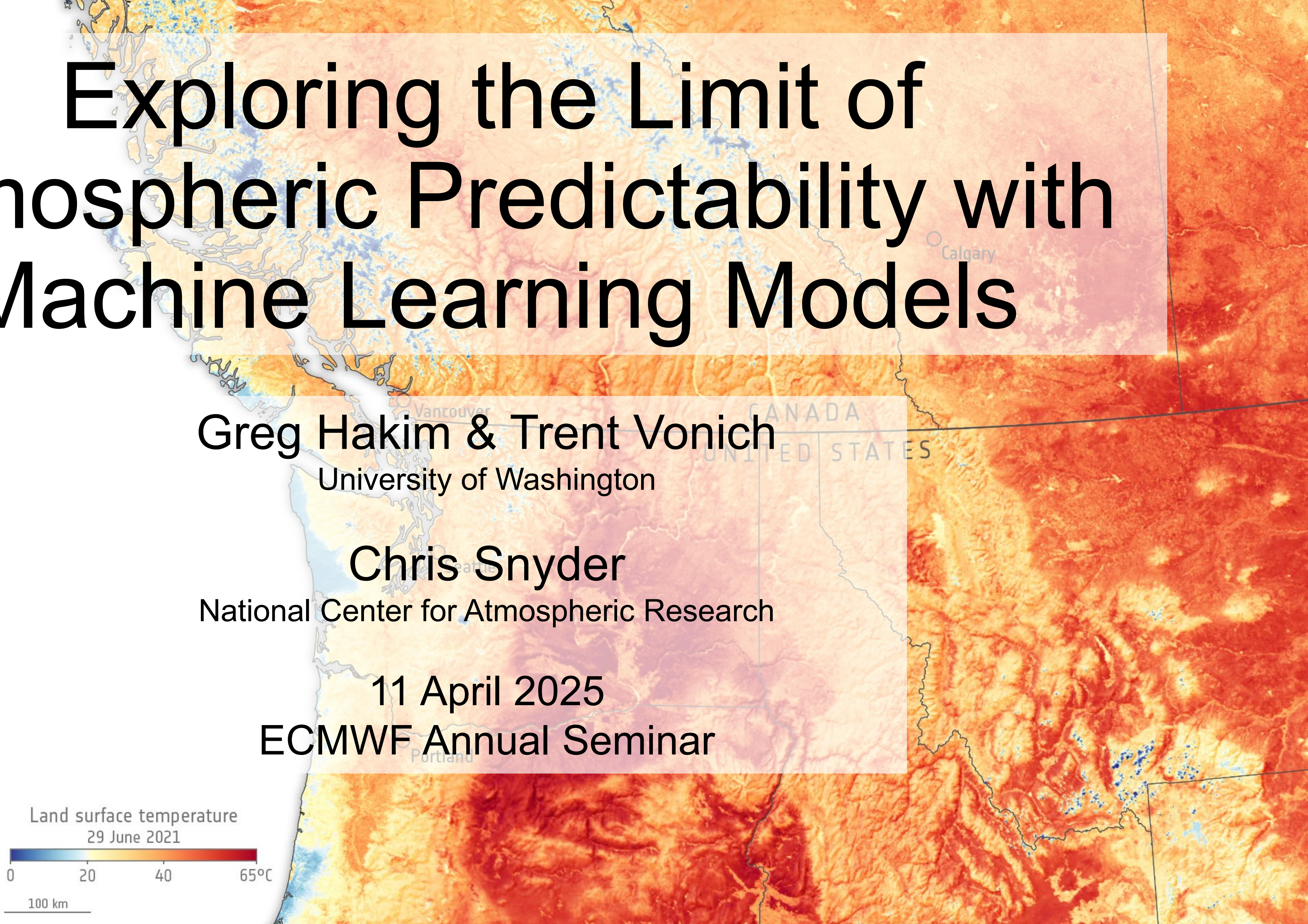
Exploring the Limit of Atmospheric Predictability with Machine Learning Models

Greg Hakim & Trent Vonich
University of Washington

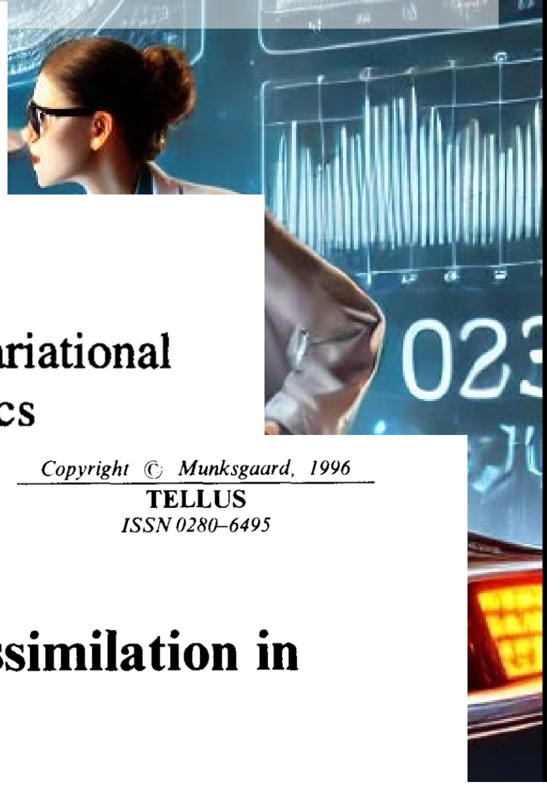
Chris Snyder
National Center for Atmospheric Research

11 April 2025
ECMWF Annual Seminar

HEISING-SIMONS FOUNDATION
NSF



4D-VAR Back to the Future!



Q. J. R. Meteorol. Soc. (1992), **118**, pp. 649–672 551.509.313

Four-dimensional assimilation in the presence of baroclinic instability

Q. J. R. Meteorol. Soc. (2000), **126**, pp. 1143–1170

The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics

By F.

Tellus (1996), 48A, 96–121
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280–6495

On extending the limits of variational assimilation in nonlinear chaotic systems

By CARLOS

Four-dimensional variational assimilation and predictability in a quasi-geostrophic model

By KYLE SWANSON* and ROBERT VAUTARD, *Laboratoire de Météorologie Dynamique, Ecole Normale Supérieure, 24 Rue Lhomond, 75231 Paris Cedex 05, France,* and CARLOS PIRES, *2 Faculdade de Ciencias, Departamento de Fisica, Edificio C1, Piso 4, Campo Grande, 1700 Lisboa, Portugal*

(Manuscript received 21 July 1997; in final form 20 February 1998)

Atmospheric Predictability

- Theory and prior results: limiting time is set by small scales
  - ~14 days, known and confirmed repeatedly since Lorenz (1969)

“the predictability speed of light”

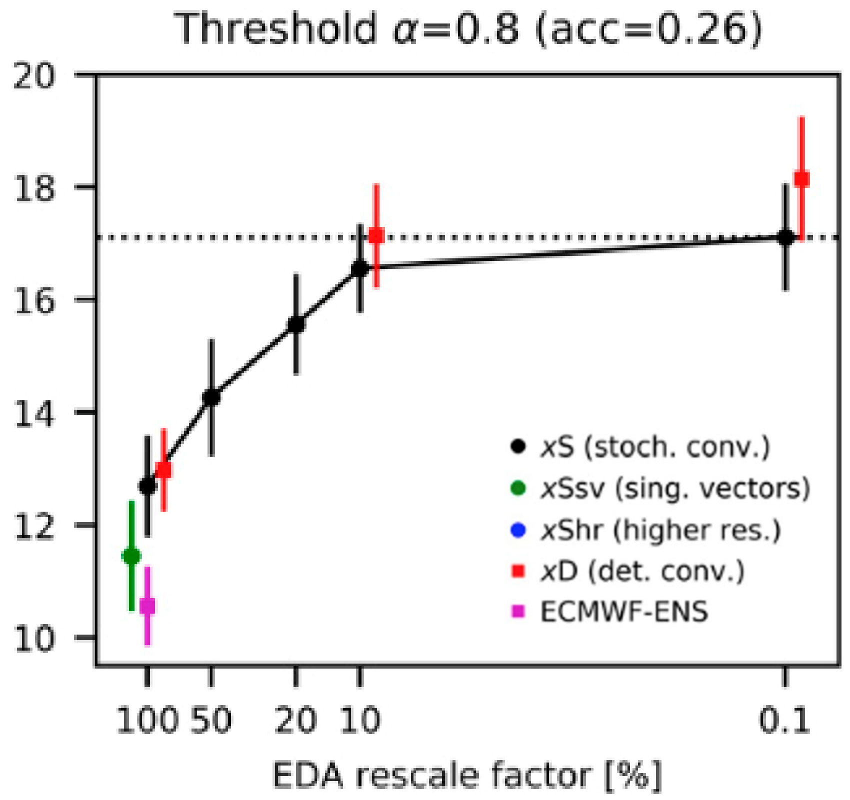
- Testing the limit with new ML models & gradient sensitivity
- Pacific Northwest heatwave of June 2021 (Vonich & Hakim, 2024)
- Large sample shows forecast skill beyond 30 days (Vonich & Hakim, 2025)
- Speculation / current work: Shadowing Lemma & model error

Atmospheric Predictability Theory (Lorenz 1969)

energy spectrum power laws control error growth

- –3 (large scales): unlimited predictability
- –5/3 (mesoscales): finite limit ~ 14 days

pred. time (days)



Selz et al. (2022)

Atmospheric Predictability Theory (Lorenz 1963)

- "chaos"
- tangent linear, $t \to \infty$, stability
- errors grow exponentially at the leading Lyapunov exponent
- very clear in ML models

New Opportunities from ML Models

- ML models provide a new approach to predictability
  - no -5/3 spectrum; strongly damped at small scales (Bonavita 2024)
  - different error growth at short leads (Selz & Craig 2023)

- ML forecast skill beyond the limit of physics models suggests that predictability of the **true system is not limited by small scales.**

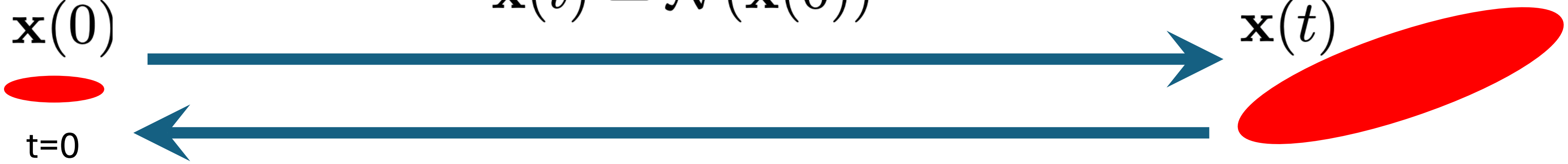
- Gradient facilities to compute derivatives of all aspects (model & state)

- Very computationally efficient (large ensembles)

Deep Learning Sensitivity Analysis

$$J(\mathbf{x}(t)) = \frac{1}{2}\left(\mathbf{x}(t) - \mathbf{x}_r\right)^{\mathrm{T}}\left(\mathbf{x}(t) - \mathbf{x}_r\right)$$

$$\mathbf{x}(t) = \mathcal{N}(\mathbf{x}(0))$$

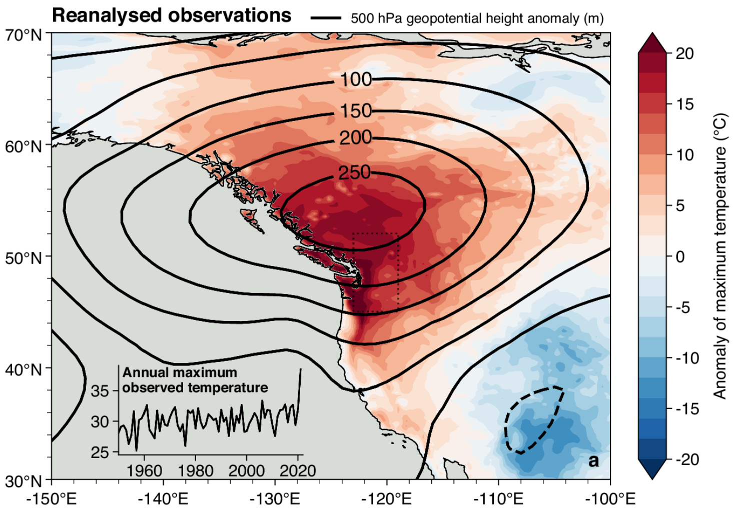


$$\mathbf{x}_{i+1} = \mathbf{x}_i - \eta \frac{\partial J(\mathbf{x}_i(t))}{\partial \mathbf{x}(0)}$$

Map large errors (future) to when they are small (analysis error)

Application: Pacific Northwest Heatwave (2021)

- Top-6 global extremes since 1960 (Thompson et al. 2022)
- Highest Canadian temperature (49.6°C, Lytton, British Columbia)
- >1400 deaths
- Not in ML training data



Leach et al. (2024)

Optimal initial condition evaluation

- GraphCast 1° model (Lam et al. 2023)
- Loss: ~weighted mean-squared error
- Derivatives: JAX framework (Bradbury et al. 2018)
- Initial conditions and verification: ERA5

10-day forecast from ERA5

~90% error reduction

Vonich & Hakim (2024)

10-day forecast from ERA5

Vonich & Hakim (2024)

Generality beyond the Heatwave Case

GraphCast initial-condition optimization daily for 2020

Improvements to Algorithm

- Optimize sequentially
  - optimize short leads first, then progress to longer leads
  - requires signal > analysis error (~2 days)
  - "quasi-static variational assimilation"
- Twice daily for 2020: n = 732

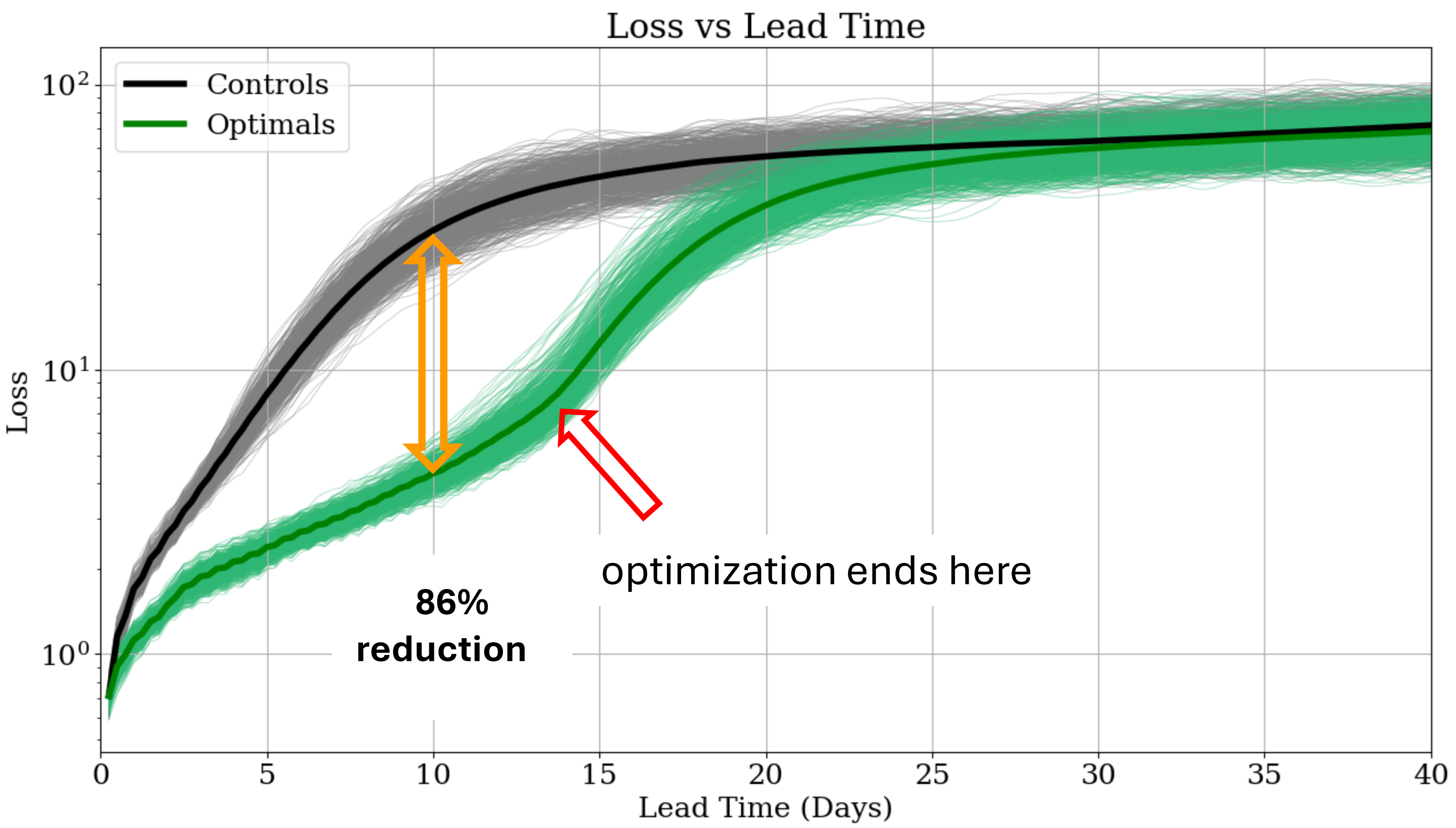
Tellus (1996), 48A, 96–121
Printed in Belgium – all rights reserved

Copyright © Munksgaard, 1996
TELLUS
ISSN 0280–6495

On extending the limits of variational assimilation in nonlinear chaotic systems

By CARLOS PIRES,* ROBERT VAUTARD and OLIVIER TALAGRAND, *Laboratoire de Météorologie Dynamique du CNRS, Paris, France*

Optimal Predictability

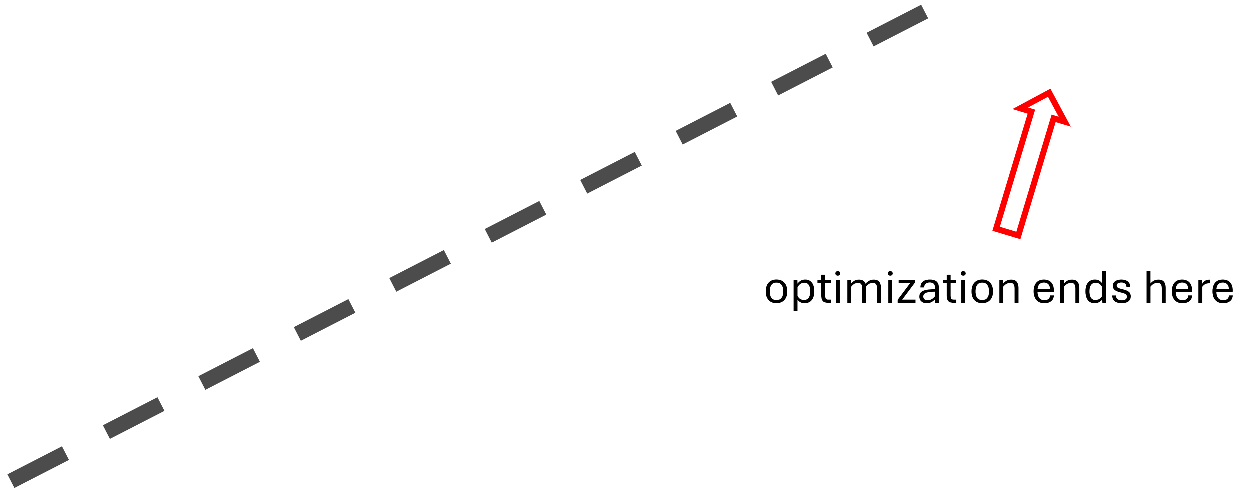


Every case has large improvement

Errors grow like the control after optimization stops

float-32 precision limit

Optimal Predictability Limit ~35+ days



Float precision matters!

- float32 fails ~14 days
- float64 fails ~32 days (out of memory)

Single error growth rate ~3-20 days

- doubling t ~5+ days
- cf. perfect twins ~1 day

Faster error growth for first 3 days

Slower error growth days 20-35

GraphCast Optimals: Time-Mean Structure

Main corrections to ERA5:
Intensification of the
Hadley circulation

GraphCast Optimal Initial Conditions

- Corrections to ERA5 initial conditions that fit a long trajectory
- Forecasts from these initial conditions have skill to ~35-40 days
- Error doubling time of ~5-6 days

Two questions:

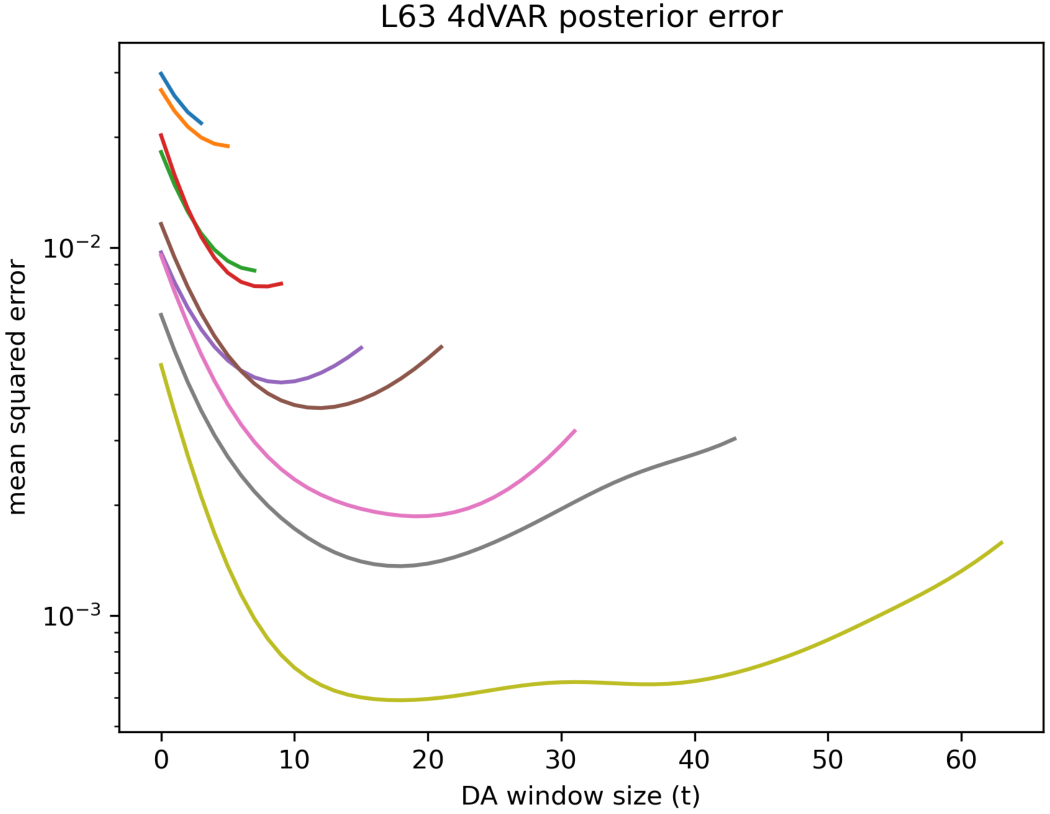
Q1) What is the role of model error?

Q2) How do the optimals avoid chaos & upscale error growth?

Pangu-Weather Forecasts from GC Optimals

- Float32 (optimized to 14 days)
- Significant improvements
- Much smaller than for GraphCast
- Suggests importance of model error

Idealized Error growth/decay on long windows



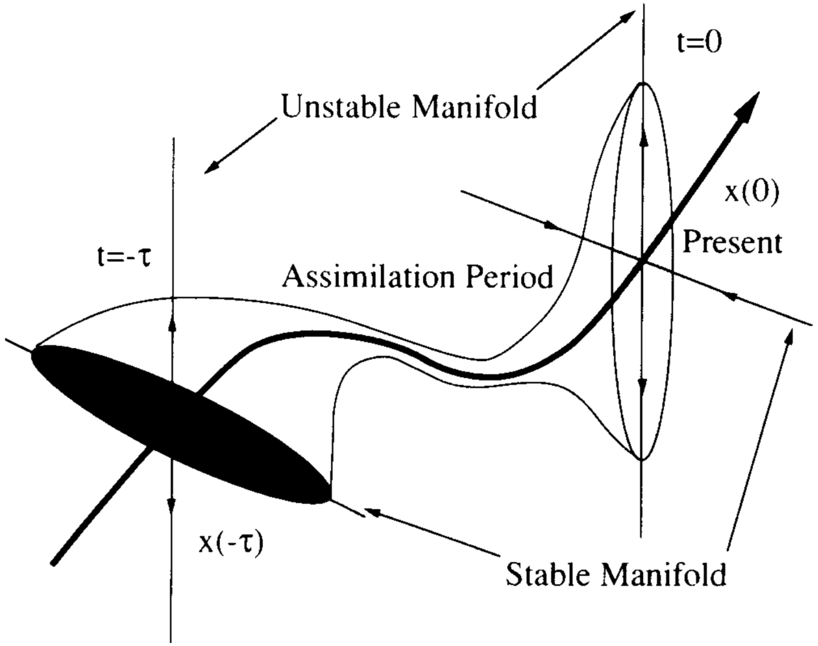
Lorenz '63 4D-VAR averaged over many DA cycles

Tellus (1996), 48A, 96–121
Printed in Belgium – all rights reserved

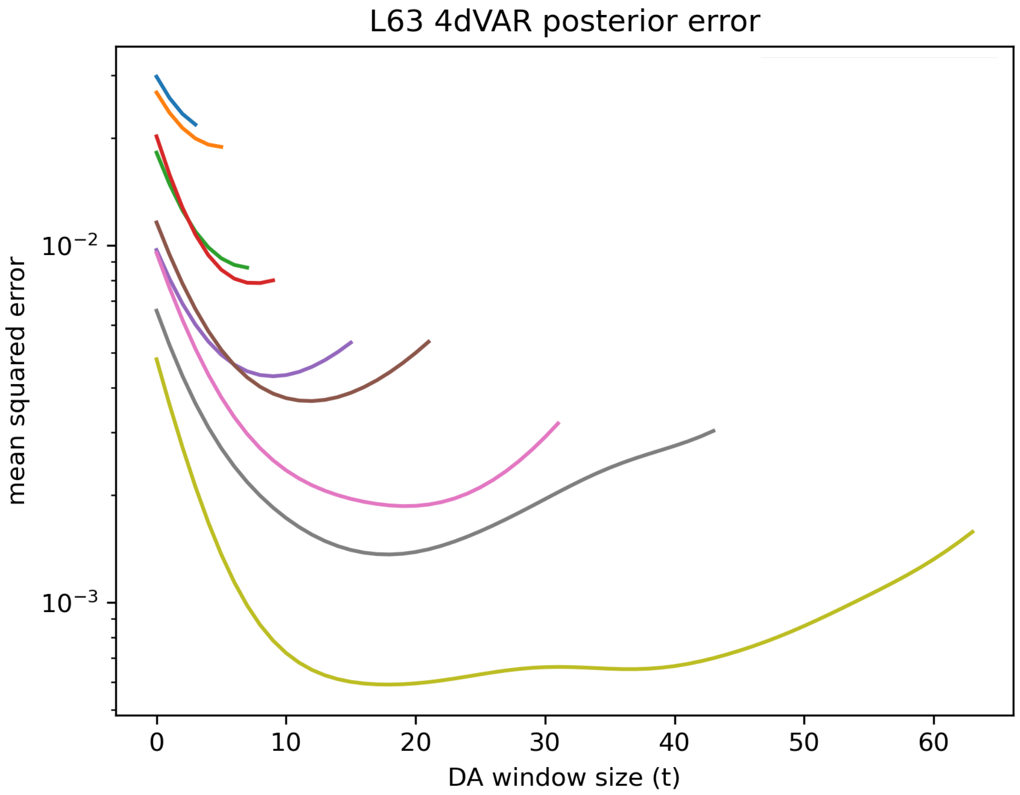
TELLUS
ISSN 0280–6495

On extending the limits of variational assimilation in nonlinear chaotic systems

By CARLOS PIRES,* ROBERT VAUTARD and OLIVIER TALAGRAND, *Laboratoire de Météorologie Dynamique du CNRS, Paris, France*



But this is not what we see in the GC optimals...



Not the same picture---why?

Q2) How do the optimals avoid chaos?

Shadowing Lemma

- For chaotic systems, how do we know if a numerical solution is valid?
  - numerical algorithm, hardware, compiler, & finite numerical precision
  - errors grow exponentially in time
- The shadowing lemma proves that numerical solutions are valid
  - pseudo-orbit: one-step predictions converge within δ of true solution
  - a true orbit ε-shadows the pseudo-orbit for nearby initial condition

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 19, Number 2, October 1988

NUMERICAL ORBITS OF CHAOTIC PROCESSES
REPRESENT TRUE ORBITS

STEPHEN M. HAMMEL, JAMES A. YORKE AND CELSO GREBOGI

THEOREM 2. *Let* $\mathbf{g}$ *be defined by* (2), *with* $\alpha = 5.5$, $\gamma = 0.85$, $\kappa = 0.4$, $R = 0.9$, *and* $\mathbf{p}_0 = (0,0)$. *For* $N = 10^6$, *the pseudo-orbit* $\{\mathbf{p}_n\}_{n=0}^{N}$ *with* $\mathbf{p}_{n+1} = \mathbf{T}(\mathbf{g}(\mathbf{p}_n))$, *is* δ_x*-shadowed by a true orbit within* $\delta_x = 10^{-8}$.

Lorenz ‘63 Long Shadows

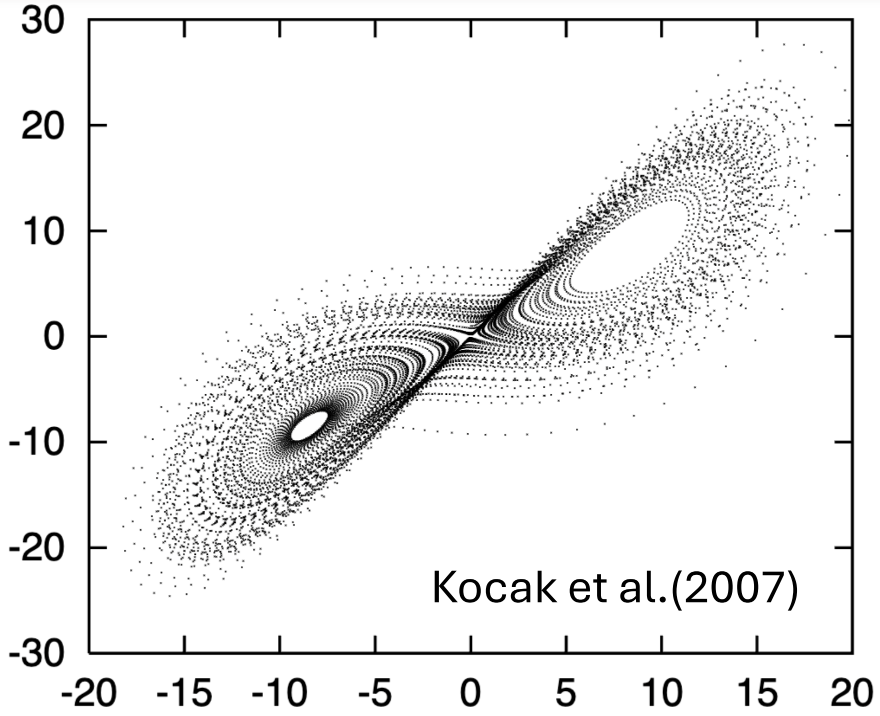
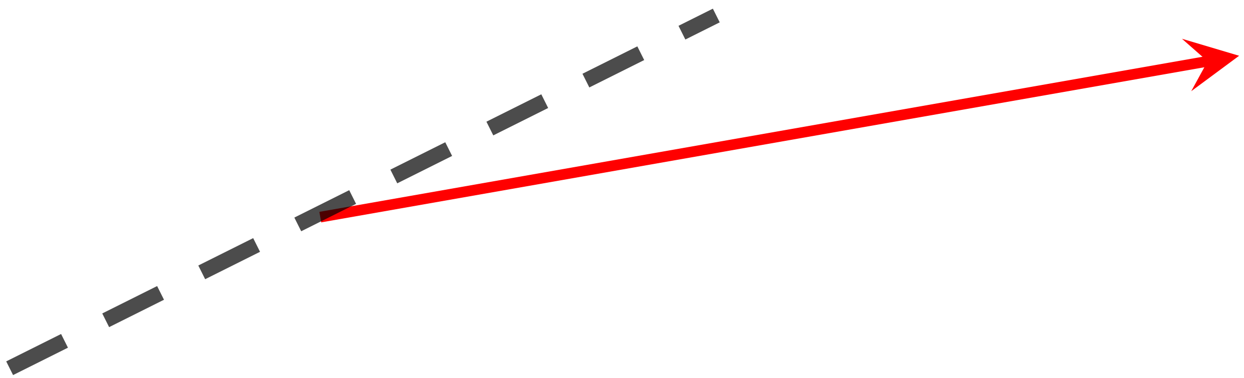


Figure 1: A pseudo orbit, with $\delta \leq 1.978 \times 10^{-12}$, of the Lorenz Equations for the classical parameter values $\sigma = 10, \beta = 8/3, \rho = 28$ and initial data $(0,\ 1,\ 0)$ projected onto the $(x,\ y)$-plane. There exists a true orbit within $\varepsilon \leq 2.562 \times 10^{-9}$ of this pseudo orbit. For clarity, we have plotted only the first 120 time units of the pseudo orbit. Shadowing of this pseudo orbit for much longer time is possible.

Application to NWP: Judd (2008)
- limited to 48h window

850,000 time units!

GC optimals quasi-shadow ERA5?



This linear growth from ~small model error (manifold drift)?

A slightly perturbed version of GC would reduce loss further?

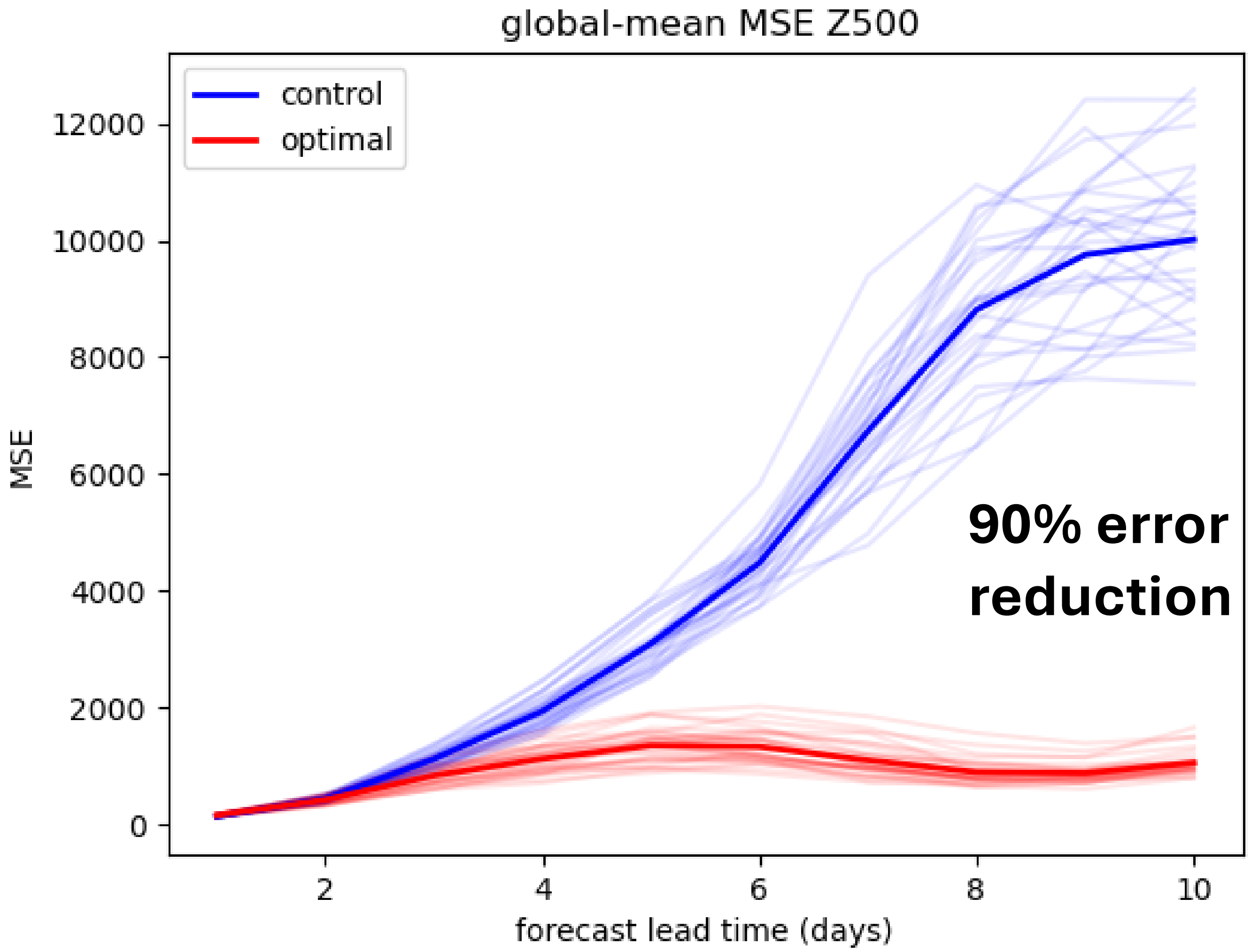
Joint IC/parameter optimization

Summary & Outlook

- Gradient sensitivity over long windows yields deterministic predictability limits
- Small corrections to ERA5: ~80—90% reduction in 10-day forecast error
    - on the order of analysis error
    - average optimal strengthens the Hadley circulation in ERA5
    - error doubling time ~5-6 days (cf ~1-1.5 days for operational forecasts; ML twins)
- **Deterministic predictability limit > 30 days**
- An interpretation: The optimals are shadowing trajectories
    - But drift from ERA5 due to model error; suggests state—model optimization
- Outlook
    - **reanalysis**: seems like a compelling application (optimal analysis & corresponding tuned models)
    - **forecasting**: less clear, but several possibilities

Happy 50th Birthday, ECMWF!

Global-mean optimal vs control MSE 500Z



loss (J)

epoch

500Z Perturbation Evolution

Vonich & Hakim (2024)