Investigating satellite data assimilation in an idealised framework using an EnKF

Luca Cantarello(1), Onno Bokhove(1), Steve Tobias(1), Gordon Inverarity(2), Stefano Migliorini(2)

3rd June 2019

(1) School of Mathematics, University of Leeds, Leeds, United Kingdom
(2) Met Office, Exeter, United Kingdom

14th international EnKF workshop, Voss (Norway), 3rd-5th June 2019

Researching satellite data assimilation

- Satellite observations are an essential ingredient in current data assimilation systems.
- They have greatly contributed to the improvement of weather forecasts over time.
- New and more precise instruments boarded on satellites are added every year to the observing system.

Aim: utilising an **idealised model** to help investigate the **impact of satellite observations** in a DA system: what is the relative impact of **large-scale or small-scale** observations? What should we focus on?

Why an idealised model?

- Idealised/simplified models have two key strengths:
  - They can capture fundamental aspects and processes;
  - They are inexpensive and easy to run.

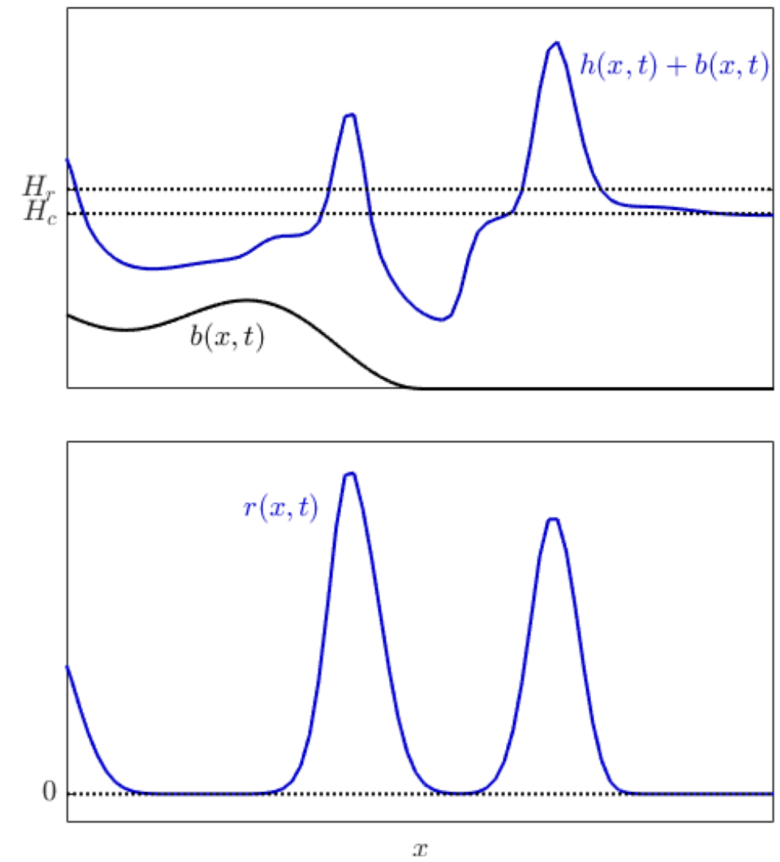
Previous work at University of Leeds, i.e. Kent et al. (2017): modified shallow water model based on the model of Würsch and Craig (2014) **suitable for data assimilation research** [1]

[1] Kent, T. et al (2017): *Dynamics of an idealized fluid model for investigating convective-scale data assimilation*. Tellus A: Dynamic Meteorology and Oceanography, 69(1), 1369332.

The modRSW model

A modified 1.5D single-layer rotating shallow water model which:

- mimics convection updrafts;
- represents idealised rain;
- includes switches.



$$h_t + (hu)_x = 0,$$

$$(hu)_t + (hu^2 + P(h))_x + c_0^2 h r_x - f hv = 0,$$

$$(hv)_t + (huv)_x + f hu = 0,$$

$$(hr)_t + (hur)_x + \beta h u_x + \alpha hr = 0\,.$$

$$P(h) = \begin{cases} \frac{1}{2} g h^2 & \text{if } h < H_c, \\ \frac{1}{2} g H_c^2 & \text{if } h \geq H_c, \end{cases}$$

$$\beta = \begin{cases} \tilde{\beta} & \text{if } h \geq H_r,\ u_x < 0, \\ 0 & \text{otherwise}\,. \end{cases}$$

Black = classic shallow water w/rotation

Red = shallow water with convection/rain

DA with modRSW model

- Twin-setting experiments:
  - A nature run at high resolution representing the truth→ used to create pseudo-observations,
  - An ensemble of forecasts generated at coarser resolution.
- All variables observed directly at evenly spaced locations (trivial observation operator).
- Deterministic EnKF (Sakov and Oke, 2008) with self-exclusion + IAU for additive inflation accounting for model error + RTPS (Whitaker and Hamill, 2012) + Gaspari-Cohn localisation

This configuration seems promising in reproducing features of operational schemes (paper by T. Kent et al. in preparation)

Code on github: https://github.com/tkent198/modRSW_EnKF

Modelling satellite observations

Idealised satellite DA will require the generation of (synthetic) satellite observations.
Our focus is on **sounding observations**.

Satellite observations	Current modRSW setup	Revised modRSW setup
Radiance (via Brightness Temperature)	✔ (*)	✔
Vertical structure	✗ single-layer	✔
Spatially varying	✗ fixed in space	✔
Non-linear observation operator	✗ linear	✔

(*) scaling issue

Modifications

The modifications to the current modRSW setup will concern three aspects:

- The mathematical formulation;
- The way the synthetic observations are generated from the truth, separating satellite observations from ground observations;
- The observation operator $\mathcal{H}$ which maps the model state into the observational space.

The revised model

Two assumptions:

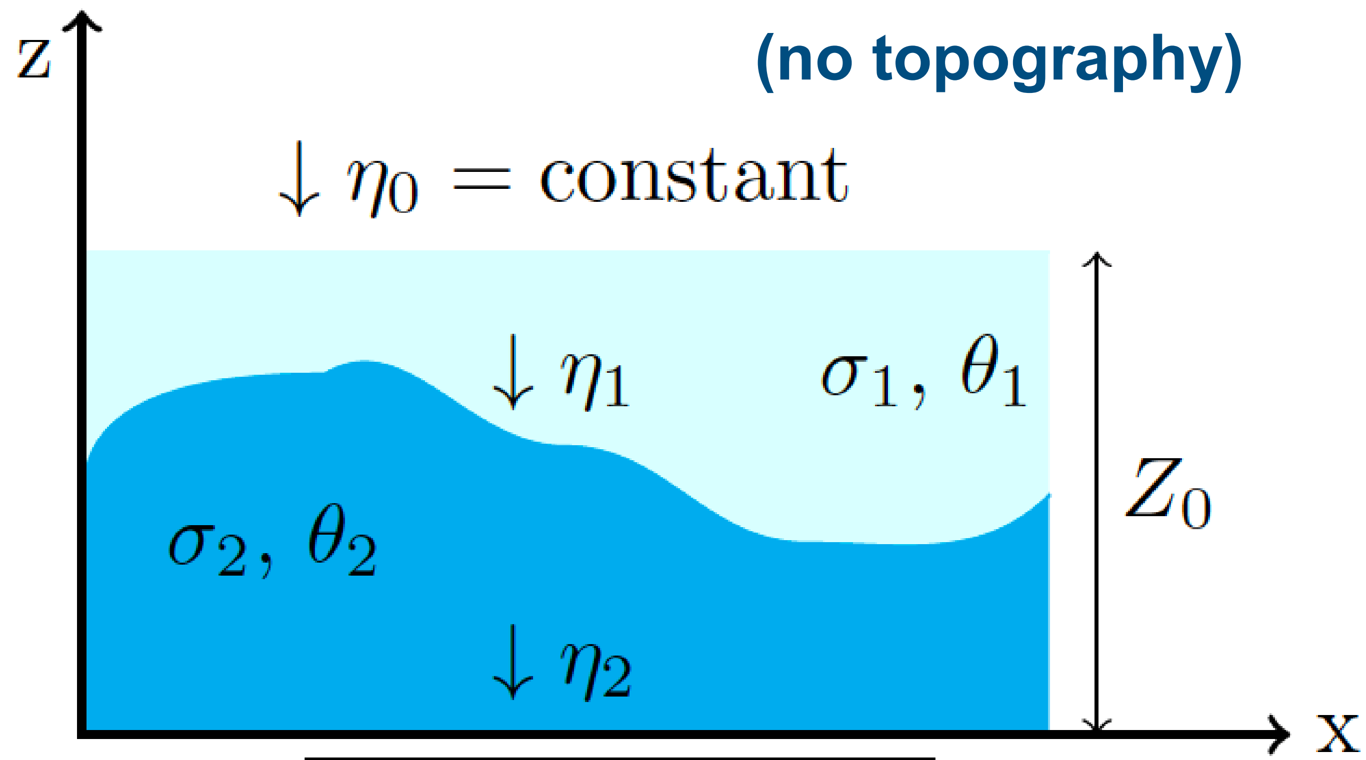
- **isentropic** fluid (robust definition of temperature):

$$T_i = \theta_i \eta_i^{\kappa}$$

$$\boxed{\eta_i = \frac{p_i}{p_r}}$$

Pressure (non-dim)

- two layers of fluid in which **the one on the top is inactive** - u_1=0 - and capped by a **rigid lid** (i.e. **1.5 layer**).



$$\boxed{\sigma_i = \frac{p_r}{g}\left(\eta_{i+1} - \eta_i\right)}$$

Pseudo-density (replaces h)

N.B. We can still solve just one set of equations (for the bottom layer).

$$\left(h, hu, hv, hr\right) \rightarrow \left(\sigma_2, \sigma_2 u, \sigma_2 v, \sigma_2 r\right)$$

The revised model

The new full set of equations reads as:

$$(\sigma_2)_t + \left(\sigma_2 u_2\right)_x = 0,$$

$$\left(\sigma_2 u_2\right)_t + \left(\sigma_2 u_2^2 + \mathscr{M}(\eta_2)\right)_x + c_0^2 \sigma_2 r_x - f\sigma_2 v_2 = 0,$$

$$\left(\sigma_2 v_2\right)_t + (\sigma u v)_x + f\sigma u = 0,$$

$$\left(\sigma_2 r\right)_t + \left(\sigma_2 u_2 r\right)_x + \beta\sigma_2 (u_2)_x + \alpha\,\sigma_2 r = 0\,.$$

$$\mathscr{M}(\eta_2) = \begin{cases} \mathscr{M}(\eta_2(\sigma_2)) & \text{if } \sigma < \sigma_c, \\ \mathscr{M}_c(\eta_c(\sigma_c)) & \text{if } \sigma \geq \sigma_c, \end{cases}$$

$$\beta = \begin{cases} \tilde{\beta} & \text{if } \sigma \geq \sigma_r, \\ 0 & \text{otherwise}\,. \end{cases}$$

Black = classic shallow water w/rotation

Red = shallow water with convection/rain

- The pseudo density is a non-linear function of the non-dim pressure η_2:

$$\sigma_2 = \eta_2 - \left(\frac{\theta_2}{\Delta\theta}\right)^{\frac{1}{\kappa}} \left(-\eta_2^\kappa + \frac{\theta_1}{\theta_2}\eta_0^\kappa + \frac{gZ_0}{c_p\theta_2}\right)^{\frac{1}{\kappa}}$$

This function is inverted online to obtain η

Checks against an analytical solution

- We derived an ODE for v from the shallow water system (**without** convection and precipitation) for stationary waves (after having defined ξ=x-ct, see Shrira papers [2],[3]):

$$v'' = \frac{1}{\mathrm{Ro}^2} \frac{v}{\tilde{c}_p \theta_2 \kappa \eta^{\kappa-1} \left(\frac{1}{\frac{\partial \sigma}{\partial \eta}} \right) \sigma_0 - \frac{1}{\mathrm{Ro}^3} \frac{1}{\left(\frac{1}{\mathrm{Ro}} + v' \right)}}$$

$$\sigma = \sigma_0 (\mathrm{Ro} + v')$$

$$u = \frac{v'}{\frac{1}{\mathrm{Ro}} + v'}$$

- We compared the solution of this equation (a stationary wave) translated in time against its evolution predicted by the numerical model (in a periodic domain).

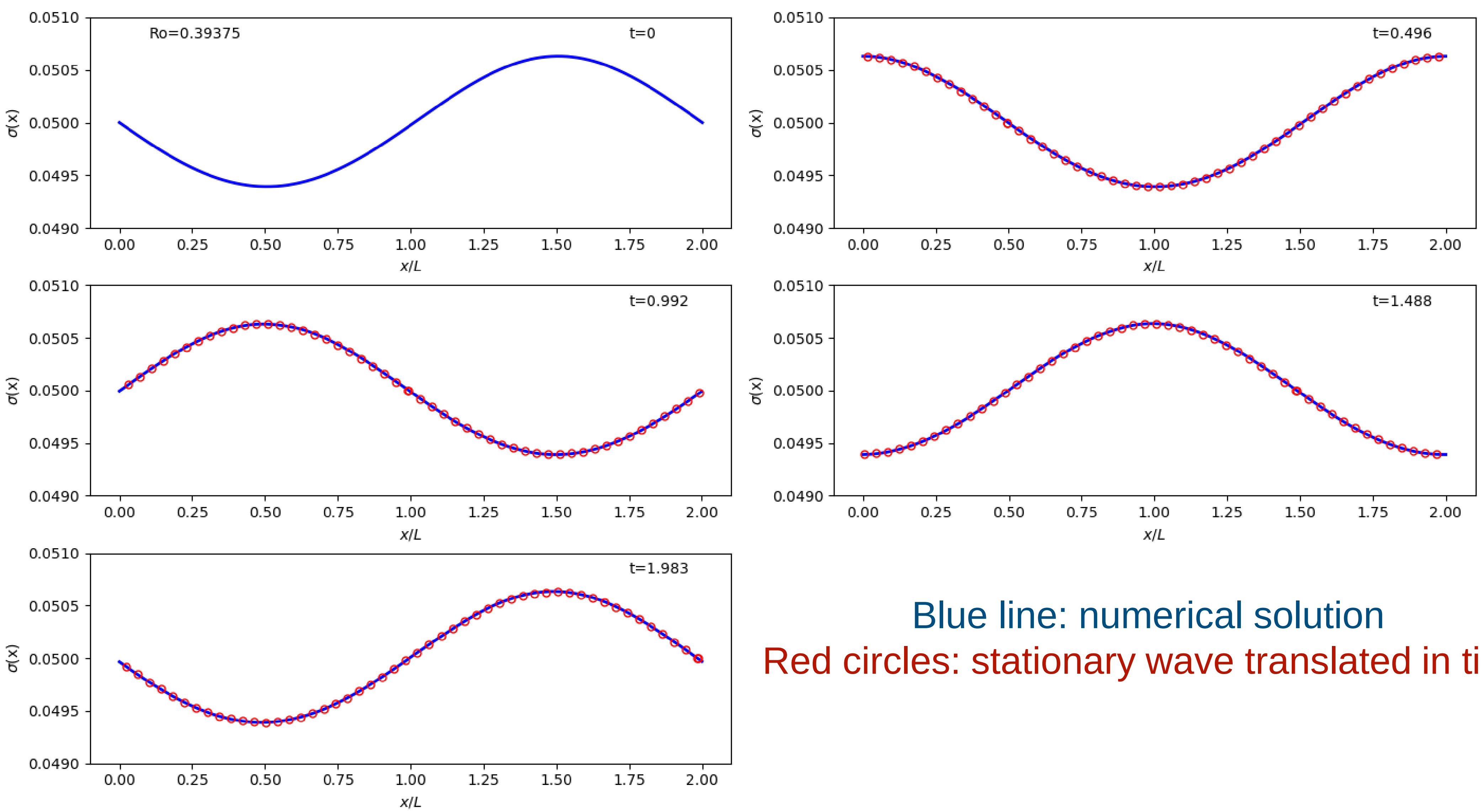
[2] Shrira, V. (1981), *Propagation of long nonlinear waves in a layer of rotating fluid*, Sov. Phys. - Izvestija, vol. 17, n. 1, pp 55-59.
[3] Shrira, V. (1986), *On the long strongly nonlinear waves in rotating ocean*, Sov. Phys. - Izvestija, vol. 22, n. 4, pp. 285-305.

Checks against an analytical solution

$$v'' = \frac{1}{\mathrm{Ro}^2}\frac{v}{\tilde{c}_p\theta_2\kappa\eta^{\kappa-1}\left(\frac{1}{\frac{\partial\sigma}{\partial\eta}}\right)\sigma_0 - \frac{1}{\mathrm{Ro}^3}\frac{1}{\left(\frac{1}{\mathrm{Ro}}+v'\right)}}$$

$$\sigma = \sigma_0\left(\mathrm{Ro} + v'\right)$$

$$u = \frac{v'}{\frac{1}{\mathrm{Ro}} + v'}$$



Blue line: numerical solution

Red circles: stationary wave translated in time

Idealised satellite observations

The radiative scheme

- Synthetic observations of radiance B are generated using the Rayleigh-Jeans law (valid for λ>50μm at T=300K):

$$B = 2\frac{k_B c}{\lambda^4}T = 2\frac{k_B c}{\lambda^4}\theta\eta^\kappa \rightarrow B' = \frac{B}{B_0} = \eta^\kappa$$

Spatially varying observations

- Let's consider polar-orbit satellites: they move and observe different portions of the Earth at different times.

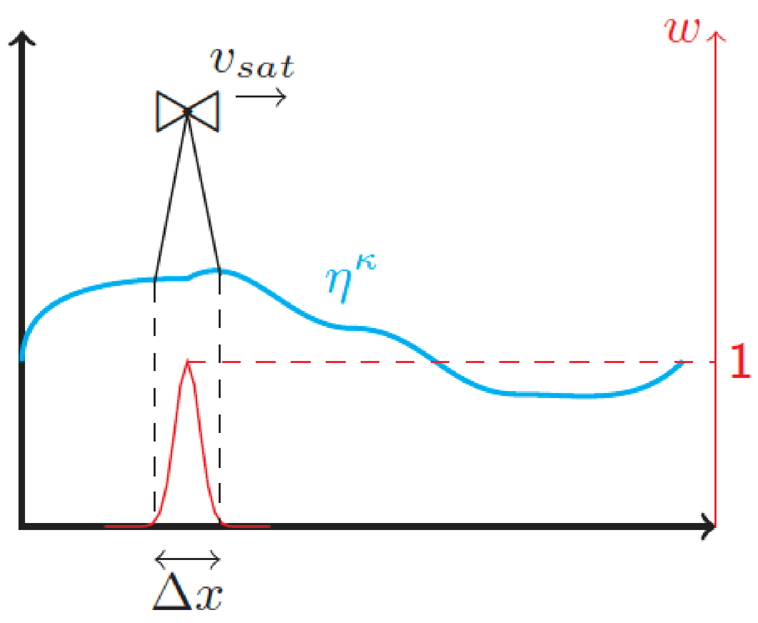
- 1 D approximation: our satellite observations move with velocity v_{sat} along a periodic domain of length L:

$$x_{sat} = v_{sat} \cdot t \mod L$$

Idealised satellite observations

Horizontally-averaged observations

- To mimic the satellite's Field Of View (FOV), a weighted mean is applied to a Δx window:



$$y^o(x_{sat}) = \int_{x_{sat}-\frac{\Delta x}{2}}^{x_{sat}+\frac{\Delta x}{2}} w(x)\eta^{\kappa}\mathrm{dx}$$

- w(x) is a Gaussian function centred on x_{sat} which is as wide as Δx.

N.B. All this is done only for σ. The other variables (u,v,r) are observed as before.

A new observation operator

- The new observation vector is split into satellite and ground observations:

$$\mathbf{y}^o = \begin{pmatrix} y^o(x_{sat}) \\ \mathbf{y}^o(\mathbf{x}_{grn}) \end{pmatrix} = \begin{pmatrix} (\eta_2^t)^\kappa(x_{sat}) \\ \mathbf{y}^o(\mathbf{x}_{grn}) \end{pmatrix},$$

in which the ground observations y^o_{grn} are direct observations of u,v,r at fixed x_{grn} positions along the domain.

- The new observation operator $\mathcal{H}$ reads as:

$$\mathscr{H}(\mathbf{x}^f) = \begin{pmatrix} y^f(x_{sat}) \\ \mathbf{y}^f(\mathbf{x}_{grn}) \end{pmatrix} = \begin{pmatrix} (\eta_2^f)^\kappa(x_{sat}) \\ \mathbf{y}^f(\mathbf{x}_{grn}) \end{pmatrix}.$$

What happens to the EnKF?

- We made no changes to the DA scheme, still an EnKF:

$$x^a = x^f + K\left(y^o - \mathcal{H}(x^f)\right) \qquad K = P^f H^T\left(HP^f H^T + R\right)^{-1}$$

- A common way of using an EnKF in the presence of a non-linear observation operator is given by Houtekamer & Mitchell (see [5]):

$$P^f H^T = \frac{1}{N-1}\sum_{i=1}^{N}\left(x^f - \overline{x^f}\right)\left(\mathcal{H}x^f - \overline{\mathcal{H}x^f}\right)^T, \qquad \overline{x^f} = \frac{1}{N}\sum_{i=1}^{N} x_i^f$$

$$HP^f H^T = \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathcal{H}x^f - \mathcal{H}\overline{x^f}\right)\left(\mathcal{H}x^f - \overline{\mathcal{H}x^f}\right)^T \qquad \overline{\mathcal{H}x^f} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{H}x_i^f$$

This, though, is not straightforward combine with the model-space localisation used in current modRSW setup.

[5] Houtekamer, P. L., & Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, *129*(1), 123-137.

What happens to the EnKF?

- Instead, we decided to linearise the observation operator $\mathcal{H}$ only for the purpose of computing the Kalman Gain:

$$H \simeq \partial_{x^f}\mathscr{H} = \left(\partial_{\sigma_2^f}\mathscr{H}, \partial_{u^f}\mathscr{H}, \partial_{v^f}\mathscr{H}, \partial_{r^f}\mathscr{H}, \right)$$

- This assumption of course is not optimal (even if we don’t know how deleterious it is), but at this stage it’s less time-consuming than moving from model-space localisation to observation-space localisation.

[5] Houtekamer, P. L., & Mitchell, H. L. (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, *129*(1), 123-137.

Conclusions

- We have modified the single-layer isopycnal ‘modRSW’ into an isentropic 1.5-layer model. We checked the new model (without convection and precipitation) against an analytical solution.

- The observations are now split into satellite and ground ones. Satellite observations are modelled as radiance measurements which take into account both the spatially varying character of polar-orbit satellite and are averaged horizontally to mimic the FOV.

- We modified the observation operator accordingly, into a new, non-linear one.

Future work

- Modify the model in order to include topography.
- Explore the possibility of defining weighting functions and using a multi-channel approach in assimilating radiance.
- Find the best strategy to define clouds.
- Explore alternative radiation schemes.
- Ultimately, use the new setup to investigate the relative impact of observing large-scale and small-scale features (what should we better focus on in the future?).

Questions?

Email: mmlca@leeds.ac.uk

Code on github: https://github.com/tkent198/modRSW_EnKF

Scaling for ‘modRSW’ in presence of temperature

- We tried to define a diagnostic equation for temperature based on hydrostatic equilibrium and the ideal gas law:

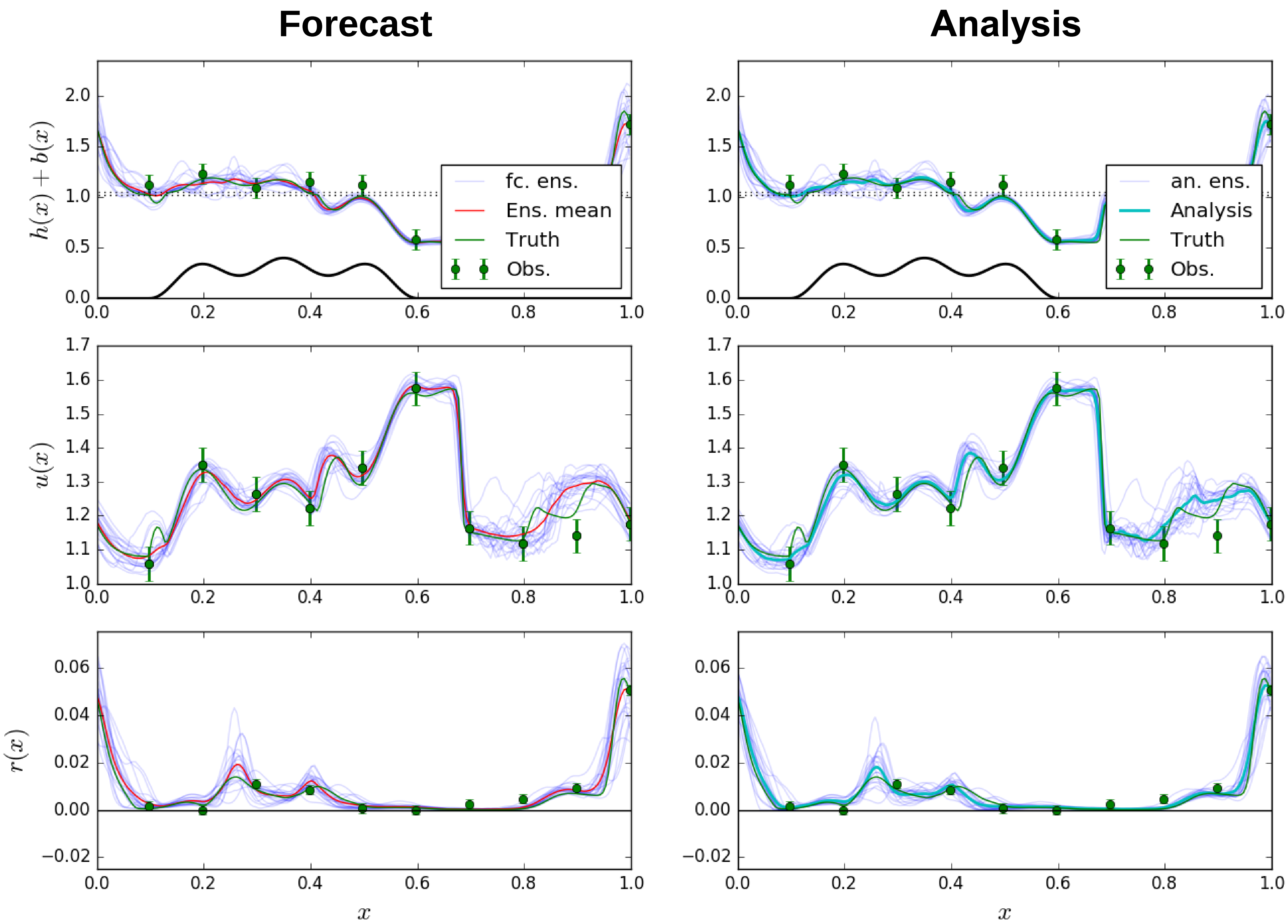
$$T = \frac{gh}{R}$$

- The scaling for gH used in [1] ($gH=330m^2s^{-2}$) leads to values of temperature of order O(1) K. But that was chosen to maintain the Froude number above 1 (with U=20m/s):

$$\mathrm{Fr} = \frac{U}{\sqrt{gH}}$$

A reminder: Fr>1 implies supercritical regime which implies traveling gravity waves (i.e. convection moving across the domain)

[1] Kent, T. et al (2017): *Dynamics of an idealized fluid model for investigating convective-scale data assimilation*. Tellus A: Dynamic Meteorology and Oceanography, 69(1), 1369332.



Forecast
Analysis
h(x) + b(x)
u(x)
r(x)
x
fc. ens.
Ens. mean
Truth
Obs.
an. ens.
Analysis
Truth
Obs.