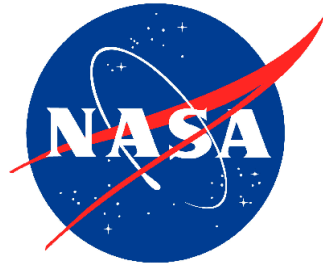




The tangent linear and adjoint of the FV3 dynamical core: development and applications



Daniel Holdaway

NASA Global Modeling and Assimilation Office

dan.holdaway@nasa.gov

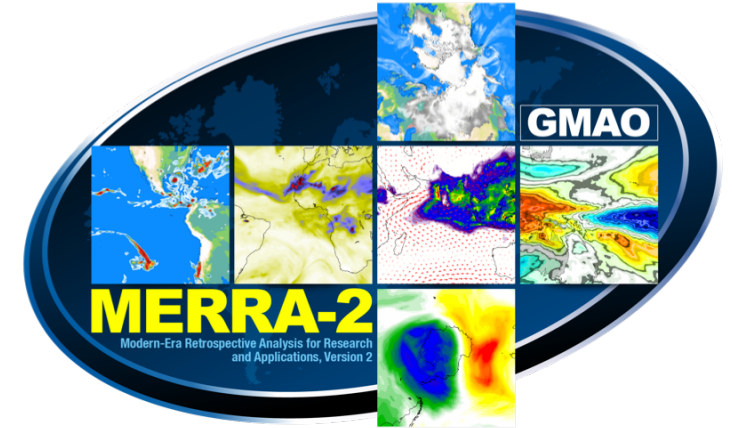


Stony Brook TAOS Seminar

Wednesday April 25nd

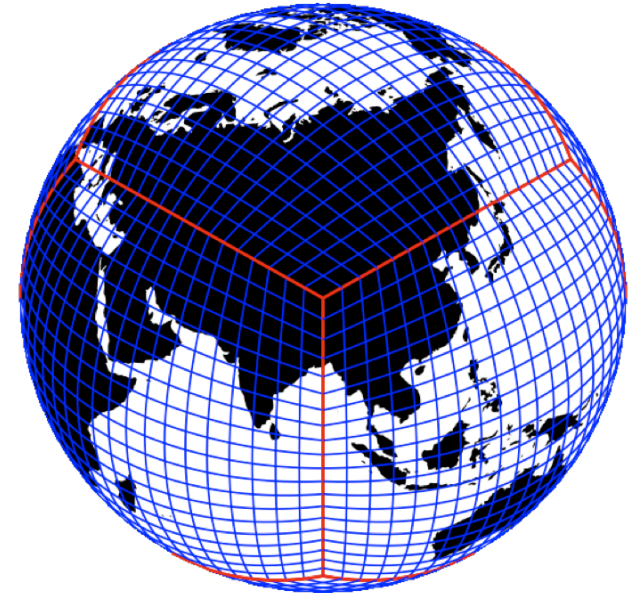
NASA's Global Modeling and Assimilation Office

- Situated in NASA's Earth Science Division and based at Goddard Space Flight Center in Greenbelt MD.
- Around 115 employees.
- Our main responsibility is to develop the Goddard Earth Observing System (GEOS).
- Central GEOS products are: Modern-Era Retrospective Analysis for Research and Applications (MERRA) (atmosphere, land, aero), S2S system, operational forecasting, providing data for processing retrievals, NASA mission support, high resolution nature runs (OSSEs).

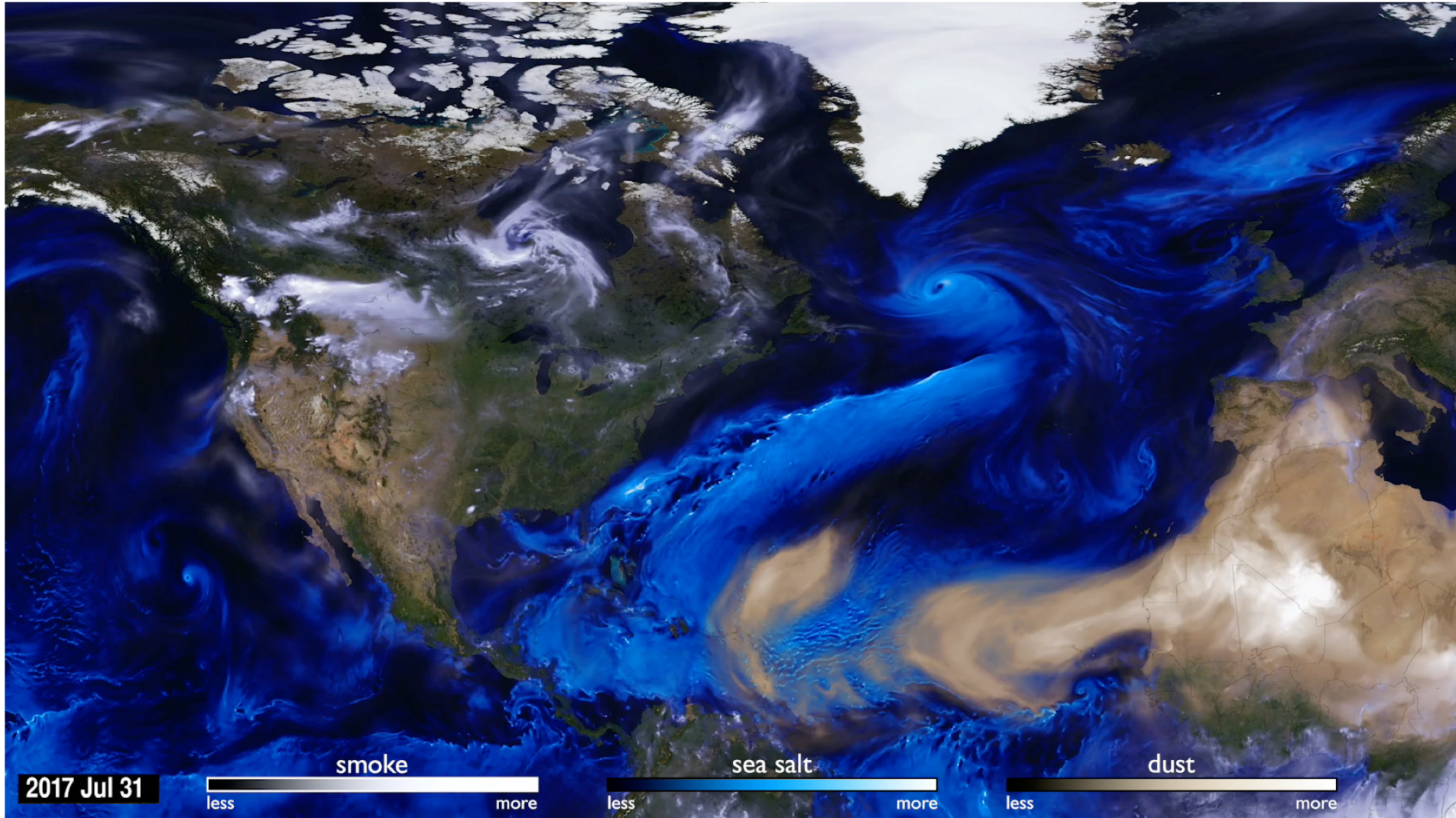


GEOS model description

- 12.5km global model for operations, 50km for reanalysis.
- Model initial conditions generated using the Gridpoint Statistical Interpolation (GSI) data assimilation system, developed in close collaboration with NOAA NCEP.
- Initialized four times daily with 10 day forecast from 00Z.
- Cubed-sphere finite volume (FV3) dynamical core (~10 year partnership with NOAA GFDL).
- Full sweep of physics.
- MOM5 ocean model for coupled simulations such as in S2S.
- Interactive aerosols using GOCART.



2017 hurricanes and aerosols simulation





Outline

1. Introduction to data assimilation and need for the adjoint
2. Development of the FV3 TLM and adjoint
3. 4DVar and the future of data assimilation
4. Using the adjoint to investigate predictability

Introduction to data assimilation

In data assimilation we essentially need to balance two pieces of information to produce a new ‘best guess’ of the current atmospheric state x_0 , winds, temperature etc.

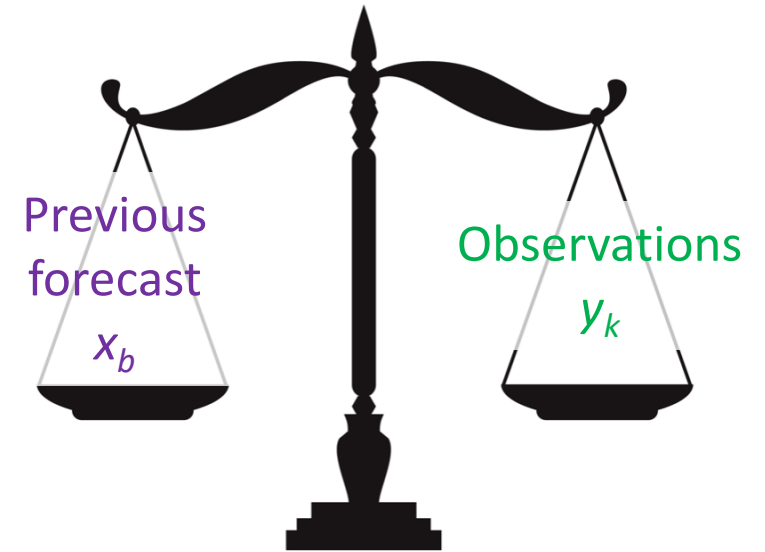
This is achieved by minimizing a cost function:

$$J(\mathbf{x}_0) = \frac{1}{2} (\mathbf{x}_b - \mathbf{x}_0)^\top \mathbf{B}^{-1} (\mathbf{x}_b - \mathbf{x}_0) + \frac{1}{2} \sum_{k=1}^K (\mathbf{y}_k - h_k [m_{t_0 \rightarrow t_k}(\mathbf{x}_0)])^\top \mathbf{R}_k^{-1} (\mathbf{y}_k - h_k [m_{t_0 \rightarrow t_k}(\mathbf{x}_0)])$$

Using its gradient:

$$\nabla J(\mathbf{x}_0) = \mathbf{B}^{-1} (\mathbf{x}_b - \mathbf{x}_0) + \sum_{k=1}^K \mathbf{H}^\top \mathbf{M}_{t_0 \rightarrow t_k}^\top \mathbf{R}_k^{-1} (\mathbf{y}_k - h_k [m_{t_0 \rightarrow t_k}(\mathbf{x}_0)])$$

↖ adjoint

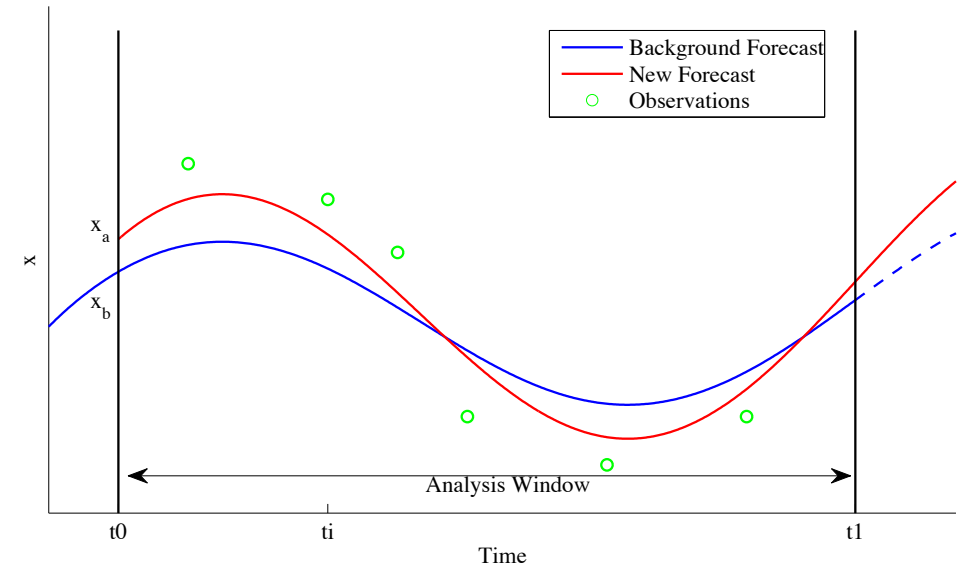


- Whole atmosphere.
- Balanced.
- But error details not well known.
- Small, well described errors.
- But under representative $\sim 150/1$.

Data assimilation formulations

The above formulation of DA above is known as 4DVar. Other forms of data assimilation are derived by making further assumptions to simplify matters:

- Incremental 4DVar: reduces resolution, replaces m with \mathbf{M} .
- 3DVar: eliminate the adjoint by assuming all observations are made at the same single time.
- 4DEnVar: replace the adjoint with covariance between time varying ensemble members, less development and cheaper.



Incremental Hybrid **4DEnVar** is used by GMAO and NCEP. Incremental Hybrid **4DVar** is used by ECMWF and the Met Office. Global forecasts from the European centers are considered most accurate so the adjoint is key.

“At the start of this work, we hoped to find a combination of ensemble size, weighting and localization length-scales which reduced the performance gap of hybrid 4D-Var over hybrid 4DEnVar seen by Lorenc et al. (2015). The combinations we tested did not achieve this” Bowler et. al. (2017, QJRMS)



Introduction to adjoints

The forecast at some grid point is given by a model:

$$y_i = m(x_1, x_2, \dots, x_j)$$

where x is the input state and y is the output state.

From a DA perspective it's all about perturbations! In particular the following questions are often raised:

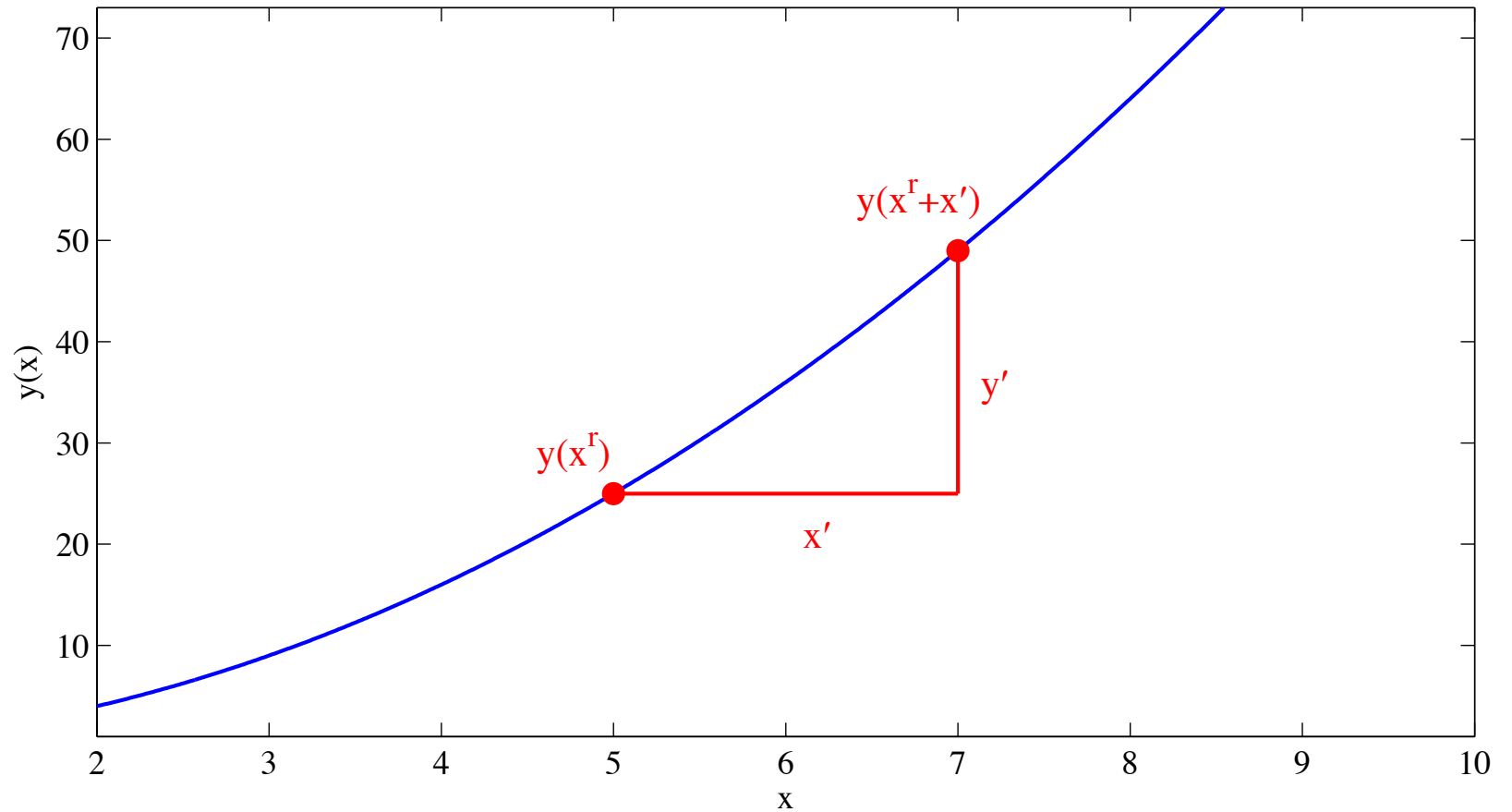
1. How does y_i change with respect to x_j ? Suppose x_j has some error. How does that error grow?
2. We're analyzing something about y_i , e.g. an error. What specifically was it about x_j that led to this?

Introduction to adjoints

The effect of a particular perturbation can always be measured with two nonlinear runs:

$$y'_i = m(x_1 + x'_1, x_2 + x'_2, \dots, x_j + x'_j) - m(x_1, x_2, \dots, x_j)$$

However, it could take many nonlinear runs to figure out why a particular result or error was seen in y .



Introduction to adjoints

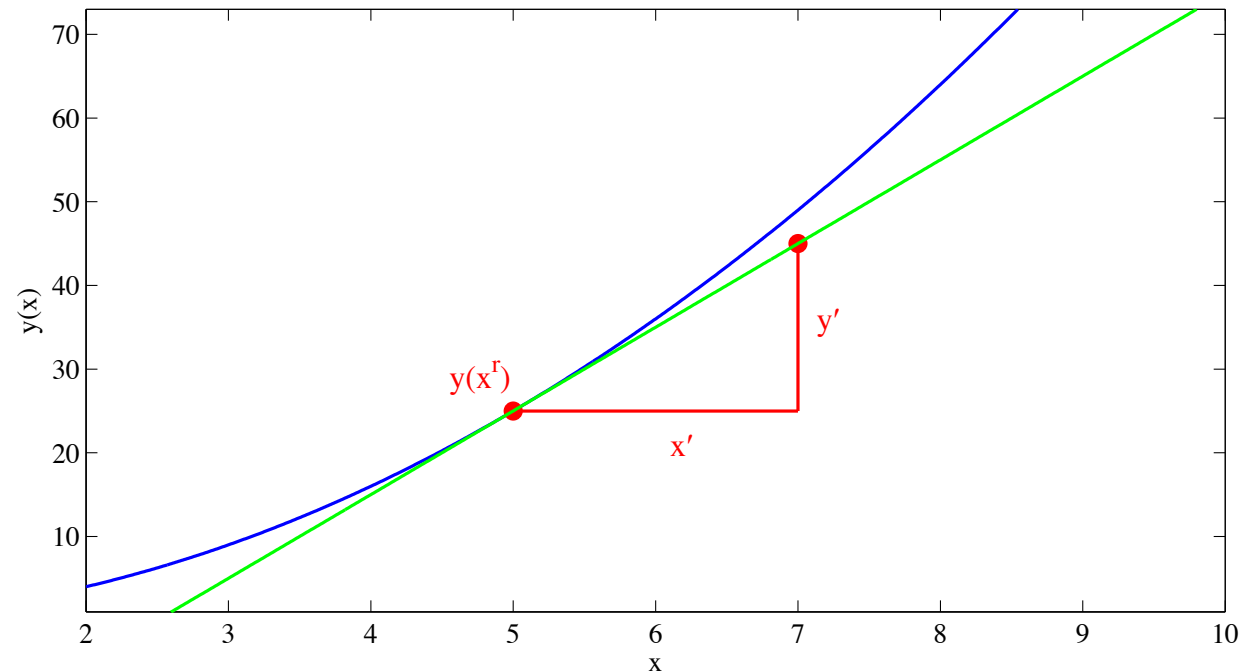
Perturbations so think Taylor series:

$$\begin{aligned}
 y'_i &= m(x_1^r + x'_1, x_2^r + x'_2, \dots, x_j^r + x'_j) - m(x_1^r, x_2^r, \dots, x_j^r) \\
 &= \cancel{m(x_1^r, x_2^r, \dots, x_j^r)} + \frac{\partial y_i}{\partial x_1} x'_1 + \frac{\partial y_i}{\partial x_2} x'_2 + \dots - \cancel{m(x_1^r, x_2^r, \dots, x_j^r)} \\
 &= \sum_j \frac{\partial y_i}{\partial x_j} x'_j
 \end{aligned}$$

Or in vector form:

$$\mathbf{y}' = \mathbf{M}\mathbf{x}'$$

This is the tangent linear model (**TLM**).



Introduction to adjoints

Introduce a scalar measure of some forecast feature or error $J = J(y) = J[m(x)]$. Again using the Taylor series an approximation for how perturbations of J can be arrived at:

$$J' = \frac{\partial J}{\partial x} x'_j$$

Gradients of J with respect to model input variables are key, this can be found using the chain rule:

$$\frac{\partial J}{\partial x_j} = \sum_i \frac{\partial y_i}{\partial x_j} \frac{\partial J}{\partial y_i}$$

Or in vector form:

$$\frac{\partial J}{\partial \mathbf{x}} = \mathbf{M}^\top \frac{\partial J}{\partial \mathbf{y}}$$

This powerful equation translates gradients at forecast time backwards in time to initial time. Rather than using the sum notation variables can be combined into vector form:

Introduction to adjoints

Perturbations and sensitivity grow fastest along the steepest gradients of the model.

The adjoint describes the sensitivity to initial conditions for a particular event provided:

- The tangent linear model provides a good approximation of the how the equivalent perturbation evolves.
- The gradient of the model(s) exists.

This means that care is required when developing adjoint models to make sure they behave well are not trying to linearize around discontinuous functions. Much testing is required.

All verification of the assumptions is done with the TLM. Correct coding of the adjoint is just confirmed using a dot product test:

$$\langle \mathbf{M}\mathbf{x}', \mathbf{M}\mathbf{x}' \rangle = \langle \mathbf{M}^T \mathbf{M}\mathbf{x}', \mathbf{x}' \rangle$$



Development of the FV3 TLM and adjoint



The GEOS adjoint system

The adjoint currently consists of:

- Cubed-sphere FV3 dynamical core (circa 2010 version).
- Convection, modified RAS scheme.
- Single moment cloud scheme.
- Boundary layer scheme.
- Radiation (research only).
- GOCART dust (research only).
- Gravity wave drag (research only).

In 2016/17 NOAA went through an extensive dynamical core comparison, which resulted in them choosing FV3 for the Next Generation Global Prediction System (NGGPS). As a result FV3 underwent significant changes which motivated redeveloping the adjoint. This would also help us to overcome some issues making implementation of operational 4DVar difficult.

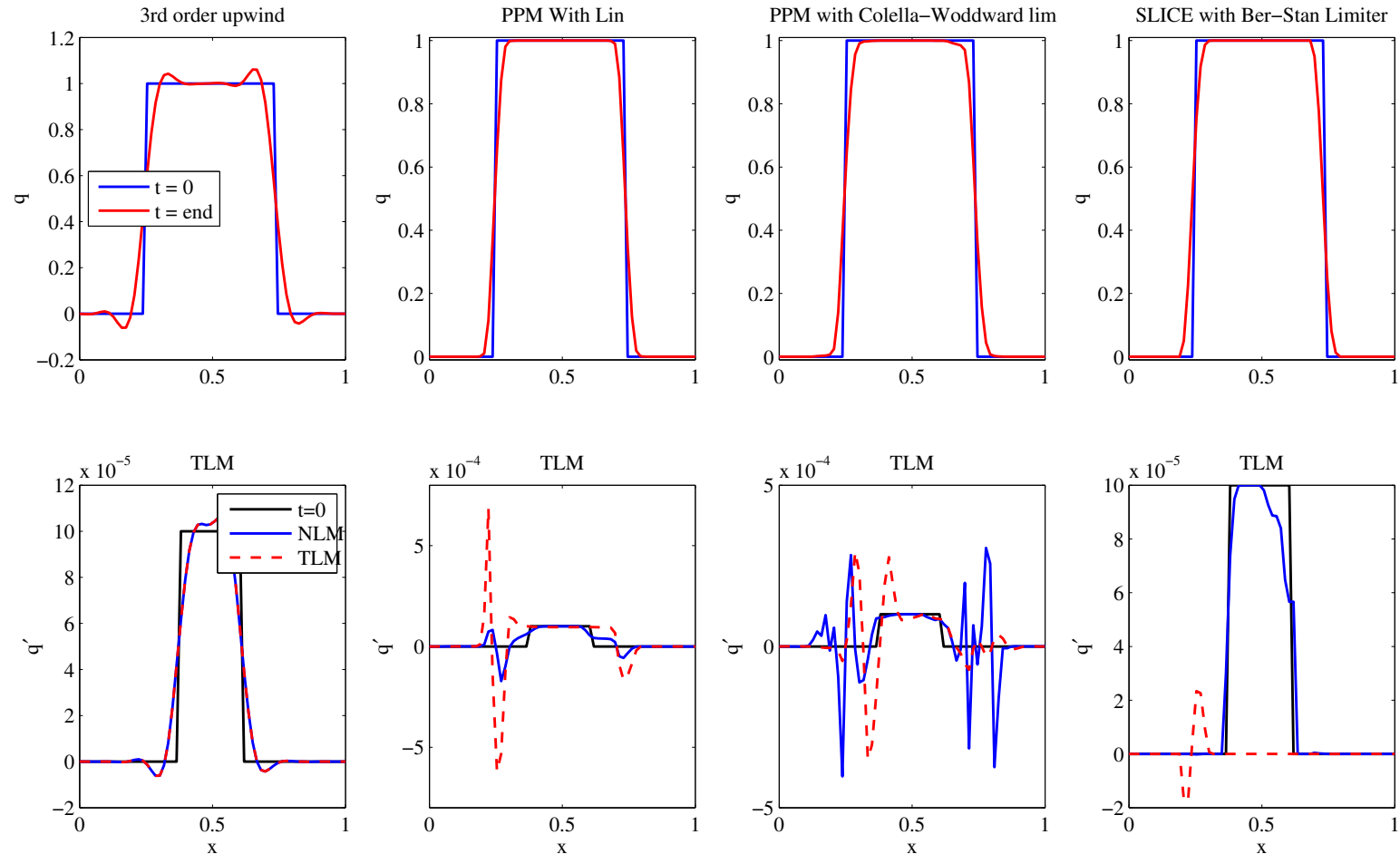
Strategy for advection and damping

Dynamical cores are based on the *weakly nonlinear* Navier-Stokes equations. For small perturbations a tangent linear model will describe the evolution well.

However, modern dynamical cores employ *strong nonlinearity* in their formulation of advection. This is to avoid spurious oscillations are avoided.

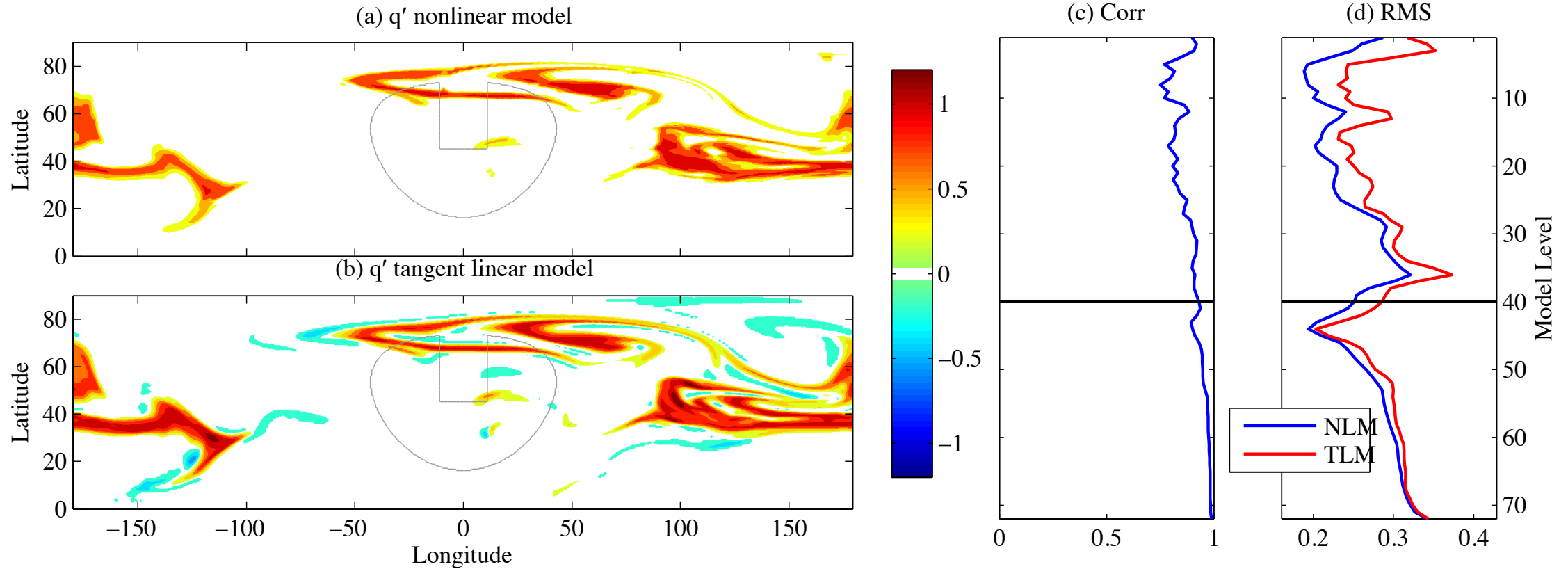
It is not possible to include these schemes in adjoint versions.

Instead linear schemes are used only for perturbations but with stronger damping.



KentHoldaway-2017-QJRM, HoldawayKent-2015-TellusA

TLM verification: passive slotted cylinder tracer



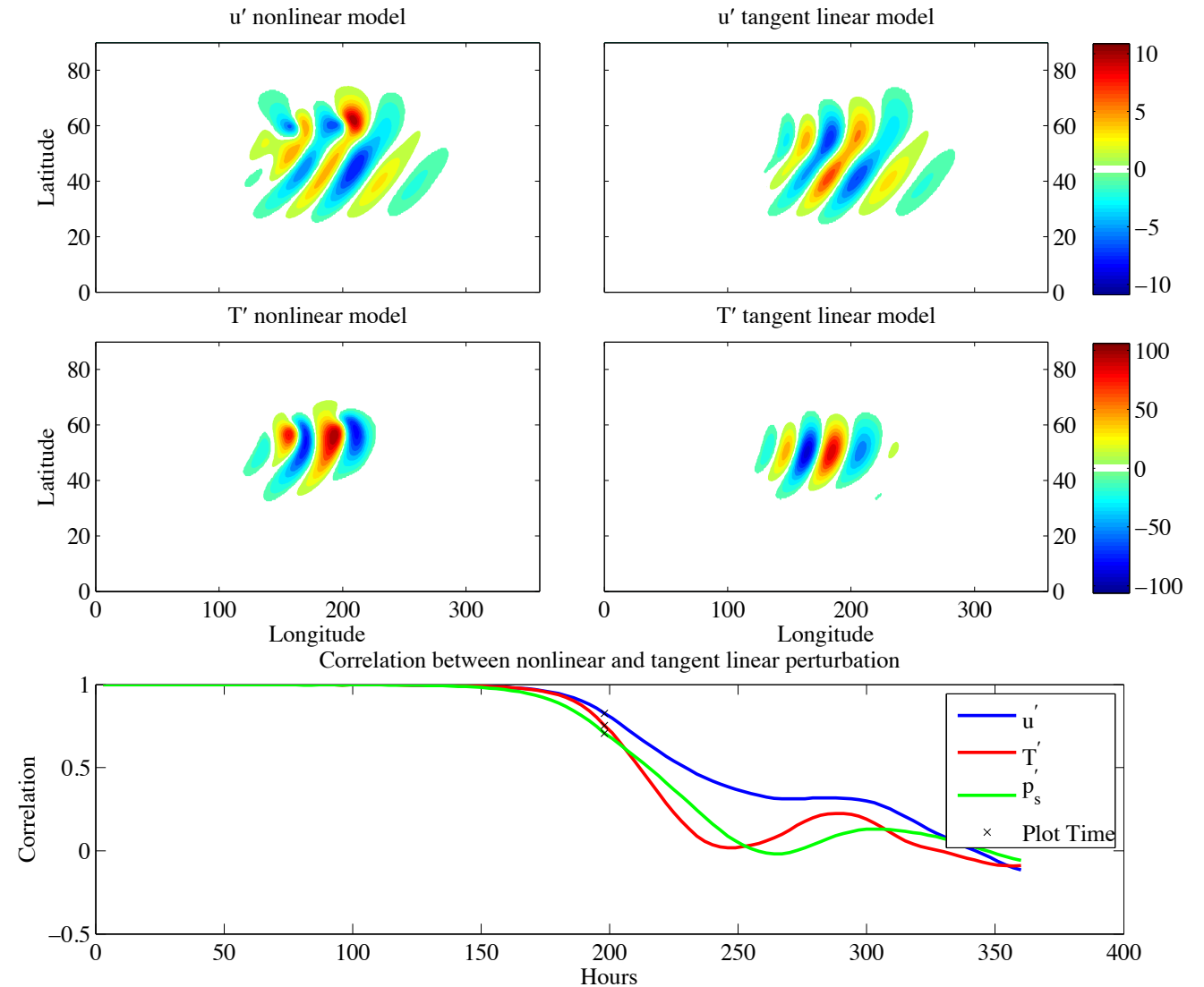
TLM verification: dynamics test case

Jablonski-Williamson dry baroclinic wave test case.

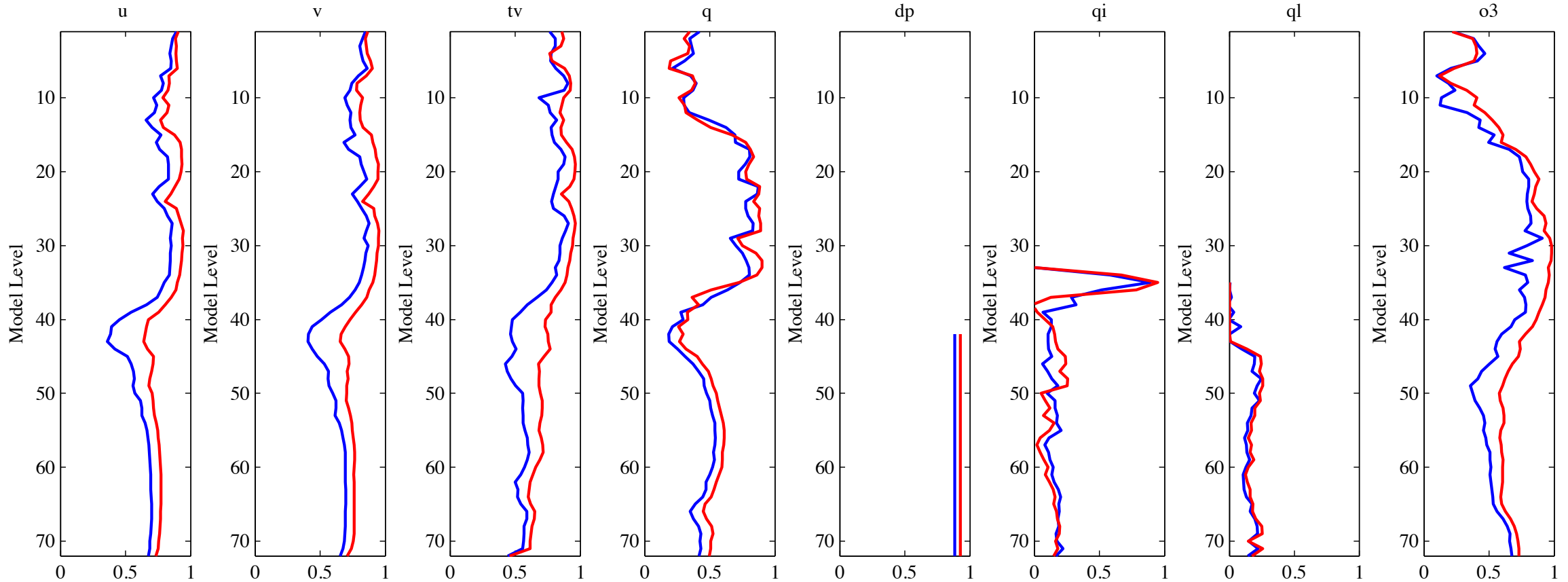
Comparing evolution of a perturbation that generates an instability over a 15 day period with the nonlinear and tangent linear models.

For the linear growth of the wave the tangent linear model performs perfectly. As the wave breaks and nonlinear restoring mechanism kicks in the TLM accuracy reduces.

In reality the model is re-linearized more frequently to avoid this issue.



TLM validation: analysis increment evolution over 24-hours



$$u' = 0.81, v' = 0.81, T_v' = 0.79, q' = 0.51, \Delta p' = 0.92$$

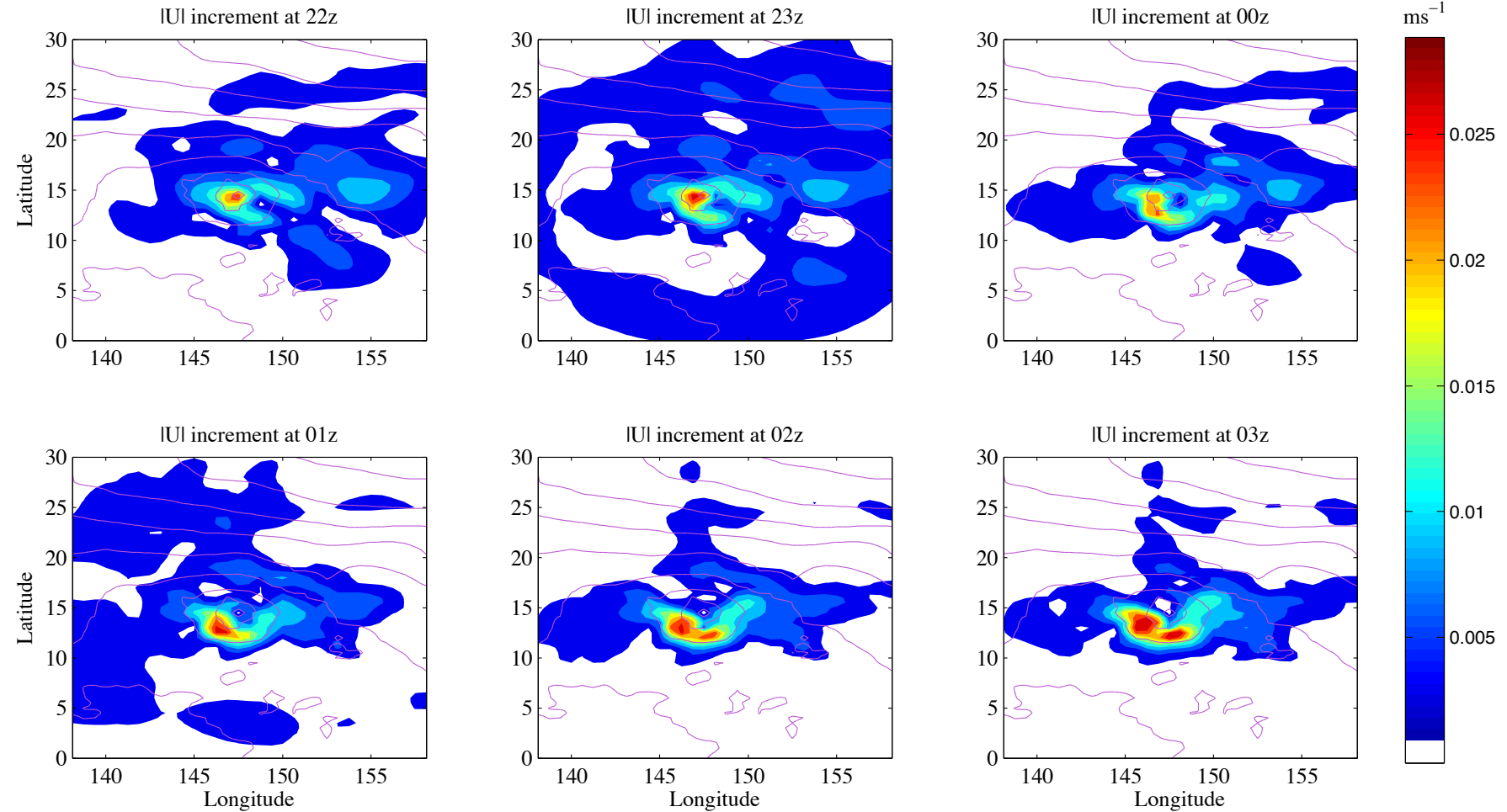


Applications of the GEOS adjoint

- 4DVar data assimilation
- Computing observations impacts.
- Sensitivity research.
- Computing singular vectors.
- Estimating surface flux emissions of constituents.
- Validation of the NASA OSSE system.

4DVar with GSI: single observation test

- Typhoon Chaba (2016)
- Too weak background
- Single observation test of TCVITAL surface pressure.
- 4DVar analysis increment of wind magnitude at 850hPa.
- Increment evolves around the storm.

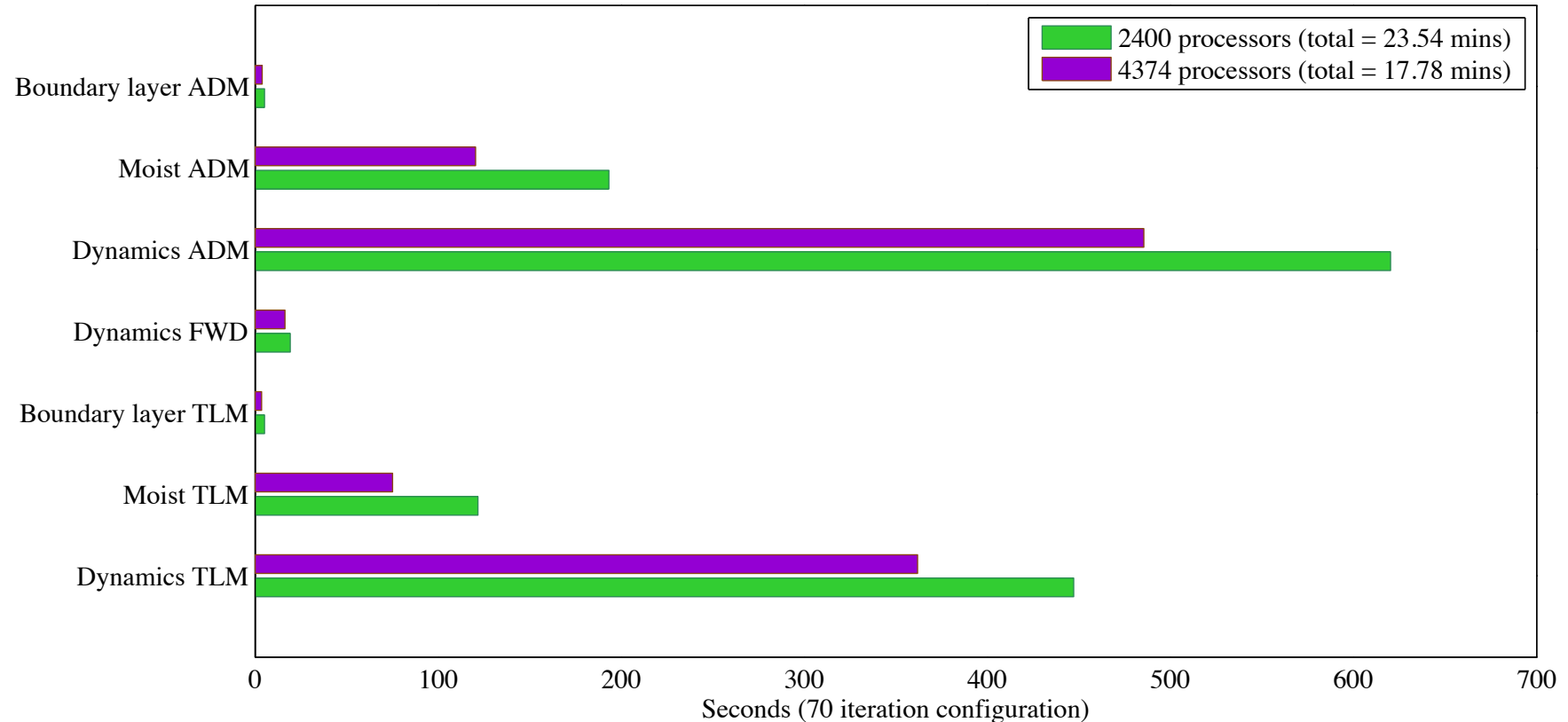


4DVar with GSI: computational cost

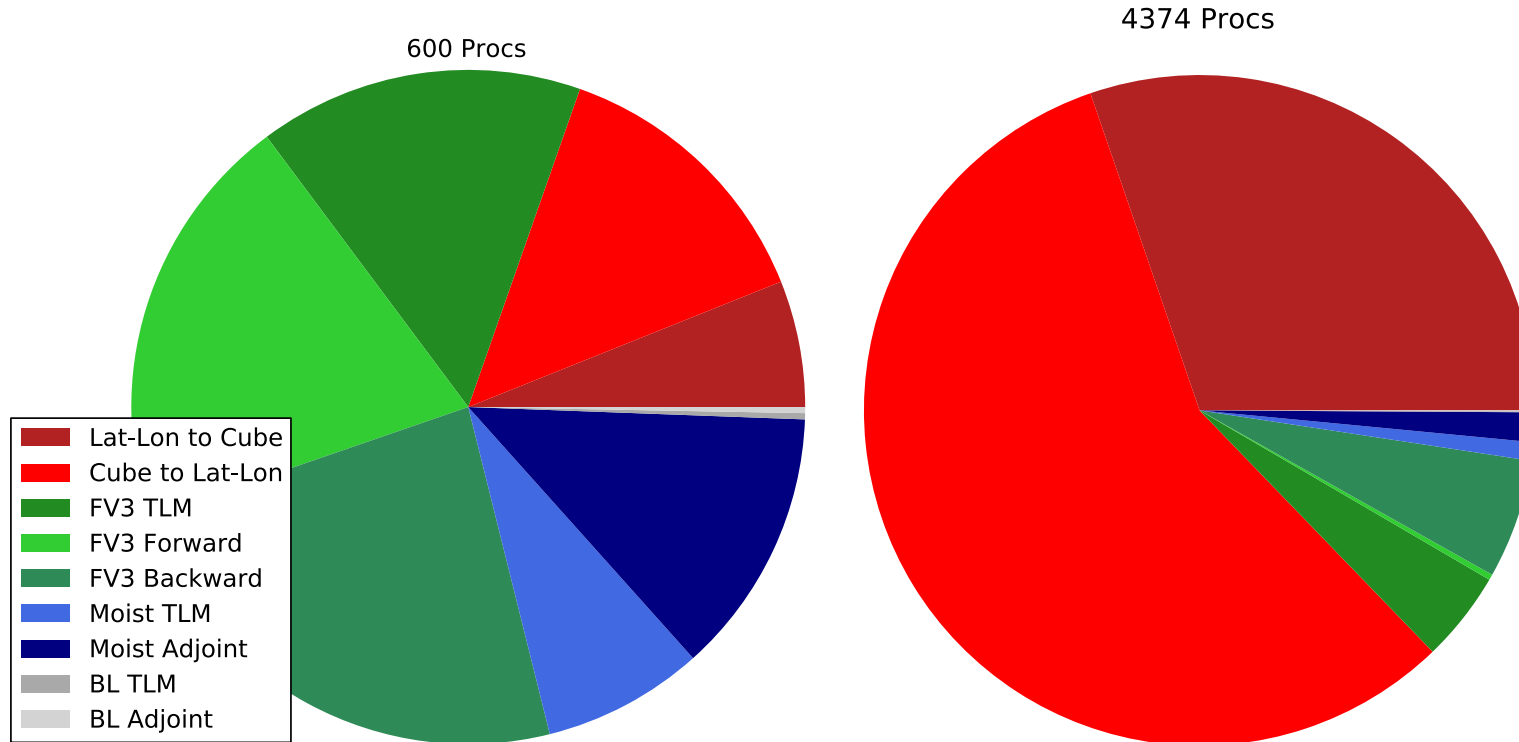
The GEOS operational system uses 5000 cores. With a similar number we can achieve reasonable performance in terms of evolving the increment using the TLM and adjoint.

The entire analysis needs to run in ~20 minutes.

Some more tuning requires but flexible code in place.



4DVar with GSI: computational cost



GSI does not run on the native model grid but on a Lat-Lon type grid.

At each inner loop this means interpolation between the grids.

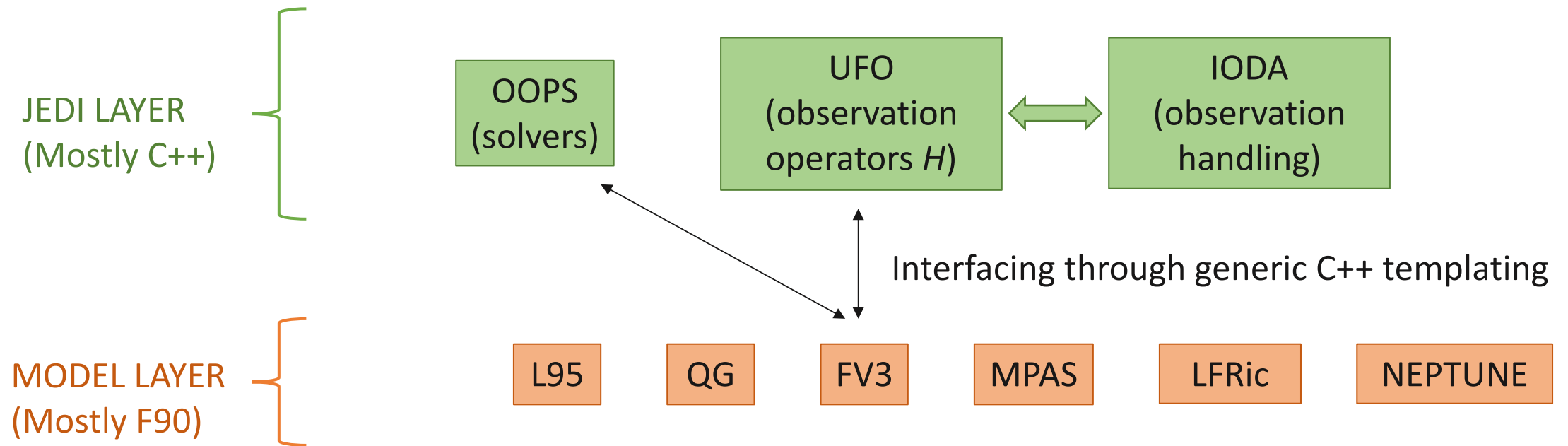
The adjoint of these routines do not currently scale well.

4DVar remains a challenge but the adjoint component is working a lot better and the code is considerably more flexible. The issues now are with the data assimilation limitations.

Aside: data assimilation with JEDI

A new tool for data assimilation is in development by the Joint Center for Satellite Data Assimilation (NASA + NOAA + DoD + UCAR).

The Joint Effort for Data Assimilation Integration (JEDI)





Advantages of JEDI

- 4DVar can be much faster by avoiding adjoint transforms from model grid to analysis grid.

In addition:

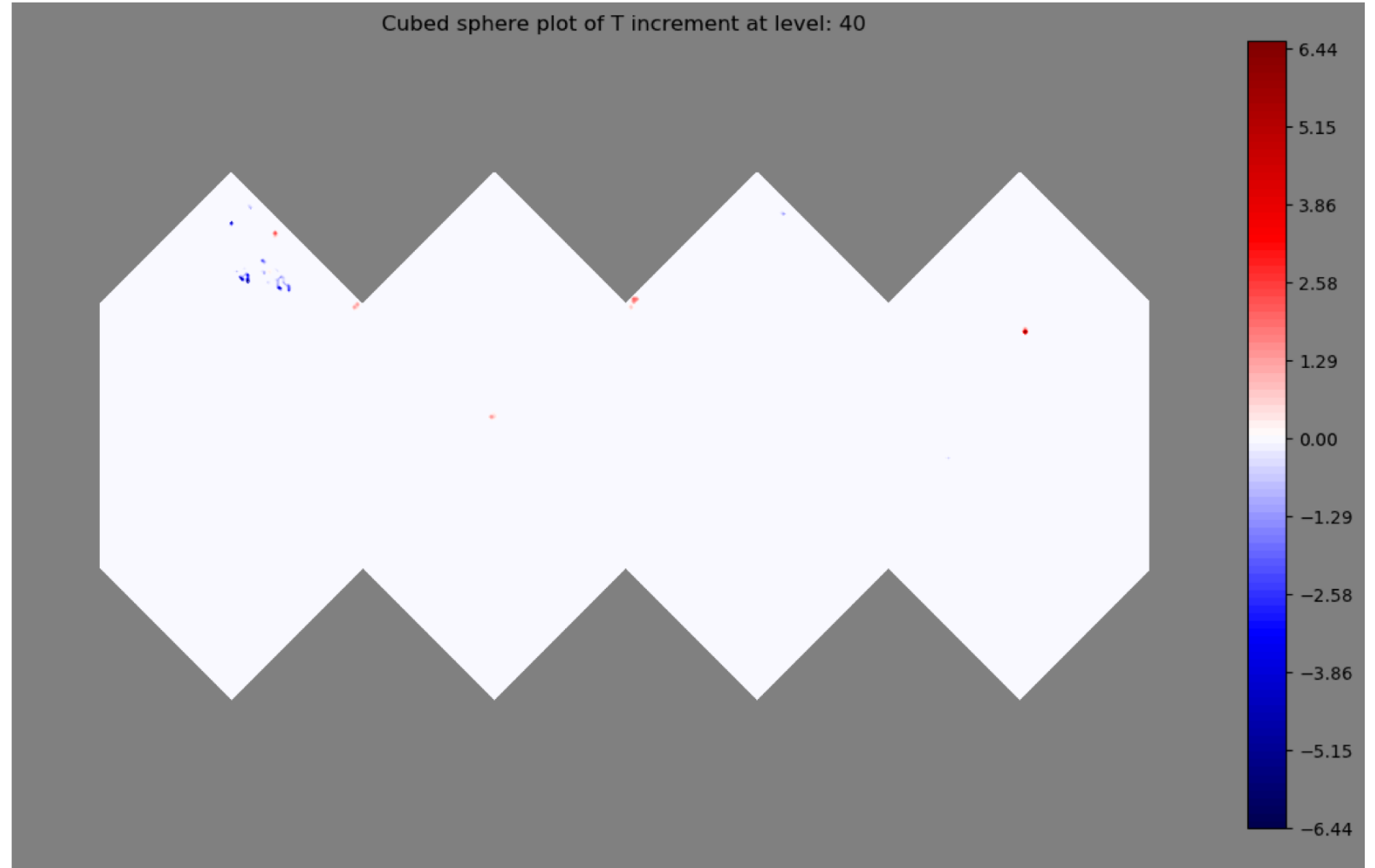
- Research to operations will be significantly improved.
- It suits a unified modeling approach well.
- It provides a clearer path to coupled data assimilation.
- Modern programming and collaboration techniques.

First increment from FV3-JEDI

Pure software demonstration. JEDI work is progressing nicely and the ability to generate an increment on the FV3 grid has been completed for 3DVar and 3DEnVar.

Lots of work to do still but progress.

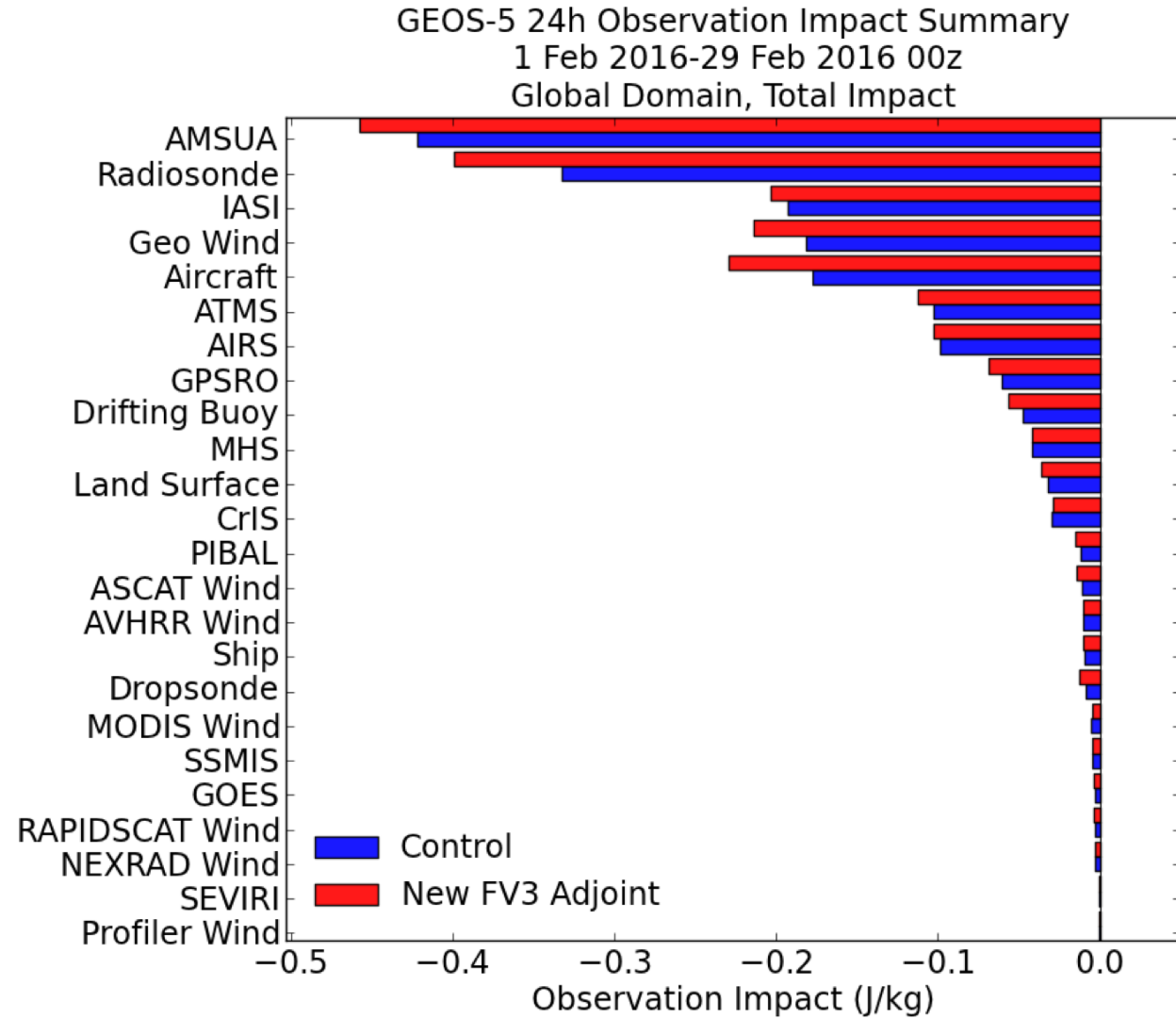
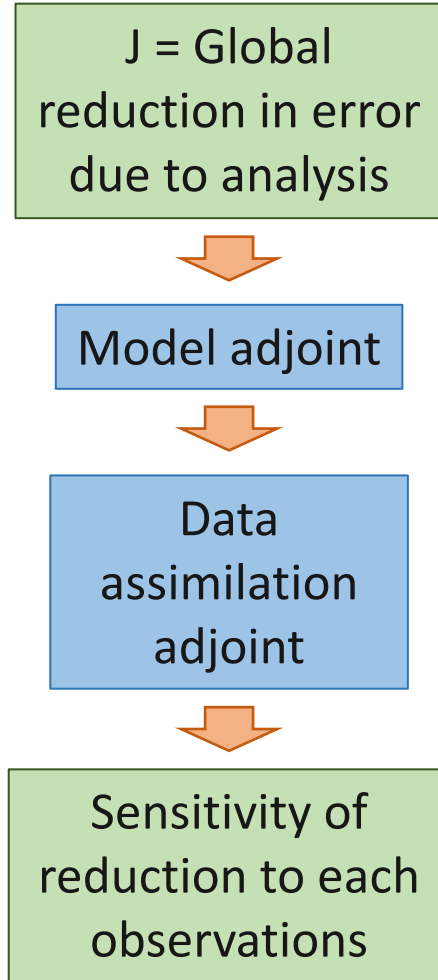
4DVar will be on the backburner until JEDI has matured.





Observation impacts

Observation impacts





Sensitivity case studies

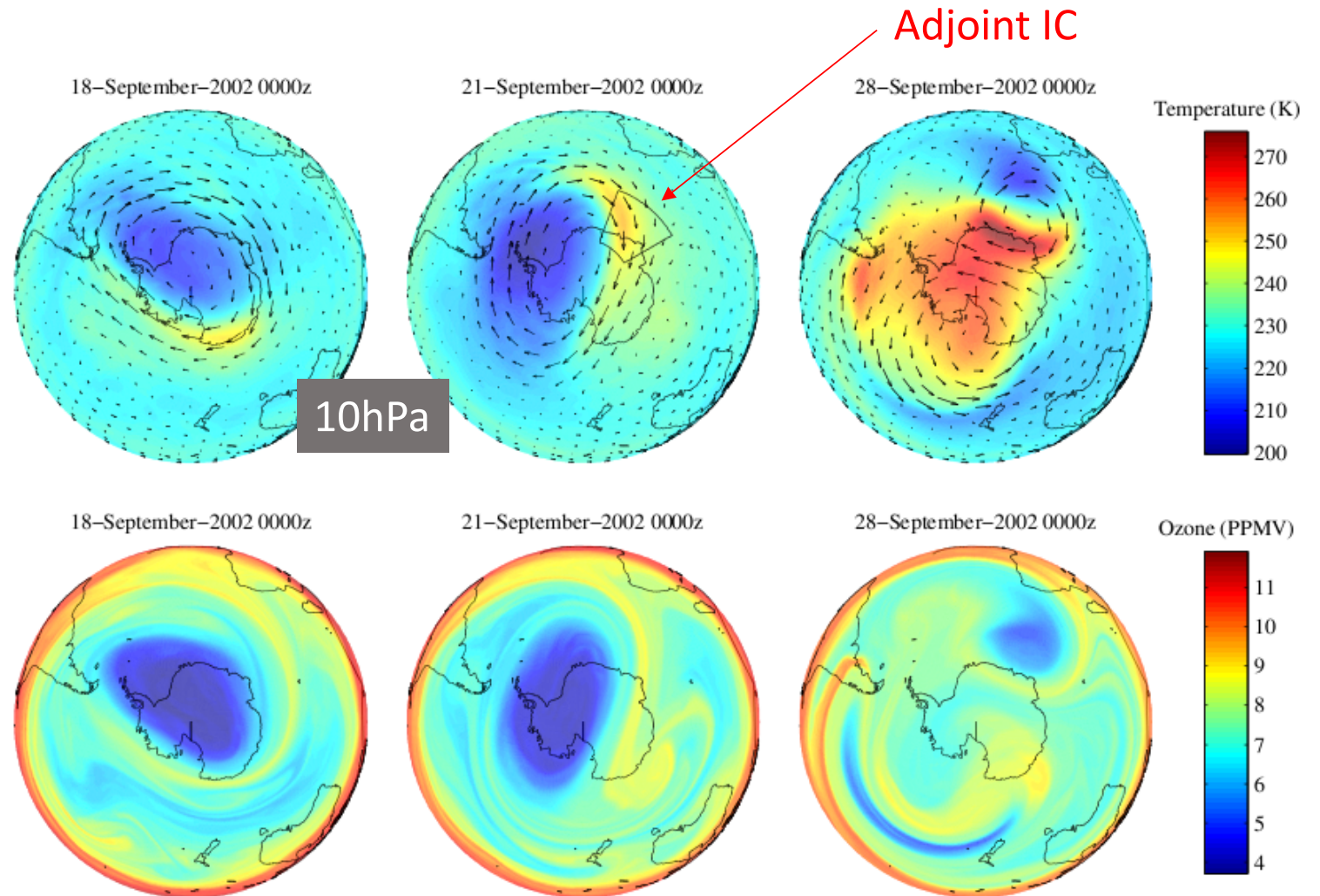
- **Sudden stratospheric warming.**
- **Hurricane Joaquin (With Jim Doyle NRL).**
- Winter storm rapid intensification (University of Wisconsin).
- Sensitivity to dust sources and radiative effects in hurricane formation.
- TC intensification rate (University of Wisconsin).
- Thinning of AIRS data around hurricanes.

SH Sudden stratospheric warming 2002

2002 saw a rare southern hemisphere major sudden stratospheric warming. The only one ever observed.

Over a period of 10 days the westerly flow in the stratosphere reversed and the ozone hole collapsed. Throughout the winter a number of minor warmings slowed the flow until the event shown to the right.

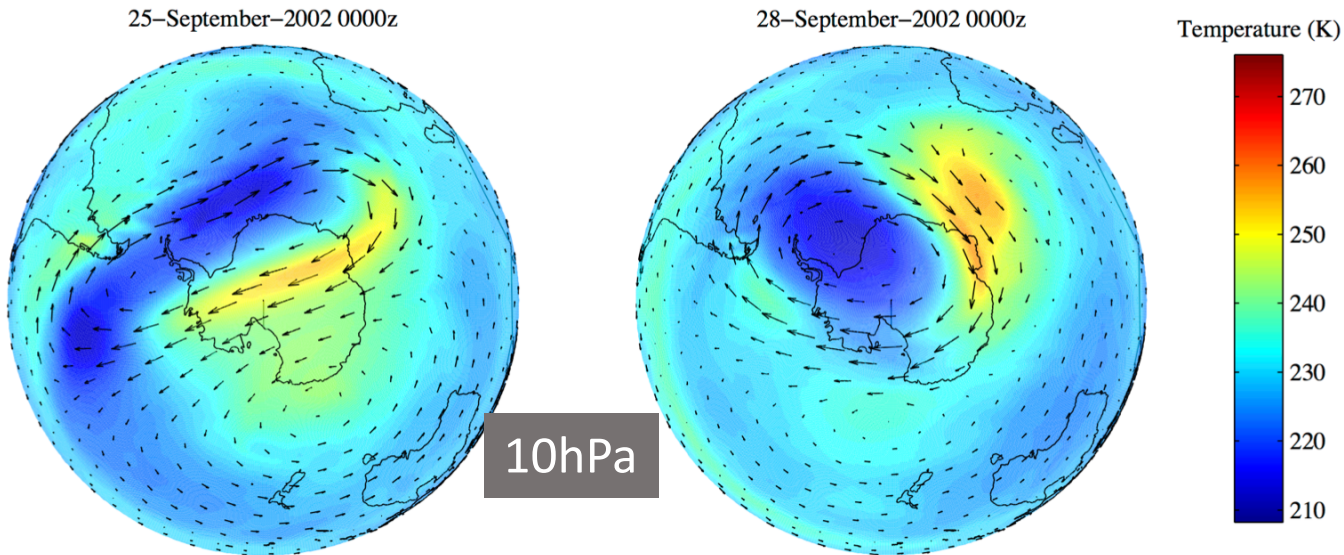
We use the adjoint to investigate the troposphere coupling to the temperature in the box.



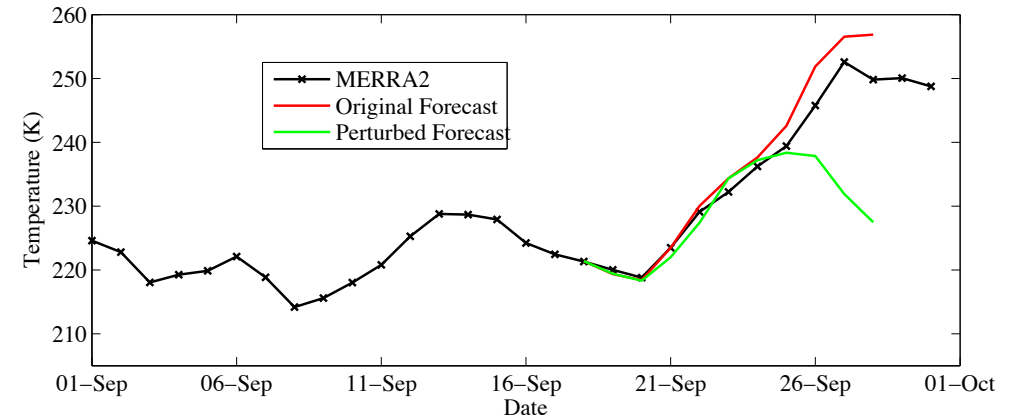
Adjoint derived perturbation.

With a targeted adjoint-derived perturbation near the surface the major SSW is avoided.

The vortex stretches but then returns to normal by September 28th.



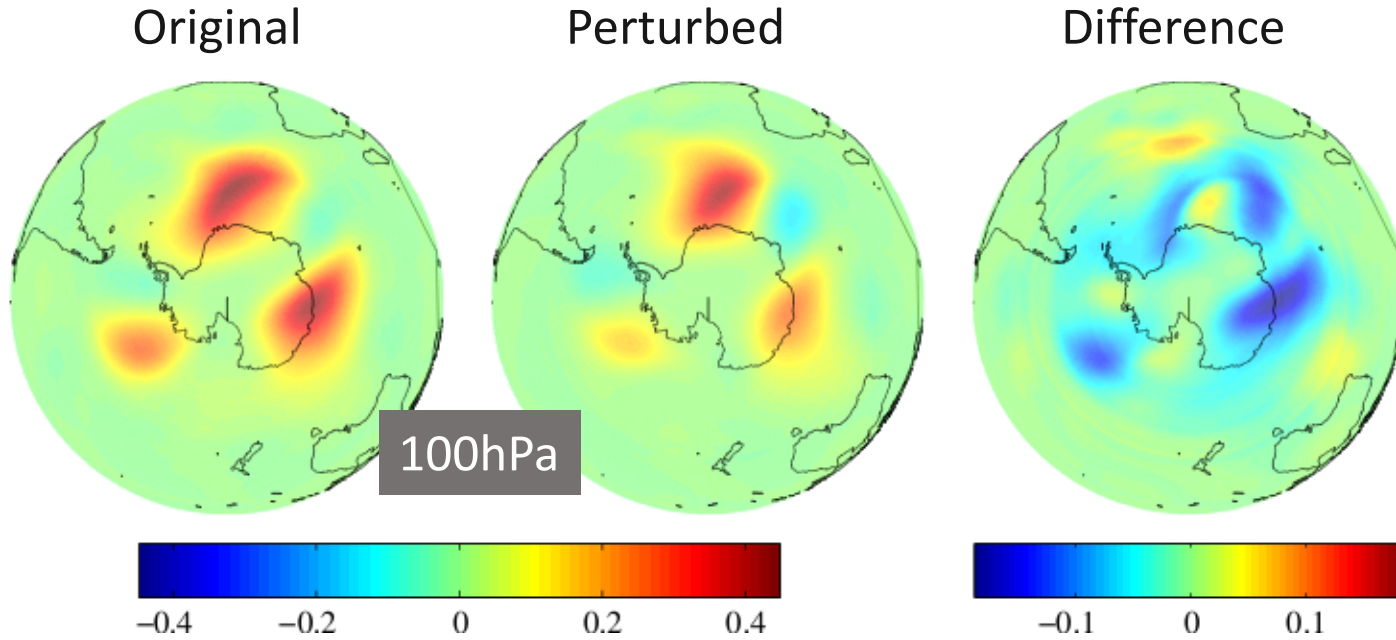
10hPa mean T, 60S – 90S



10 Day forecasts match MERRA-2 well for this metric.

High predictability case means longer adjoint runs.

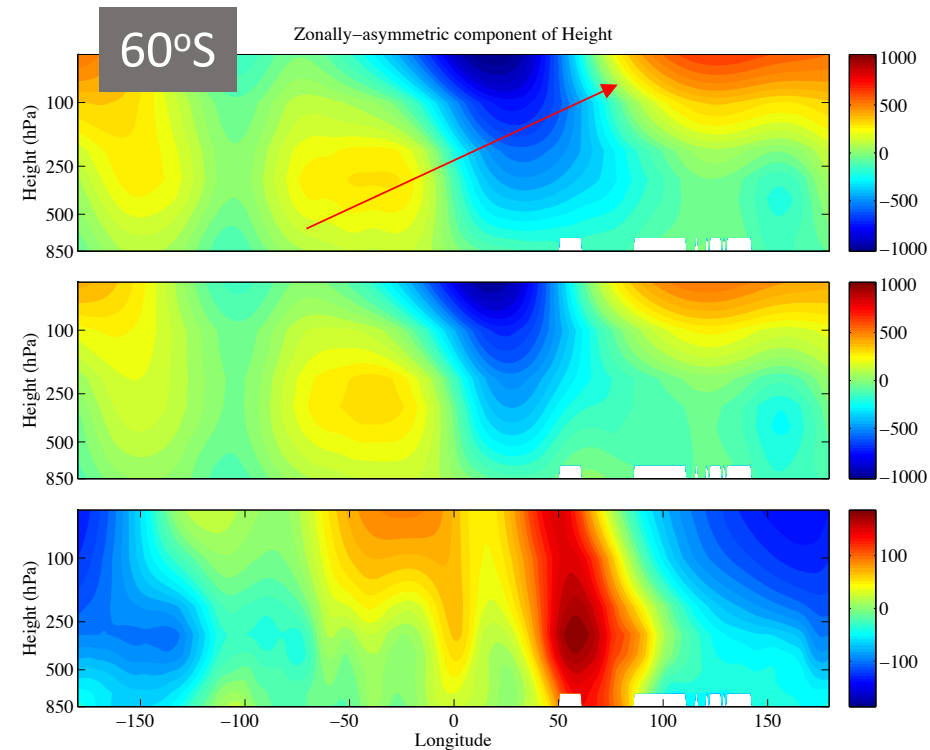
Tracking the perturbation



5 day running mean vertical wave flux (Plumb). Large reduction in amplitude of vertically propagating Rossby waves.

Next step is to track the origin of these waves (in the tropics?) and investigate what was unique about 2002.

5 day running mean of height just ahead of the SSW. Perturbation decreases amplitude of wave.

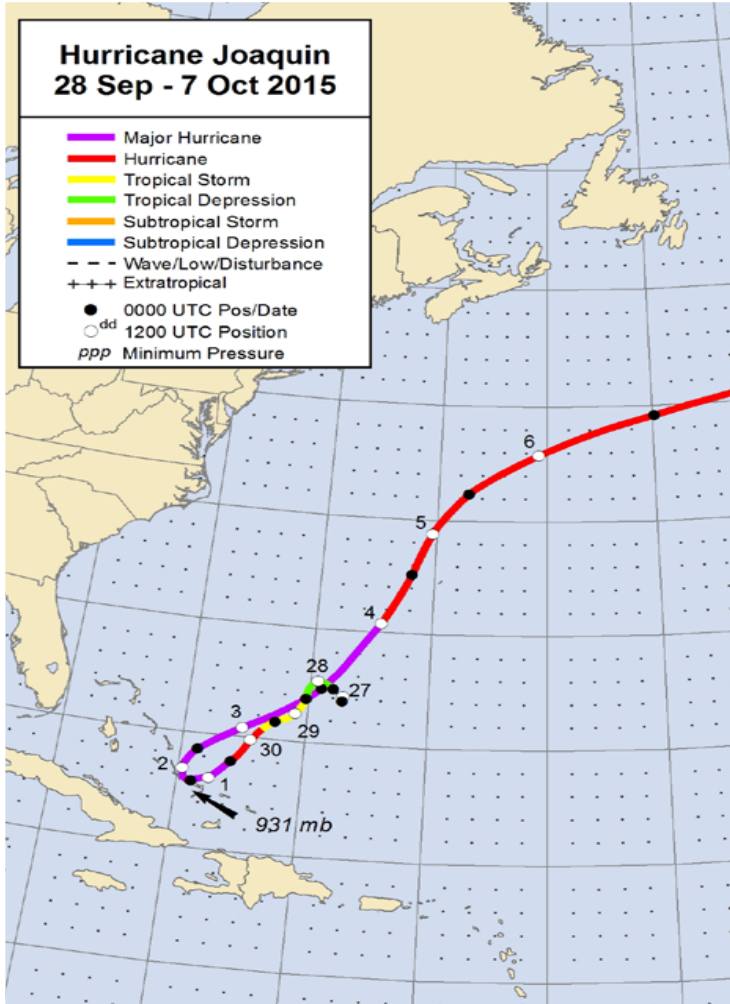




Sensitivity case studies

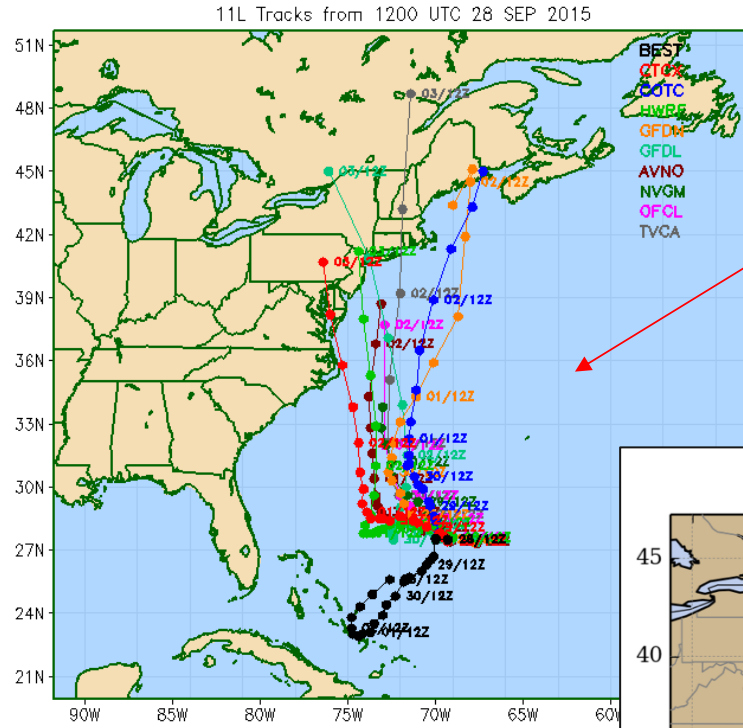
- Sudden stratospheric warming.
- **Hurricane Joaquin (With Jim Doyle NRL).**
- Winter storm rapid intensification (University of Wisconsin).
- Sensitivity to dust sources and radiative effects in hurricane formation.
- TC intensification rate (University of Wisconsin).
- Thinning of AIRS data around hurricanes.

Hurricane Joaquin 2015

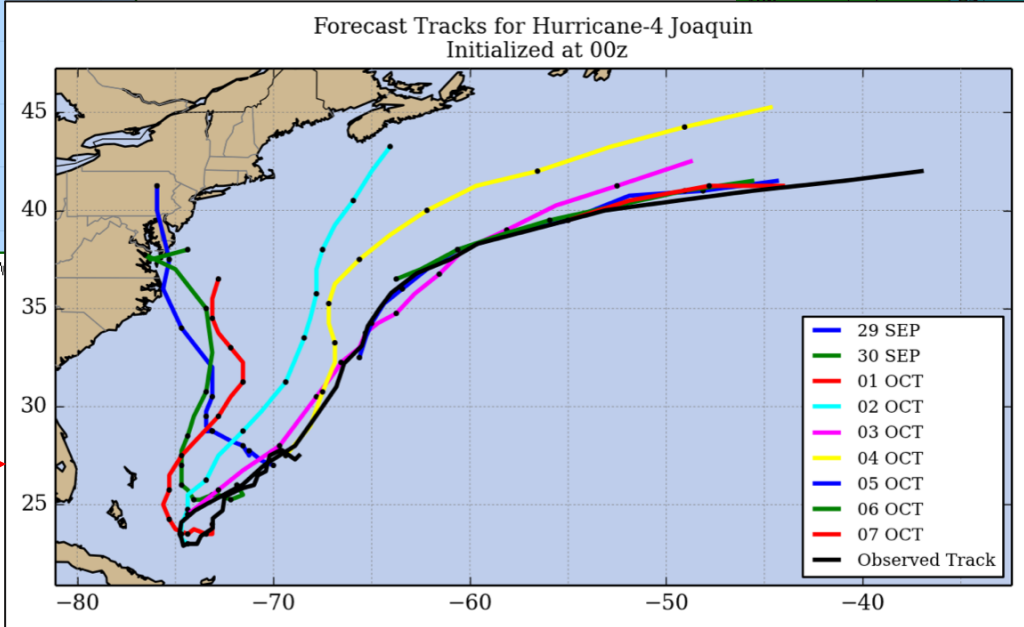
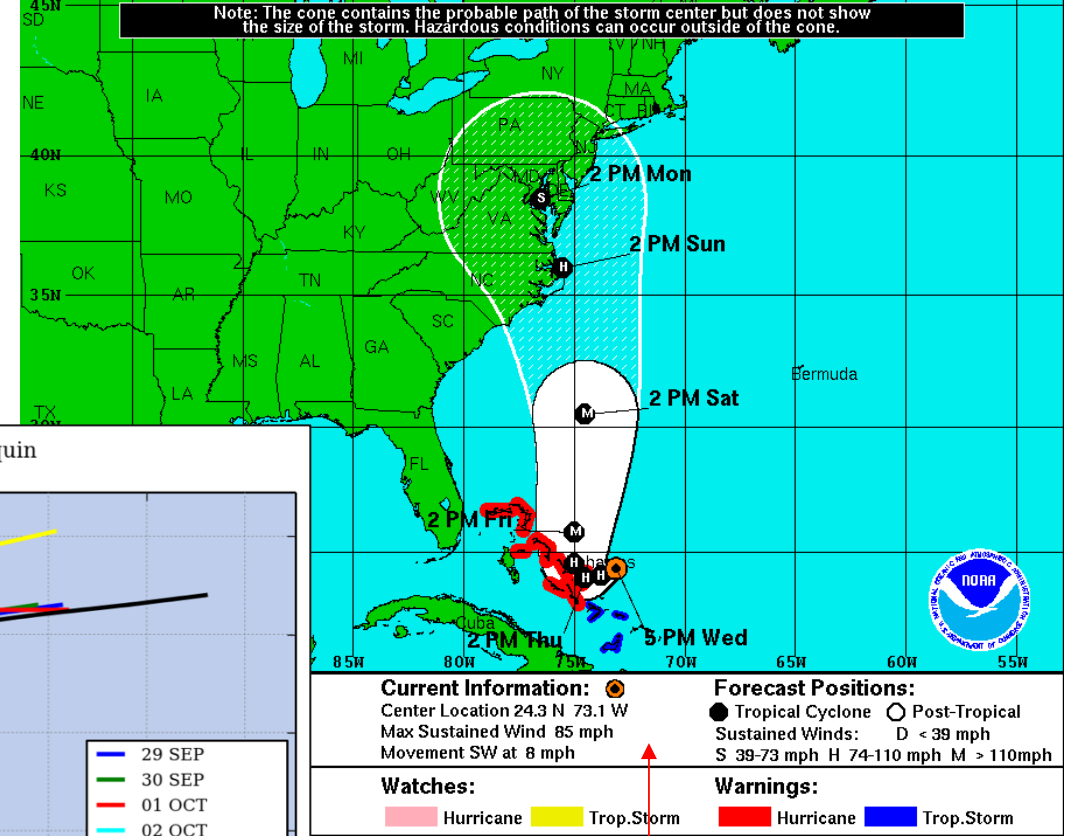


- Category 4 major hurricane that devastated the Bahamas and impacted Bermuda, Turks and Caicos and the Greater Antilles. El Faro went down in eyewall.
- Strongest Atlantic hurricane of non-tropical origin in the satellite era.
- Initially moved south west before making sharp U-turn and traveled back to the north east.
- Models did not accurately predict the southwest motion (except ECMWF).
- Models did not accurately predict the full turn back to the north east (except ECMWF).

Forecasting difficulties



Forecasts of the early part of the development were especially poor

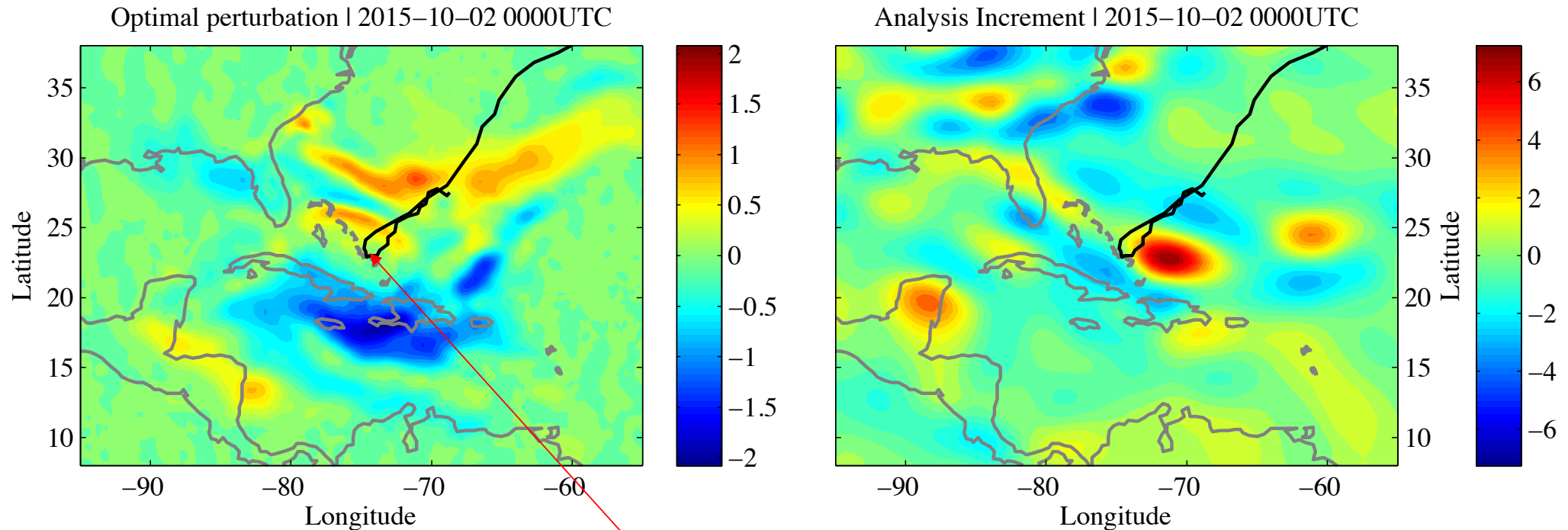


NASA GEOS

NHC forecast issued Wednesday September 30th

Turn phase: comparison with analysis increment

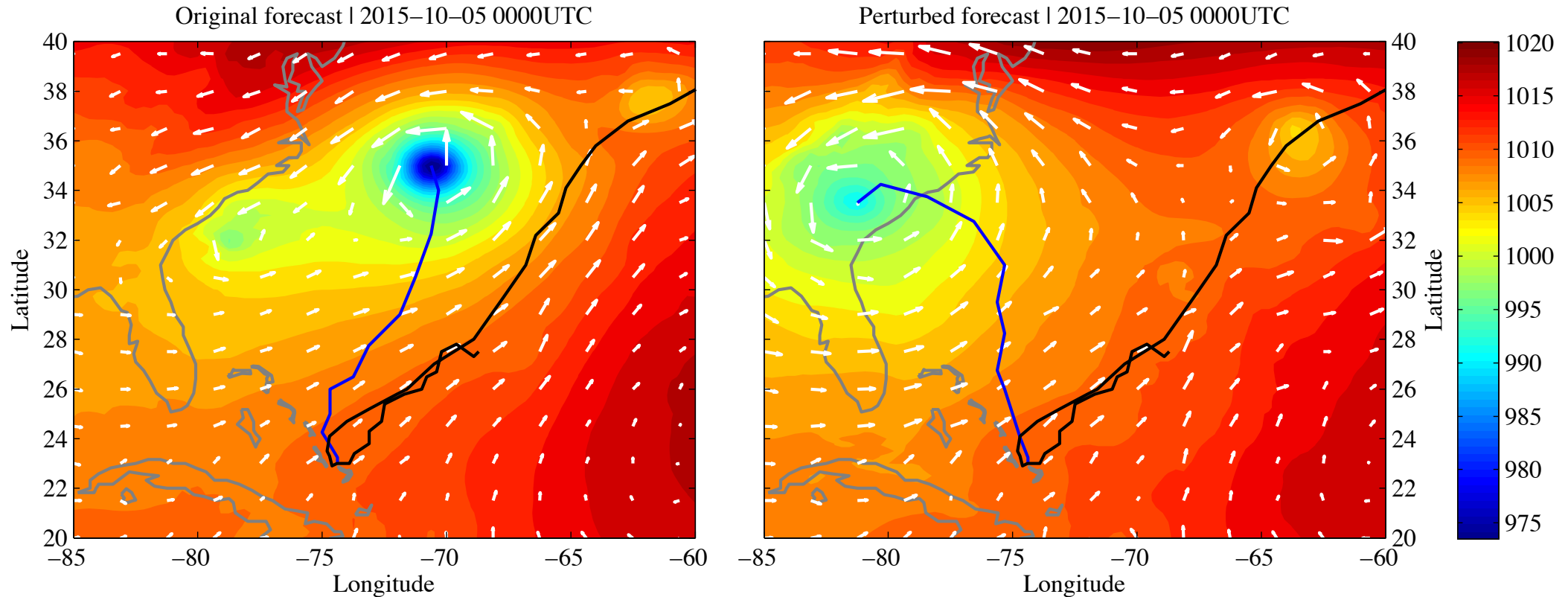
Perturbation of u-wind (ms^{-1}) at 850hPa scaled optimal (left), analysis increment (right)



Center of storm

The adjoint derived perturbation has smaller magnitude than the analysis increment and similar scales.

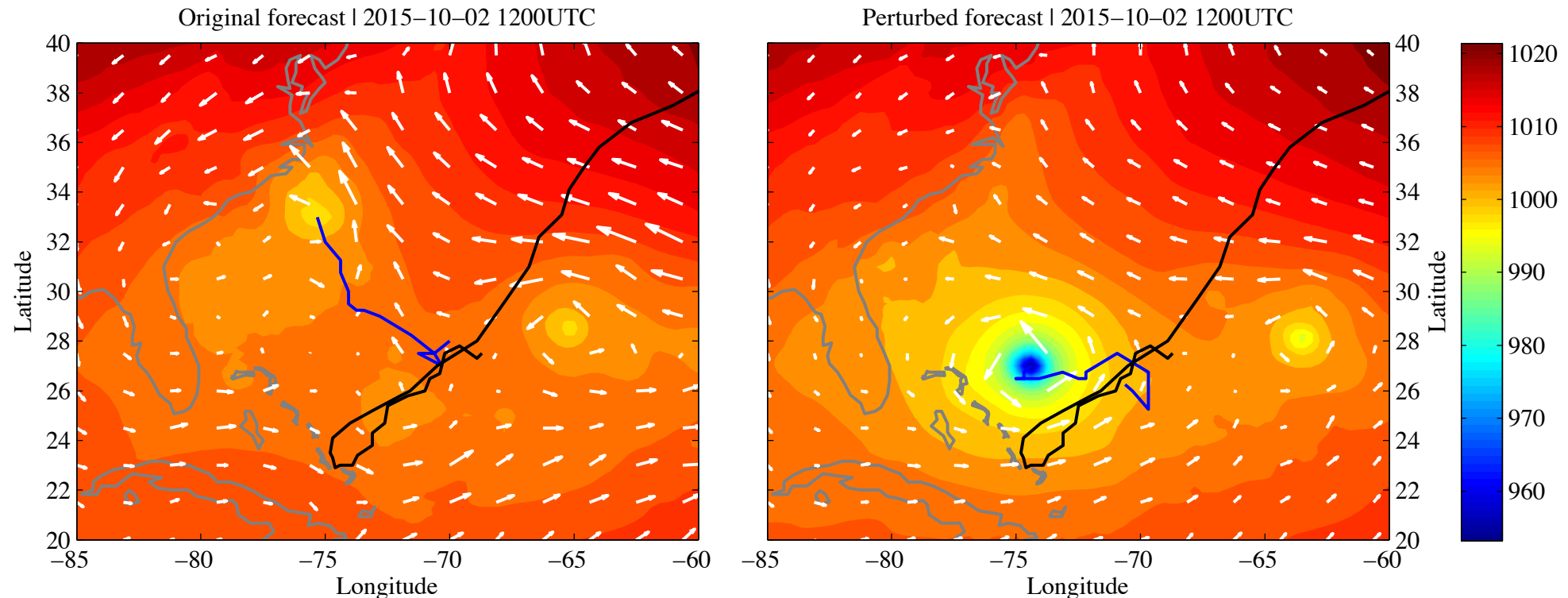
Turn phase: perturbation to weaken



The forecast weakens significantly by making a very small perturbation that slightly weakens the storm.

Early phase: perturbation to strengthen

The general consensus was that models did not sufficiently intensify Joaquin during the early development. So here we optimally perturb to increase strength.

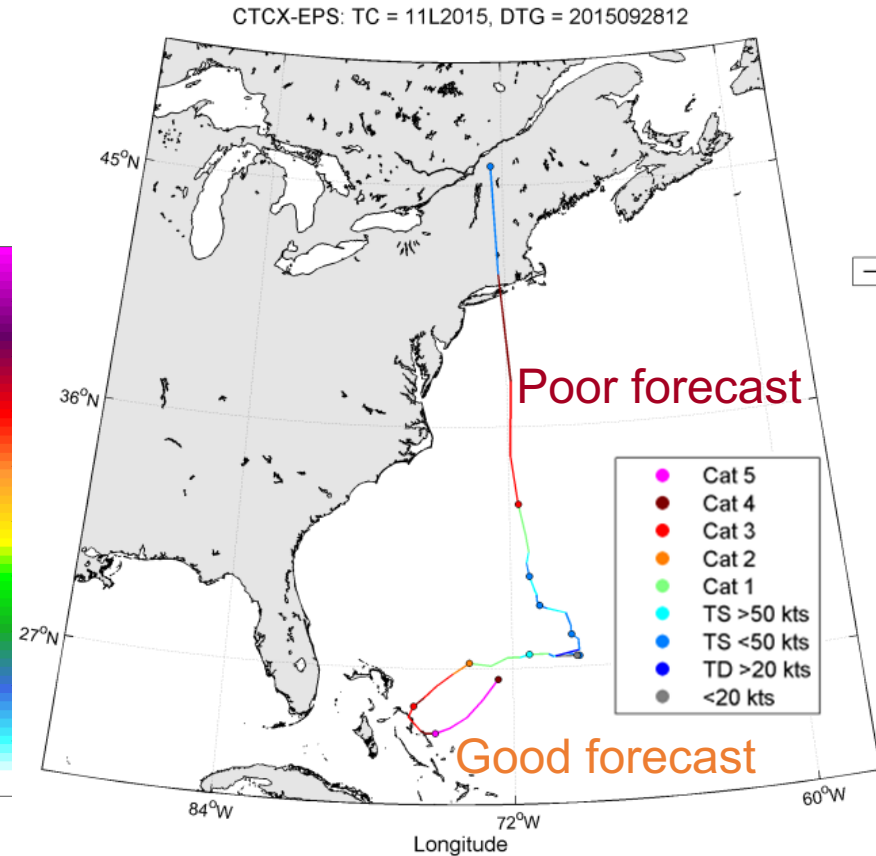
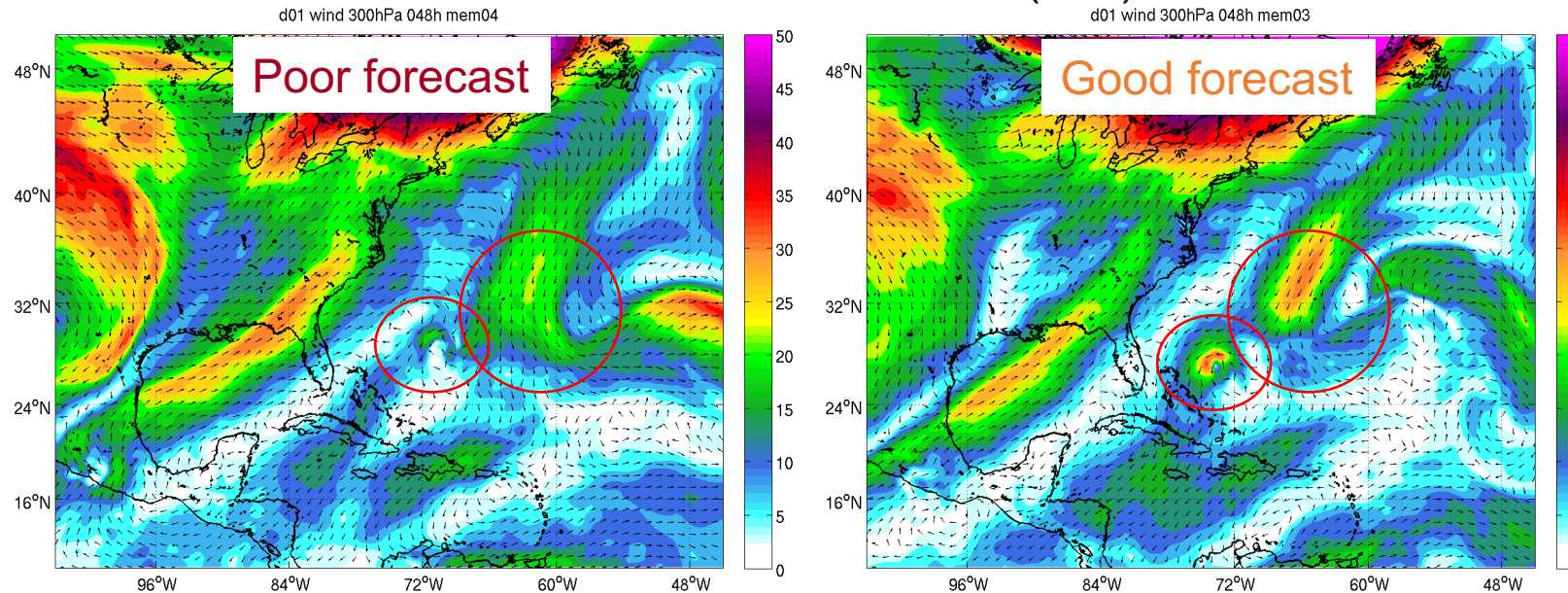


This perturbation results in a better forecast in the sense that the storm intensified and moved on a more south westerly path.

COAMPS-TC: ensemble forecast early development (Jim Doyle NRL)

- 11 member ensemble initialized from 12z on September 26th.
- 3 of the members correctly forecast the south westerly track.

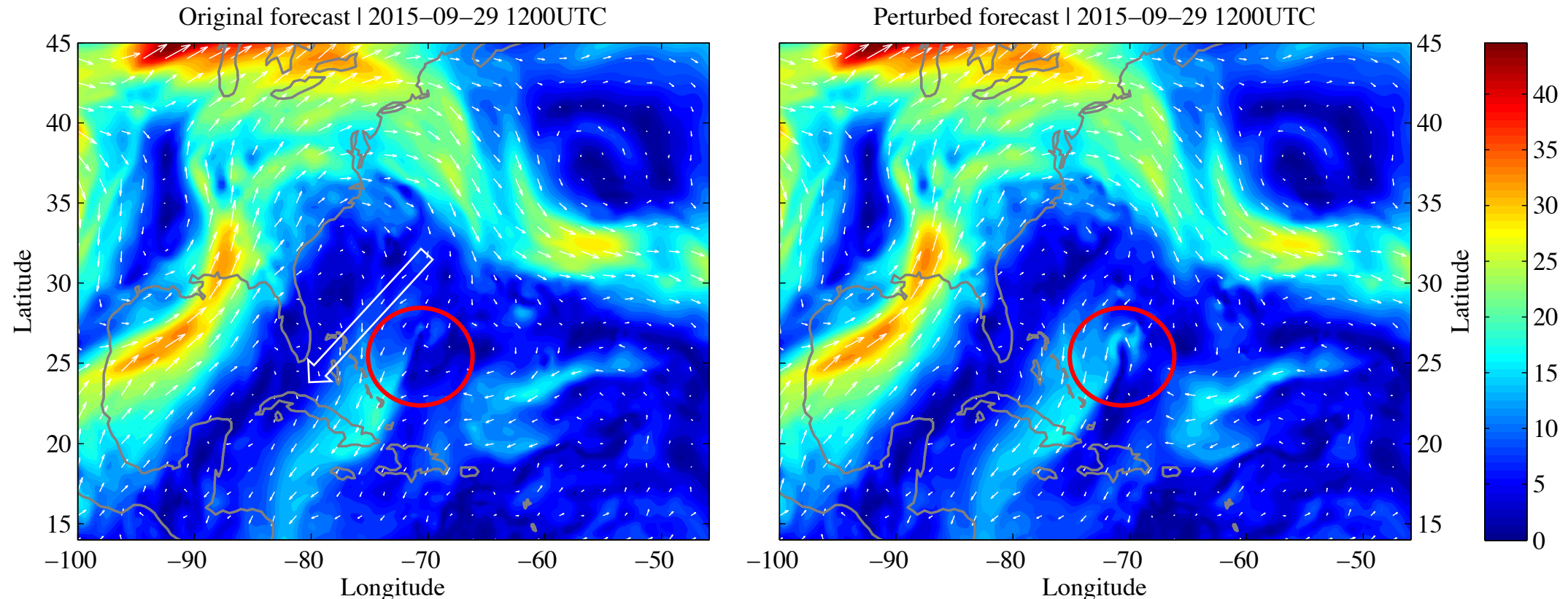
48h forecasts of 300 hPa wind (m/s)



- All 'good' members had stronger upper level trough NE or Joaquin, which develops lower level circulation.
- 'Good' members also developed a stronger hurricane.

Early phase: deepening through the steering flow

24h forecasts of higher level winds (ms^{-1}), original and optimally perturbed, Sep 28th 00z analysis



With the perturbation the circulation reaches higher, through the SW steering flow. But we're still lacking the strong winds seen in the 'good' COAMPS ensemble members.



Conclusions

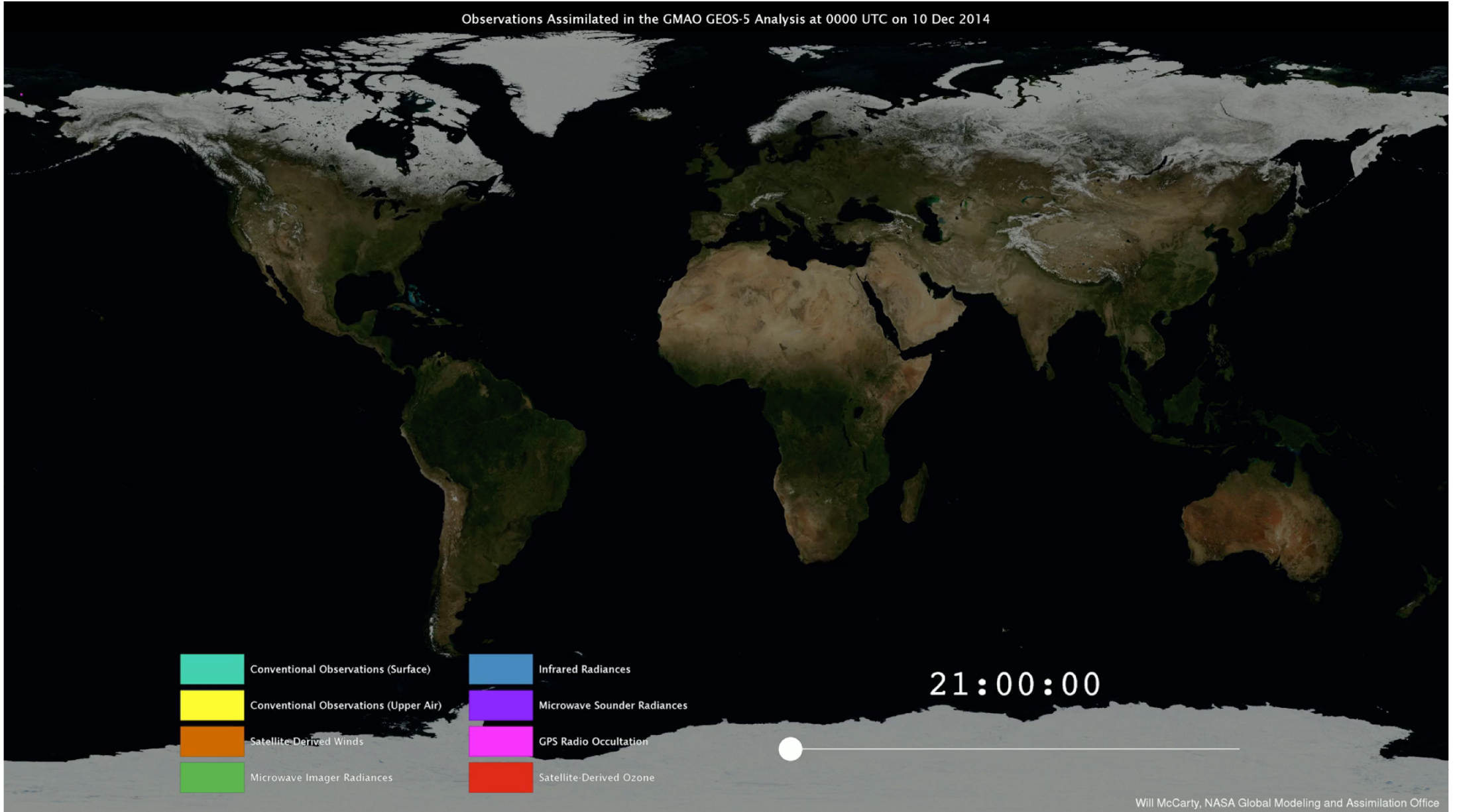
- Over the past 5 years or so GMAO has developed a detailed adjoint version of the GEOS model. The model now uses the latest version of FV3 and has various physics components.
- The main use of the adjoint is in 4DVar but there are many useful operational and research based applications of the tool.
- 4DVar has proven difficult due to the relative cost of FV3 but a move to JEDI looks to alleviate some of this, and bring many other advantages as well.
 - *“Numerical weather prediction is, at its core, an initial value problem, and to increase our predictive skill, we will be addressing three key areas: model code, observations, and computational resources. We need both better and more frequent observations, as well as the model code to assimilate these observations”* – Neil Jacobs (NOAA deputy administrator, WPOST, 2018/04/23).
- Sensitivity studies of all kinds are being examined with the GEOS adjoint within GMAO and by academic partners. These can reveal interesting phenomena and provide insight into the onset of extreme weather and why forecasts sometimes struggle.



BACKUP SLIDES



Observations Assimilated in the GMAO GEOS-5 Analysis at 0000 UTC on 10 Dec 2014

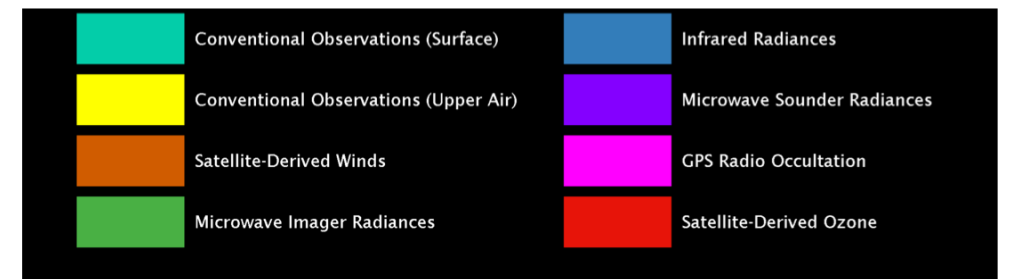
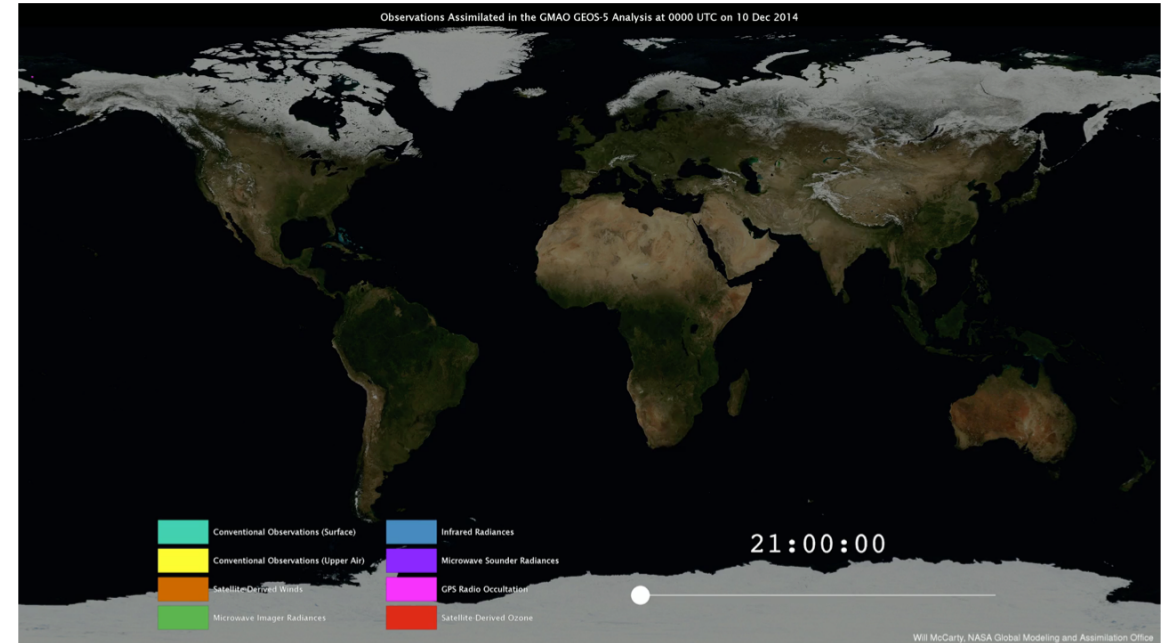


Will McCarty, NASA Global Modeling and Assimilation Office

Data assimilation at GMAO

Gridpoint statistical analysis system:

- EnKF with 32 ensemble members.
- Hybrid so B is mixture of climatology and 'errors of the day'.
- Initialized four times daily.
- 6 hour window to gather observations.
- 25km inner loop in incremental formulation.
- Analysis added using Incremental Analysis Update (IAU).
- SST analysis.



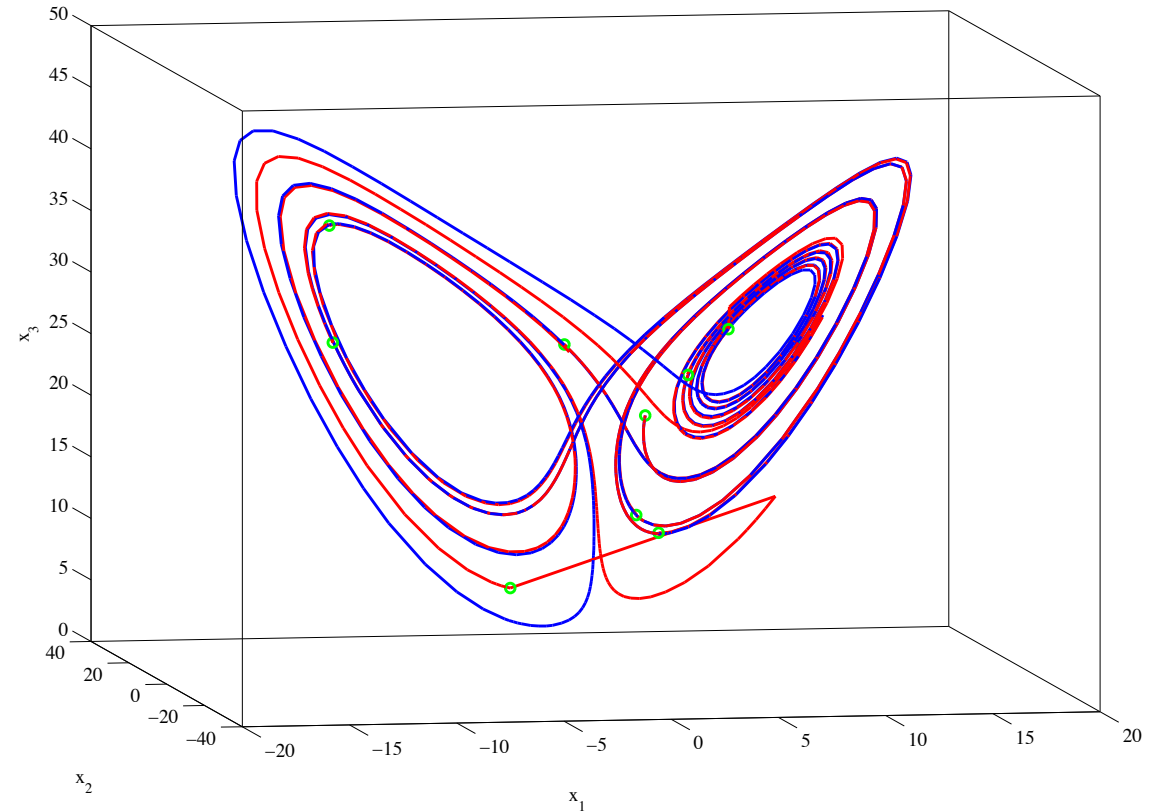
Introduction to data assimilation

In his seminal papers of the 1960s Ed Lorenz showed that the atmosphere exhibits chaotic behavior.

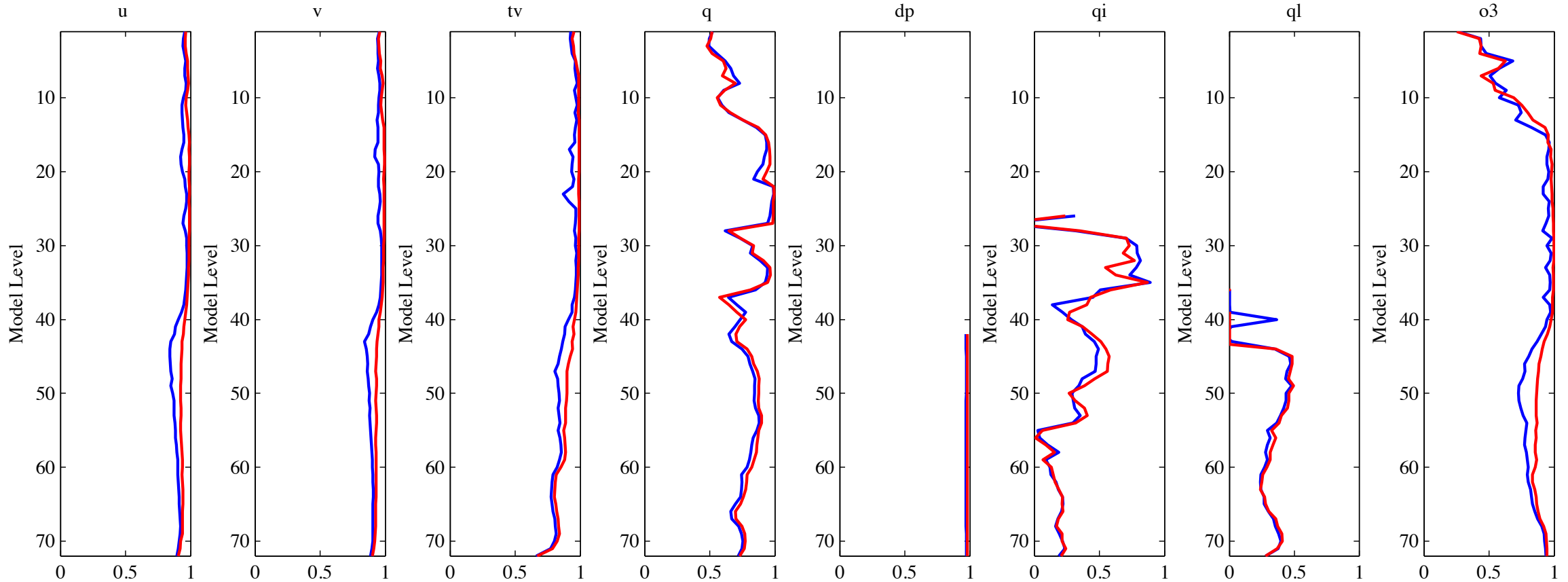
He demonstrated the so-called 'butterfly effect' with a simple three variable problem, shown on the right.

Infinitesimal changes to the initial conditions can grow into very large errors within a relatively short amount of time.

With data assimilation the solution can be kept in check with reality via observations.



TLM validation: analysis increment evolution over 6-hours



$$u' = 0.96, v' = 0.96, T_v' = 0.93, q' = 0.81, \Delta p' = 0.98$$

Observation impacts

Recently the WMO has led a round of Forecast Sensitivity Observation Impacts (FSOI) comparisons between multiple centers.

These compare adjoint and ensemble techniques and show the most important observation types across all centers.

