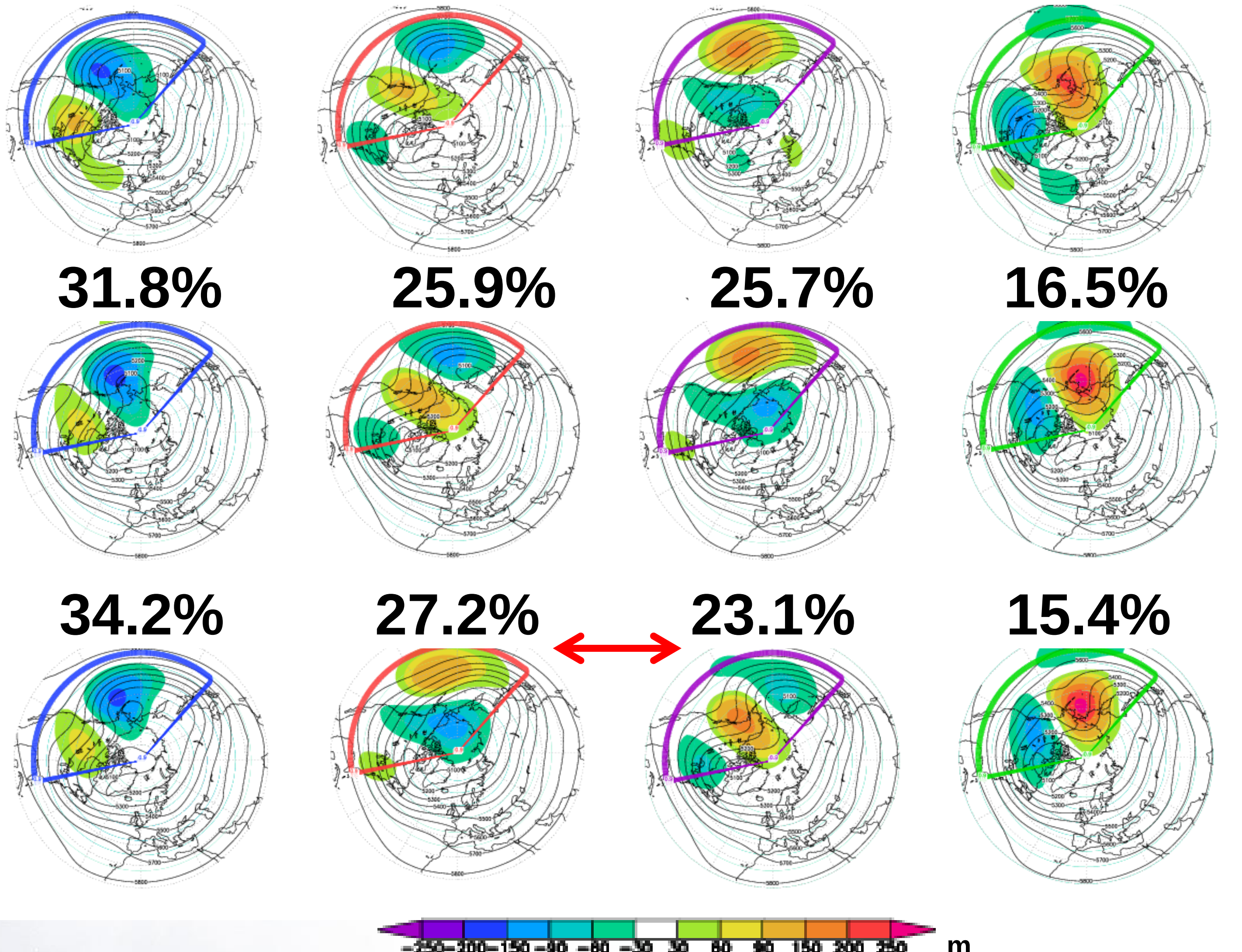
- **Increased frequency of events**
- **Improved amplitude distribution**





Impact of SPPT & SKEB in S4: Pacific North America (PNA) circulation regimes

	Cluster 1 “Pacific Trough”	Cluster 2 “PNA+”	Cluster 3 “PNA-”	Cluster 4 “Alaskan Ridge”
ERA	28.6%	28.1%	27.0%	16.3%
System 4	31.8%	25.9%	25.7%	16.5%
stochphysOFF	34.2%	27.2%	23.1%	15.4%



sarah-jane.lock@ecmwf.int

Model uncertainty representation: brief outlook for IFS

- **Exploring alternative stochastic perturbations**:

Currently, in SPPT, we perturb:

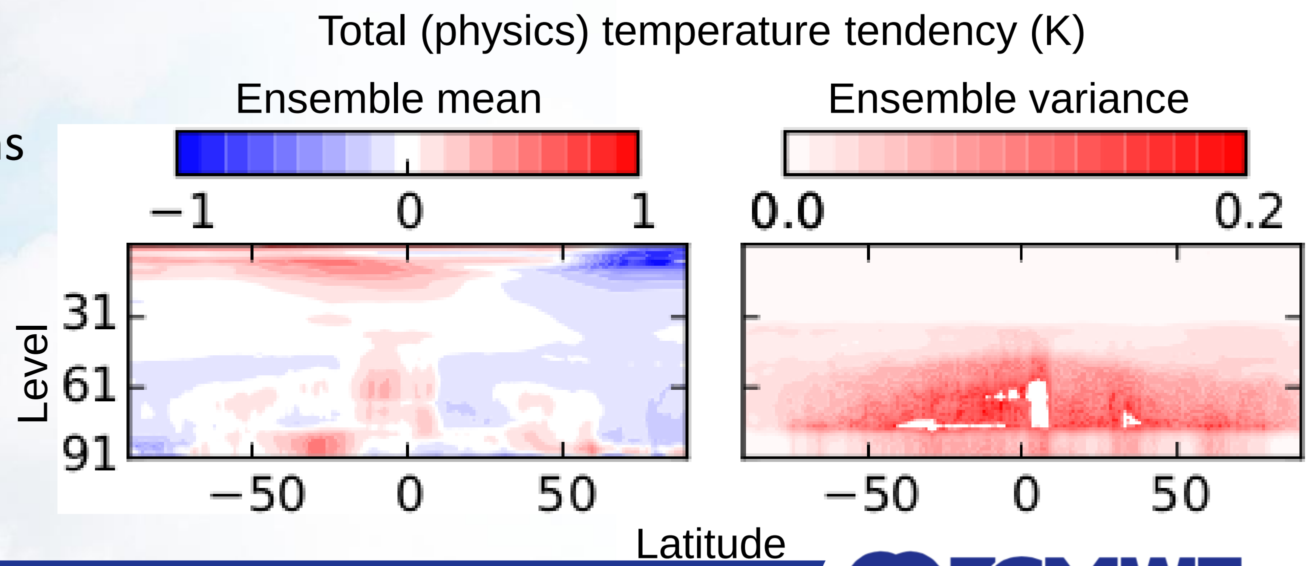
$$\boldsymbol{X} = X_{RAD} + X_{GWD} + X_{MIX} + X_{CON} + X_{CLD}$$

with zero perturbations near the surface and in the stratosphere.

- **Instead**, identify and perturb individual uncertain parameters.

e.g. **Boundary layer**:

SPPT: no BL perturbations



sarah-jane.lock@ecmwf.int

Model uncertainty representation: brief outlook for IFS

- **Exploring alternative stochastic perturbations**:

Currently, in SPPT, we perturb:

$$\boldsymbol{X} = X_{RAD} + X_{GWD} + X_{MIX} + X_{CON} + X_{CLD}$$

with zero perturbations near the surface and in the stratosphere.

- **Instead**, identify and perturb individual uncertain parameters.

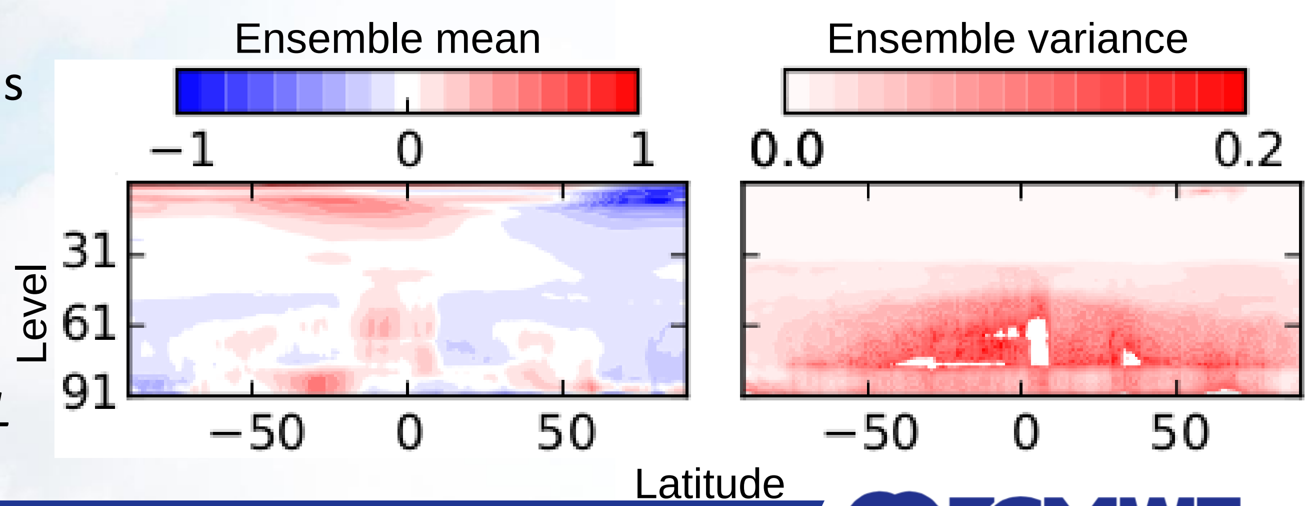
e.g. **Boundary layer**:

SPPT: no BL perturbations

New: 4 BL parameters perturbed

→ *Greater variance in BL*

Total (physics) temperature tendency (K)



sarah-jane.lock@ecmwf.int

Representing model uncertainty: summary

- Model uncertainty arises due to unresolved and misrepresented processes
  - finite-resolution of a discrete numerical model
  - parametrisations must describe multi-scale sub-grid processes in bulk
- Difficult to characterise sources of model errors due to lack of observations
- Without representing model uncertainty, ensembles are under-dispersive
- ECMWF ensembles include 2 stochastic model uncertainty schemes:
  - SPPT: representing uncertainty due to sub-grid physics parameterisations
  - SKEB: simulating upscale transfer of kinetic energy from unresolved scales
- Medium-range: increased ensemble spread, greater probabilistic skill
- Seasonal: reduction in biases; better representation of MJO, PNA regimes
- Outlook: Seeking to focus perturbations on individual uncertain parameters

sarah-jane.lock@ecmwf.int