Representing model uncertainty

Stochastic perturbations

Sarah-Jane Lock

Model Uncertainty, Earth System Modelling, Research Department

sarah-jane.lock@ecmwf.int

ECMWF

Introduction: Model Uncertainty

- Ensemble forecasts enable a quantification of the confidence in a forecast, e.g. 10% chance of rain

- An ensemble forecast is made from multiple forecasts or "members", each member perturbed with respect to the others

- The perturbations comprise

  a) different initial conditions for each member, to sample the uncertainty in our description of the initial state (***Simon Lang's lecture***); and

  b) a different forecast model for each member, to sample the uncertainty due to the model integrations or the "***model uncertainty***"

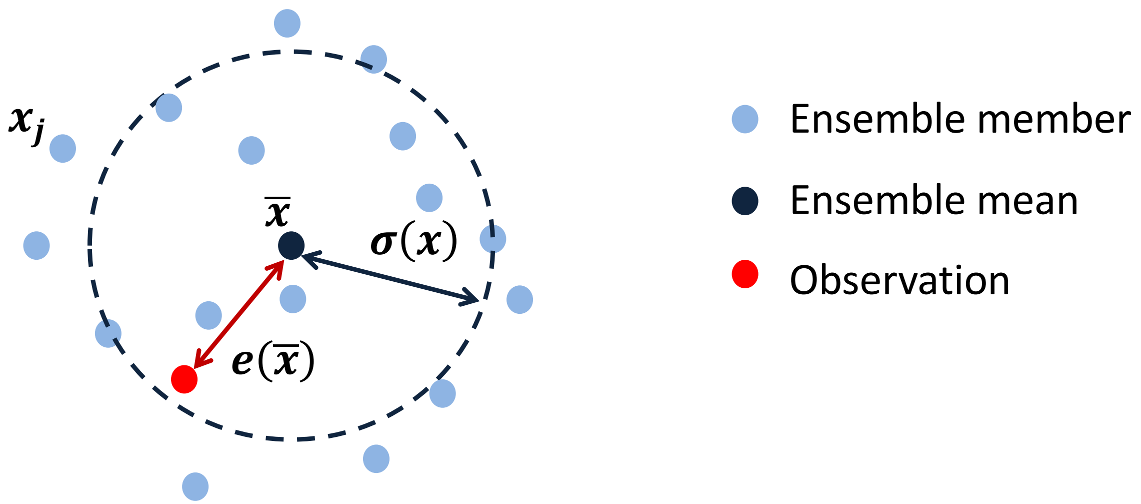
- To date, much effort has been focused on **model uncertainty** due to the parametrization schemes that describe sub-grid atmospheric physics --- representing this with stochastic perturbations gives rise to "***stochastic physics***"

Using stochastic physics to represent model uncertainty

- Why do we represent model uncertainty in an ensemble forecast?
- What are the sources of model uncertainty?
- How do we currently represent model uncertainty in the IFS?
- Ongoing work towards process-level simulation of model uncertainty

Ensemble reliability

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



i.e. averaged over many ensemble forecasts,

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

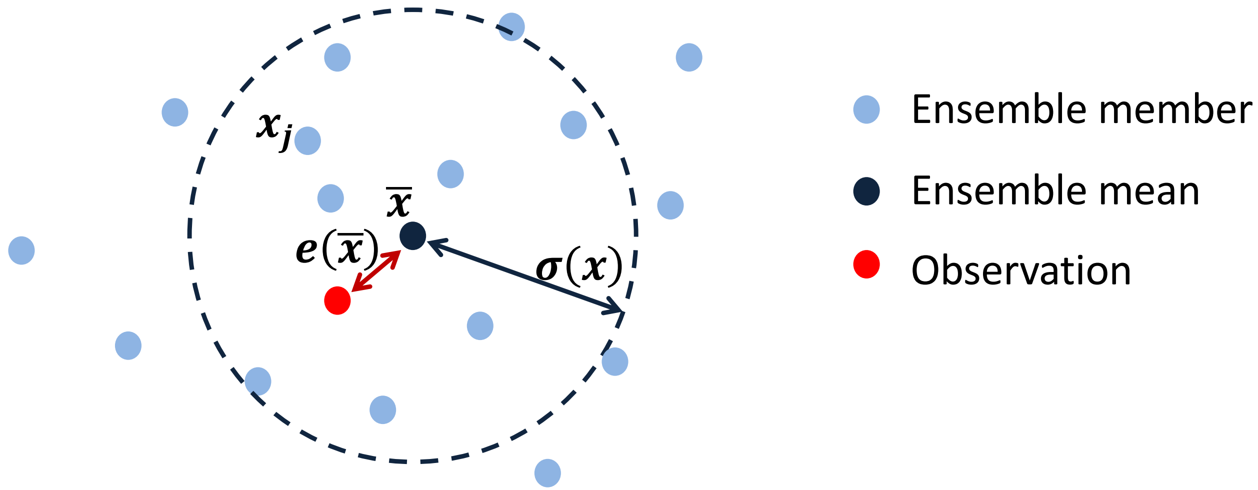
For a thorough discussion of this relationship:

Martin Leutbecher's lectures

Ensemble reliability

- In an **over**-dispersive ensemble,

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



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

The relatively large spread implies large uncertainty and hence, likely large error:

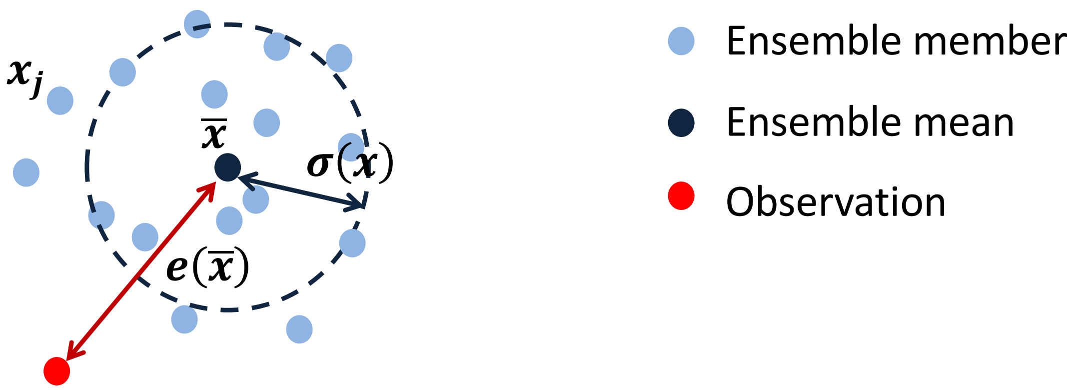
an “under-confident forecast”

ECMWF

Ensemble reliability

- In an **under**-dispersive ensemble,

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



The small spread implies low uncertainty and hence, small errors:

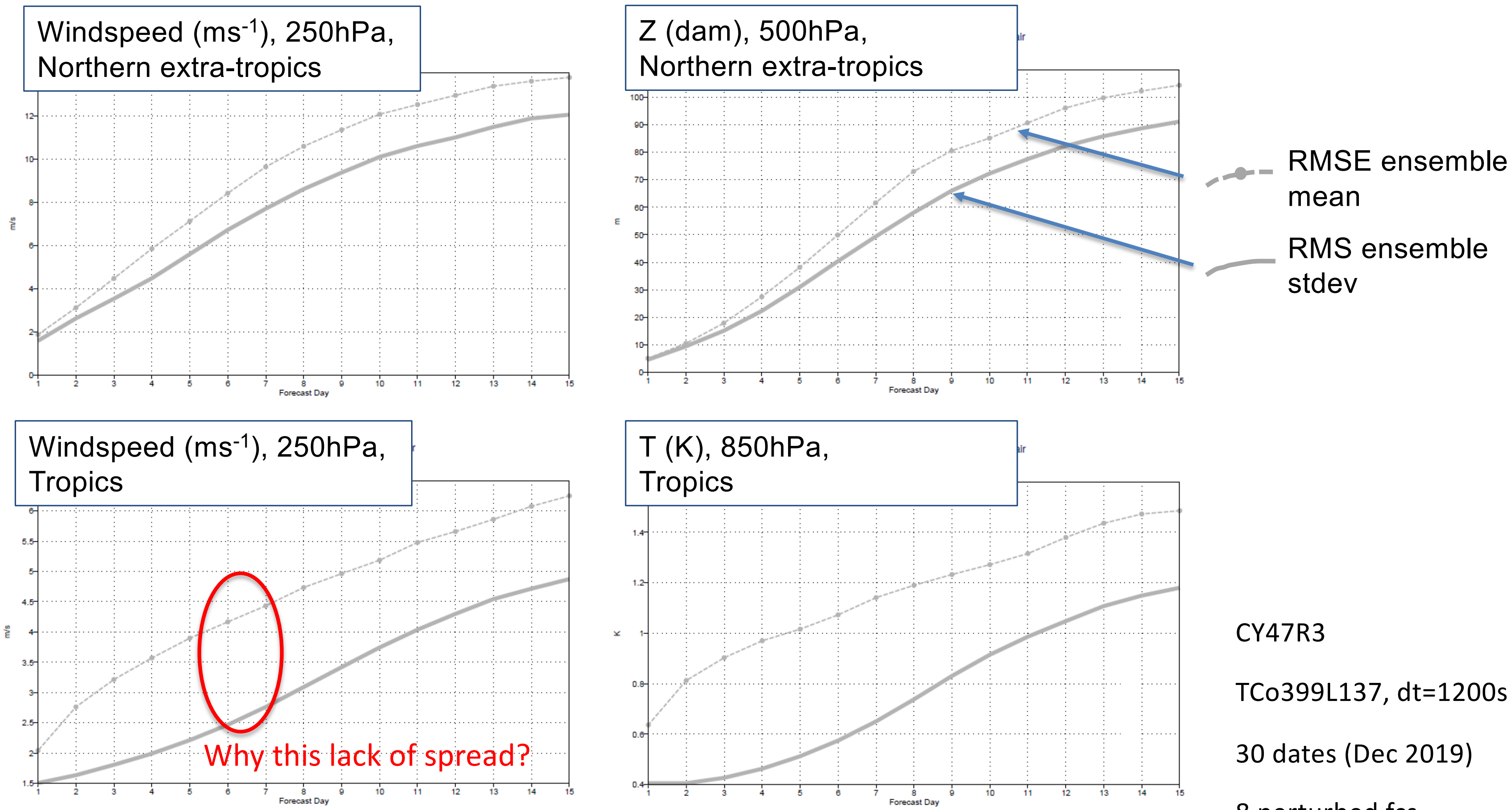
an “over-confident forecast”

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

ECMWF

Ensemble forecasts with only initial conditions perturbations

Ensemble mean RMSE (“Error”) & standard deviation (“Spread”)



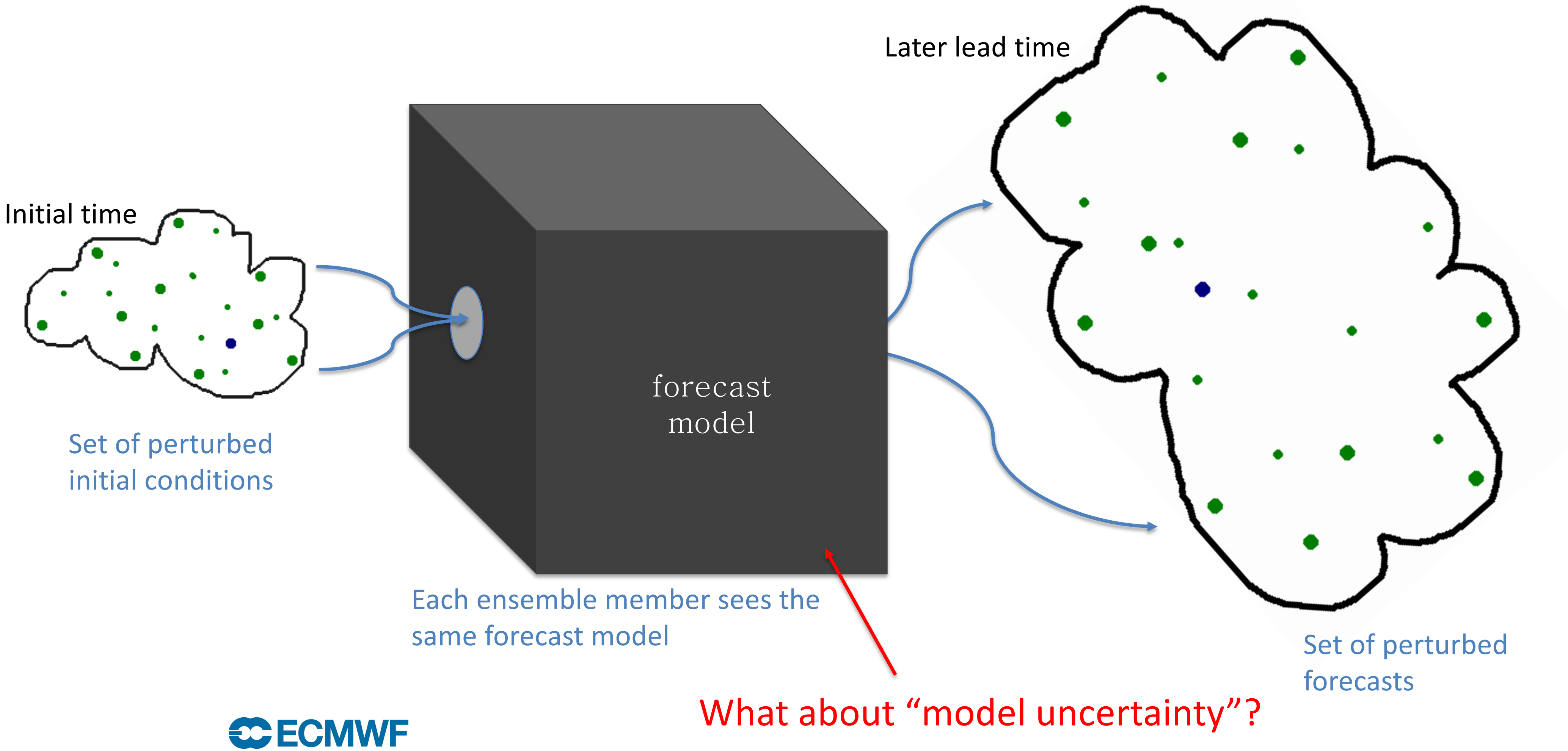
CY47R3

TCo399L137, dt=1200s

30 dates (Dec 2019)

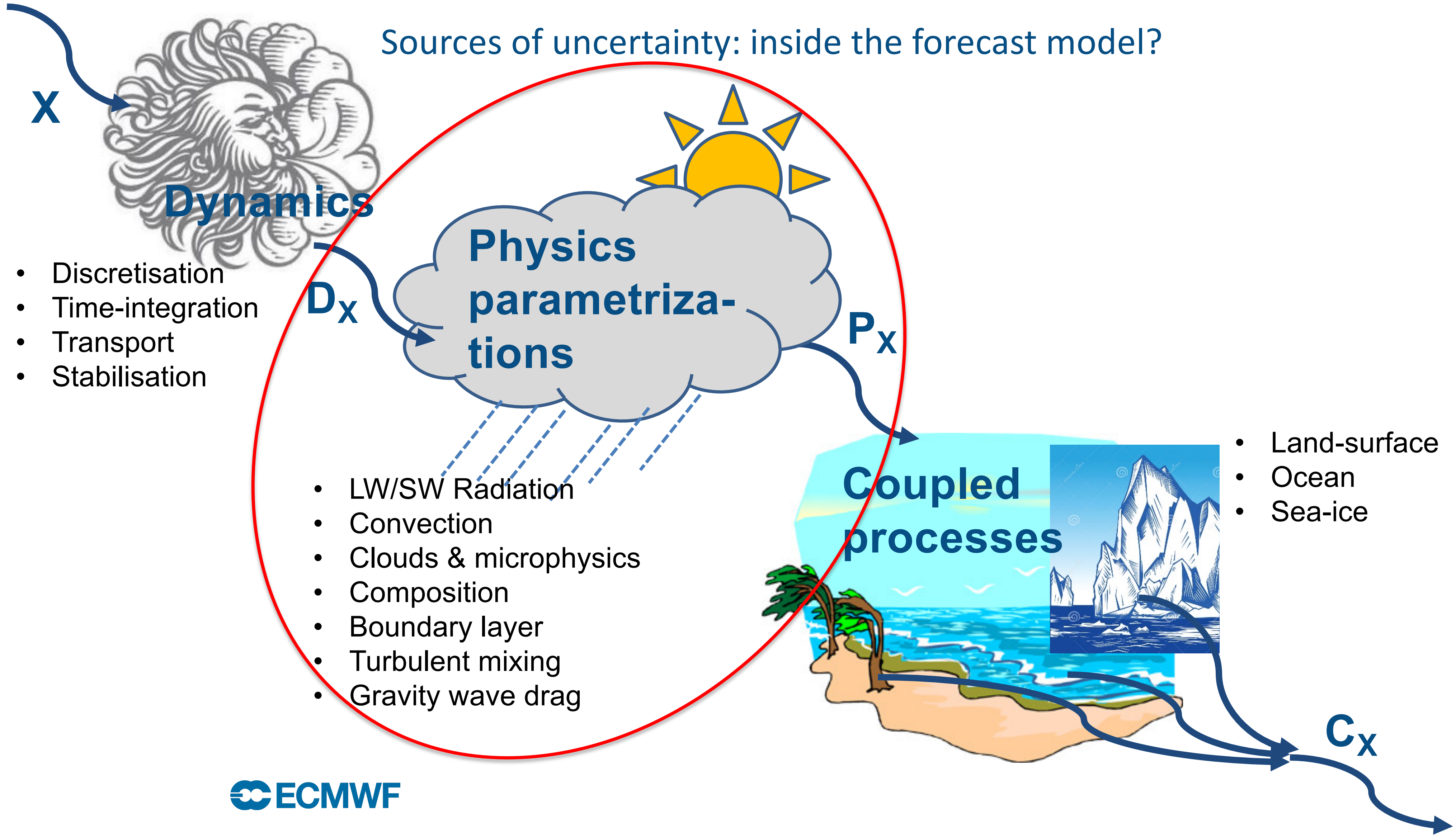
8 perturbed fcs

Sources of uncertainty: initial conditions



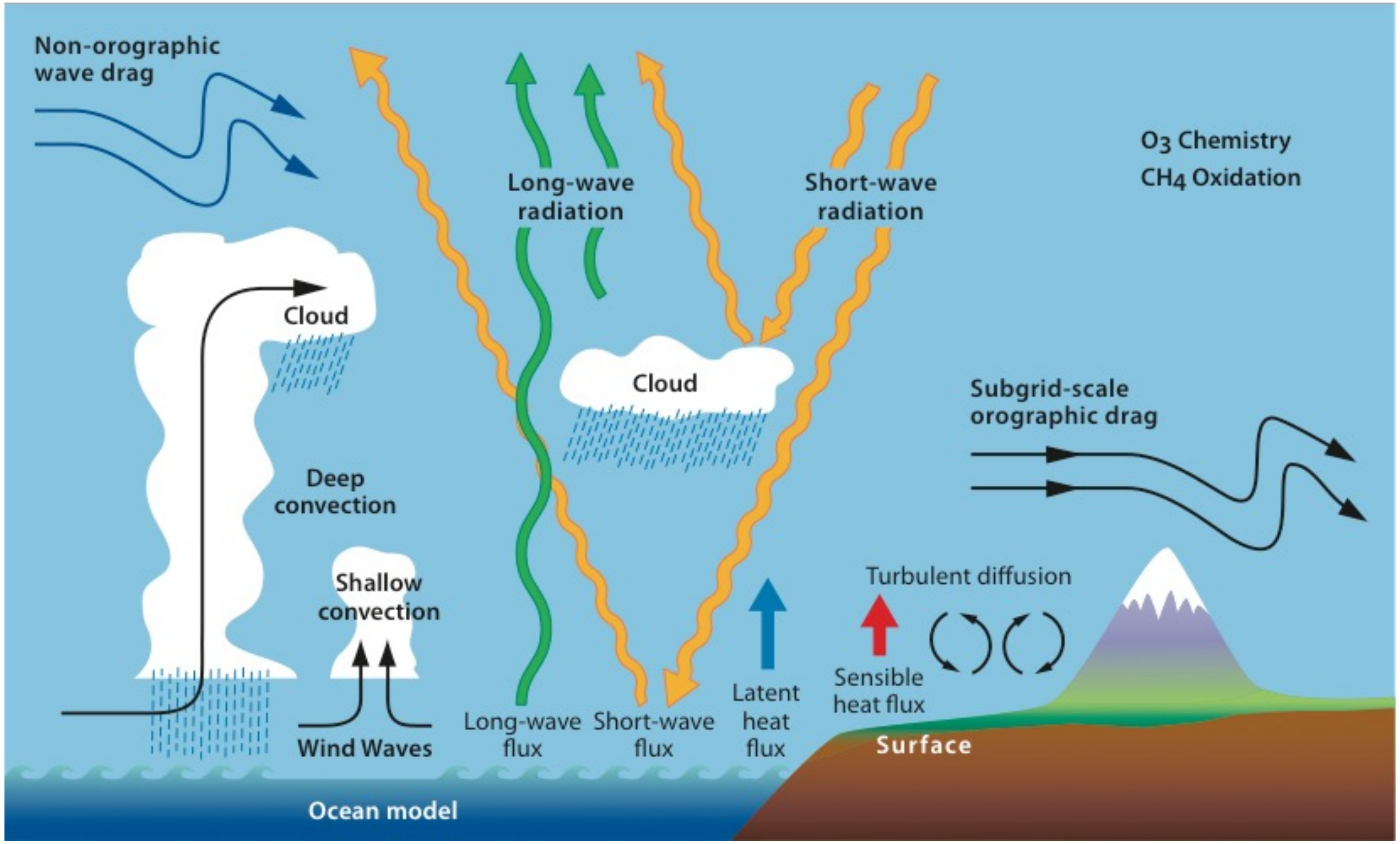
ECMWF

Sources of uncertainty: inside the forecast model?



ECMWF

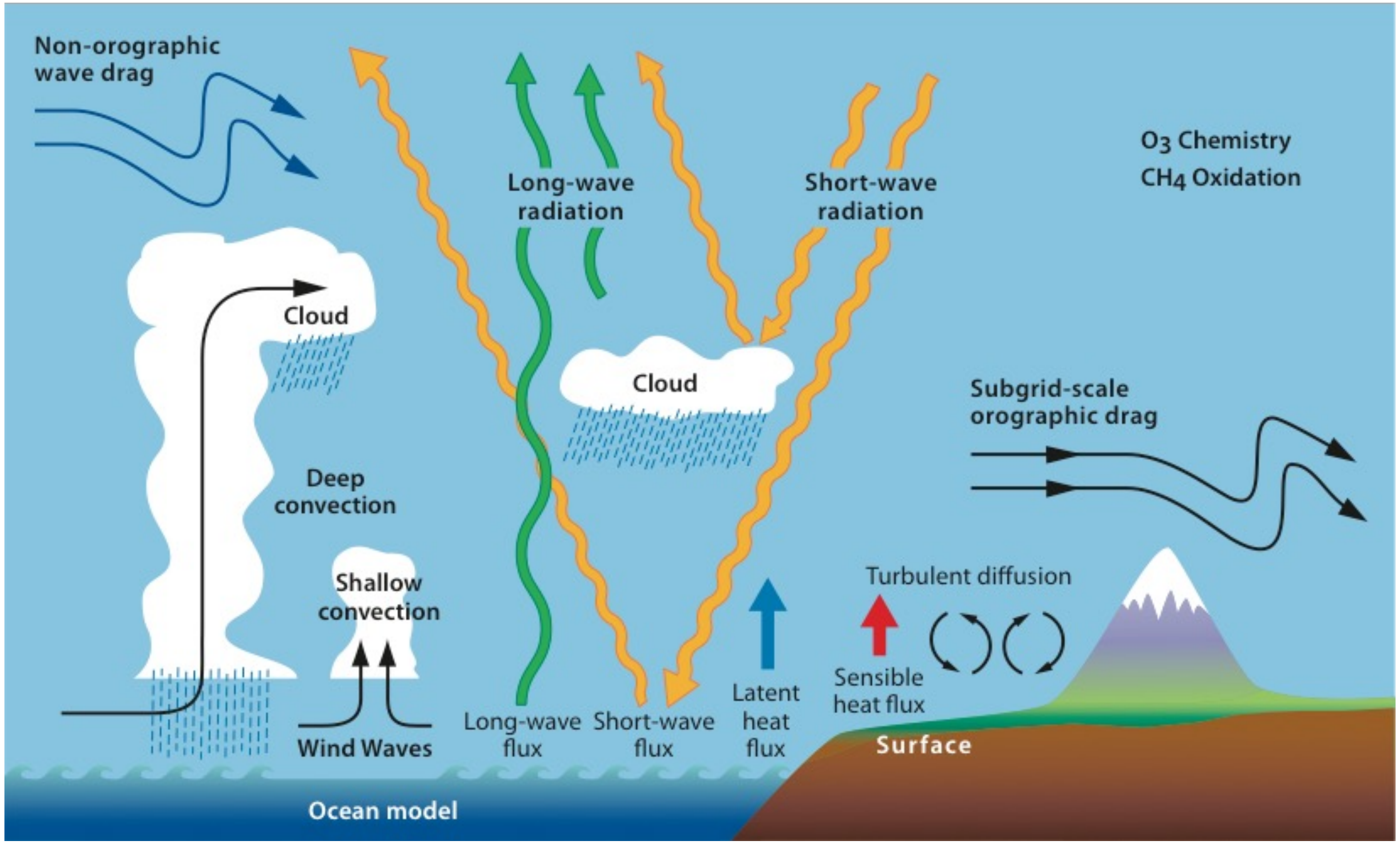
Model uncertainty: parametrized atmospheric physics processes



Uncertainties arise due to:

- Inability to resolve sub-grid scales, e.g.
  - Surface drag (orography/waves)
  - Convection rates (occurrence / en/detrainment)
  - Phase transitions
  - Radiation transfer in cloudy skies
- Poorly constrained parameters, e.g.
  - Vertical cloud-overlap (radiation)
  - Composition
  - Non-orographic drag

Model uncertainty: parametrized atmospheric physics processes



"Let's take the positives"

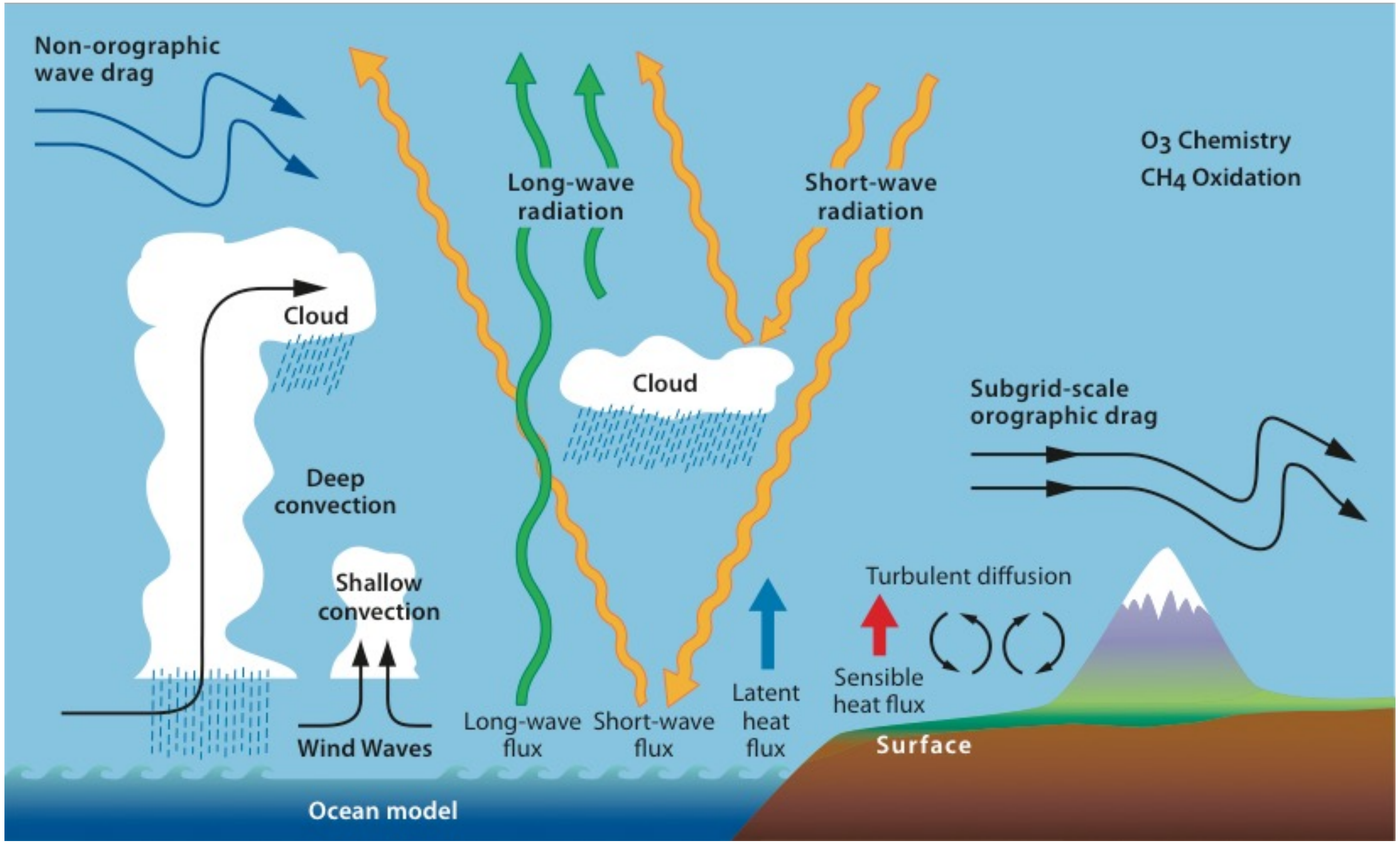
Parametrisation schemes:

- developed/operate together
- highly tuned for best performance

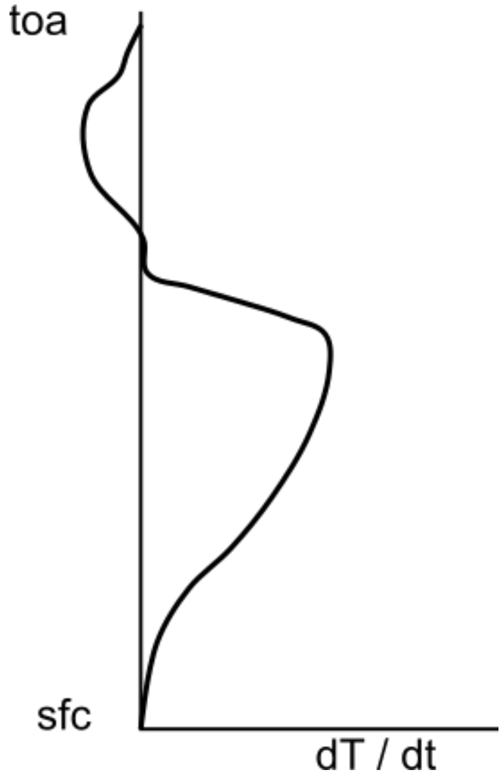
Seek a description of uncertainty that retains consistencies of the representation of the physical processes.

ECMWF

Model uncertainty: parametrized atmospheric physics processes

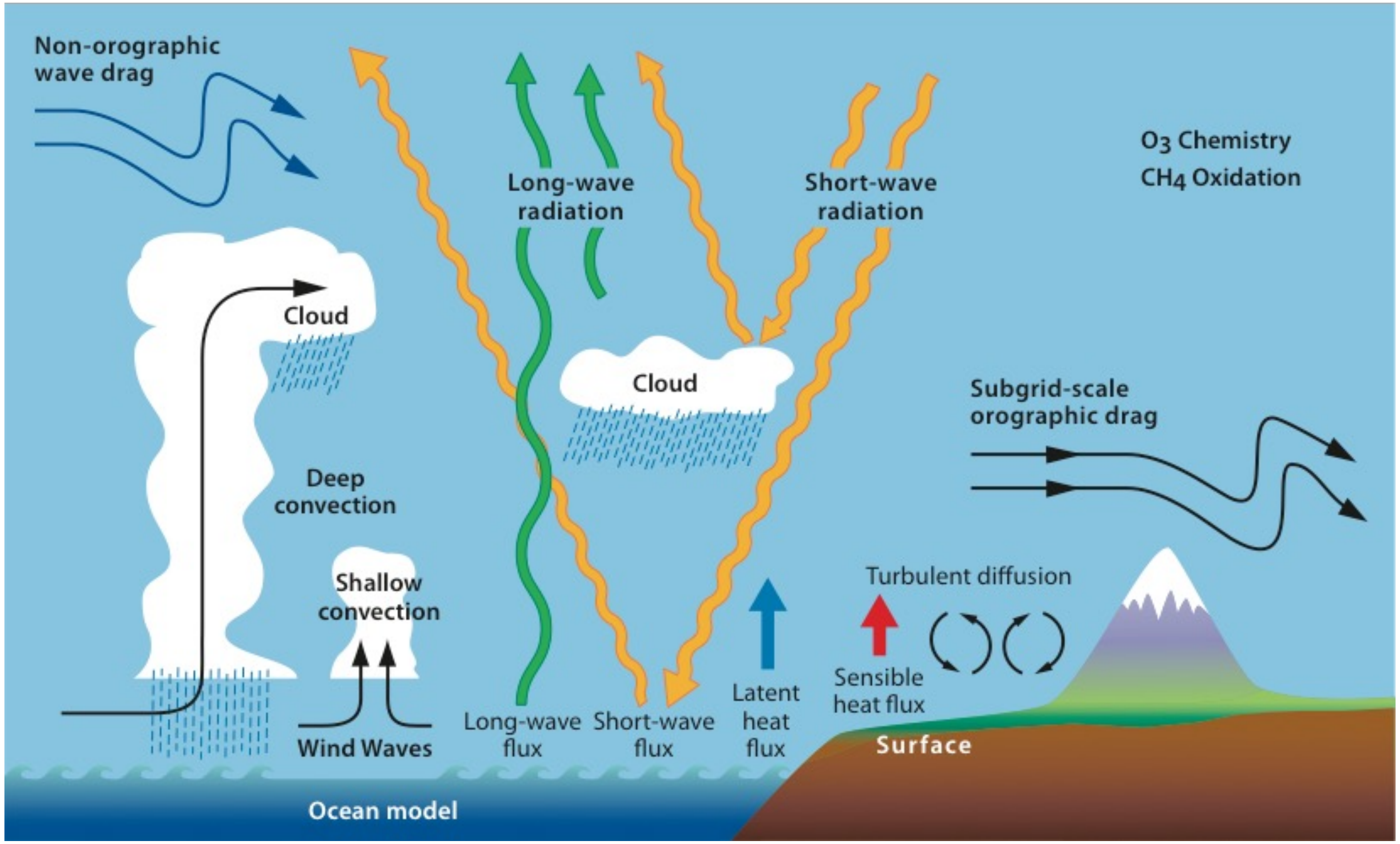


e.g. profile of heating rates from physics parametrisations:

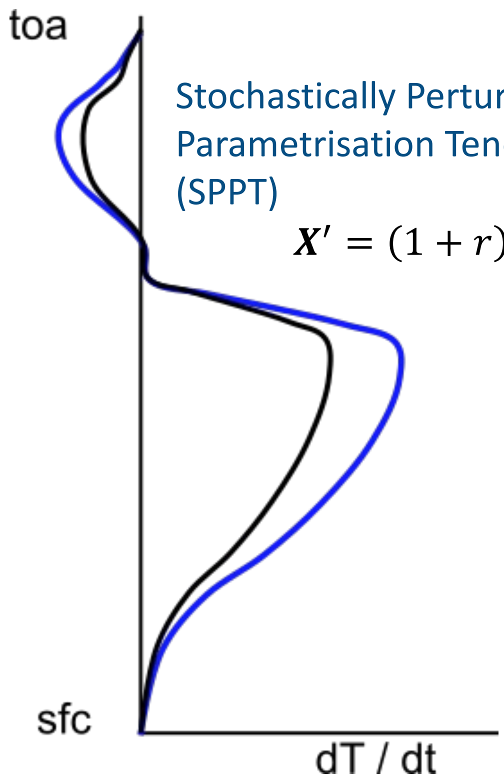


ECMWF

Model uncertainty: parametrized atmospheric physics processes



Proposal: represent uncertainties with a perturbation proportional to the profile of net physics tendencies

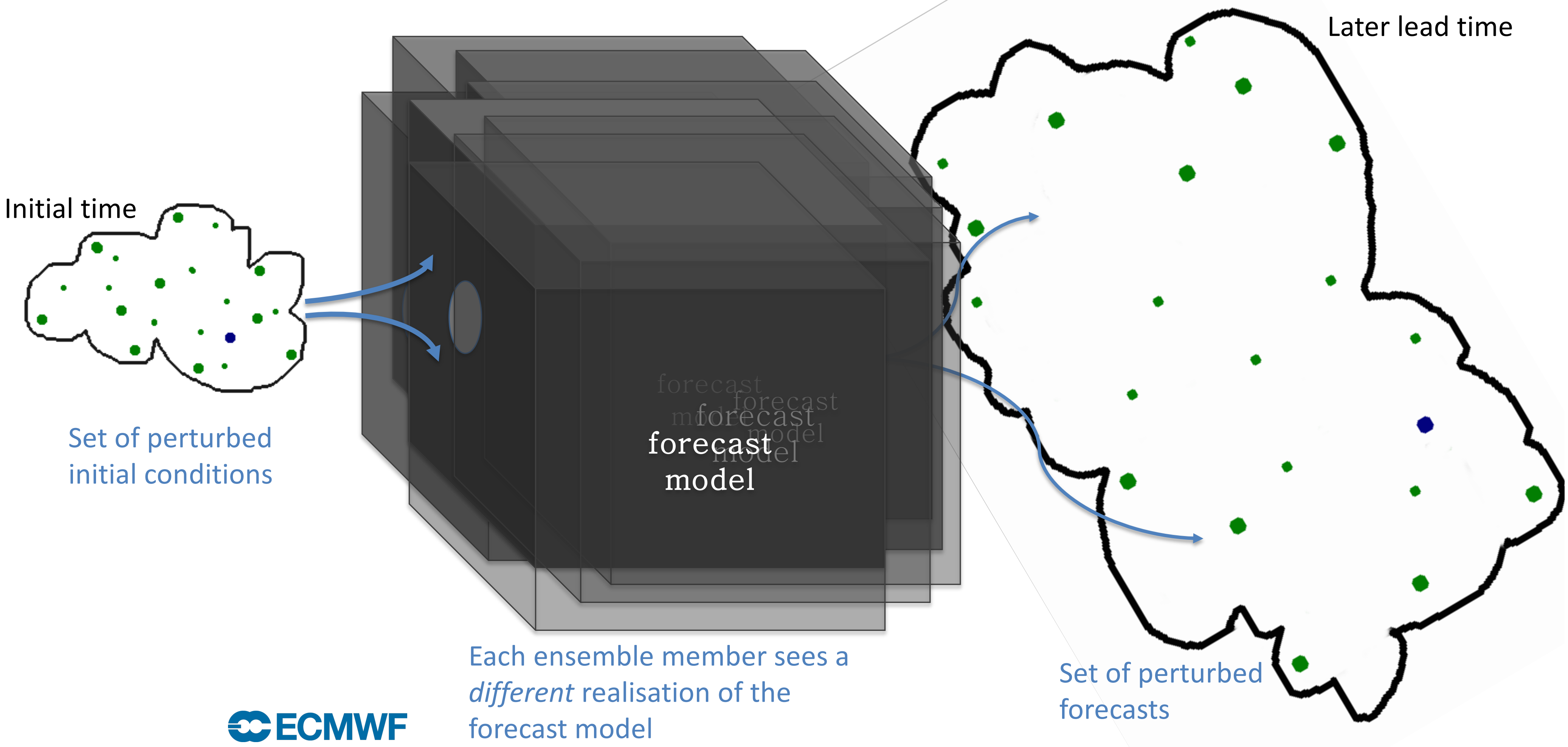


Stochastically Perturbed Parametrisation Tendencies (SPPT)

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

ECMWF

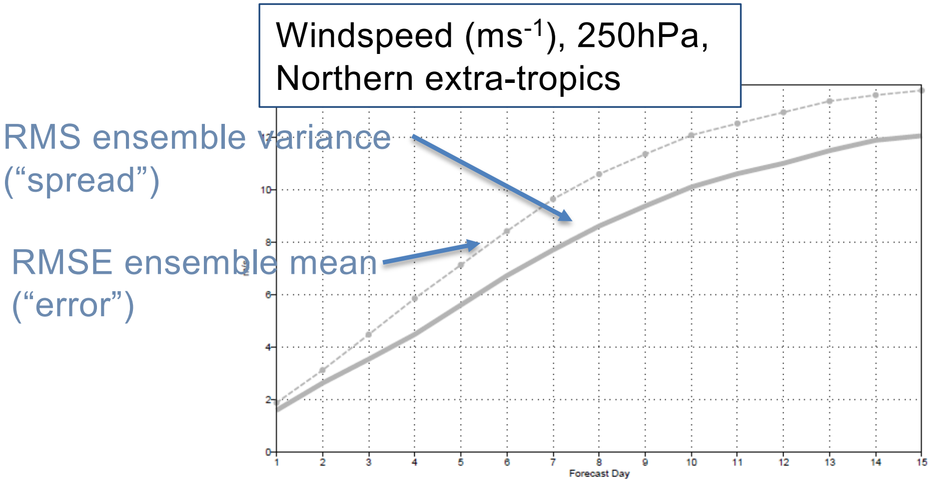
Sources of uncertainty: accounting for model uncertainty

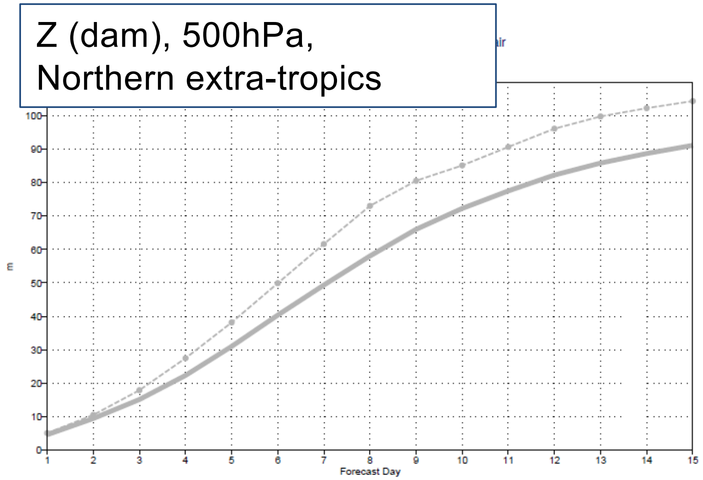


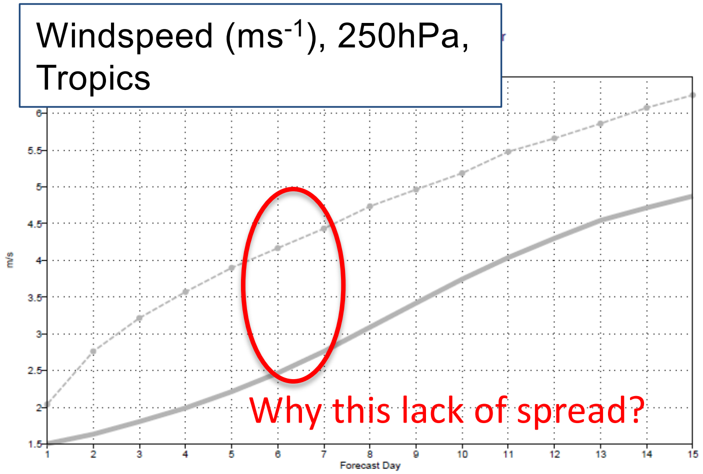
ECMWF

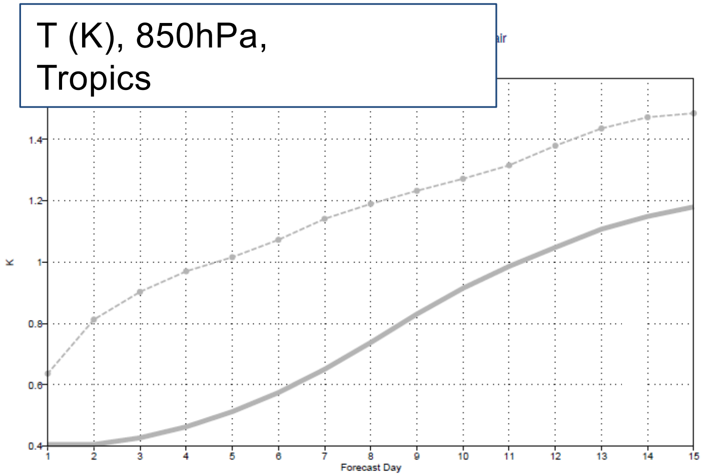
Recall: Ensemble forecasts: with initial conditions perturbations (IP) only

Ensemble mean RMSE (“Error”) & standard deviation (“Spread”)







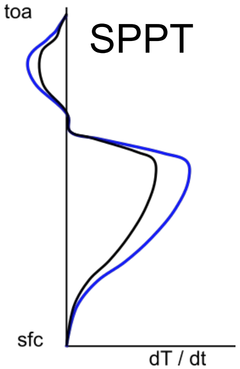


CY47R3

TCo399L137, dt=1200s

30 dates (Dec 2019)

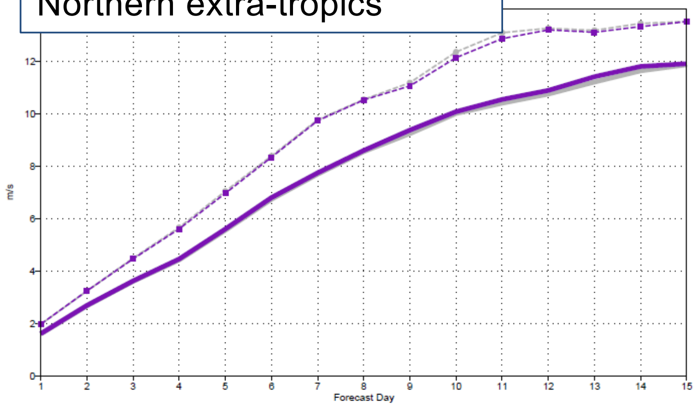
8 perturbed fcs



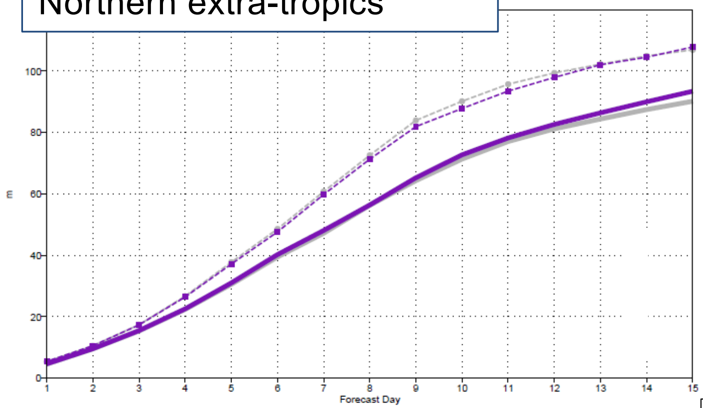
Ensemble forecasts: with **grid-scale** model uncertainty perturbations (SPPT)

Ensemble mean RMSE (“Error”) & standard deviation (“Spread”)

Windspeed (ms^{-1}), 250hPa, Northern extra-tropics



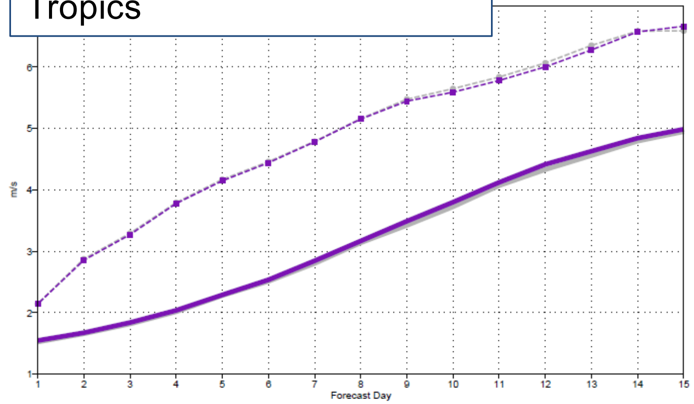
Z (dam), 500hPa, Northern extra-tropics



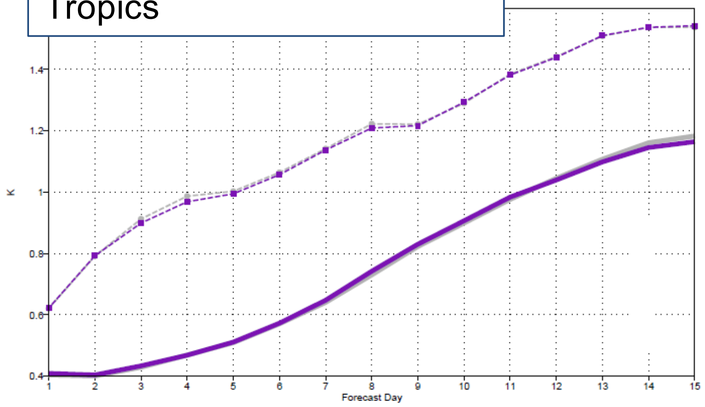
IP only

IP + SPPT*
(*white noise wrt time/horizontal)

Windspeed (ms^{-1}), 250hPa, Tropics



T (K), 850hPa, Tropics



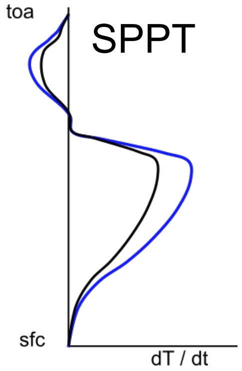
Uncorrelated noise yields little benefit

CY47R3

TCo399L137, dt=1200s

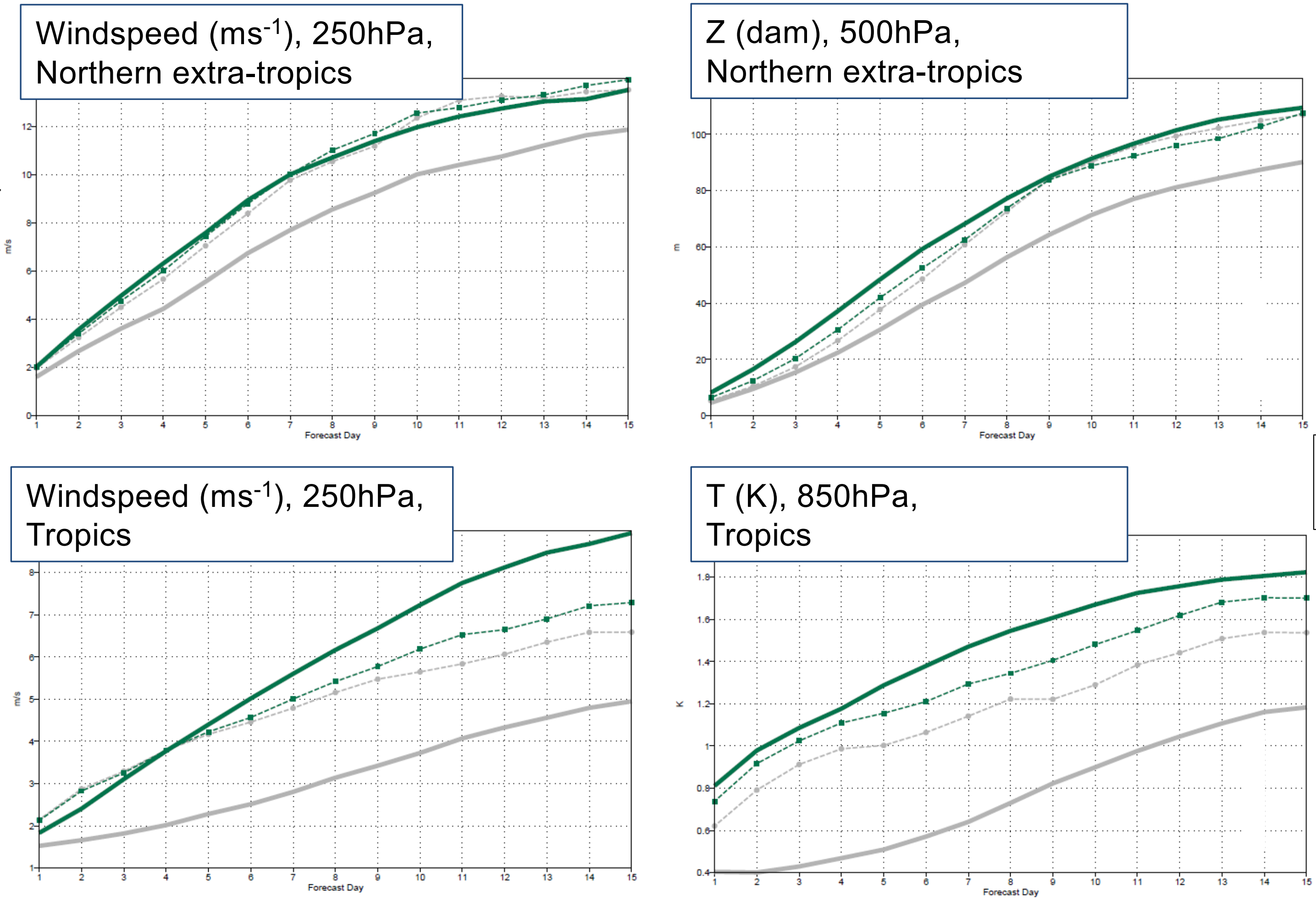
11 dates (Dec 2019/Jan 2020)

8 perturbed fcs



Ensemble forecasts: with **fixed** model uncertainty perturbations (SPPT)

Ensemble mean RMSE (“Error”) & standard deviation (“Spread”)



IP only

IP + SPPT*
(*fixed perts wrt time/horizontal)

Fixed perturbations yield increased errors

CY47R3

TCo399L137, dt=1200s

11 dates (Dec 2019/Jan 2020)

8 perturbed fcs

Stochastically Perturbed Parametrisation Tendencies (SPPT) scheme

- History (IFS): implemented, 1998 (**Buizza et al., 1999**); revised, 2009 (**Palmer et al., 2009**), 2019 (**Lock et al., 2019**):
- Simulates model uncertainty due to ***physics parameterisations*** by
  - taking the net tendencies from the physics parametrisations (excl. clear-sky heating rates):

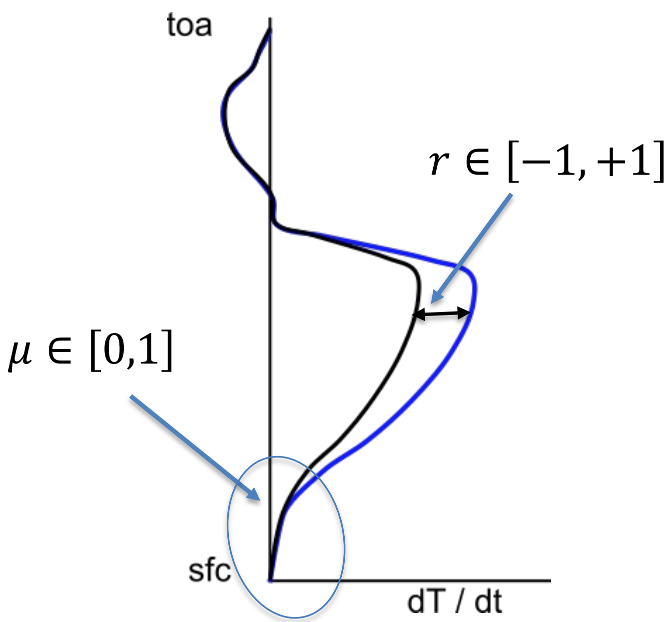
$$\boldsymbol{X} = [X_U, X_V, X_T, X_Q]$$

coming from [*radiation (cloudy-skies)*, *gravity wave drag*, *vertical mixing*, *convection*, *cloud physics*] 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.

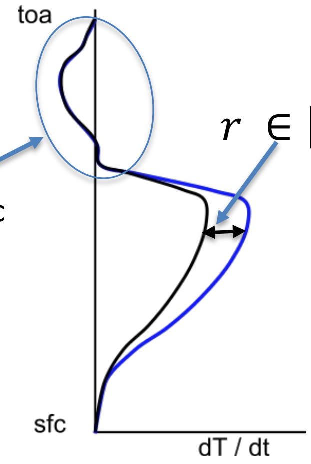


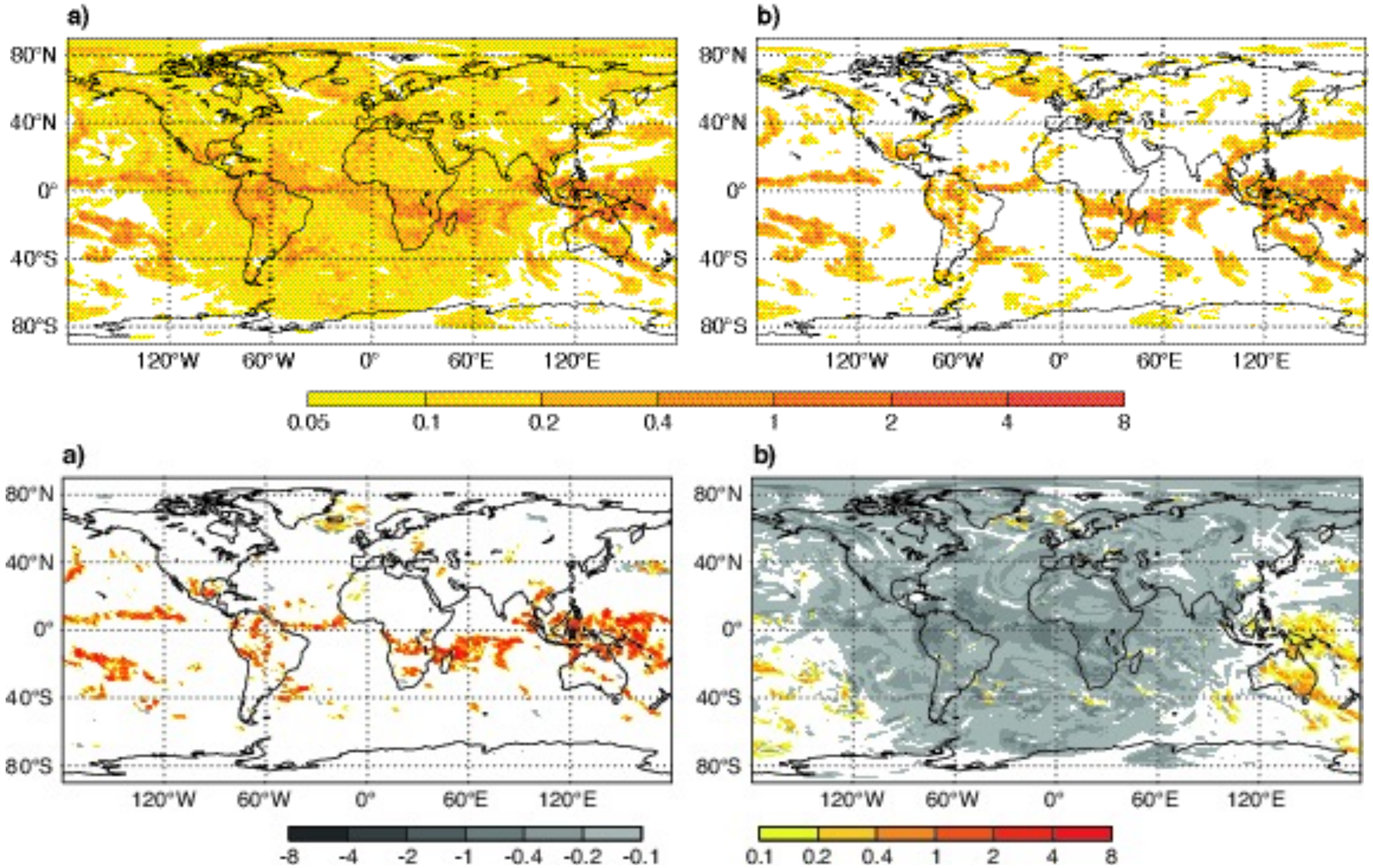
Shutts et al. (2011, ECMWF Newsletter); Palmer et al., (2009, ECMWF Tech. Memo.); Lock et al., (2019, QJRMS)

SPPT perturbations: why exclude clear-sky heating rates?

- SPPT perturbs net physics tendencies (***excluding clear-sky heating rates***) with multiplicative noise $r \in [-1, +1]$ as:

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





T tendencies from a model level in mid-troposphere accumulated during t+0-3h (K/3h):

Top: Ensemble stdev with SPPT perturbations with (a) clear-sky HRs (a) included & (b) excluded.

Bottom: From control forecast, from (a) convection & (b) radiation schemes

Figure 2 & Figure 1,from Lock et al. (2019, QJRMS)

SPPT random 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 around the globe) for each wavenumber

define ρ for random numbers, η

- Resulting pattern mapped into grid-point space r:

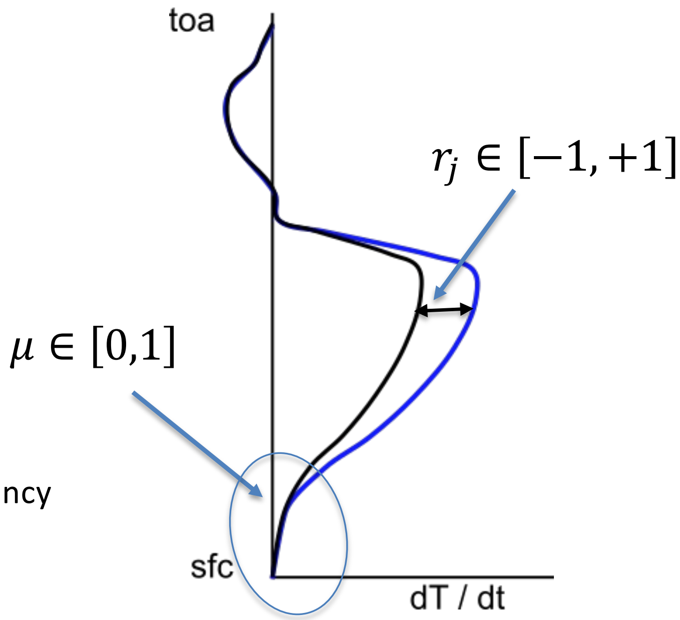
– clipped such that $r \in [-1, +1]$ --- prevents perturbation ***changing the sign*** of the tendency

– same pattern is applied to T, Q, U, V (***excluding clear-sky heating rates from radiation***)

– applied at all model levels to preserve vertical structures**

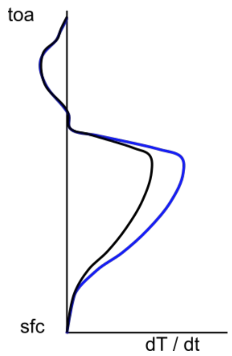
– ***Except*: tapered to zero at model bottom, to avoid:

  - excessive spread in the boundary layer caused by applying perturbations to large wind tendencies.



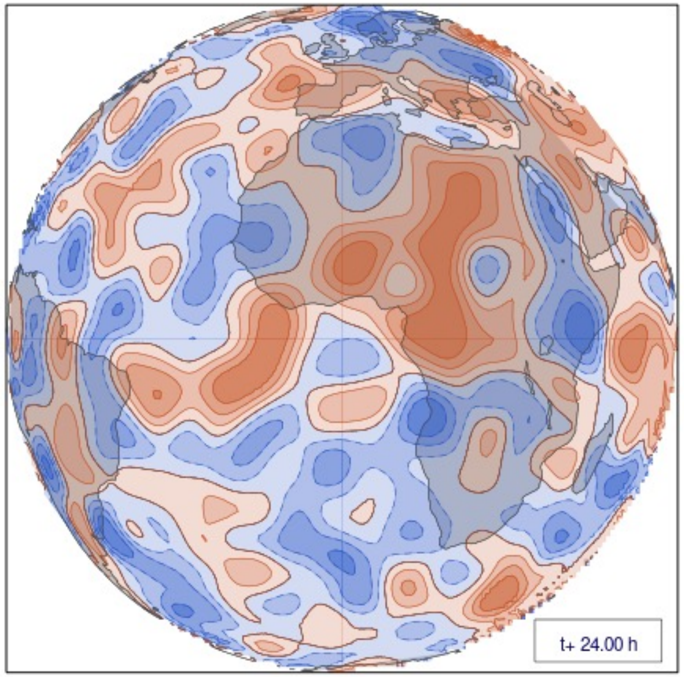
SPPT random pattern

- 2D random pattern, r:
  - Time-correlations: AR(1)
  - Spatial-correlations: Gaussian shape around the globe
  - Clipped such that $r \in [-1, +1]$



- Applied at all model levels to preserve vertical structures**

  ***Except*: tapered to zero at model bottom

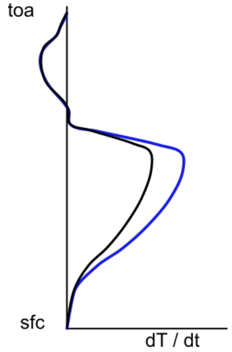


Example random pattern:

- Perturbed member, number 1
- Pattern at t = 24h
- Colours: blues = [-1,0), reds = (0,1]

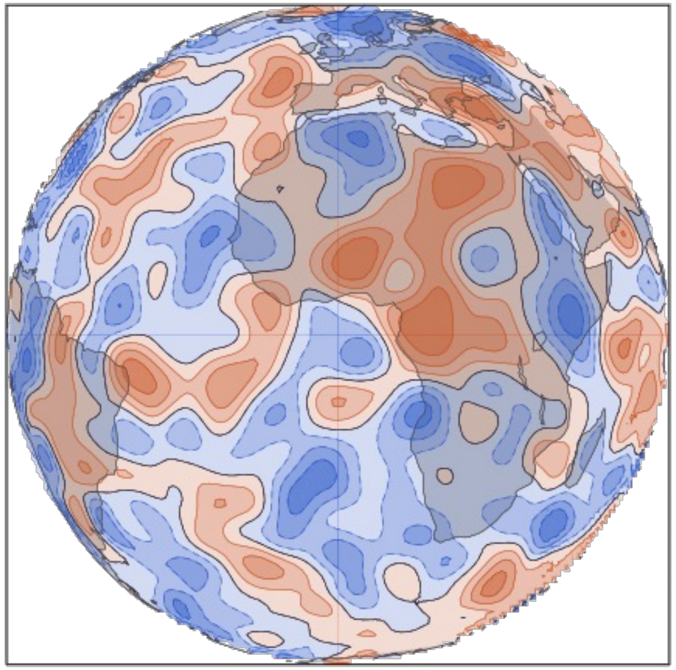
SPPT random pattern

- 2D random pattern, r:
– Time-correlations: AR(1)
– Spatial-correlations: Gaussian shape around the globe
– Clipped such that $r \in [-1, +1]$



- Applied at all model levels to preserve vertical structures**

  ***Except*: tapered to zero at model bottom

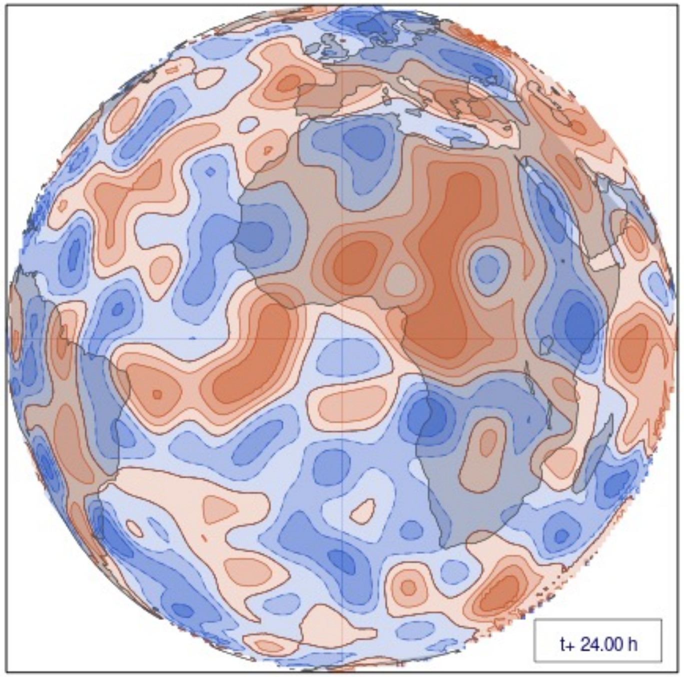


Example random pattern:

- Perturbed member, number 1
- Pattern at t = 0 .. 48h (dt = 15 min)
- Colours: blues = [-1,0), reds = (0,1]

SPPT random pattern

- 2D random pattern, r:
  – Time-correlations: AR(1)
  – Spatial-correlations: Gaussian shape around the globe
  – Clipped such that $r \in [-1, +1]$



- Applied at all model levels to preserve vertical structures**

  ***Except*: tapered to zero at model bottom

- Multi-scale pattern:
  - 3 time/space scales
  - Shortest scales dominate
  - $\sigma_{3-scale} = 0.4453$

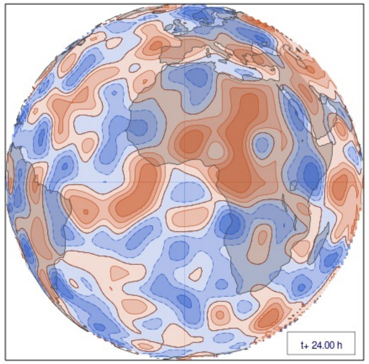
sppt1	6 hours,	500 km,	$\sigma = 0.42$
sppt2	3 days,	1 000 km,	$\sigma = 0.14$
sppt3	30 days,	2 000 km,	$\sigma = 0.048$

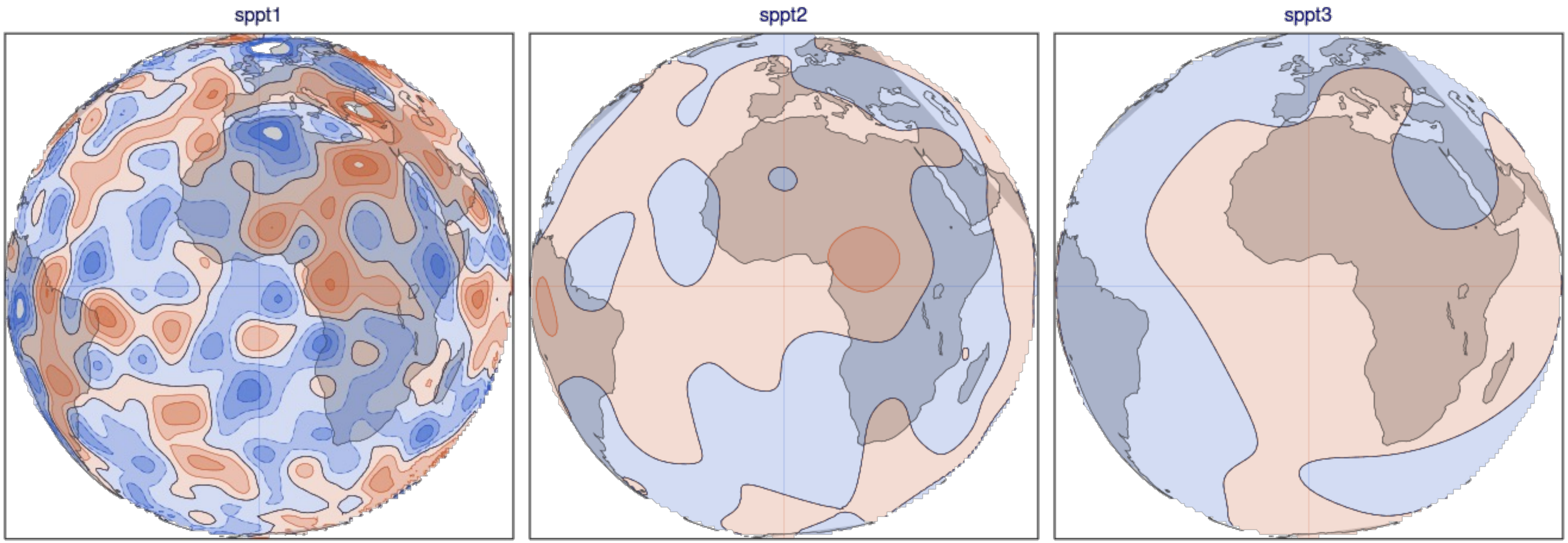
SPPT random pattern: multi-scale

sppt1	6 hours,	500 km,	$\sigma = 0.42$
sppt2	3 days,	1 000 km,	$\sigma = 0.14$
sppt3	30 days,	2 000 km,	$\sigma = 0.048$

Example random patterns:

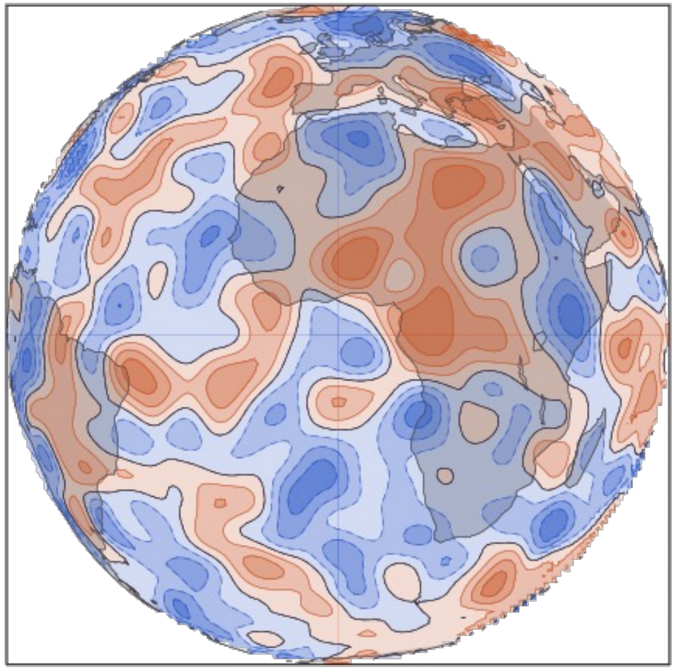
- Perturbed member, number 1
- Patterns at t = 0 .. 48h (dt = 15 min)
- Colours: blues = [-1,0), reds = (0,1]

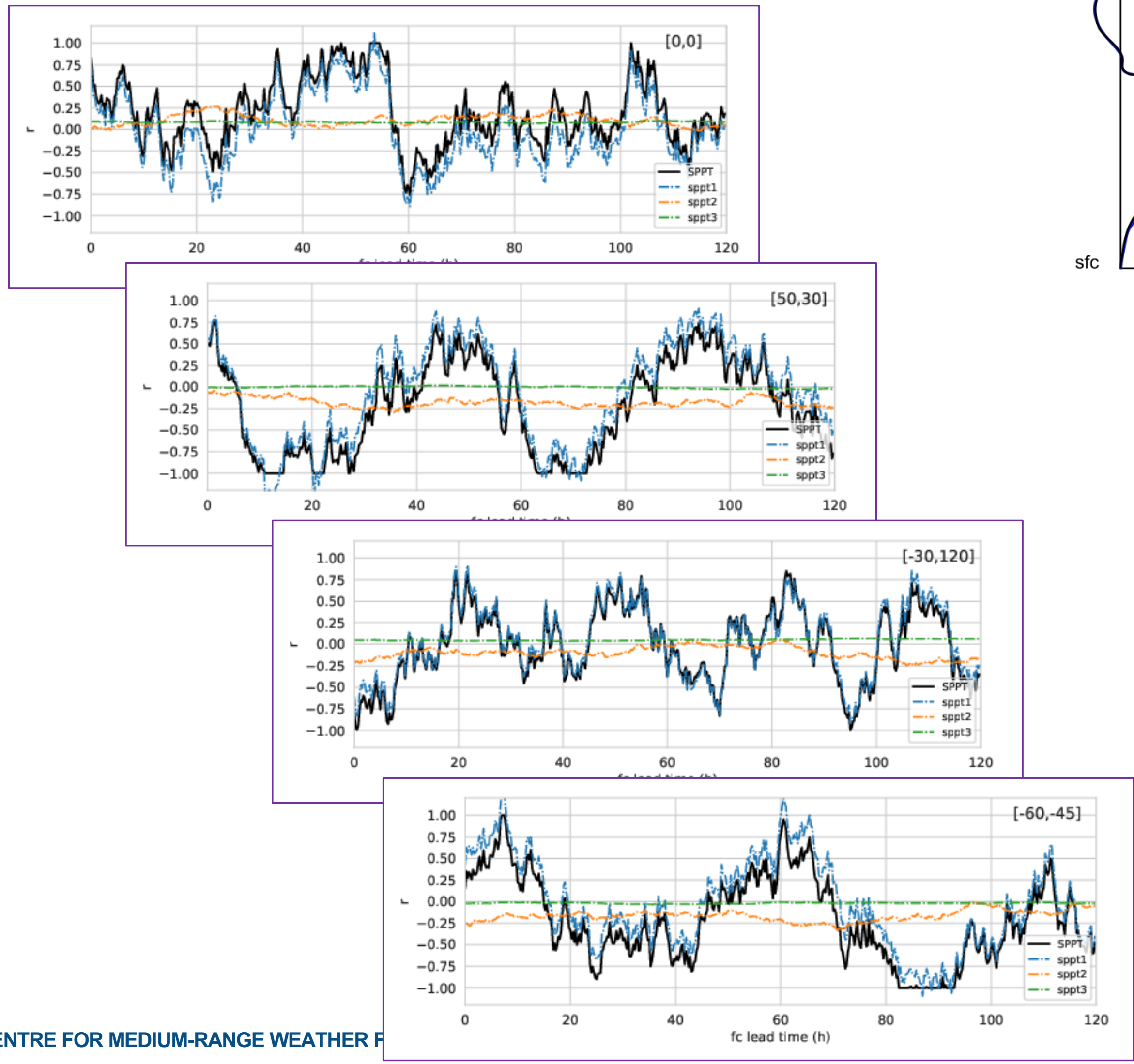


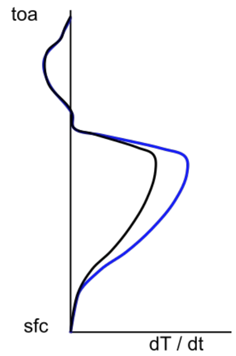


SPPT random pattern

sppt1	6 hours,	500 km,	$\sigma = 0.42$
sppt2	3 days,	1 000 km,	$\sigma = 0.14$
sppt3	30 days,	2 000 km,	$\sigma = 0.048$

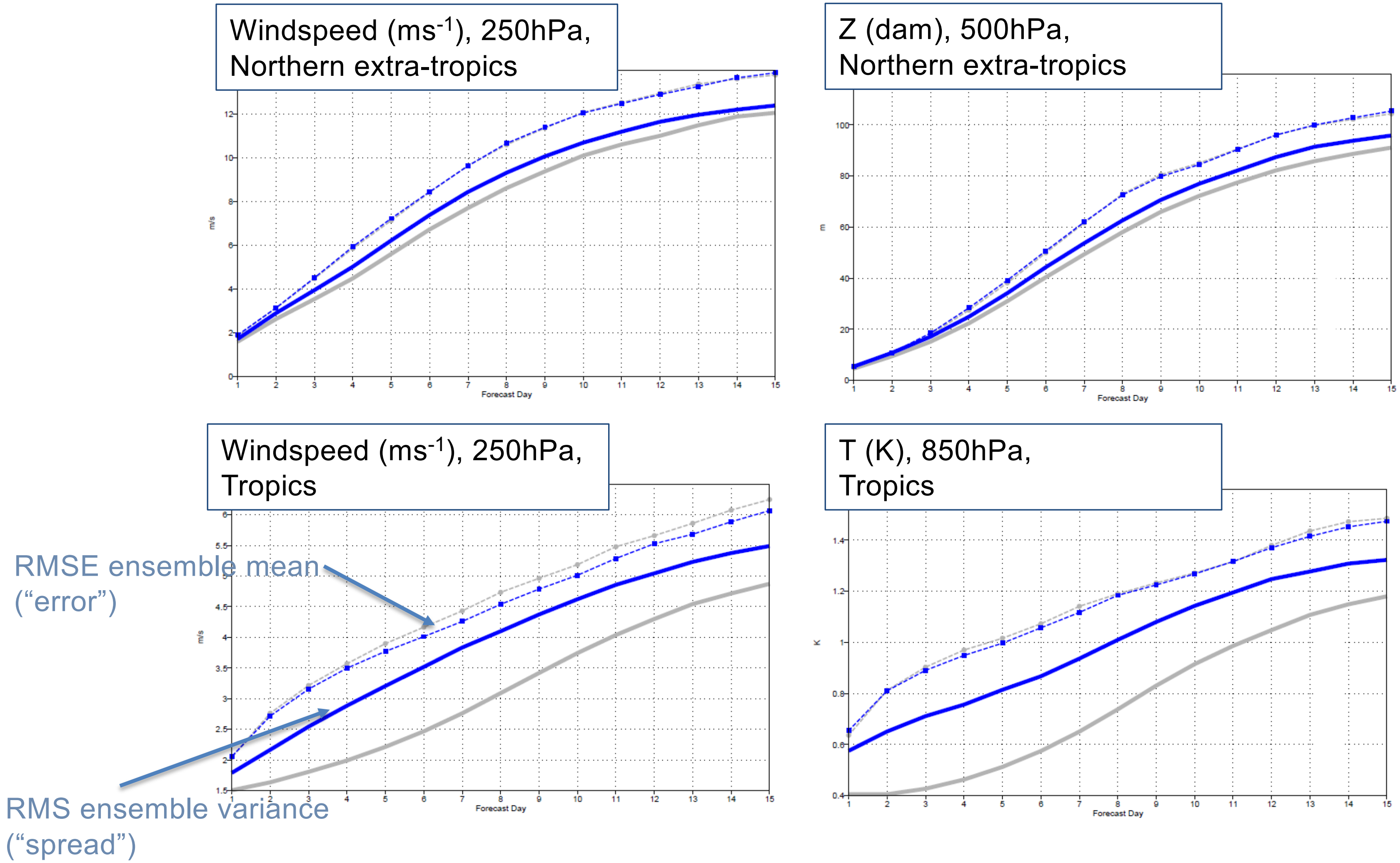






Ensemble forecasts: with **multi-scale** model uncertainty perturbations (SPPT)

Ensemble mean RMSE (“Error”) & standard deviation (“Spread”)



CY47R3

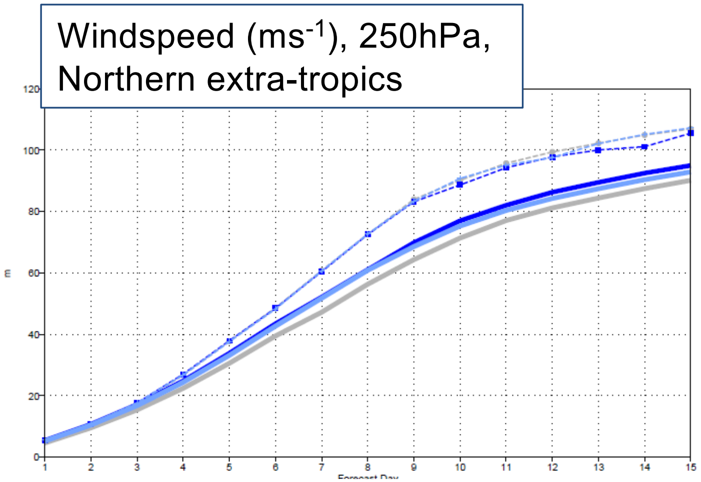
TCo399L137, dt=1200s

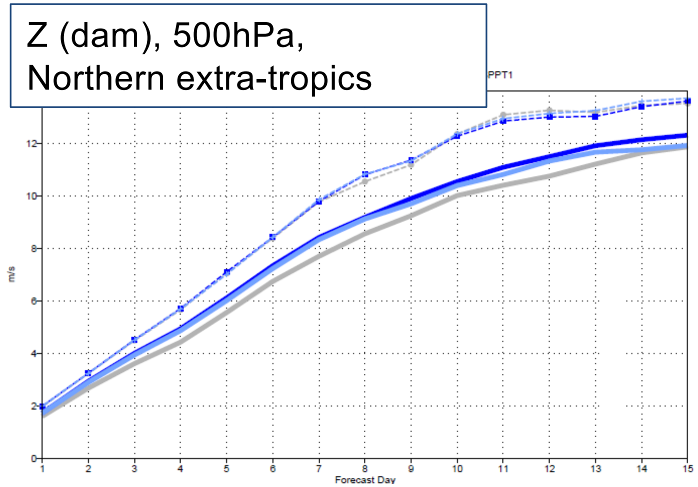
30 dates (Dec 2019)

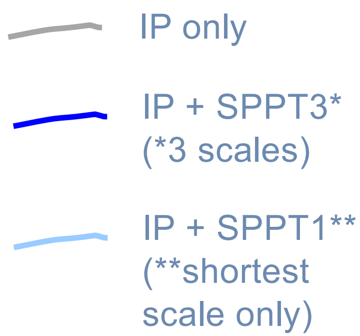
8 perturbed fcs

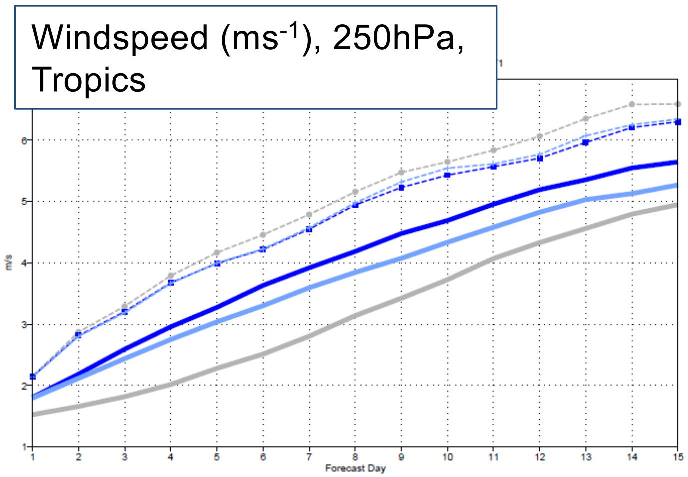
Ensemble forecasts: with **multi-scale** model uncertainty perturbations (SPPT)

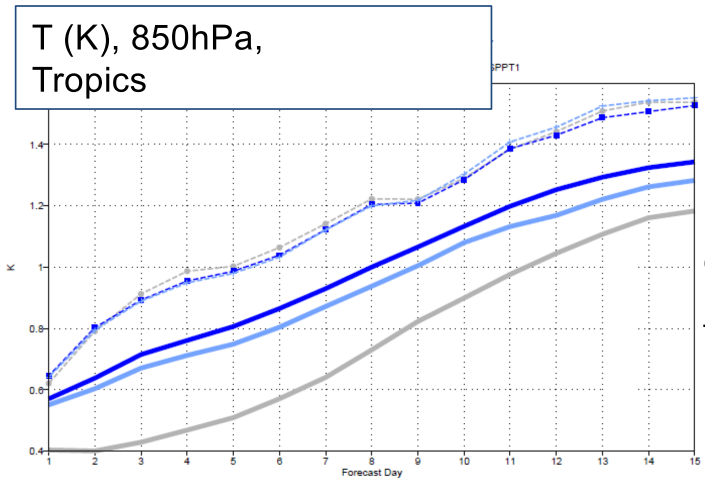
Ensemble mean RMSE ("Error") & standard deviation ("Spread")











Some additional spread from SPPT3
- 3rd scale deemed important for longer-range forecasts

CY47R3

TCo399L137, dt=1200s

11 dates (Dec 2019/Jan 2020)

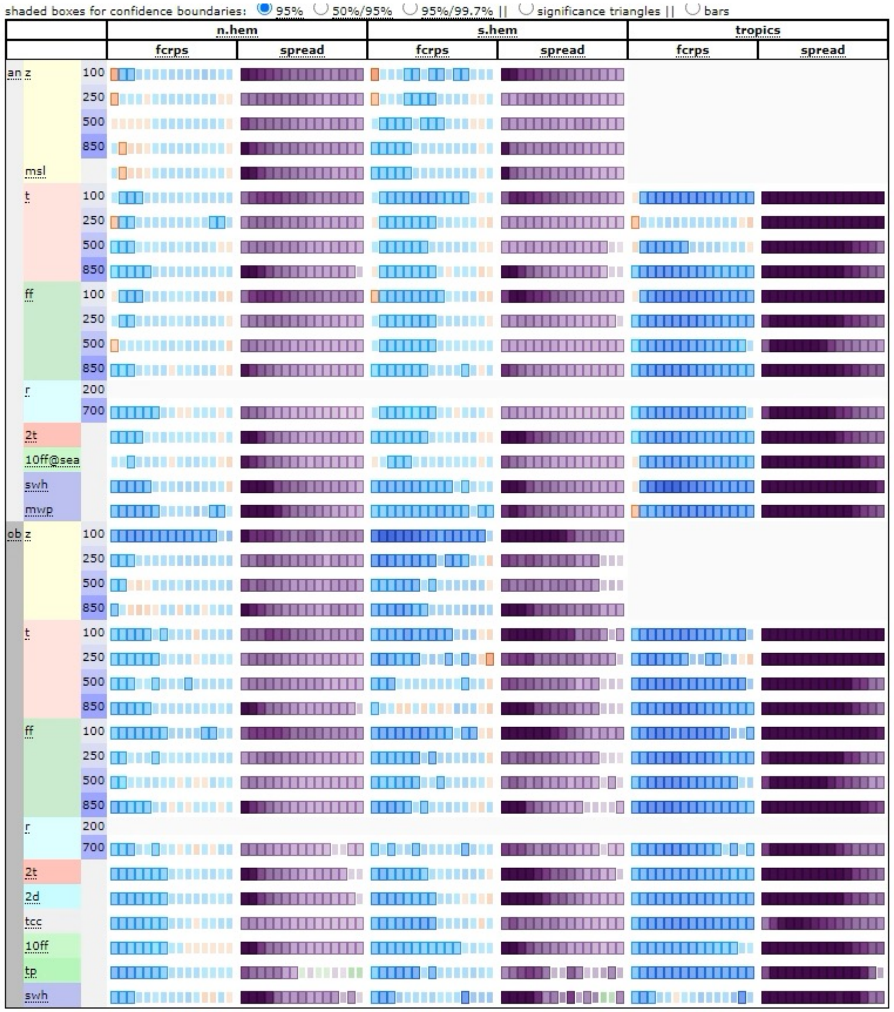
8 perturbed fcs

Ensemble forecasts: with **multi-scale** model uncertainty perturbations (SPPT)

Scorecard of probabilistic skill (“fCRPS”) & ensemble standard deviation (“Spread”)

verified against analysis

verified against observations



Scorecard (summary):

IP + SPPT3* (*3 scales) *versus* IP only

Spread:
Purple = more spread / Green = less spread

fCRPS:
Blue = more skillful / Red less skillful

Framed cell indicates statistically significant differences at the 95% confidence interval

CY47R3

TCo399L137, dt=1200s

30 dates (Dec 2019)

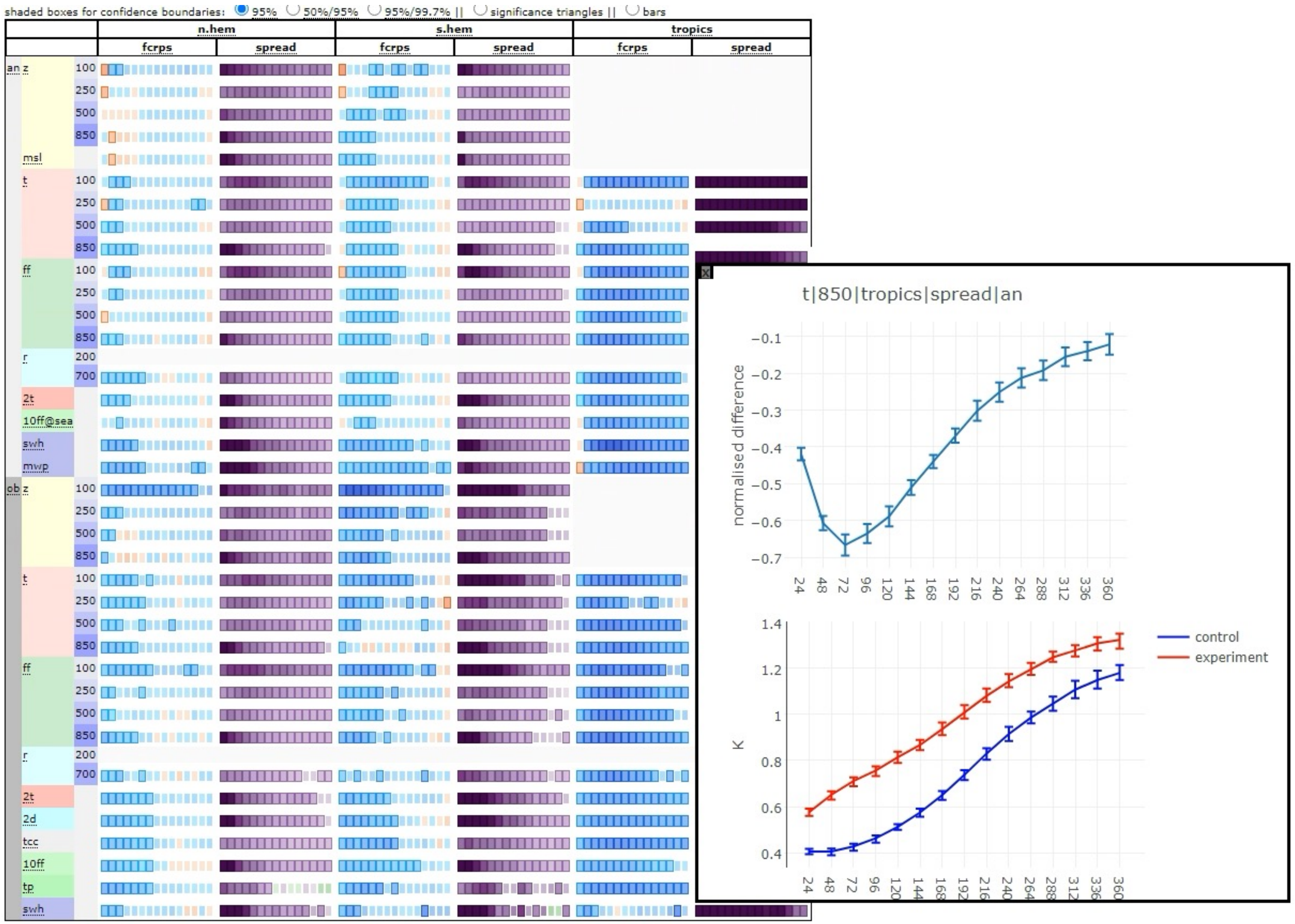
8 perturbed fcs

Ensemble forecasts: with **multi-scale** model uncertainty perturbations (SPPT)

Scorecard of probabilistic skill (“fCRPS”) & ensemble standard deviation (“Spread”)

verified against analysis

verified against observations



CY47R3

TCo399L137, dt=1200s

30 dates (Dec 2019)

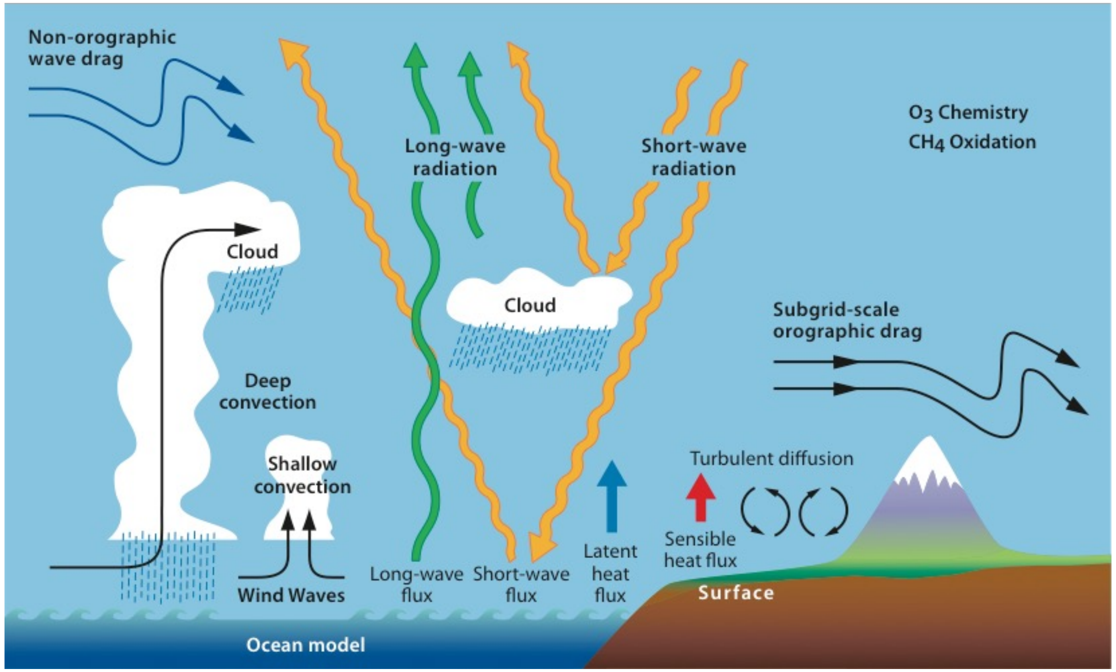
8 perturbed fcs

Summary: stochastic representation of model uncertainty in IFS

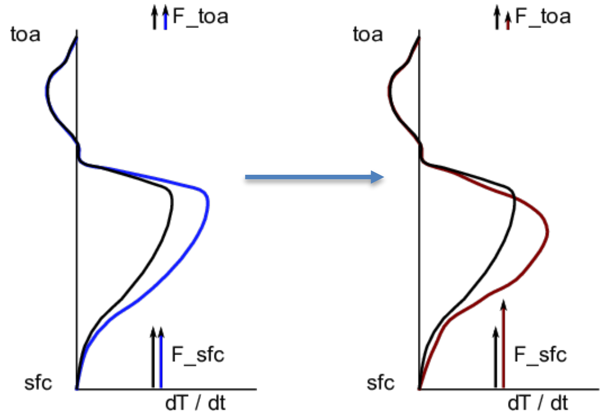
- Model uncertainty (MU) due to unresolved and misrepresented processes
- Without representing MU, ensemble forecasts are under-dispersive => over-confident
- Stochastic representations of model uncertainty can improve ensemble reliability
- SPPT: represents uncertainty due to sub-grid atmospheric physics parameterisations
  - Medium-range: increased ensemble spread, greater probabilistic skill
  - Seasonal: reduction in biases; better representation of MJO, ENSO, PNA regimes (Weisheimer et al., 2014, Phil. Trans. R. Soc. A)
- *Difficult to characterise sources of model uncertainty due to their small scales*

Stochastic representations of model uncertainty: **outlook for IFS**

Towards process-level model uncertainty representation

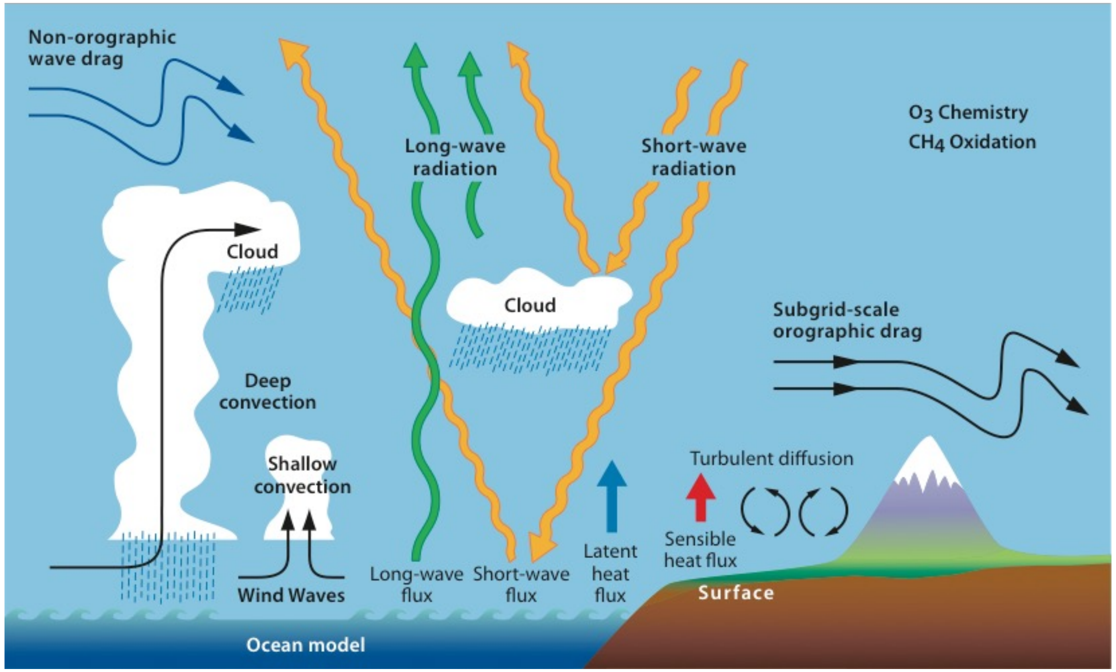


- **Aim**: to improve the physical consistency
- Local conservation of moisture, momentum, energy
- Generate flux perturbations at the top of atmosphere (TOA) and surface that are consistent with tendency perturbations within the atmospheric column
- Remove ad hoc tapering in boundary layer
- Include multi-variate aspects of uncertainties



Stochastic physics: outlook for IFS

Towards process-level model uncertainty representation



Stochastically Perturbed Parametrisations (SPP)

(Lang et al., 2021, QJRMS; Ollinaho et al., 2017, QJRMS)

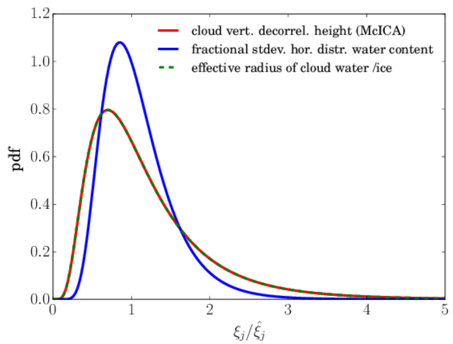
- Embed stochasticity inside IFS parametrisations
- Perturb parameters/variables directly
- Specify spatial/temporal correlations
- Target uncertainties that matter (level of uncertainty and impact)
- Require that stochastic schemes converge to deterministic schemes in limit of vanishing variance

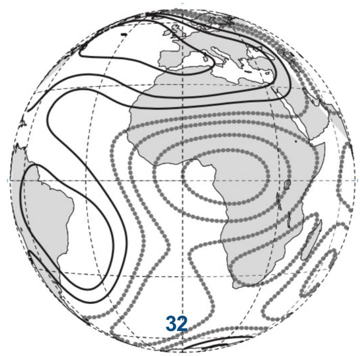
Stochastically perturb parameters/variables in the physics parametrisations ($\hat{\xi}_j$):

$$\xi_j = \hat{\xi}_j \exp(\Psi_j)$$

where

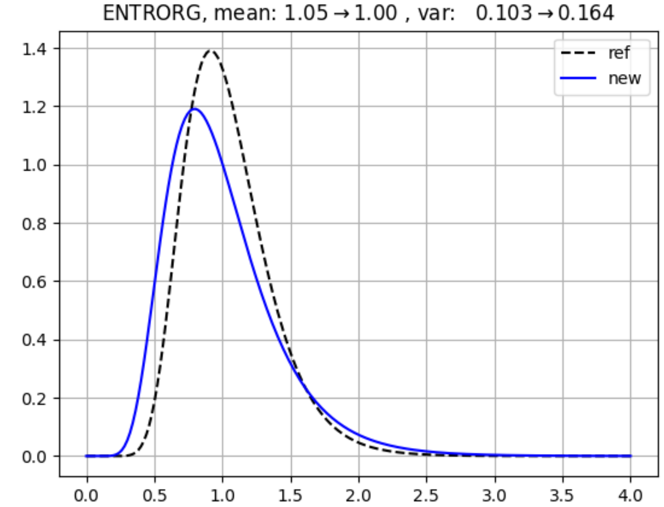
$$\Psi_j \sim \mathcal{N}(\mu_j, \sigma_j^2)$$



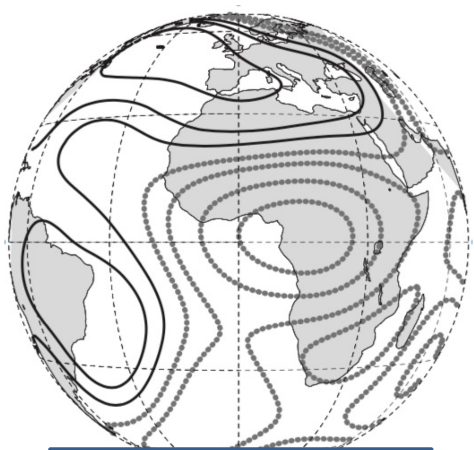


Revisions to SPP

- Work on a revision of SPP completed: Lang et al (2021), https://doi.org/10.1002/qj.3978

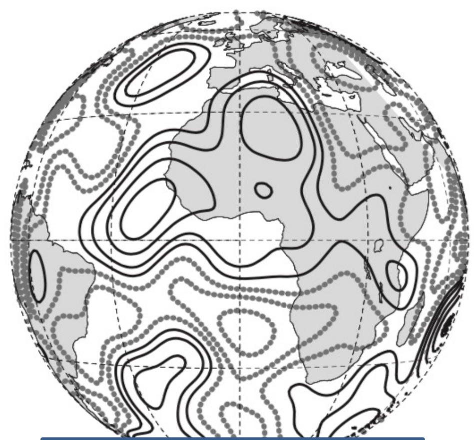


- Summary of changes
  - probability distributions (mean and variance)
  - correlation scale
  - additional perturbed quantities (total 27):
    - Cloud scheme
    - Convection scheme
    - BL scheme



2000 km

→

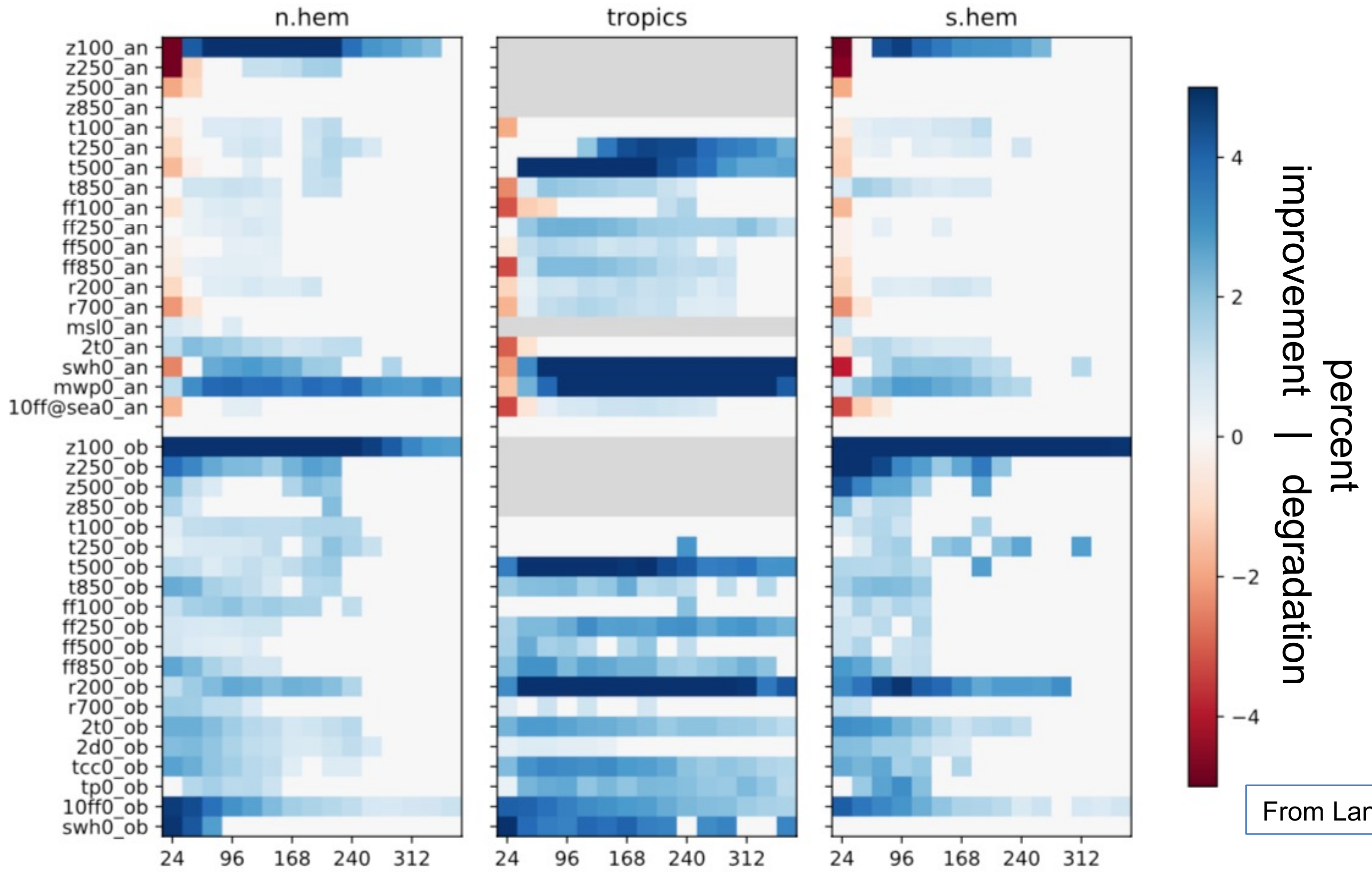


1000 km

SPP-new versus SPP-ref scorecard showing fCRPS changes

verified against analysis

verified against observations

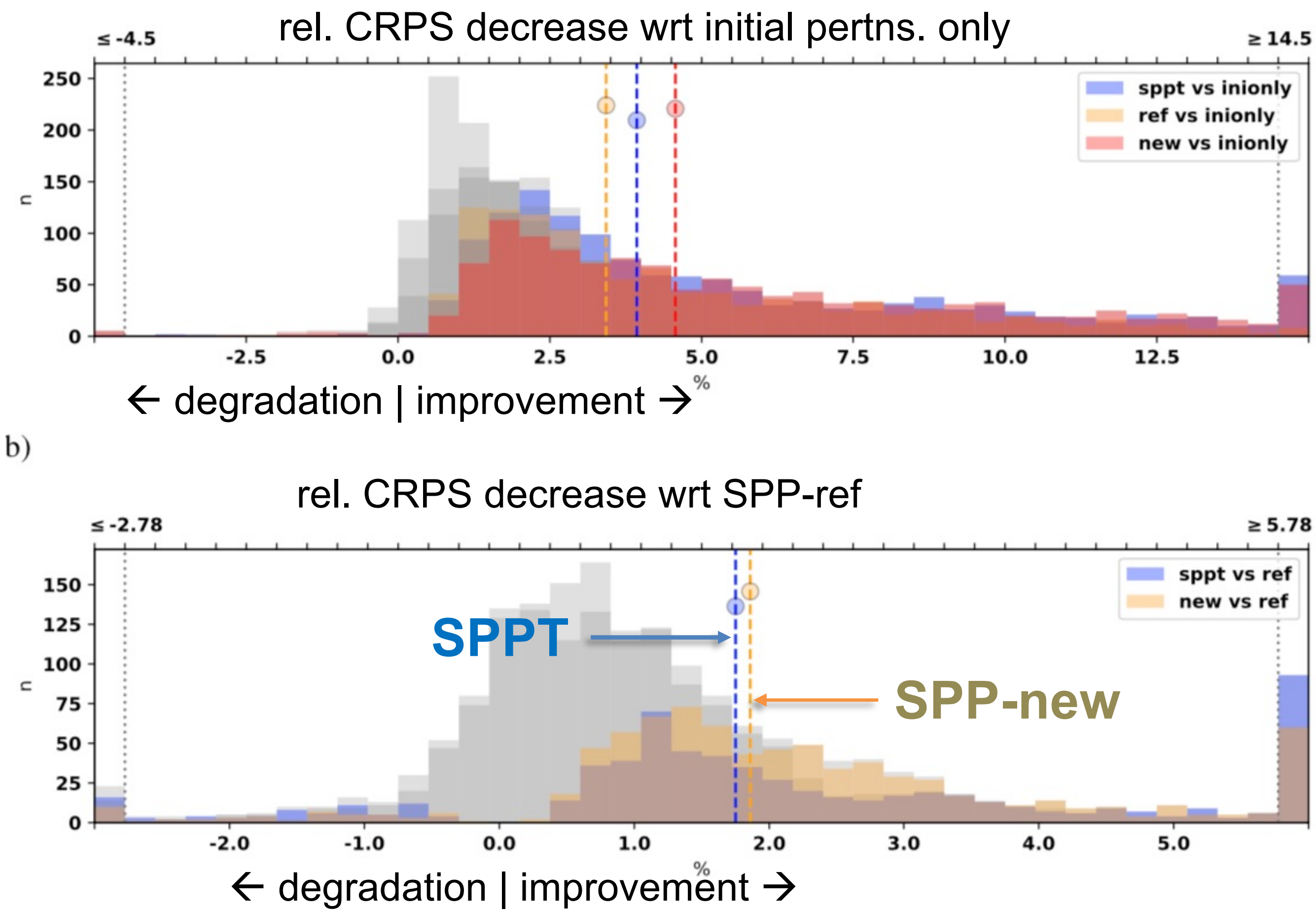


From Lang et al. (2021)

lead time (h)

Coloured: 99.7% sig. level

Histogram of relative CRPS changes: SPPT, SPP



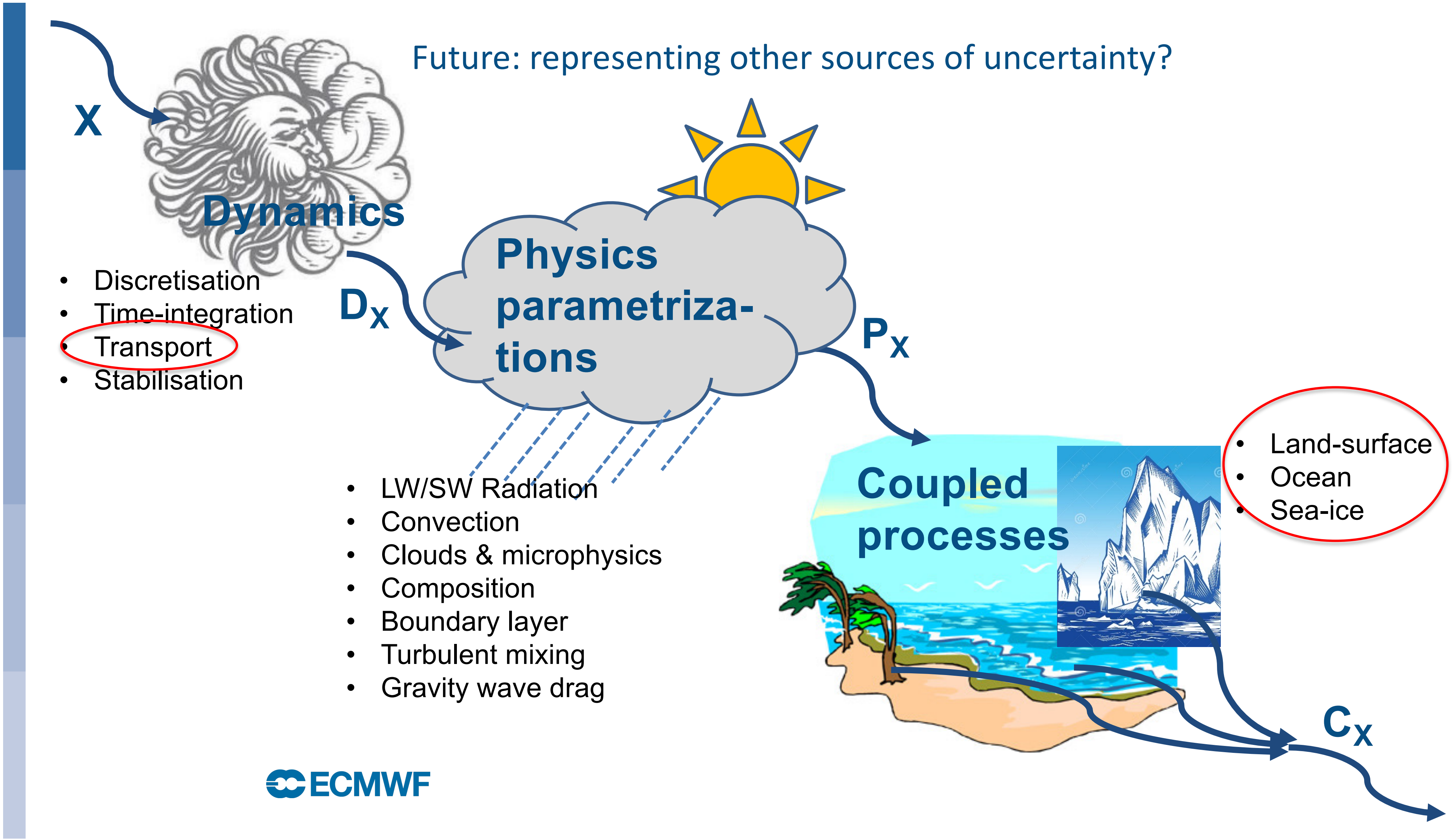
lead times 24 h ... 360 h
combination of variables and levels evaluated in scorecard
3 regions: N-Hem, S-Hem, tropics

vertical lines: median change

based on 8 members and fair CRPS, boreal summer + boreal winter, 212 start dates, TCo399

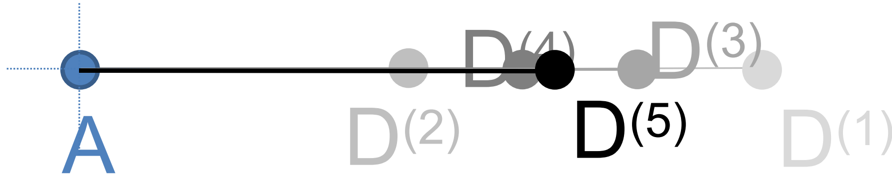
colour: stat. significance 99.7%, grey otherwise

Enters operations in CY49R1 (2024)!


Future: representing other sources of uncertainty?
X
Dynamics
• Discretisation
• Time-integration
• Transport
• Stabilisation
Dx
Physics parametriza-tions
• LW/SW Radiation
• Convection
• Clouds & microphysics
• Composition
• Boundary layer
• Turbulent mixing
• Gravity wave drag
Px
Coupled processes
• Land-surface
• Ocean
• Sea-ice
Cx
ECMWF

STOCHDP:

Stochastically perturbed semi-Lagrangian (SL) departure point (DP) estimates



Diamantakis & Magnusson (2016):

- Explored convergence rate of the iterative DP estimate
- Slowest convergence ←→ most complex flow (strong shear / curvature)
- Example: Typhoon Neoguri:
  - HRES forecast: initialised: 2014-07-05, 00UTC

Fig. 1c: t+96h, 850hPa windspeeds

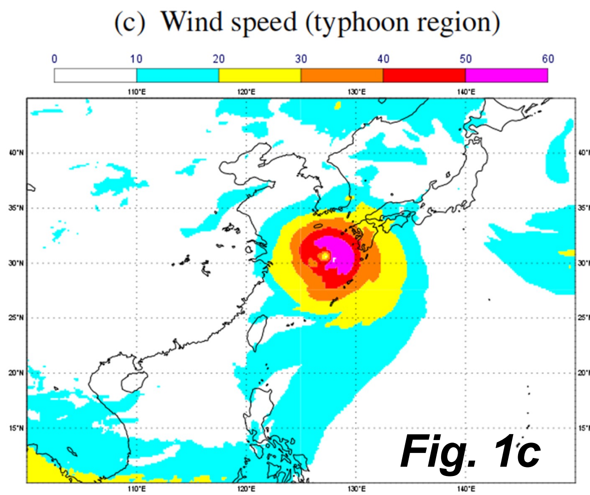
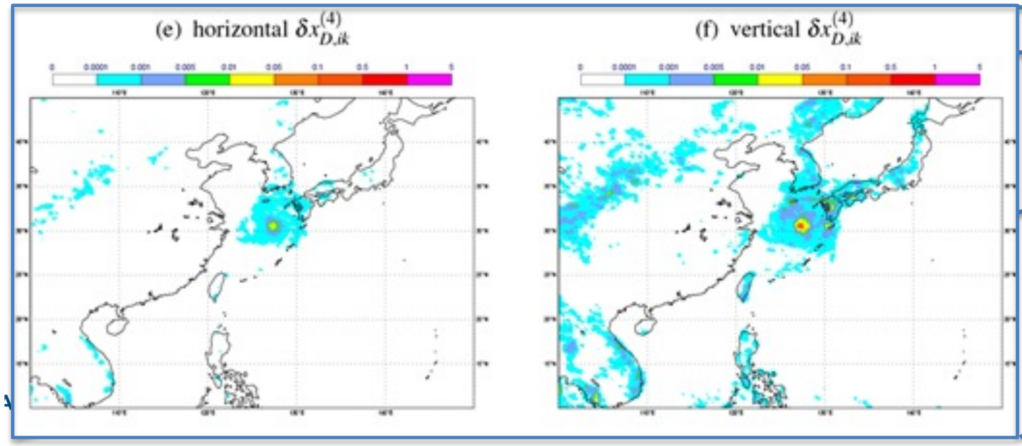
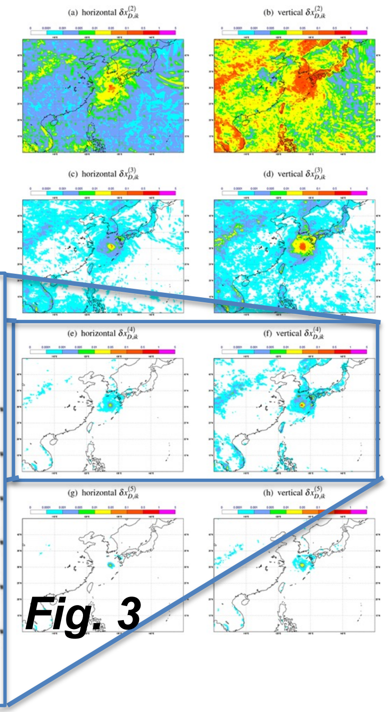


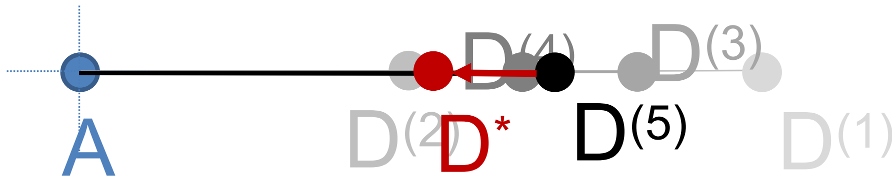
Figure 3: difference in DP estimate between consecutive iterations (scaled)





STOCHDP:

Stochastically perturbed semi-Lagrangian (SL) departure point (DP) estimates



Model uncertainty scheme, "STOCHDP":

- use the DP estimate convergence rate to attribute MU:

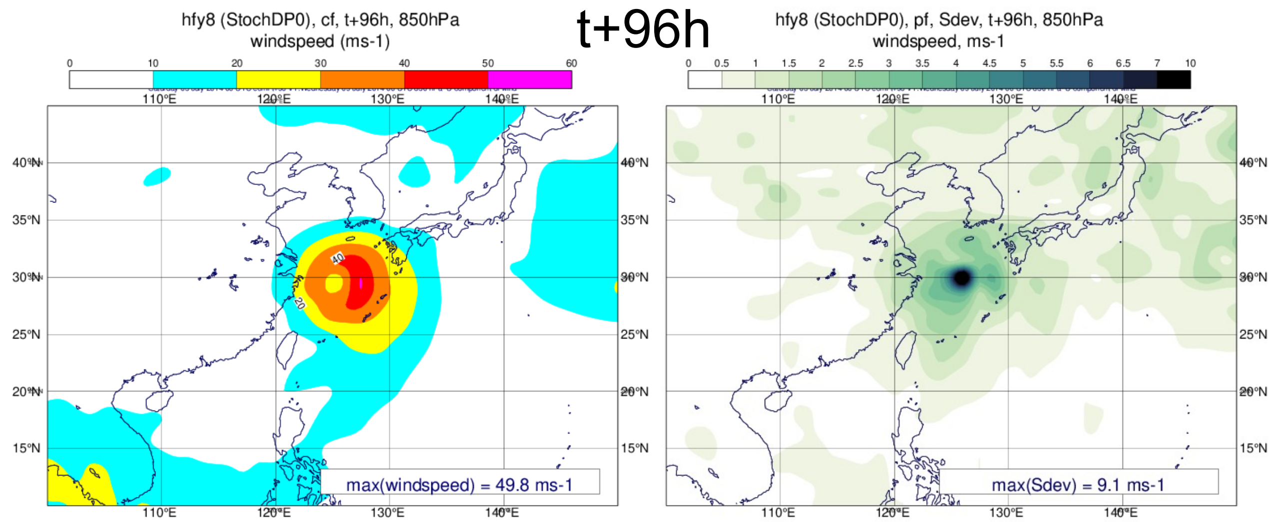
$$D^* = D^{(5)} + r\left(D^{(5)} - D^{(5-i)}\right), i = 1..4$$

where $D*$ is the perturbed DP and r is a random number

➢ STOCHDP represents MU from SL advective winds

Early results, e.g.:

- Typhoon Neoguri case
- ENS: STOCHDP only
- TCo639L91, dt=720s
- 20+1 members
- Peak ENS stdev develops and tracks with TC



Control forecast

Ensemble stdev

ECMWF EUROPE

Summary

- Including a representation of model uncertainty can improve the reliability of ensemble forecasts
- ”Model uncertainty” describes inaccuracies due to the model integrations
- Using stochastic physics schemes enables representation of the model uncertainty arising from the parametrization of unresolved atmospheric physics
- Current stochastic physics scheme used in the IFS: SPPT
- Outlook: new scheme ”SPP” improves the physical consistency of the stochastic physics perturbations
- Ongoing: exploring stochastic perturbations to represent model uncertainty in the dynamics – STOCHDP

References

- Buizza, R., et al. (1999), Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. QJRMS, 125: 2887-2908. https://doi.org/10.1002/qj.49712556006
- Diamantakis, M., & Magnusson, L. (2016). Sensitivity of the ECMWF Model to Semi-Lagrangian Departure Point Iterations, MWR, *144*(9), 3233-3250. https://doi.org/10.1175/MWR-D-15-0432.1
- Lang, STK, et al. (2021), Revision of the Stochastically Perturbed Parametrisations model uncertainty scheme in the Integrated Forecasting System. QJRMS, 147: 1364–1381. https://doi.org/10.1002/qj.3978
- Lock, S-J, et al. (2019), Treatment of model uncertainty from radiation by the Stochastically Perturbed Parametrization Tendencies (SPPT) scheme and associated revisions in the ECMWF ensembles. QJRMS, 145 (Suppl. 1): 75- 89. https://doi.org/10.1002/qj.3570
- Ollinaho, P., et al. (2017), Towards process-level representation of model uncertainties: stochastically perturbed parametrizations in the ECMWF ensemble. QJRMS, 143: 408-422. https://doi.org/10.1002/qj.2931
- Palmer et al., 2009: Stochastic parametrization and Model Uncertainty, ECMWF Tech. Mem., **598**, pp. 42
- Weisheimer A., et al. (2014), Addressing model error through atmospheric stochastic physical parametrizations: impact on the coupled ECMWF seasonal forecasting system. *Phil. Trans. R. Soc. A.* **372:** 20130290. https://doi.org/10.1098/rsta.2013.0290

Further reading & upcoming workshop

In 2016, we undertook an extensive review of existing and future efforts in model uncertainty representation – as a Special Topic paper for our Scientific Advisory Committee

Report covers:

- Literature review of model uncertainty work
- Descriptions/discussions of SPPT / SKEB / SPP schemes
- Impacts of the schemes in the IFS (EDA; short / medium / extended / longer forecast ranges)
- Proposals for future directions – improvements to SPPT; extensions to SPP; new approaches

Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S.T.K., et al. (2017), Stochastic representations of model uncertainties at ECMWF: state of the art and future vision. Q.J.R. Meteorol. Soc, 143: 2315-2339. https://doi.org/10.1002/qj.3094