COUPLED DATA ASSIMILATION:

What we need from observations and modellers to make coupled data assimilation the new standard for prediction and reanalysis.

Prof. Stephen G. Penny
University of Maryland
NOAA/NCEP, RIKEN AICS, ECMWF
US CLIVAR Summit, 9 August 2017

DIDACTIC DIATRIBE

- A message to modellers:

Reanalysis $\neq$ observations

DIDACTIC DIATRIBE

- A message to modellers:

Reanalysis ≠ observations

- A message to observationalists:

Reanalysis ≠ models

DIDACTIC DIATRIBE

- A message to modellers:

Reanalysis ≠ observations

- A message to observationalists:

Reanalysis ≠ models

A little better:

$$\mathbf{x}_{\text{analysis}} = w_1 \mathbf{x}_{\text{observed}} + w_2 \mathbf{x}_{\text{modelled}}$$

DIDACTIC DIATRIBE

- A message to modellers:

Reanalysis ≠ observations

- A message to observationalists:

Reanalysis ≠ models

A little better:

$$\mathbf{x}_{\text{analysis}} = w_1 \mathbf{x}_{\text{observed}} + w_2 \mathbf{x}_{\text{modelled}}$$

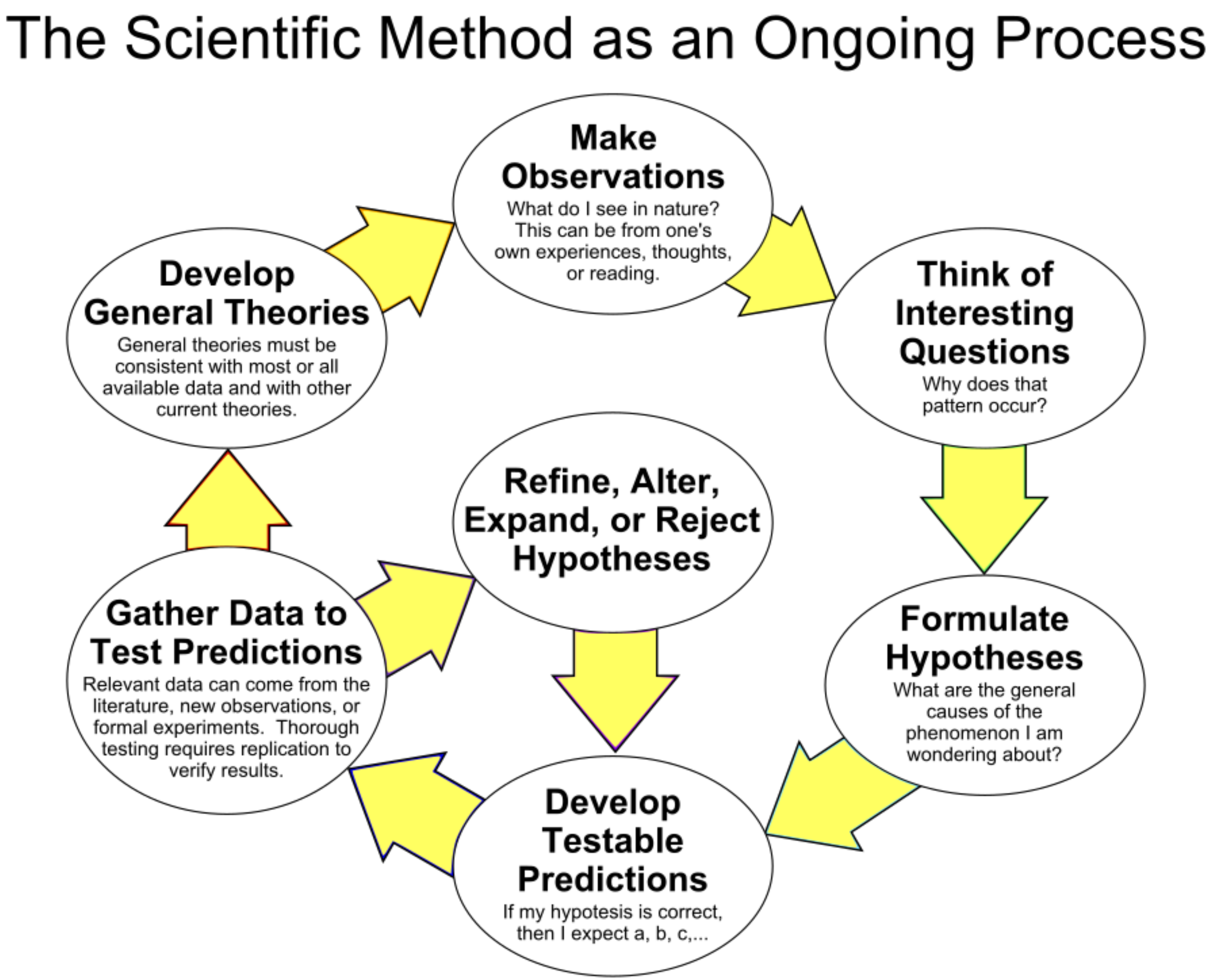
A little better than that:

$$\mathbf{x}_{\text{analysis}}^{t_i} = w_1 \mathbf{x}_{\text{observed}}^{t_i} + w_2 M(\mathbf{x}_{\text{observed}}^{t_0:t_{i-1}})$$

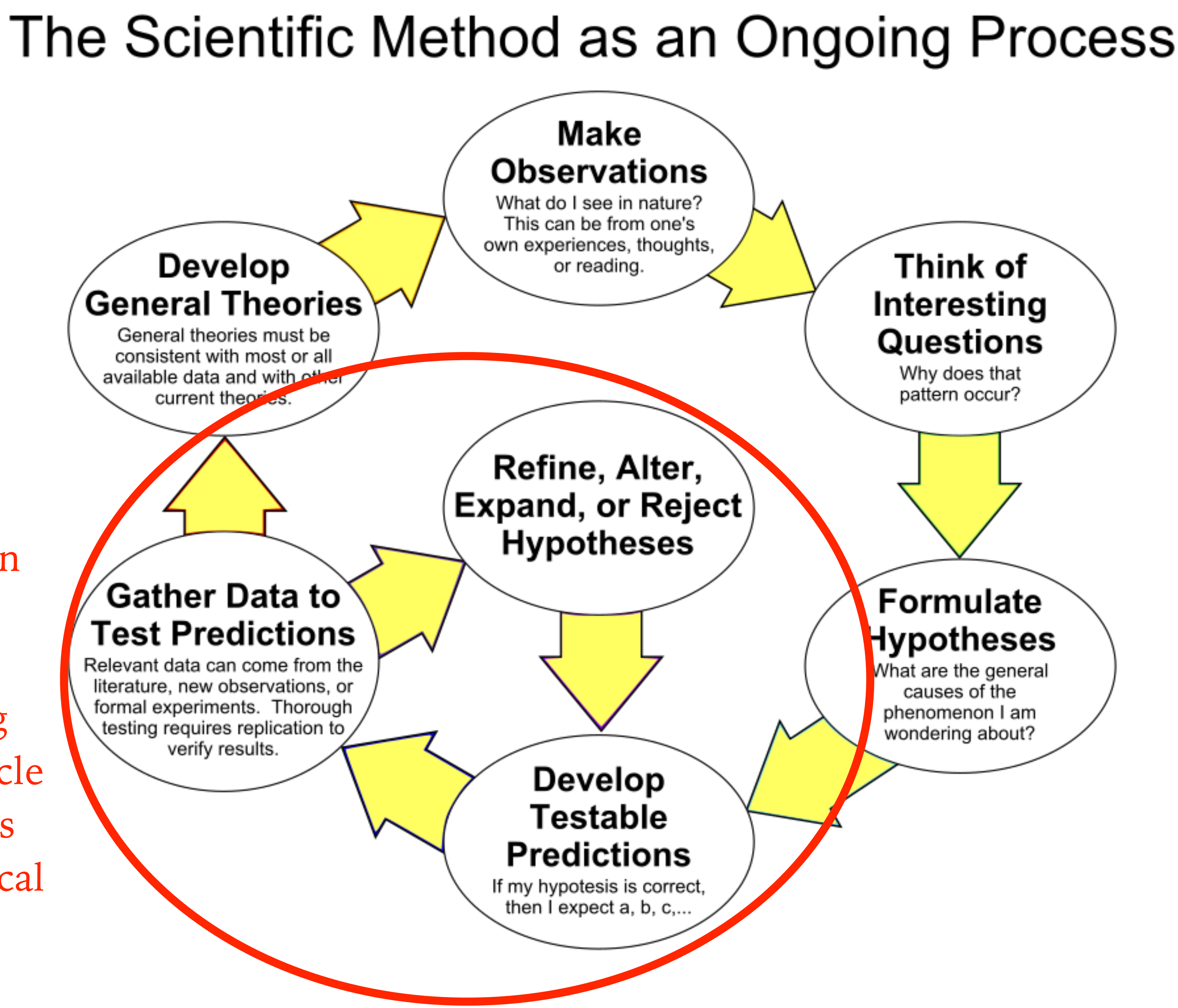
WHAT WE NEED FIRST FROM DEVELOPERS OF OBSERVING SYSTEMS AND GEOPHYSICAL MODELS

- 1) An understanding of how data assimilation fits in the scientific process - how data assimilation can help you:
  - Make better observations (where to observe, what to observe, when to observe, how to prioritize, validate observing systems, identify instrument errors)
  - Make better models (where are there biases, what processes are missing, how to identify systematic errors as early as possible, tune model parameterizations)

DATA ASSIMILATION AND THE SCIENTIFIC METHOD

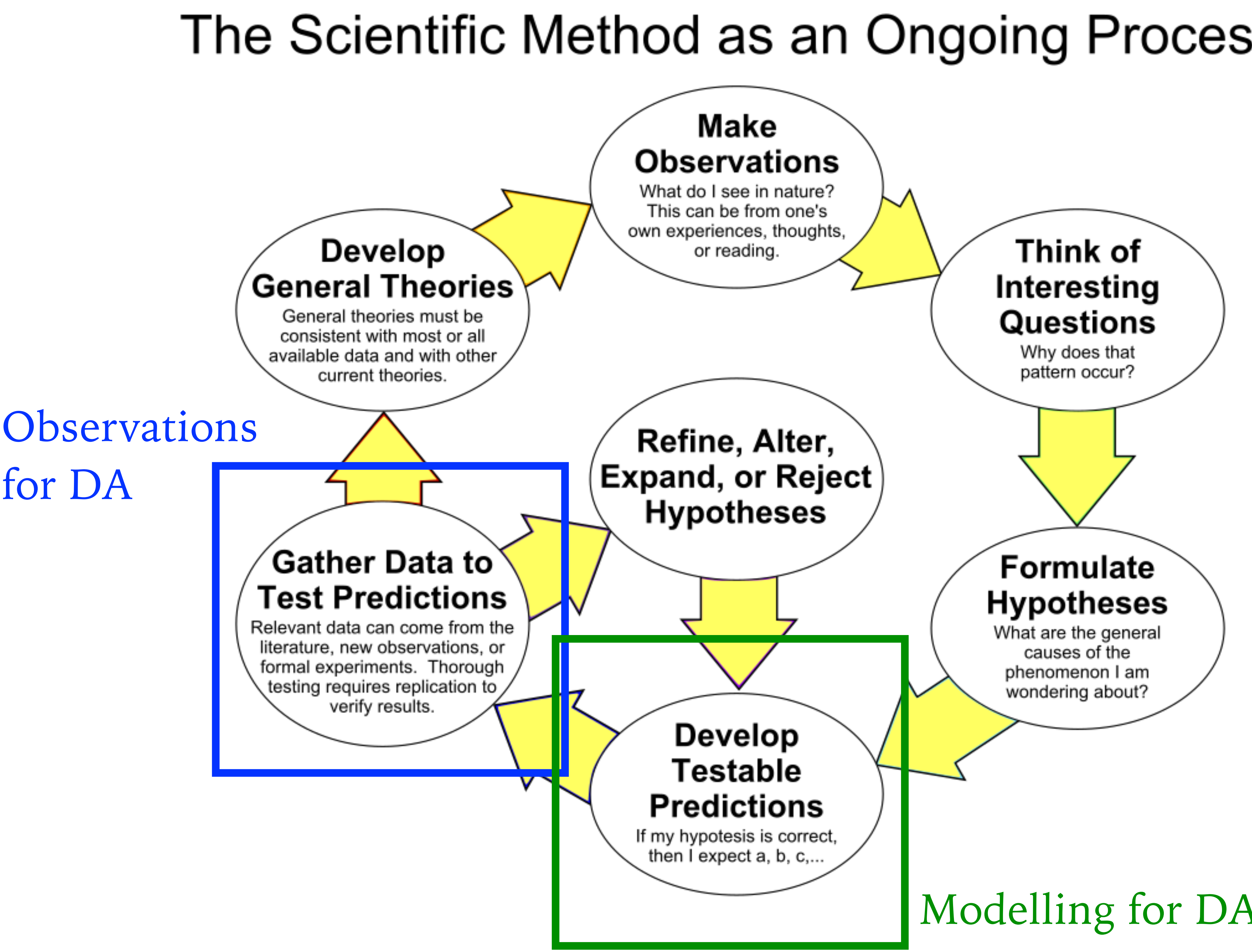


DATA ASSIMILATION AND THE SCIENTIFIC METHOD

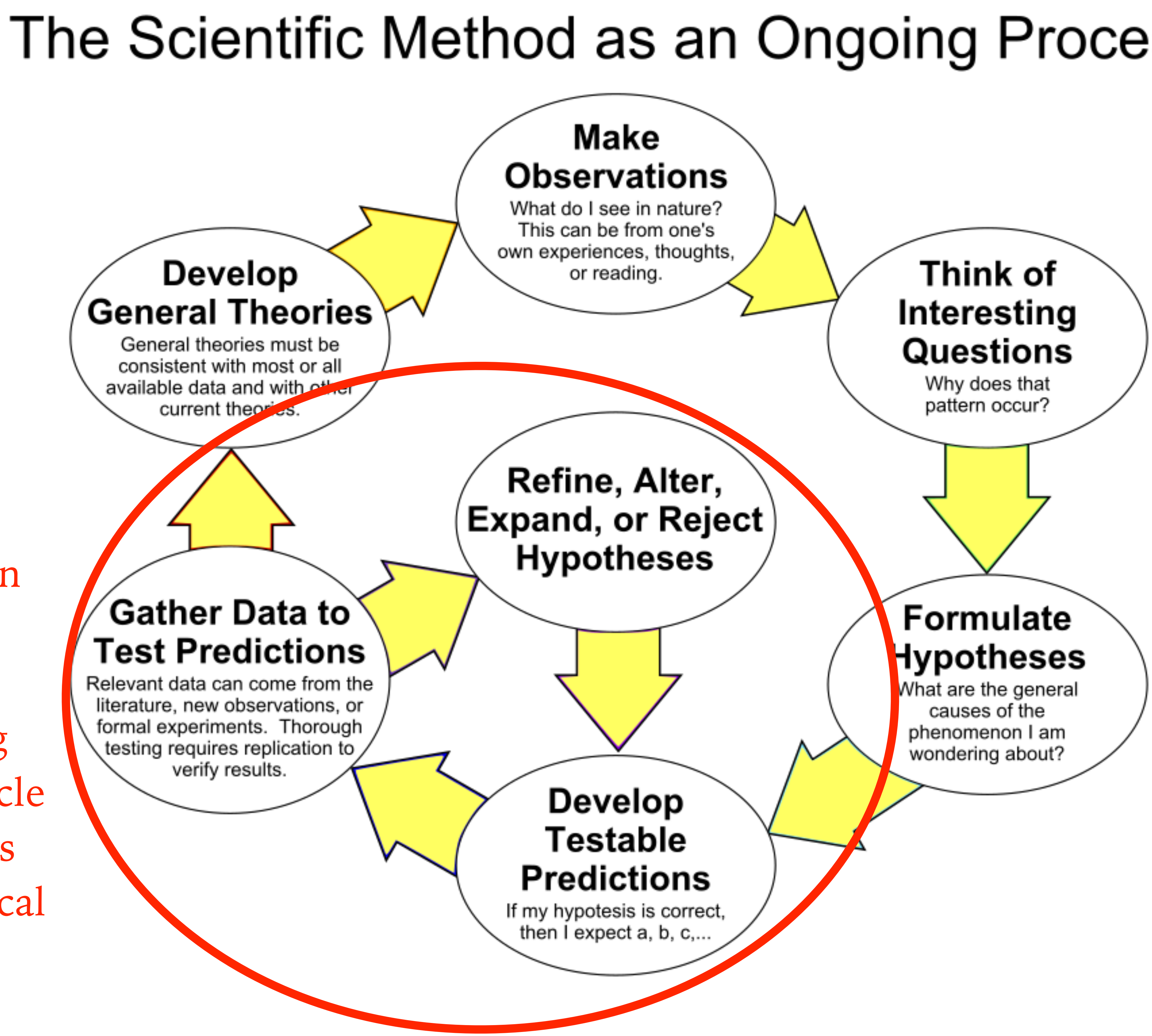


Data assimilation largely focuses on automating this sub cycle via rigorous mathematical methods

DATA ASSIMILATION AND THE SCIENTIFIC METHOD

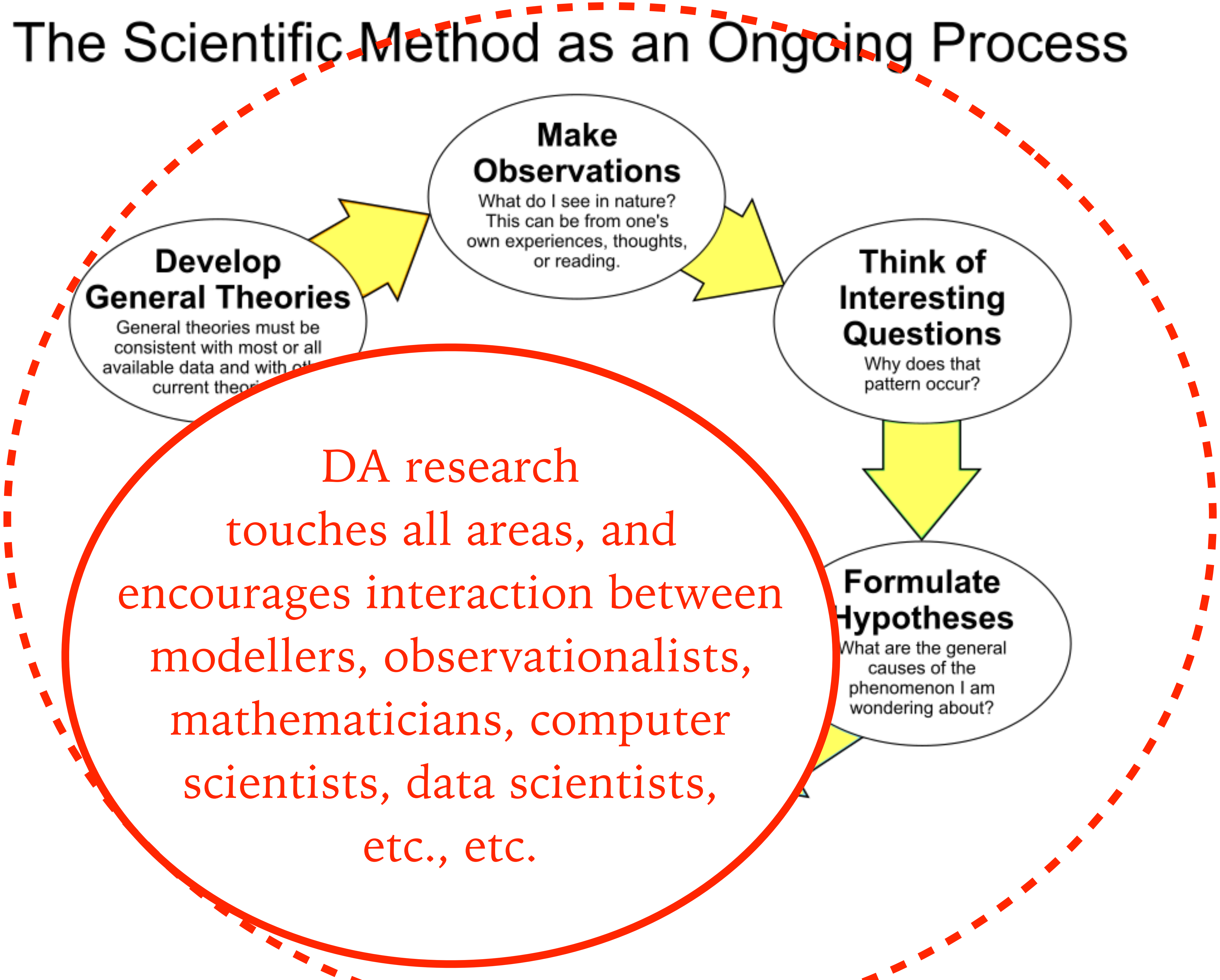


DATA ASSIMILATION AND THE SCIENTIFIC METHOD



Data assimilation largely focuses on automating this sub cycle via rigorous mathematical methods

DATA ASSIMILATION AND THE SCIENTIFIC METHOD



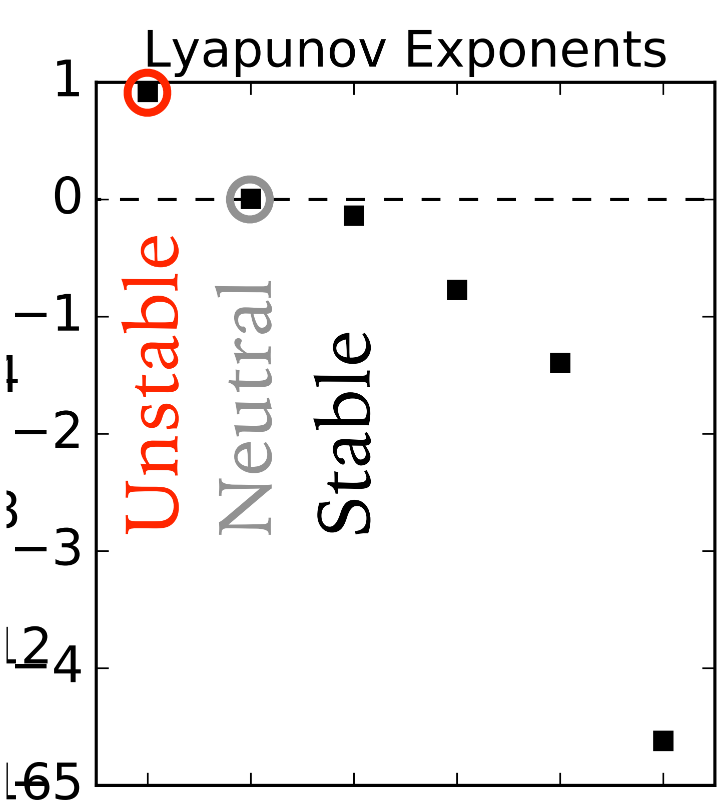
DATA ASSIMILATION AND:

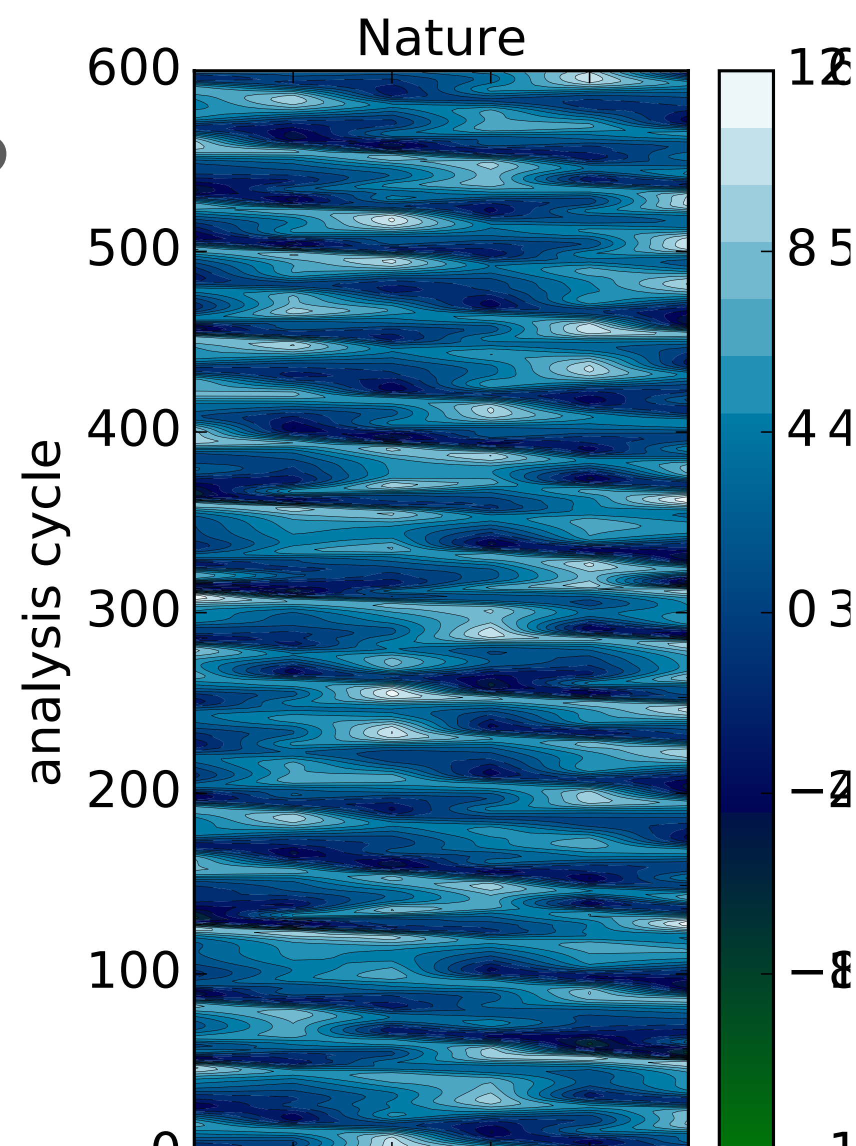
- Dynamical Systems theory
- Machine learning / AI
- Optimization / Minimization / Calculus of Variations
- Probability theory
- Control theory
- Synchronization
- Impulsive differential equations
- Nonlinear filtering
- Signal Processing
- Big Data
- Neuroscience

DYNAMICAL SYSTEMS - UNSTABLE MODES

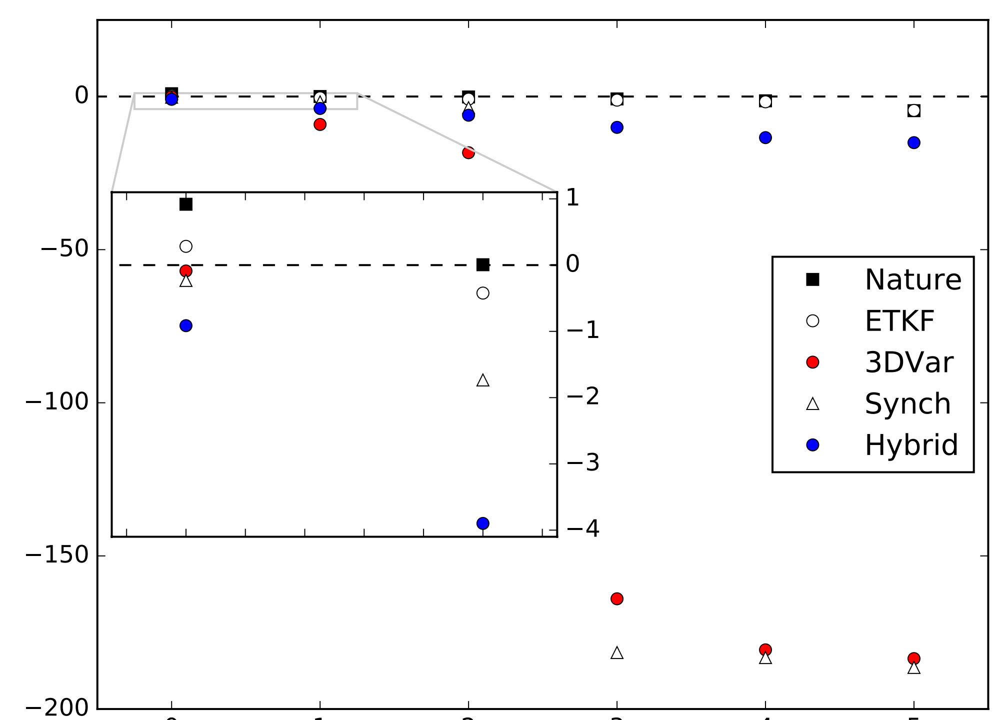
- Dynamical systems have identifiable modes that are either Unstable, Neutral, or Stable in the long-time average.

E.g. Lorenz-96 system restricted to 6 dimensions:





Computed via integration of the system Jacobian



DA **reduction of LEs of error system** (2-member ensemble for ETKF and Hybrid)

DYNAMICAL SYSTEMS - COUPLED SYSTEMS

- Extended slow manifold in coupled systems $\left(\text{LEs} \approx 0\right)$
- Example: Simplified QG model, 2-layer atmos, 1-layer ocean (Vannitsem and Lucarini; 2016)

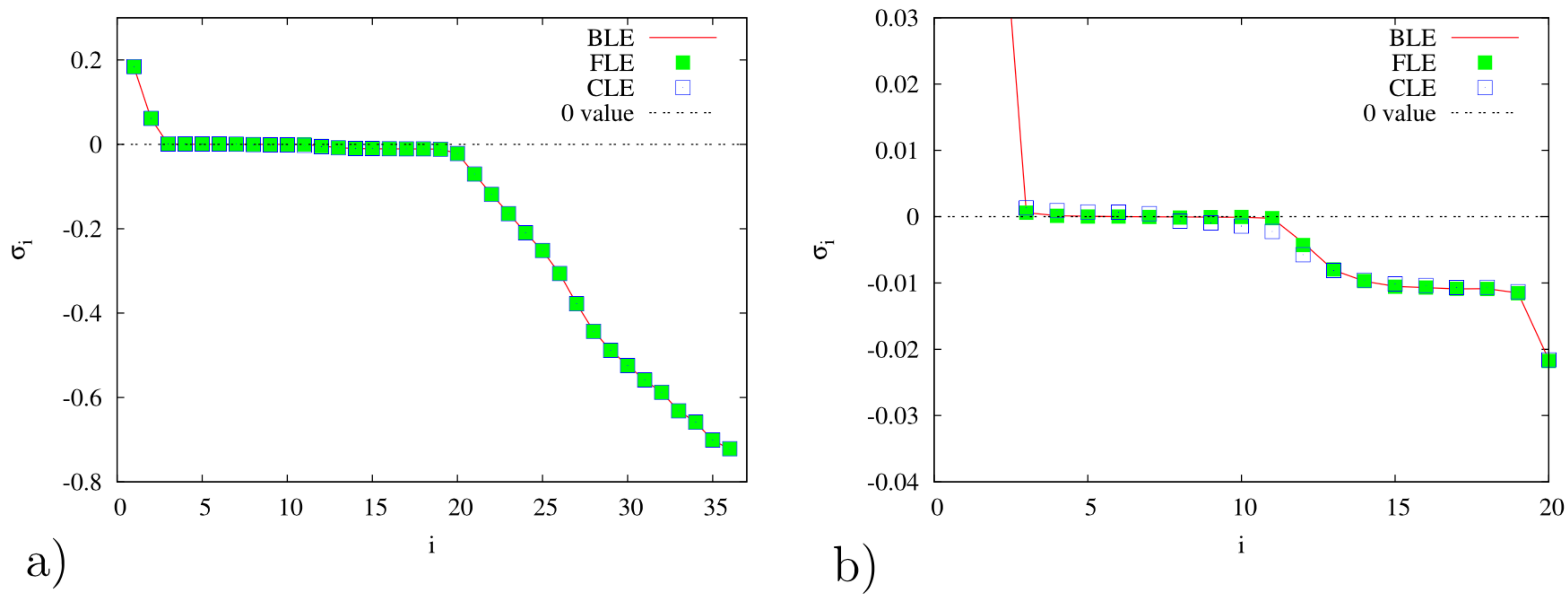


Figure 1. Lyapunov exponents, given in day^{-1}, based on the forward integration (FLEs, red line), backward integration (BLEs, full green squares), and on the CLVs (CLEs, empty blue squares). Panel (a): the full spectrum of LEs. Panel (b) a zoom around 0. Parameters' value: $C_o = 350\ \mathrm{W\ m^{-2}}$ and $d = 1 \times 10^{-8}\,\mathrm{s}^{-1}$.

DYNAMICAL SYSTEMS - OBSERVABILITY

- Unstable dynamics (positive and neutral Lyapunov exponents) indicate modes that must be controlled by observations (e.g. Trevisan, Uboldi, Carrassi, et al.)

Uboldi, F. and Trevisan, A. 2006: Detecting unstable structures and **controlling error growth** by **assimilation of** standard and **adaptive observations** in a primitive equation ocean model. *Nonlin. Processes Geophys.* **13**, 67–81.

Carrassi, A., Trevisan, A. and Uboldi, F. 2007: **Adaptive observations** and **assimilation in the unstable subspace** by Breeding on the Data Assimilation System. *Tellus,* **59A**, 101-113. DOI: 10.1111/j.1600-0870.2006.00210.x

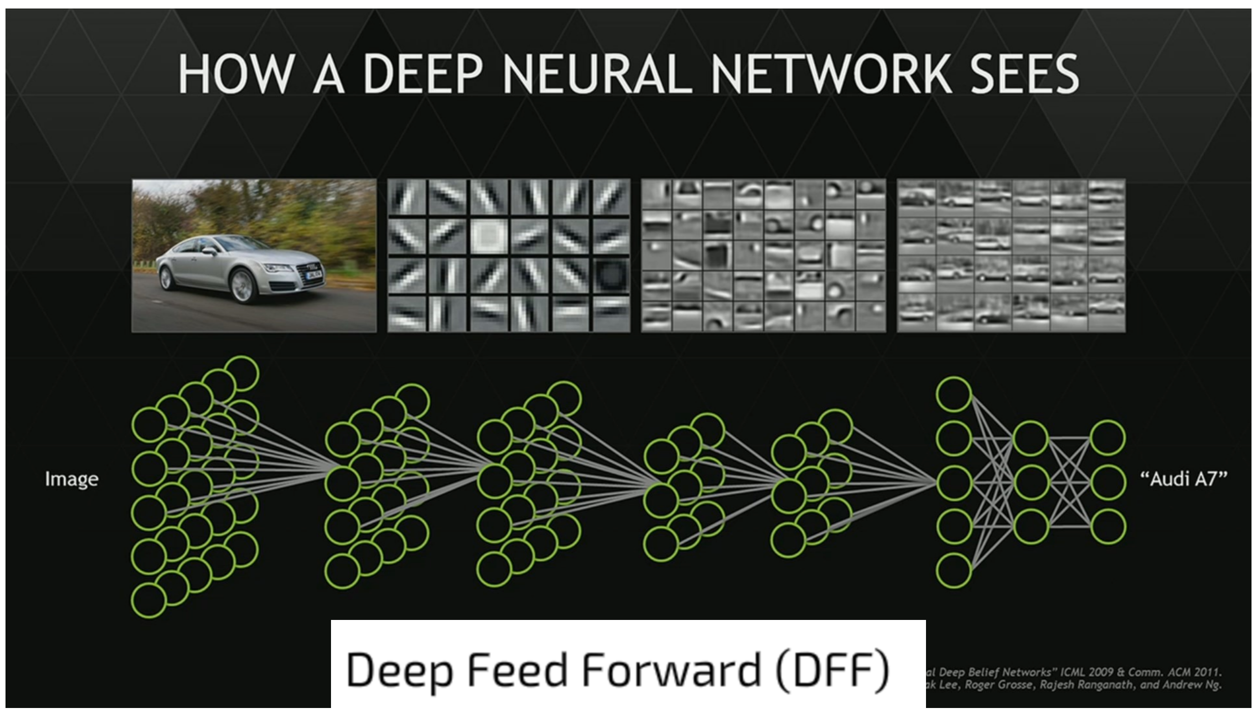
- Synchronization studies search for minimum observational coverage to achieve stability (e.g. Whartenby et al. 2013)

An, Z, Rey D, Ye JX, Abarbanel HDI. 2017: Estimating the state of a geophysical **system with sparse observations**: time delay methods to achieve accurate initial states for prediction. *Nonlinear Processes in Geophysics*. **24**:9-22. 10.5194/npg-24-9-2017

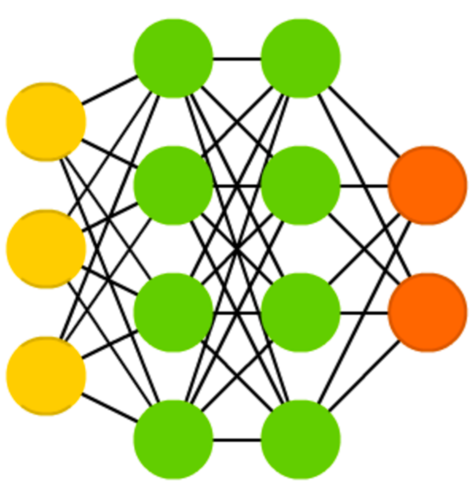
Whartenby, WG, Quinn JC, Abarbanel HDI. 2013: The **number of required observations in data assimilation** for a shallow-water flow. *Monthly Weather Review*. **141**:2502-2518. 10.1175/mwr-d-12-00103.1

MACHINE LEARNING

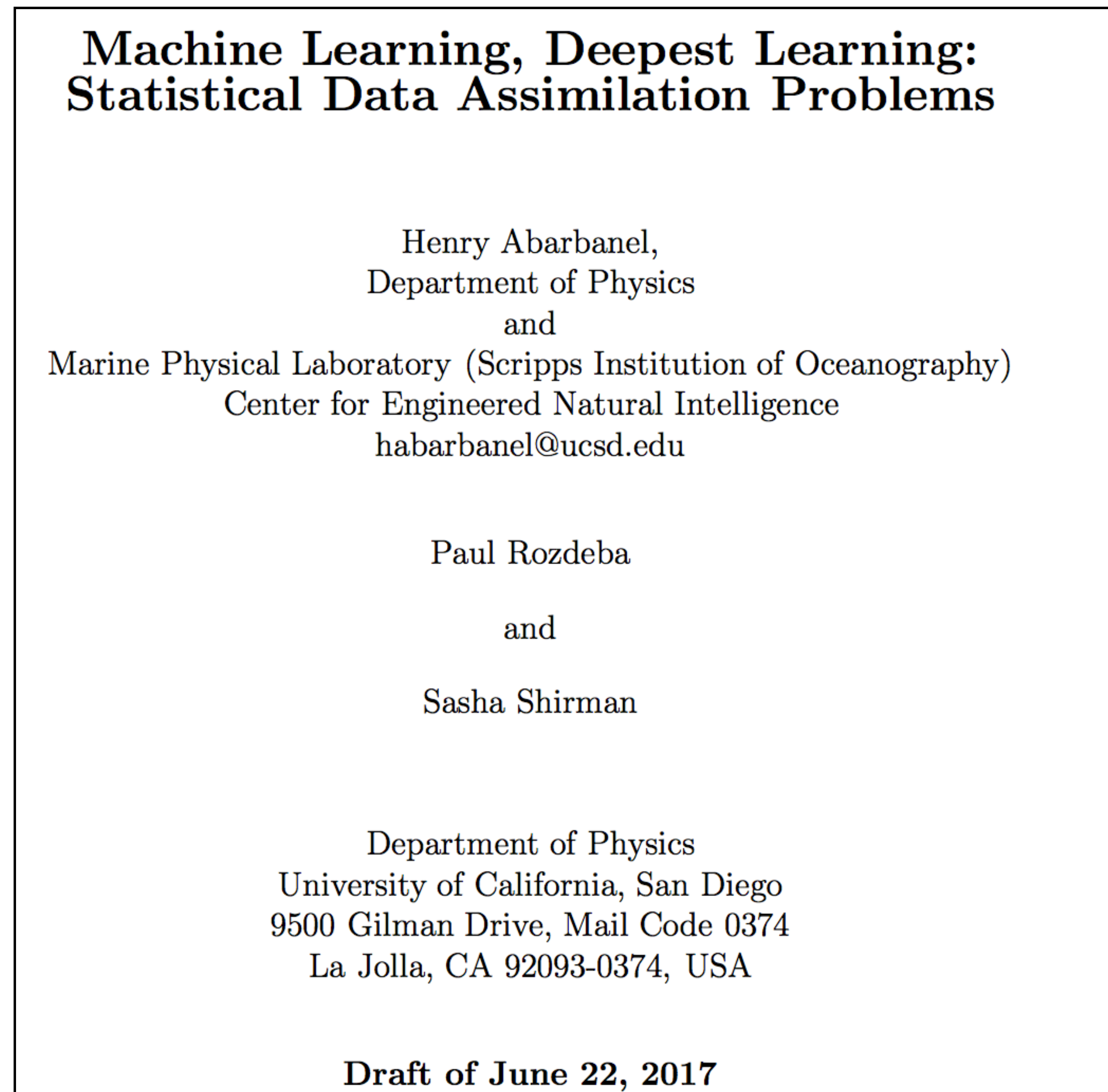
- Equivalence between machine learning algorithms, in the field of artificial intelligence, and data assimilation methods.



Deep Feed Forward (DFF)



Interpret hidden layers as time



Machine Learning, Deepest Learning: Statistical Data Assimilation Problems

Henry Abarbanel,
Department of Physics
and
Marine Physical Laboratory (Scripps Institution of Oceanography)
Center for Engineered Natural Intelligence
habarbanel@ucsd.edu

Paul Rozdeba

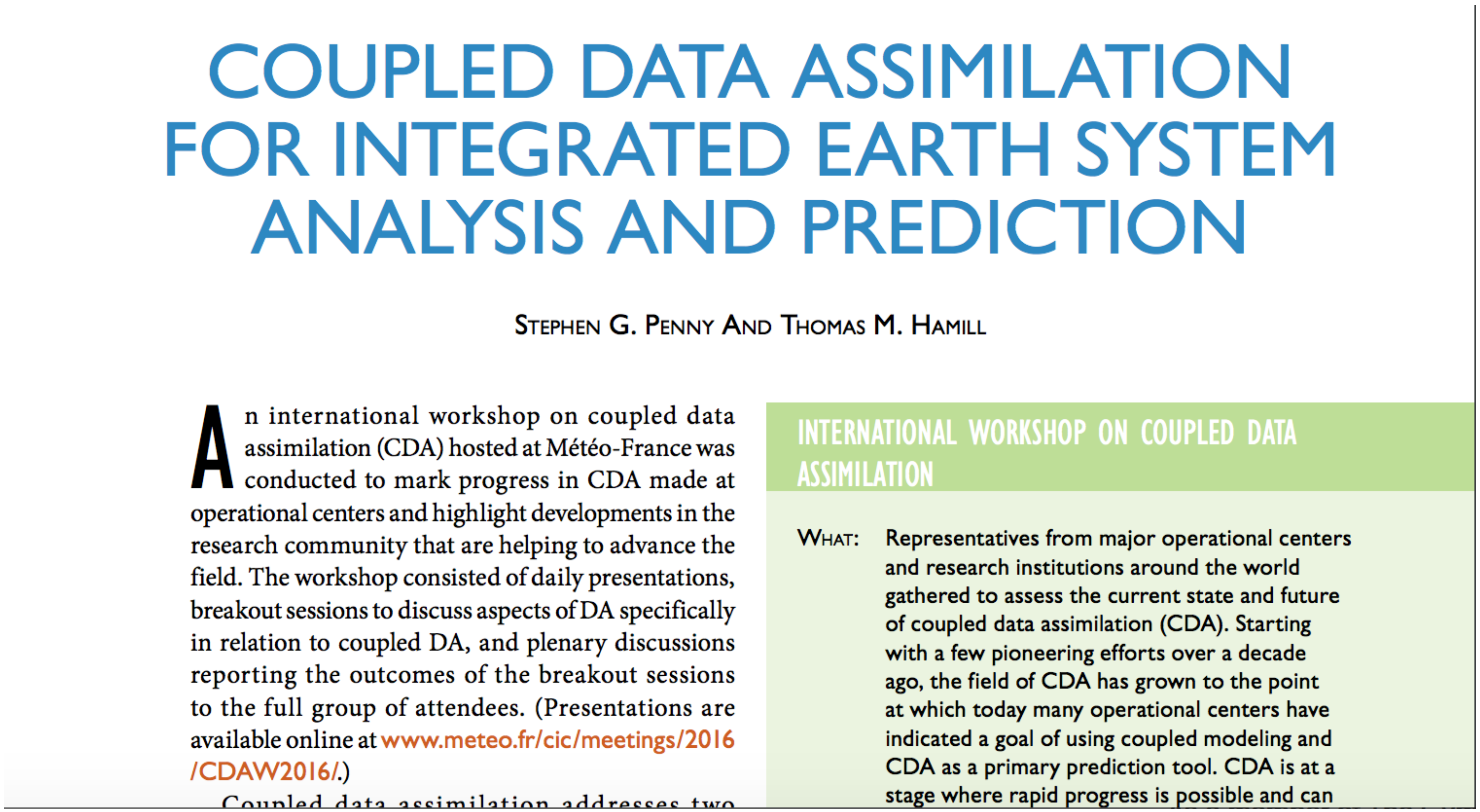
and

Sasha Shirman

Department of Physics
University of California, San Diego
9500 Gilman Drive, Mail Code 0374
La Jolla, CA 92093-0374, USA

Draft of June 22, 2017

INTERNATIONAL WORKSHOP ON COUPLED DATA ASSIMILATION



COUPLED DATA ASSIMILATION FOR INTEGRATED EARTH SYSTEM ANALYSIS AND PREDICTION

Stephen G. Penny and Thomas M. Hamill

An international workshop on coupled data assimilation (CDA) hosted at Météo-France was conducted to mark progress in CDA made at operational centers and highlight developments in the research community that are helping to advance the field. The workshop consisted of daily presentations, breakout sessions to discuss aspects of DA specifically in relation to coupled DA, and plenary discussions reporting the outcomes of the breakout sessions to the full group of attendees. (Presentations are available online at www.meteo.fr/cic/meetings/2016/CDAW2016/.)

Coupled data assimilation addresses two

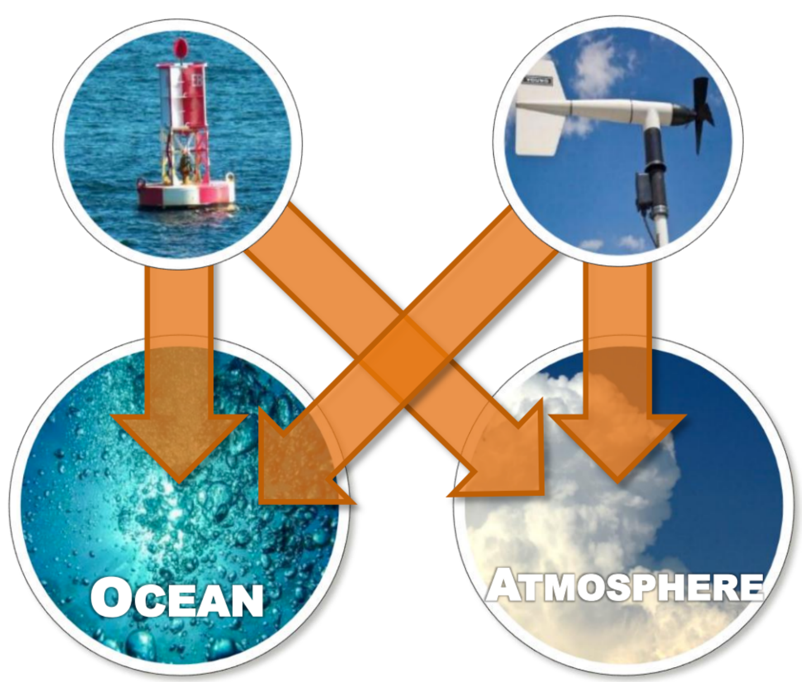
INTERNATIONAL WORKSHOP ON COUPLED DATA ASSIMILATION

What: Representatives from major operational centers and research institutions around the world gathered to assess the current state and future of coupled data assimilation (CDA). Starting with a few pioneering efforts over a decade ago, the field of CDA has grown to the point at which today many operational centers have indicated a goal of using coupled modeling and CDA as a primary prediction tool. CDA is at a stage where rapid progress is possible and can

- Penny and Hamill, 2017: Coupled Data Assimilation for Integrated Earth System Analysis and Prediction. *BAMS Workshop Report*, 2017. http://journals.ametsoc.org/doi/full/10.1175/BAMS-D-17-0036.1

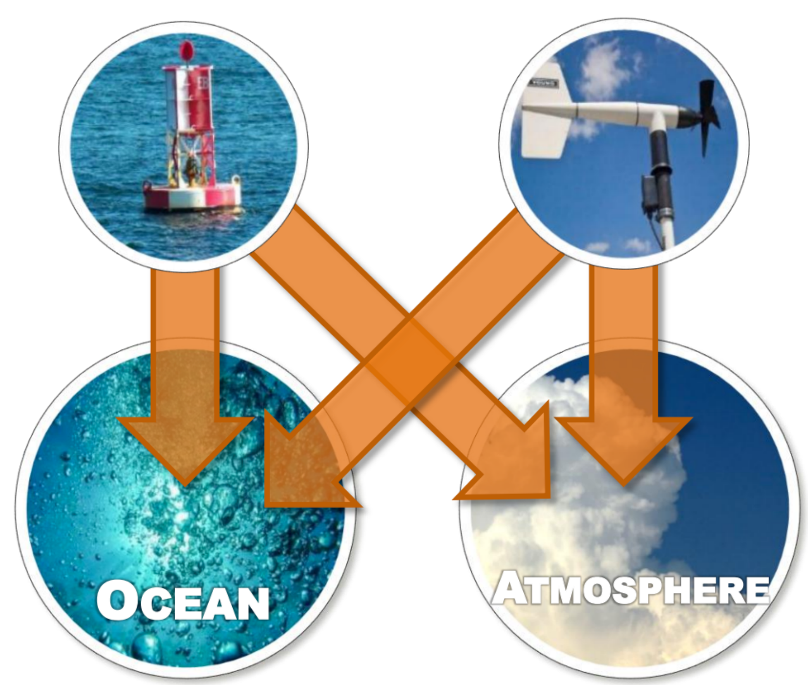
DEGREES OF COUPLING IN COUPLED DA

- Various degrees of coupling in CDA, for example:
  - Quasi-Weakly Coupled (Assimilation in only 1 domain)
  - Weakly Coupled (Coupling in Forecast model, not DA)
  - Quasi-Strongly Coupled (Observation influence across domains in DA analysis)
  - Strongly Coupled (Full cross-domain error covariances)



DEGREES OF COUPLING IN COUPLED DA

- Various degrees of coupling in CDA, for example:
  - Quasi-Weakly Coupled (Assimilation in only 1 domain)
  - **Weakly Coupled** (Coupling in Forecast model, not DA)
  - Quasi-Strongly Coupled (Observation influence across domains in DA analysis)
  - Strongly Coupled (Full cross-domain error covariances)



CLIMATE FORECAST SYSTEM REANALYSIS (CFSR)

➤ Good example of an operational weakly coupled DA system

Deterministic data assimilation (3DVar applied to each component), coupled forecast

THE NCEP CLIMATE FORECAST SYSTEM REANALYSIS

BY SURANJANA SAHA, SHRINIVAS MOORTHI, HUA-LU PAN, XINGREN WU, JIANDE WANG, SUDHIR NADIGA, PATRICK TRIPP, ROBERT KISTLER, JOHN WOOLLEN, DAVID BEHRINGER, HAIXIA LIU, DIANE STOKES, ROBERT GRUMBINE, GEORGE GAYNO, JUN WANG, YU-TAI HOU, HUI-YA CHUANG, HANN-MING H. JUANG, JOE SELA, MARK IREDELL, RUSS TREADON, DARYL KLEIST, PAUL VAN DELST, DENNIS KEYSER, JOHN DERBER, MICHAEL EK, JESSE MENG, HELIN WEI, RONGQIAN YANG, STEPHEN LORD, HUUG VAN DEN DOOL, ARUN KUMAR, WANQIU WANG, CRAIG LONG, MUTHUVEL CHELLIAH, YAN XUE, BOYIN HUANG, JAE-KYUNG SCHEMM, WESLEY EBISUZAKI, ROGER LIN, PINGPING XIE, MINGYUE CHEN, SHUNTAI ZHOU, WAYNE HIGGINS, CHENG-ZHI ZOU, QUANHUA LIU, YONG CHEN, YONG HAN, LIDIA CUCURULL, RICHARD W. REYNOLDS, GLENN RUTLEDGE, AND MITCH GOLDBERG

A new coupled global NCEP Reanalysis for the period 1979–present is now available, at much higher temporal and spatial resolution, for climate studies.

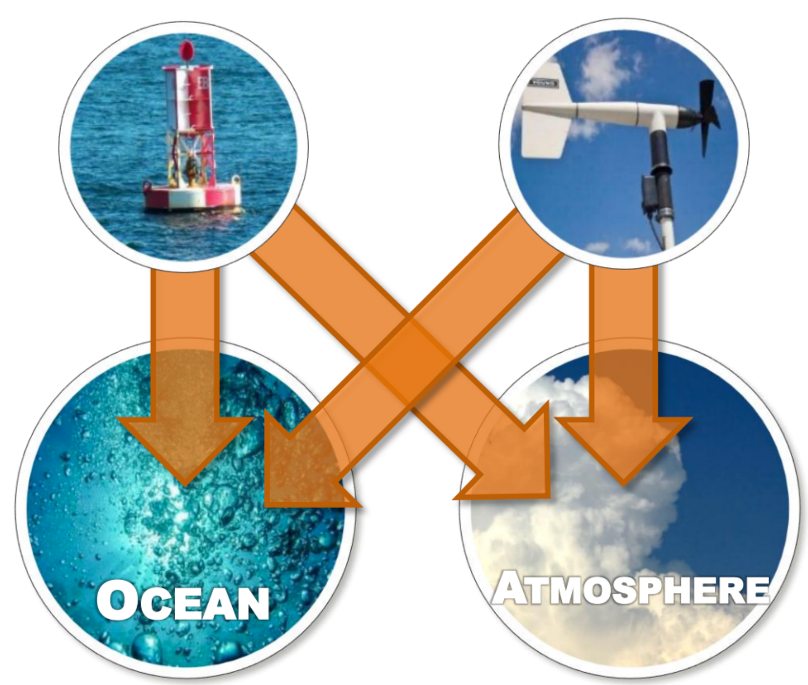
AFFILIATIONS: SAHA, MOORTHI, PAN, BEHRINGER, STOKES, GRUMBINE, HOU, CHUANG, JUANG, SELA, IREDELL, TREADON, KEYSER, DERBER, EK, AND LORD—Environmental Modeling Center, NCEP/NWS/NOAA, Camp Springs, Maryland; VAN DEN DOOL, KUMAR, W. WANG, LONG, CHELLIAH, XUE, SCHEMM, EBISUZAKI, XIE, CHEN, AND HIGGINS—Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, Maryland; WU, JI. WANG, NADIGA, KISTLER, WOOLLEN, H. LIU, GAYNO, JU. WANG, KLEIST, VAN DELST, MENG, WEI, AND YANG—Science Applications International Corporation (SAIC), McLean, Virginia; ZOU, CHEN, HAN, CUCURULL, AND GOLDBERG—Center for Satellite Applications and Research, NESDIS/NOAA, Camp Springs, Maryland;

The first reanalysis at NCEP (all acronyms are defined in the appendix), conducted in the 1990s, resulted in the NCEP–NCAR reanalysis (Kalnay et al. 1996), or R1 for brevity, and ultimately covered many years, from 1948 to the present (Kistler et al. 2001). It is still being executed at NCEP, to the benefit of countless users for monthly, and even daily, updates of the current state of the atmosphere. At the same time, other reanalyses were being conducted, namely, ERA-15 (Gibson et al. 1997) was executed for a more limited period (1979–93) at the ECMWF,

http://journals.ametsoc.org/doi/pdf/10.1175/2010BAMS3001.1

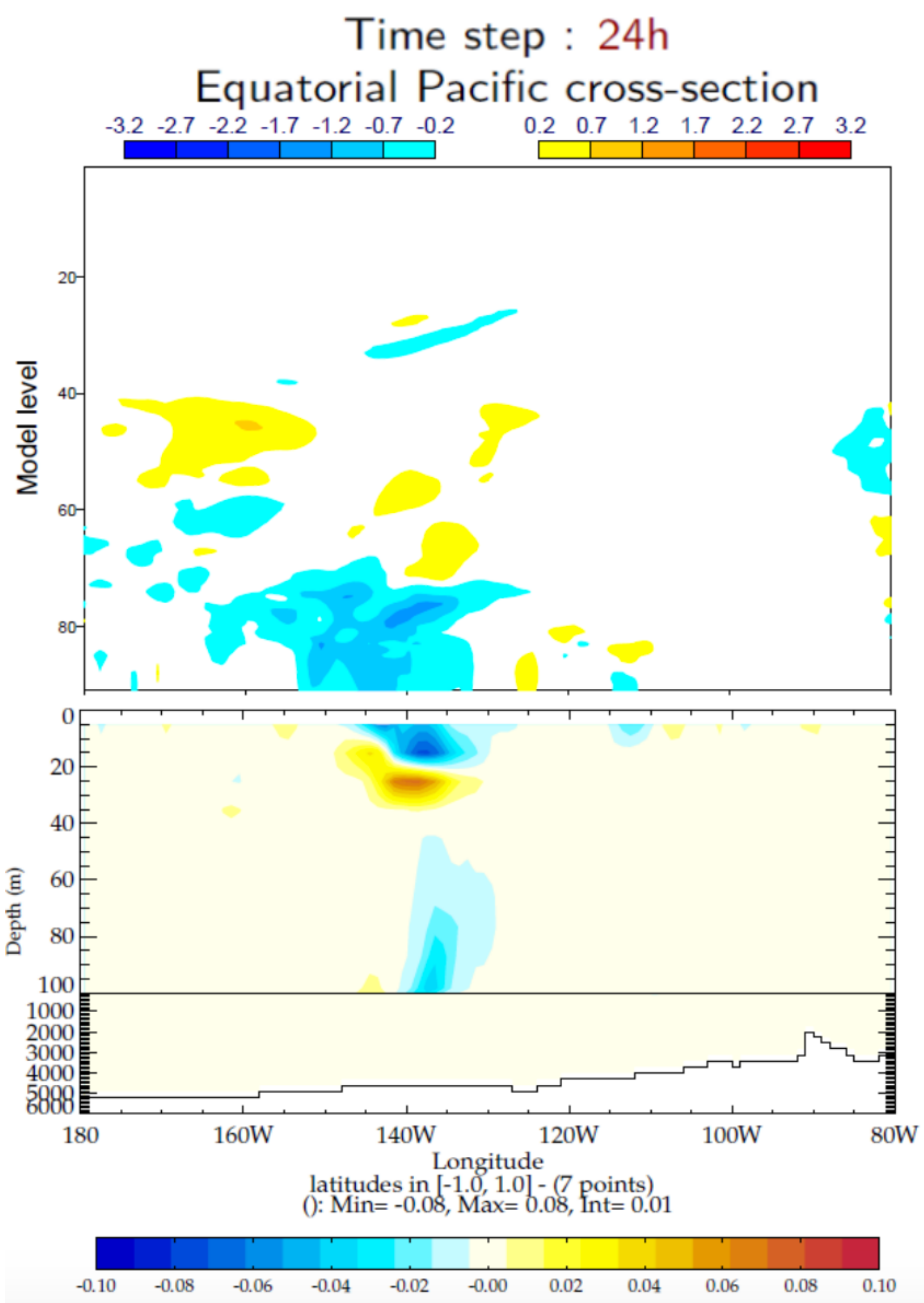
DEGREES OF COUPLING IN COUPLED DA

- Various degrees of coupling in CDA, for example:
  - Quasi-Weakly Coupled (Assimilation in only 1 domain)
  - Weakly Coupled (Coupling in Forecast model, not DA)
  - **Quasi-Strongly Coupled** (Observation influence across domains in DA analysis)
  - Strongly Coupled (Full cross-domain error covariances)



SELECTED NOTABLE RESULTS: REANALYSIS

- ECMWF CERA System



Atmosphere-ocean cross-section (wind and temperature)

Atmospheric wind increment (one station with hourly measurements of a 10m/s westward wind) spreads in the ocean as a temperature increment during the model integration (outer loop)

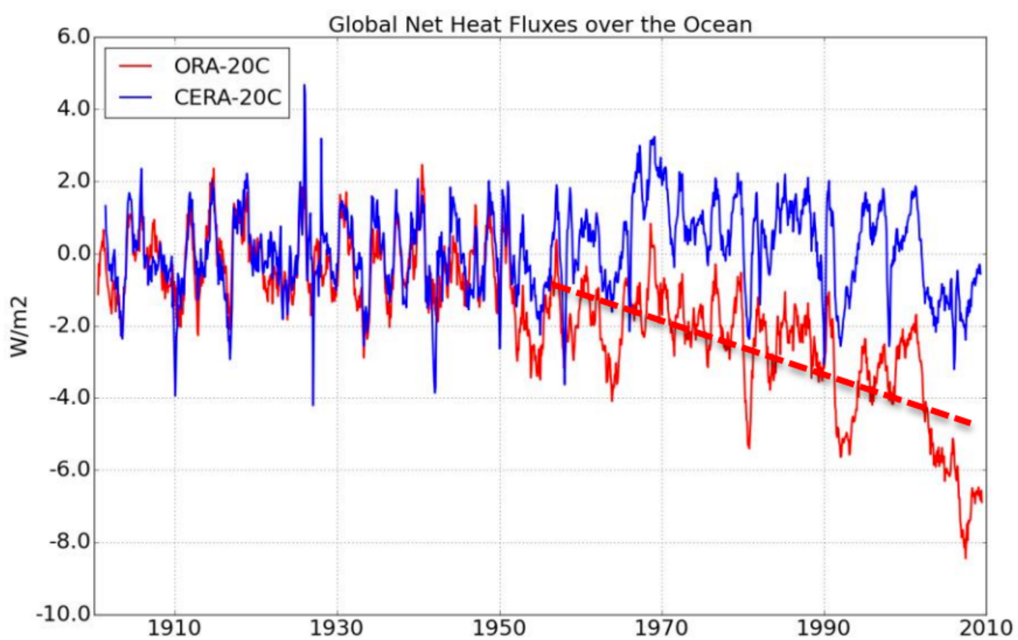
Ocean-atmosphere correlations are generated within the CERA incremental variational approach

A coupled data assimilation system for climate reanalysis. P. Laloyaux, M. Balmaseda, D. Dee, K. Mogensen and P. Janssen. QJRMS, 142: 65–78, 2016.

P. Laloyaux

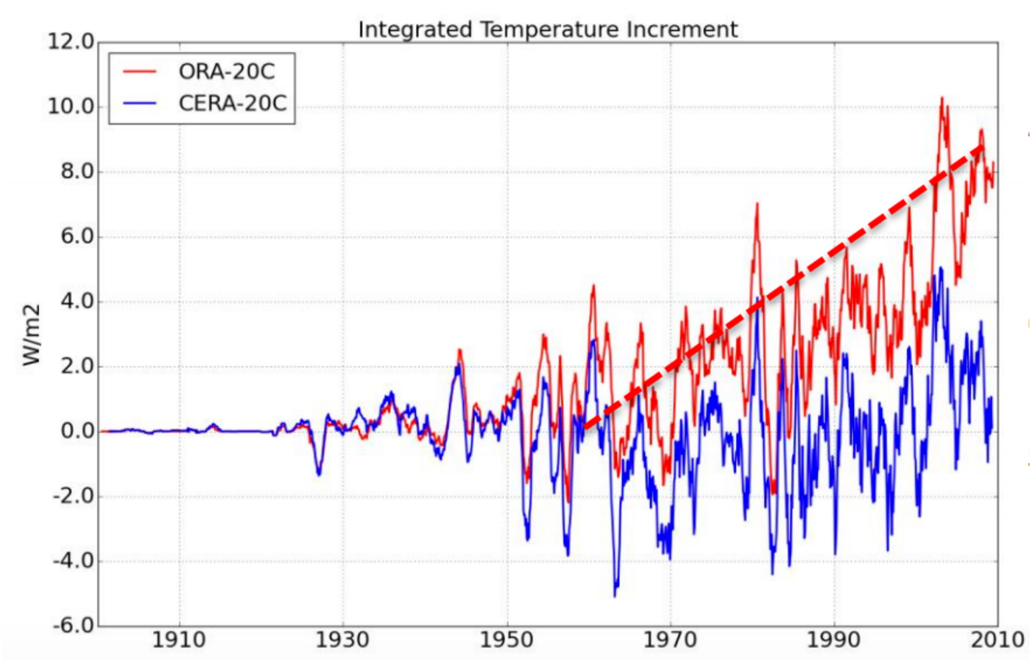
SELECTED NOTABLE RESULTS: REANALYSIS

➤ ECMWF 20th Century Coupled System Reanalysis (CERA-20C)



Global net **air-sea fluxes** toward the ocean in CERA-20C and ORA-20C.

→ spurious trend in ORA-20C probably due to shift in wind forcing in ERA-20C (heat lost)



Ocean temperature increment in CERA-20C and ORA-20C.

→ increment in ORA-20C is trying to compensate the heat lost
→ CERA-20C fluctuates around zero suggesting a more balanced air-sea interface

E. de Boisseson

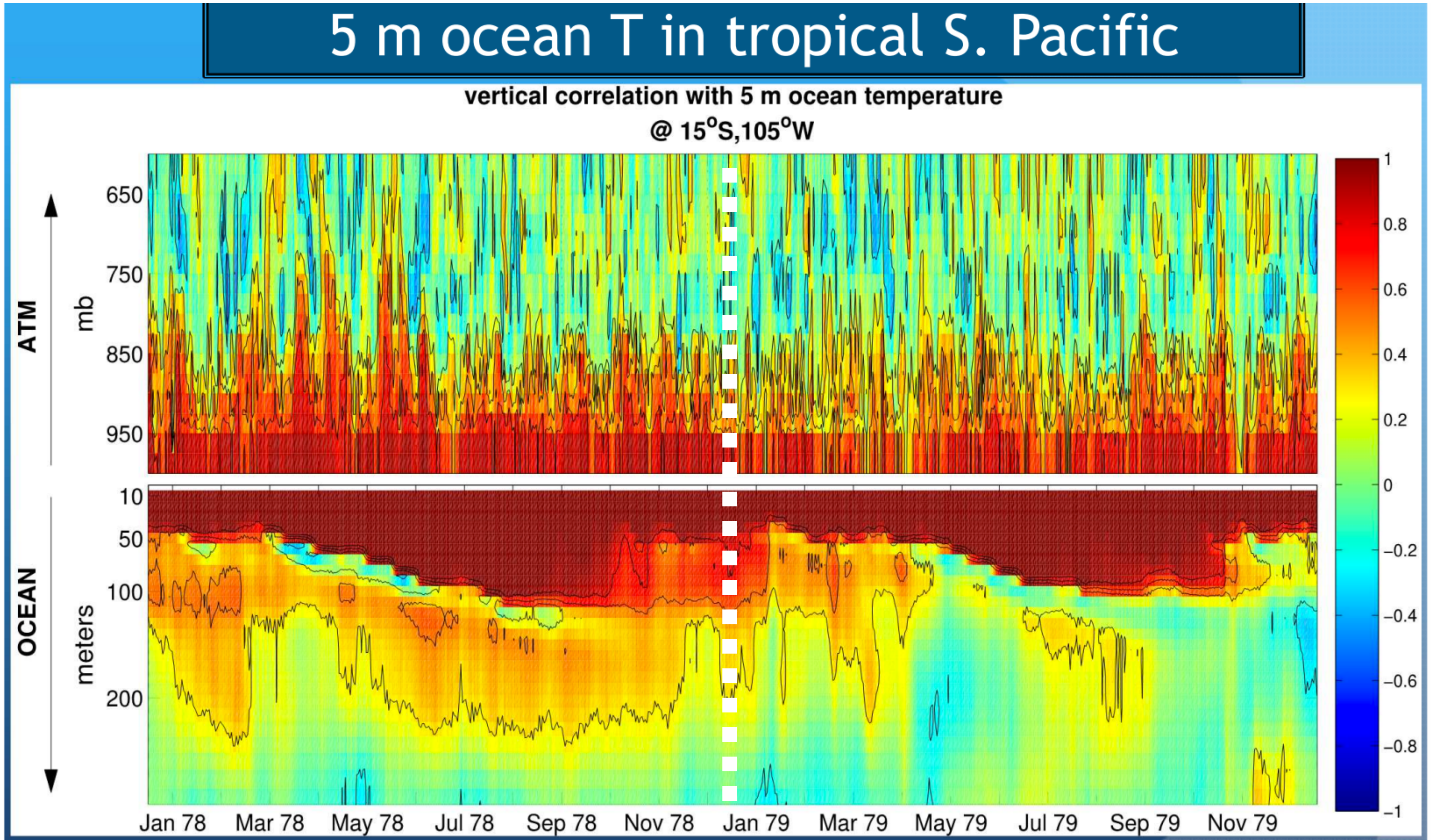
http://www.meteo.fr/cic/meetings/2016/CDAW2016/presentations/1_3.pdf

INVESTIGATIONS OF CROSS-DOMAIN ERROR COVARIANCE

- Using a weakly coupled **ensemble** DA system, we can investigate the error covariance between domains

CROSS-DOMAIN ERROR COVARIANCE ESTIMATION

➤ Vertical cross-domain error correlations (ensemble-estimated)



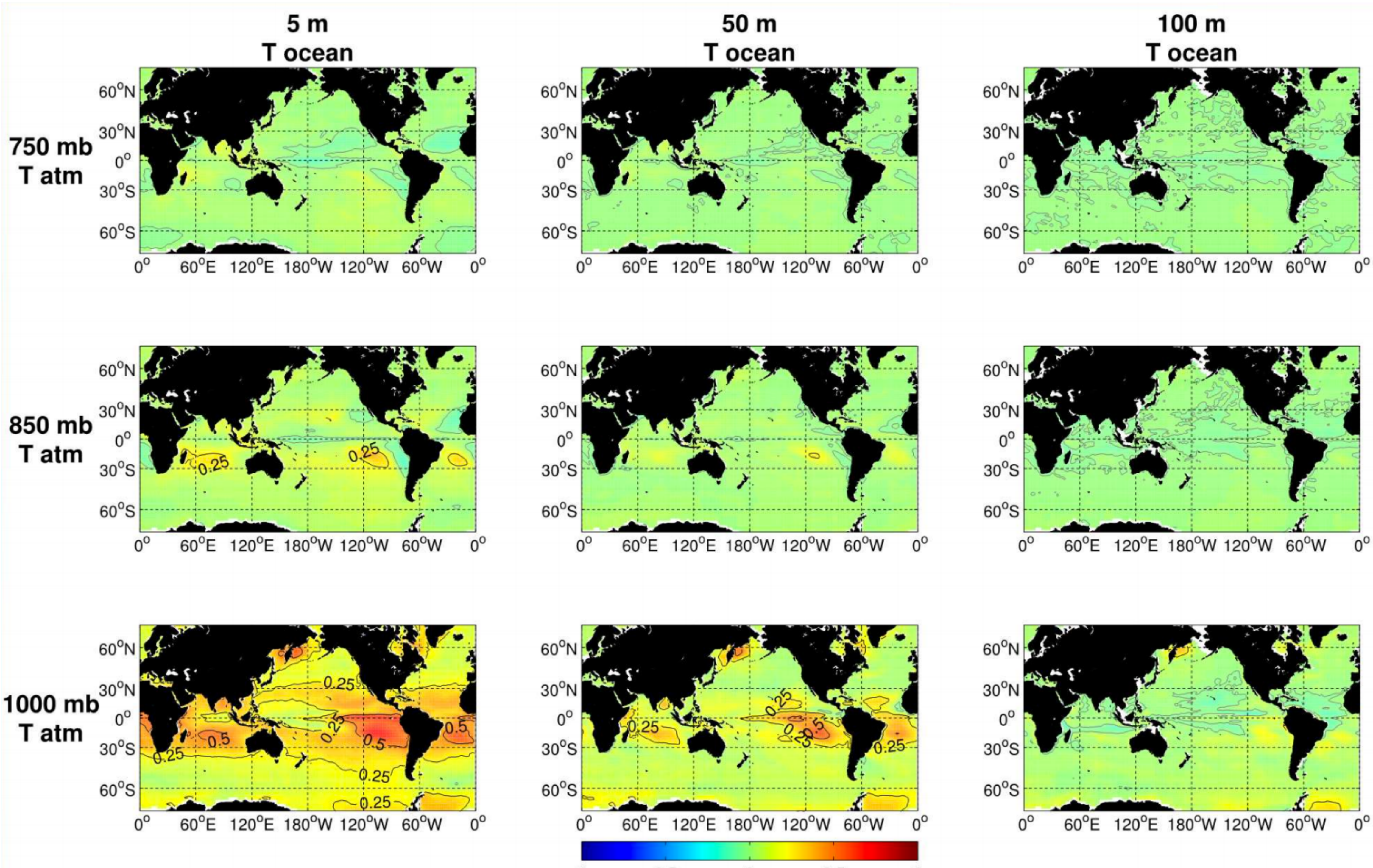
Instantaneous ensemble correlations reveal the different time scales and dynamics within the ocean and atmosphere and highlight the non-stationarity of coupled covariances.

A. Karspeck

http://www.data-assimilation.riken.jp/risda2017/program/abstracts/pdf/06_2_A.Karspeck.pdf

CROSS-DOMAIN ERROR COVARIANCE ESTIMATION

➤ Time-averaged correlations

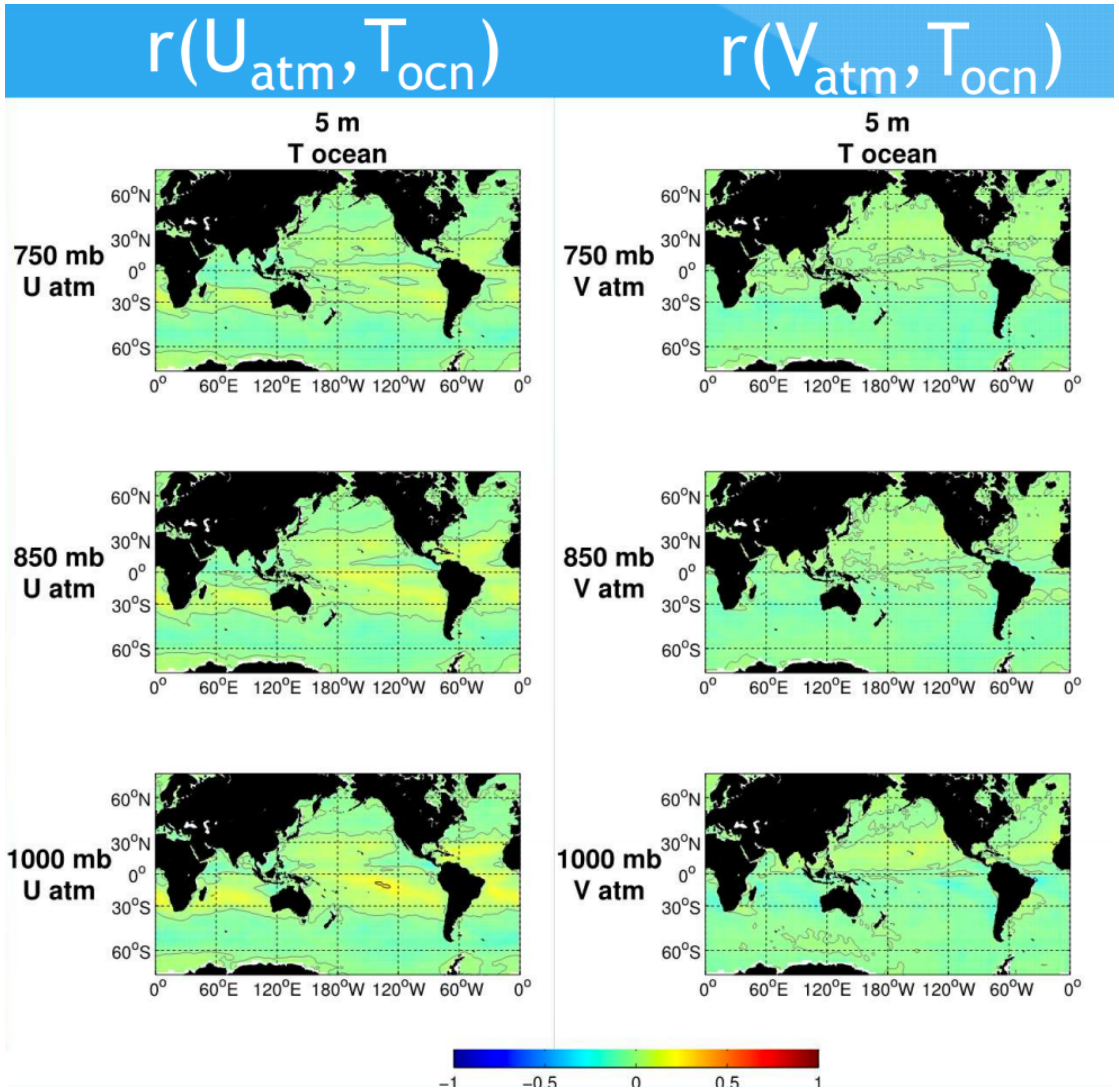


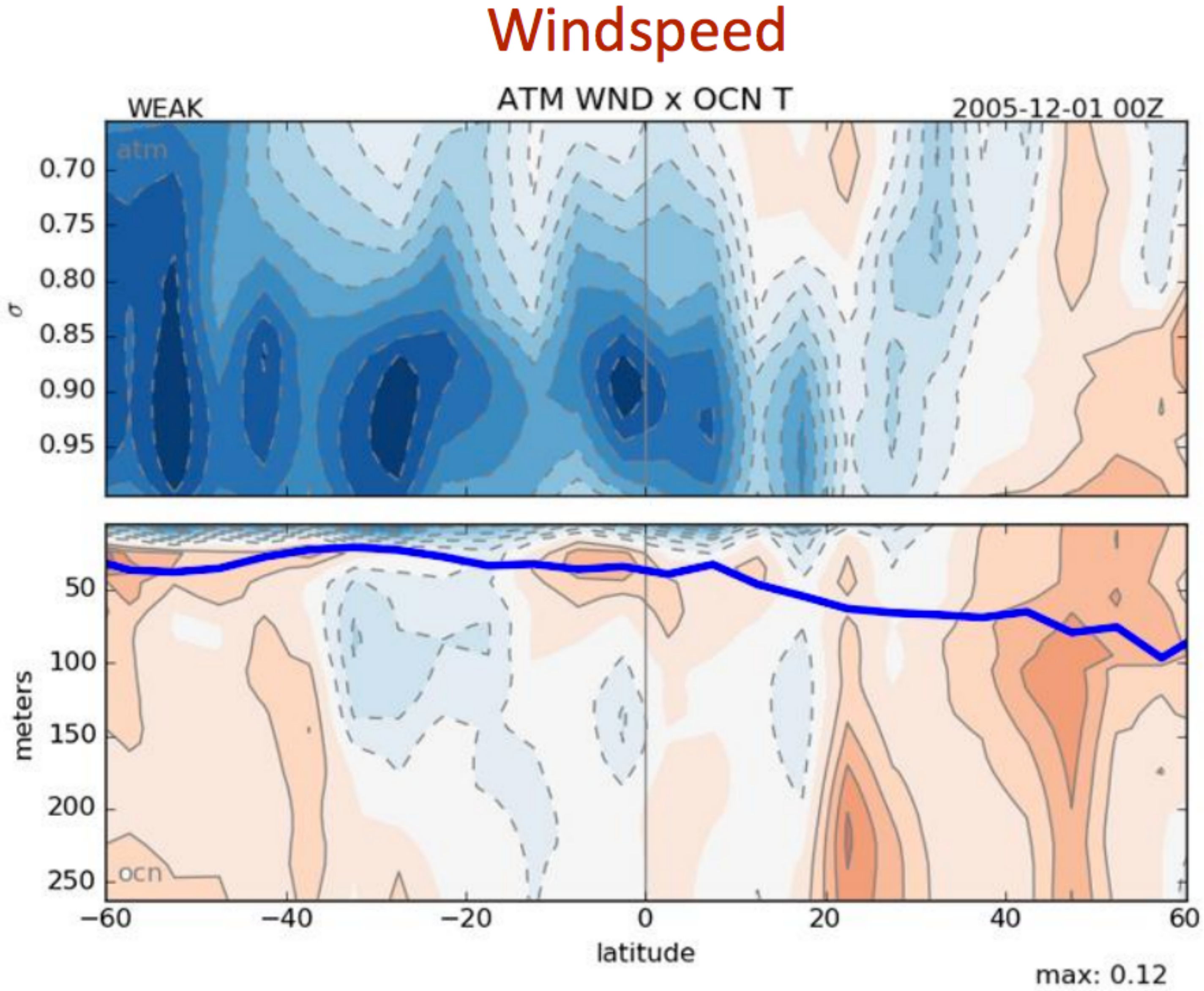
A. Karspeck

http://www.data-assimilation.riken.jp/risda2017/program/abstracts/pdf/06_2_A.Karspeck.pdf

CROSS-DOMAIN ERROR COVARIANCE ESTIMATION

- Time-averaged correlations between atmospheric winds and SST
- Stronger wind-speeds associated with cooler SST





A. Karspeck http://www.data-assimilation.riken.jp/risda2017/program/abstracts/pdf/06_2_A.Karspeck.pdf

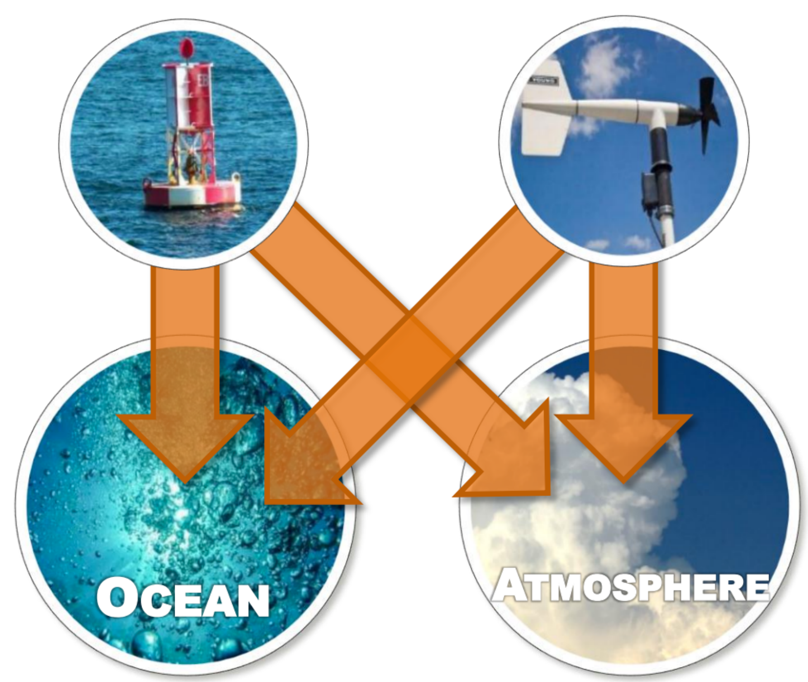
T. Sluka http://www.meteo.fr/cic/meetings/2016/CDAW2016/presentations/3_1.pdf

INVESTIGATIONS OF CROSS-DOMAIN ERROR COVARIANCE

- Using a weakly coupled ensemble DA system, we can investigate the error covariance between domains
- Assuming an accurate model, using the cross-domain error covariance in the DA should improve analysis accuracy
- This enables the whole coupled Earth system model to be effectively treated as a single dynamical system
- And leads to what is called 'strongly coupled' data assimilation…

DEGREES OF COUPLING IN COUPLED DA

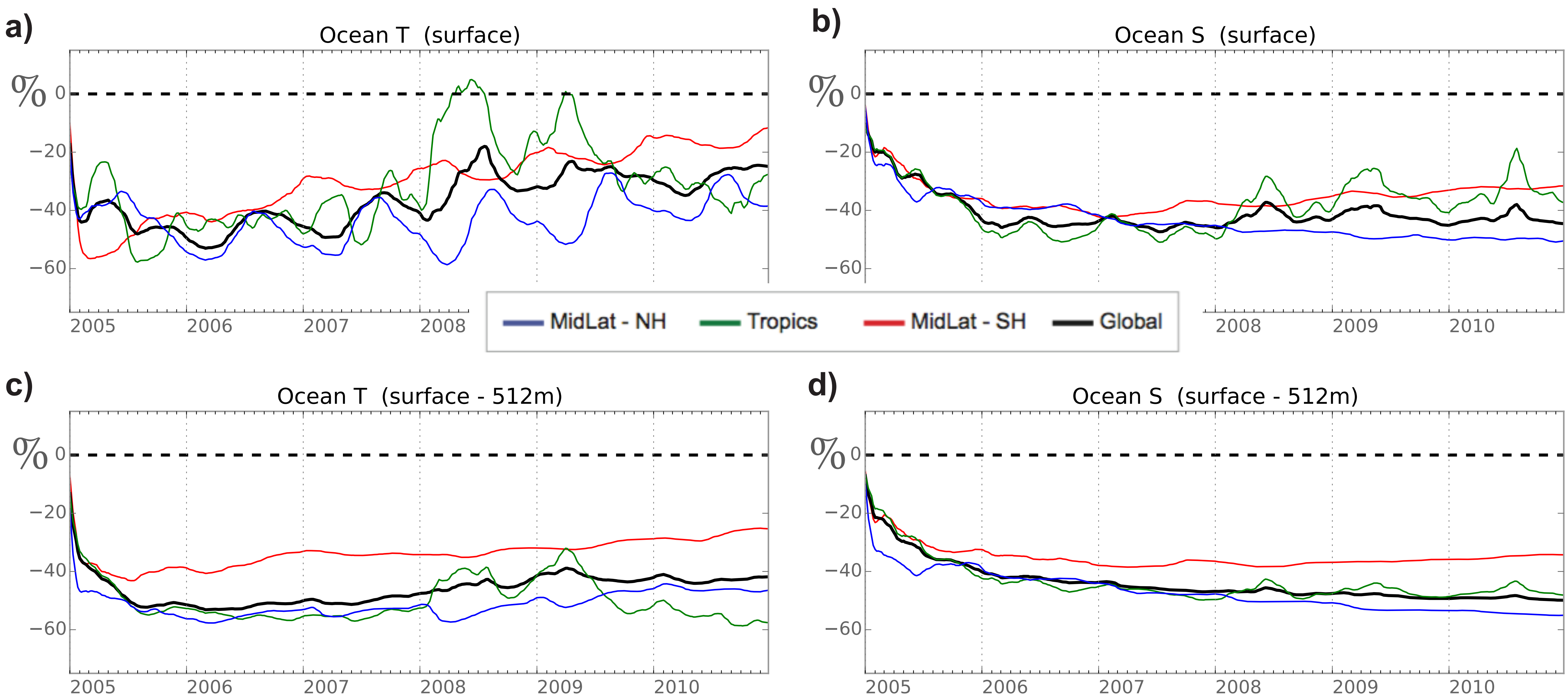
- Various degrees of coupling in CDA, for example:
  - Quasi-Weakly Coupled (Assimilation in only 1 domain)
  - Weakly Coupled (Coupling in Forecast model, not DA)
  - Quasi-Strongly Coupled (Observation influence across domains in DA analysis)
  - **Strongly Coupled** (Full cross-domain error covariances)



SELECTED NOTABLE RESULTS: STRONGLY COUPLED DA

➤ Cross-domain correlations improve the analysis in OSSEs

Strongly minus Weakly coupled DA **RMSE**. Assimilating only atmospheric observations:

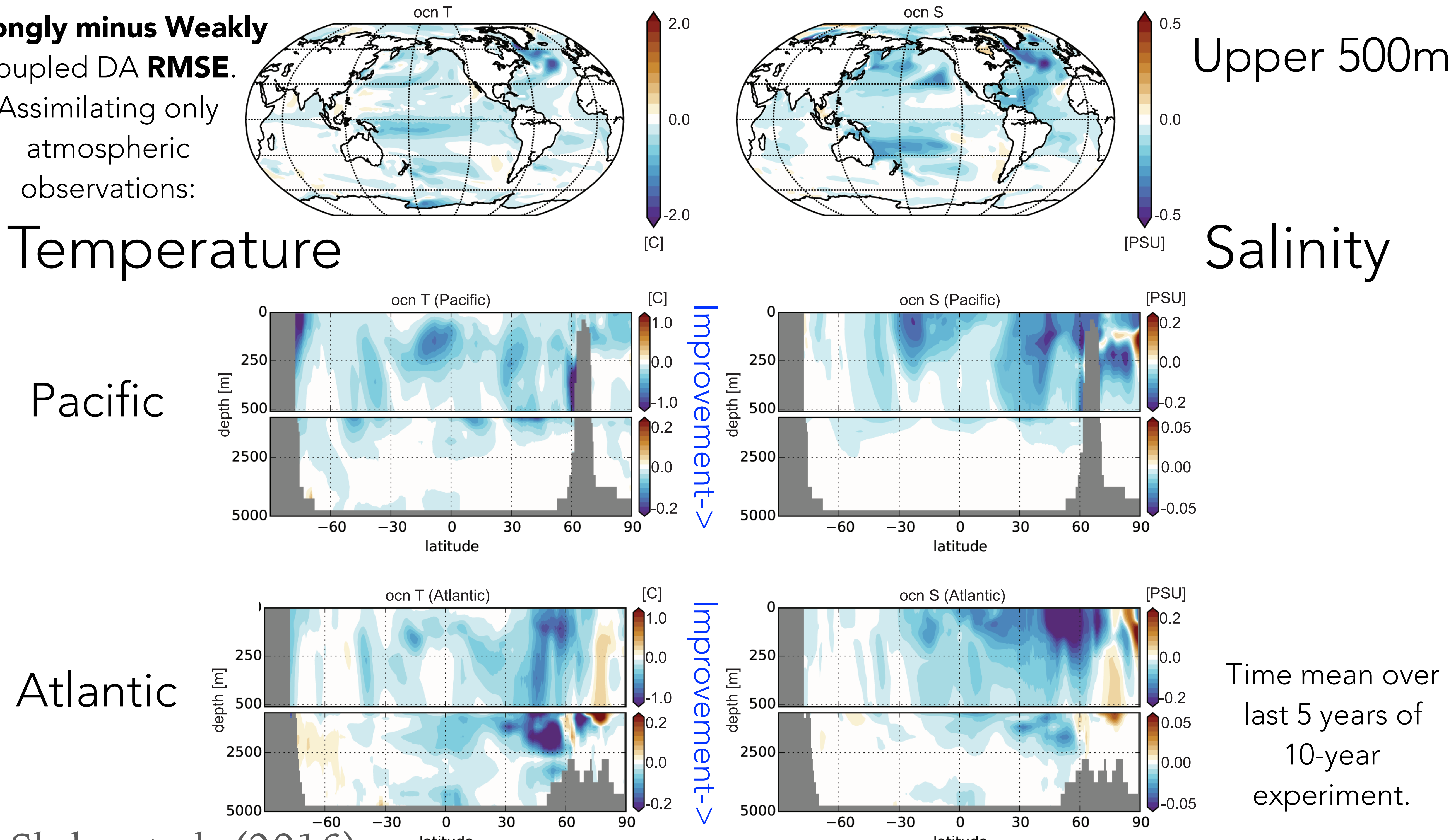


Sluka, T. C., S. G. Penny, E. Kalnay, and T. Miyoshi (2016), Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation, Geophys. Res. Lett., 43, 752–759, doi: 10.1002/2015GL067238.

SELECTED NOTABLE RESULTS: STRONGLY COUPLED DA

➤ Cross-domain correlations improve the analysis in OSSEs

Strongly minus Weakly coupled DA **RMSE**. Assimilating only atmospheric observations:



Time mean over last 5 years of 10-year experiment.

T. Sluka et al. (2016)

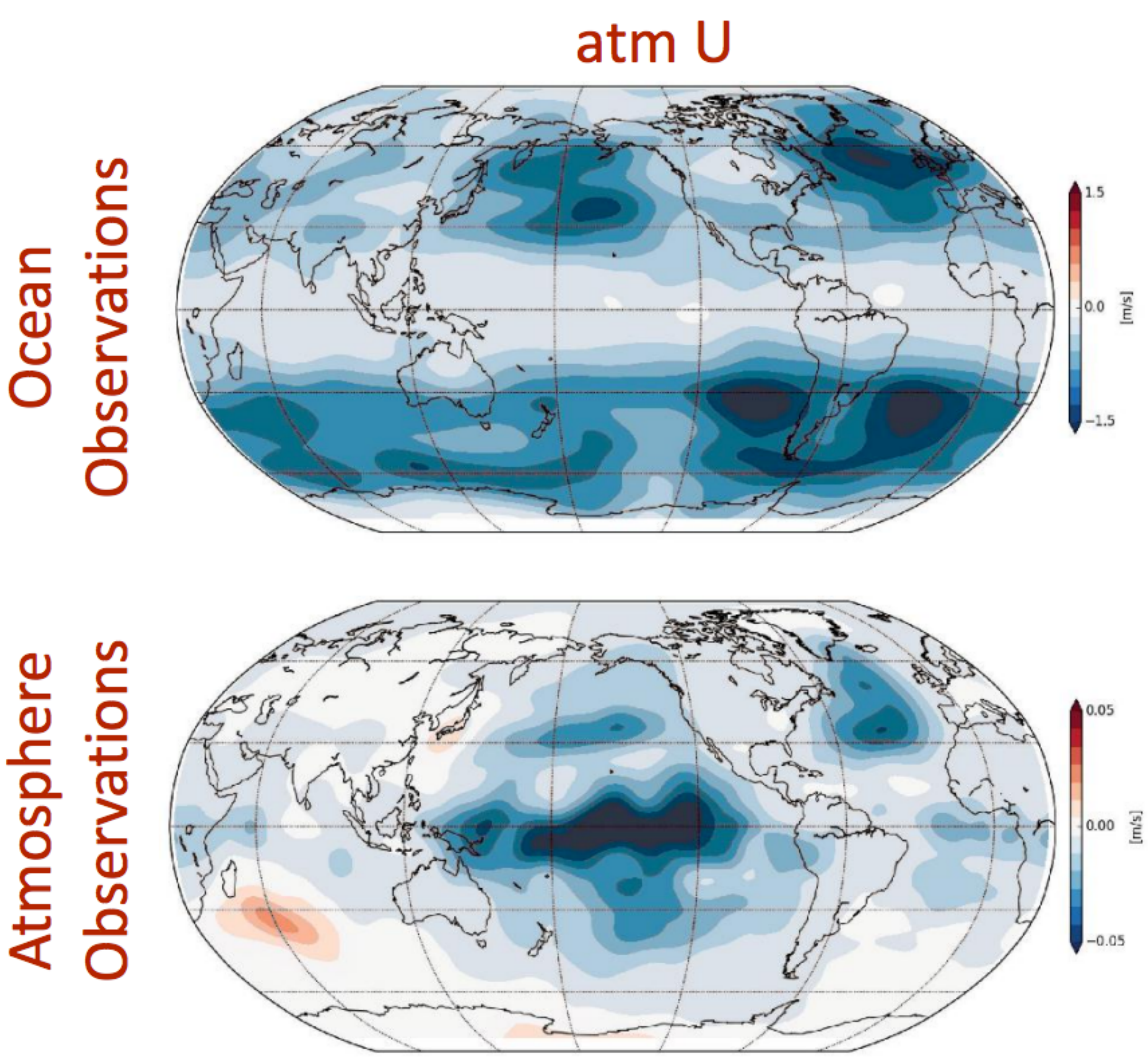
http://www.meteo.fr/cic/meetings/2016/CDAW2016/presentations/3_1.pdf

SELECTED NOTABLE RESULTS: STRONGLY COUPLED DA

➤ Cross-domain correlations improve the analysis in OSSEs

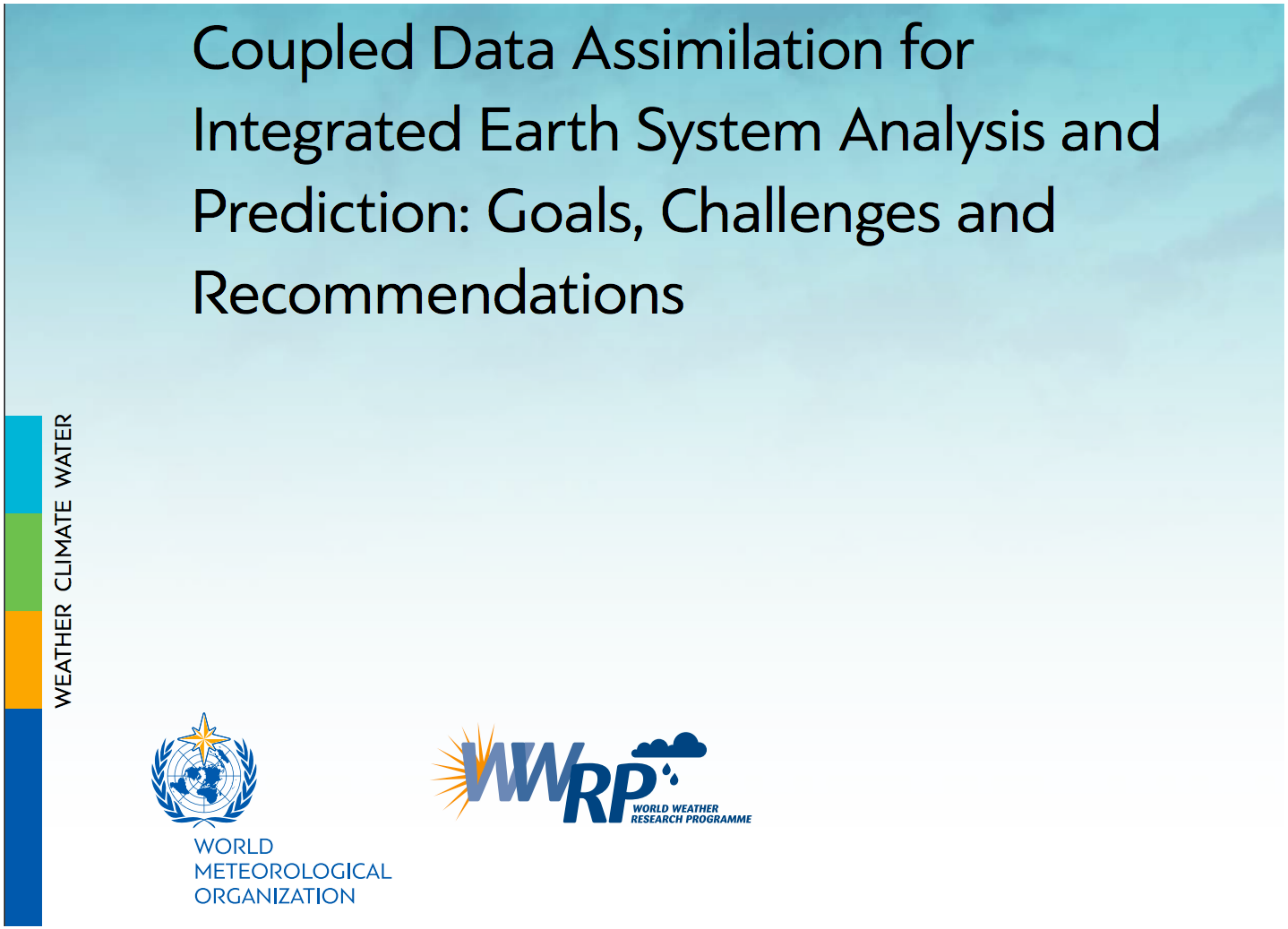
STRONG-WEAK, blue is good

- The opposite experiment (assimilating OCN obs into the atmosphere) shows improvement as well



T. Sluka et al. (2016)

http://www.meteo.fr/cic/meetings/2016/CDAW2016/presentations/3_1.pdf



- Penny et al., 2017: Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations. *WMO*, WWRP 2017 - 3. http://www.wmo.int/pages/prog/arep/wwrp/new/documents/Final_WWRP_2017_3_27_July.pdf

RECOMMENDATIONS (**OBSERVING SYSTEM**)

- (a) Standardize the observing network for all Earth system domains in order to meet the timeliness and quality control requirements of NWP;
- (b) identify gaps in the observing system that are essential for constraining CDA applications, including fluxes at the domain interfaces, and traditionally under-observed regions

RECOMMENDATIONS (**OBSERVING SYSTEM**)

- **Increase the observing effort of the cross-domain interfaces.** This includes measurements of air-sea fluxes, ice-ocean fluxes, air-land fluxes, etc.
- **Encourage field campaigns that plan for co-located observations spanning multiple domains**.
- **Increase collaboration between field campaigns, modellers, and the CDA community** for conducting process studies that cross-domain boundaries.
- Regional field campaigns may be valuable for performing focused tests of CDA on a limited scale.
- **Increase the observing effort for areas of the Earth system that are under-observed and under-constrained**. As we shift from forced single-domain models to coupled Earth system models, the predominant impacts from biases in one domain can shift to another. To constrain these biases there must be a concerted effort to constrain regions **such as the deep ocean, sea ice, and the ocean under sea ice**.
- **Establish a mechanism for developing countries to securely contribute local observations** to major global NWP operational forecasts. Further, establish a visiting scientist plan for providing expert training in DA and NWP to forecasters, researchers, and students from these countries.
- Dedicated field campaigns should be identified that can improve earth-system predictability through better formulation of either forecast, observation, or CDA methods due the insights derived from those campaigns.
- **Establish a mechanism to contribute observations from temporary and/or experimental observing systems that focus on taking measurements across domains**, or that complement an existing observing system in a different domain to be used for CDA studies.

RECOMMENDATIONS (**MODELING**)

- (g) promote improved representation of model uncertainty in the coupled forecast system using stochastic physics and other advanced methods;
- (h) perform research to increase knowledge on how to best represent evolving errors in non-atmospheric model components (e.g. sea ice, land and ocean) on the timescales of NWP;
  - **Improved representation of cross-domain interfaces**
  - **Better separation of scales in coupled Earth System Models**
  - **Revisit the 'domain decomposition' of Atmos/Ocean/etc.**
  - **Better tuning of short-term sensitivities (CLVs)**

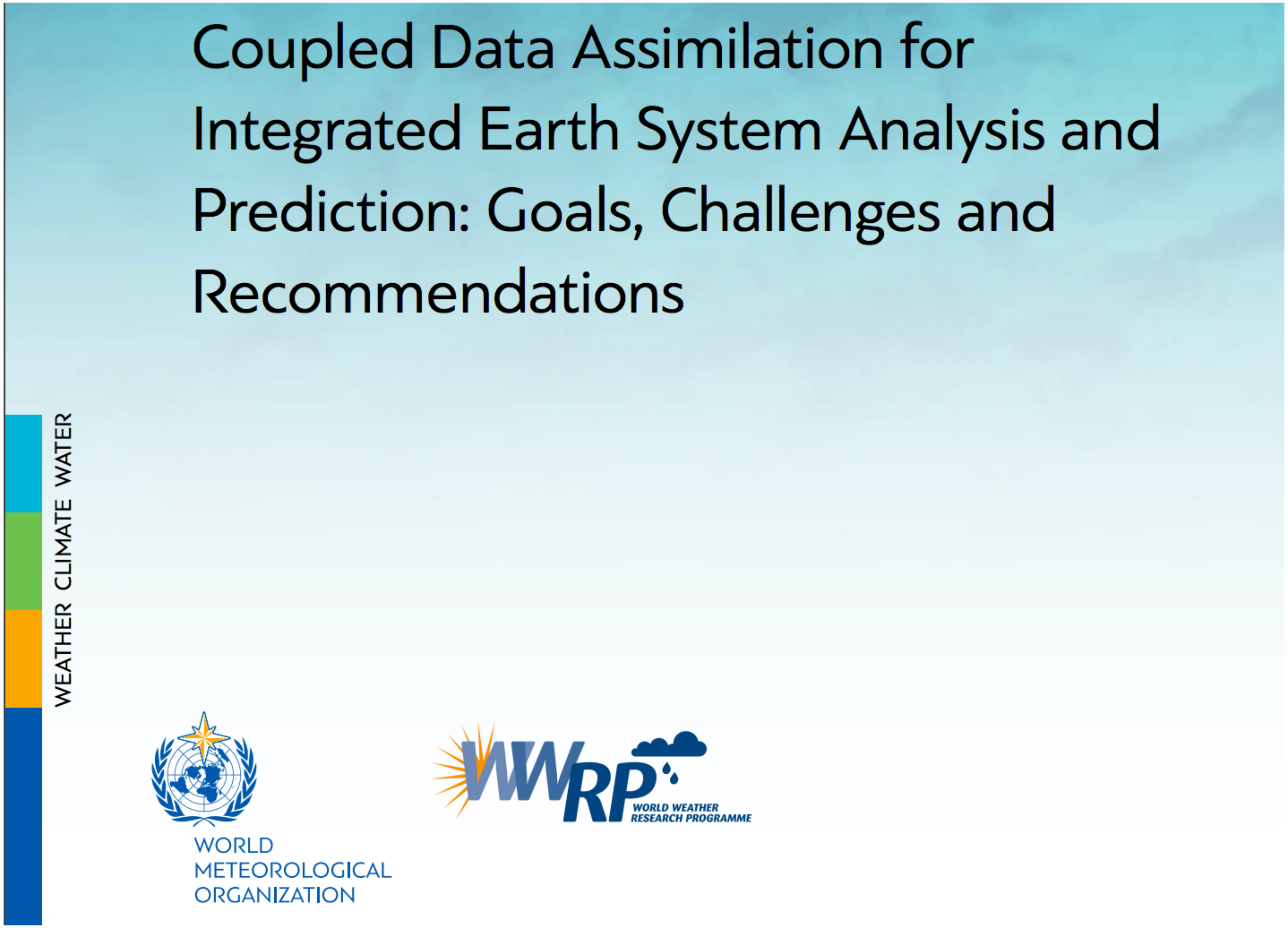
RECOMMENDATIONS (**MODELING**)

- **Representations of surface fluxes must be improved,** whether through improvements to the bulk formulae or improved resolution and modelling of the near surface boundary layers. A more sophisticated 'surface interface model' should be considered, to better match observed surface quantities and better resolve modelled surface fluxes.
- **A better understanding of the impact of representing model error in CDA systems is necessary.** This includes, for example, the impact of stochastic parameterizations, bias corrections, surface flux relaxations, and other frequently used tools used to correct model errors in DA systems.
- **An improved understanding of the dynamical instabilities that dominate coupled systems** is needed in order to develop CDA methods that target unstable modes.
- For any coupled modelling system configuration, **the saturation time scales and amplitudes of each component should be quantified**, and it should be evaluated how these affect the other (faster and slower) model components.
- Coherence between initial conditions of slow and fast modes relies on cross-domain error covariances… **It is possible that deficiencies in the modelling of cross-domain interface dynamics may render large errors in these estimates. New methods and diagnostics are needed to evaluate the cross-domain error covariance derived from coupled model ensembles.**
- **Investigate the time-varying nature of cross-domain error covariances in detail,** showing how the cross-domain error covariance structure changes based on diurnal, seasonal, and longer timescales
- **Evaluate the impact of model resolution on cross-domain error covariances.**
- **Grid choice, choice of vertical coordinate, and choice of interpolation method to interface** multiple **domains** can impact the accuracy of coupled models and CDA systems. Research is needed to understand the amount of error introduced due to these sources, and quantify how much that error degrades the coupled model forecast.

RECOMMENDATIONS (**PROGRAMMATIC**)

- Due to the highly interdisciplinary nature of CDA, it is recommended that CDA workshops have a significant training and tutorial sessions on coupled Earth system processes relevant to CDA. It is also highly encouraged that **future CDA workshops encourage participation of experts in Earth system modelling and** experts focusing on **observations and process studies** of exchange processes **between Earth system domains**.
- …

CONCLUSION



Coupled Data Assimilation for Integrated Earth System Analysis and Prediction: Goals, Challenges and Recommendations

WEATHER CLIMATE WATER

WORLD METEOROLOGICAL ORGANIZATION

WWRP WORLD WEATHER RESEARCH PROGRAMME

- Contact: Steve.Penny@noaa.gov