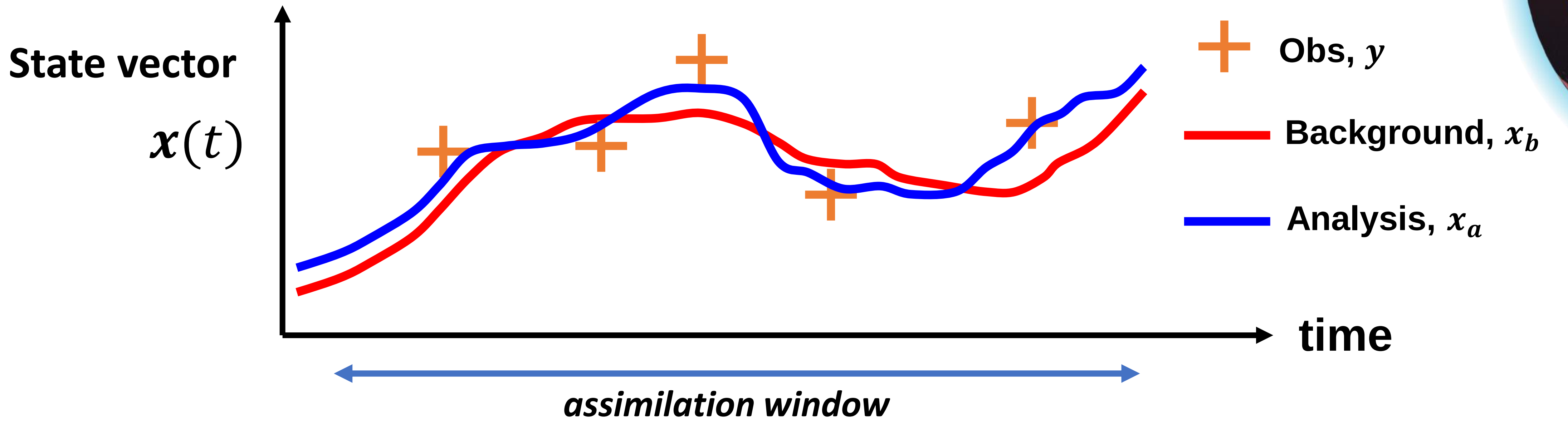
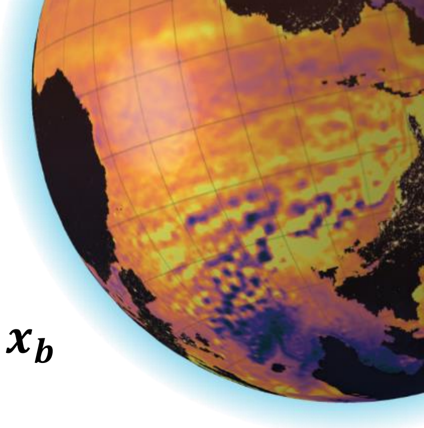
Ocean Predict
SYMPOSIUM OP'24
In partnership with unesco Intergovernmental Oceanographic Commission
2021 2030 United Nations Decade of Ocean Science for Sustainable Development

Weak Constraint 4D-Var Data Assimilation in the Regional Ocean Modeling System (ROMS) using a Saddle-Point Algorithm

Andy Moore, UC Santa Cruz
Hernan Arango, Rutgers University
John Wilkin, Rutgers University
Chris Edwards, UC Santa Cruz

4-Dimensional Variational (4D-Var) Data Assimilation





4D-Var employs variational calculus to minimize a cost function.

Forecast model: $\underbrace{x_k}_{\text{state vector}} = \underbrace{\mathcal{M}_k}_{\text{forecast model}}(x_{k-1}) + \underbrace{q_k}_{\text{model error}} \qquad k = 1, n$

Prior assumption: "perfect" model ($q_k = 0$) → Strong constraint DA

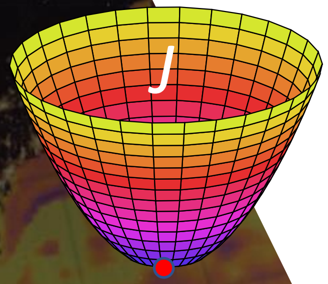
"imperfect" model ($q_k \neq 0$) → Weak constraint DA

Motivation: Increase the efficiency of the iterative 4D-Var cost function minimization.

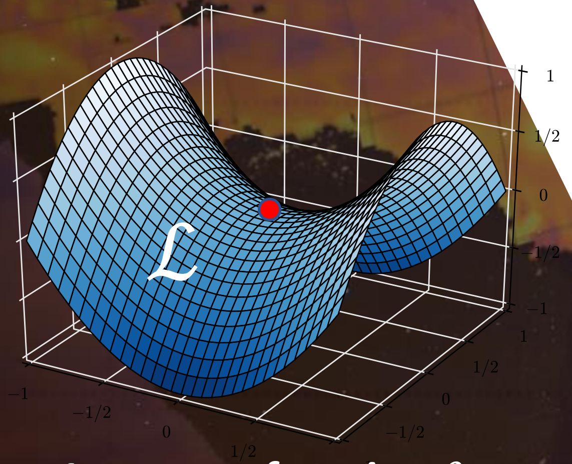
Incremental Weak Constraint 4D-Var

Fisher and Gürol (2017)

Incremental → linearize about the background x_b → $\delta x_k = x_k - (x_b)_k$; $\delta q_k = q_k - (q_b)_k$



Cost function J



Lagrange function ℒ

Forcing Formulation

↓

Minimize the usual cost function J

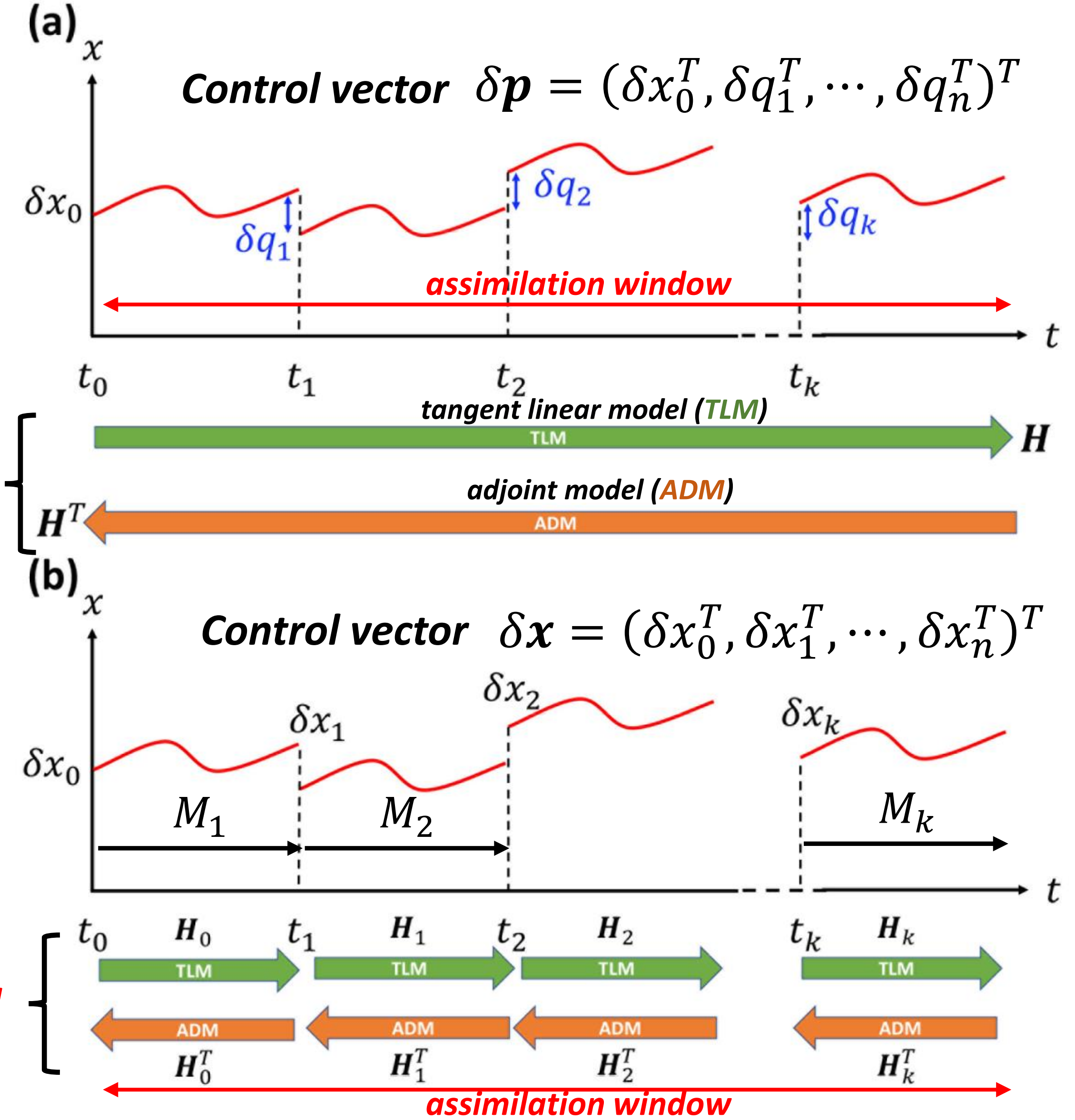
Time sequential

Saddle-point Formulation

↓

Minimize J subject to additional constraints → Lagrange function

Time parallel



Saddle-Point 4D-Var in the California Current System using the Regional Ocean Modeling System (ROMS)

Two ROMS configurations:

- 1/3rd degree resolution, 42 σ-levels
- COAMPS surface forcing
- ECCO open boundary conditions
- Observations:
  - satellite SST
  - Aviso altimetry
  - Argo profiling floats
- 4-day 4D-Var windows
- Standard test case (WC13)

- 1/10th degree resolution, 42 σ-levels
- ERA surface forcing
- Global HYCOM open boundary conditions
- Observations:
  - satellite SST
  - Aviso altimetry
  - Argo profiling floats
- 8-day 4D-Var windows

Forcing formulation: RBCG (Restricted B-preconditioned CG)

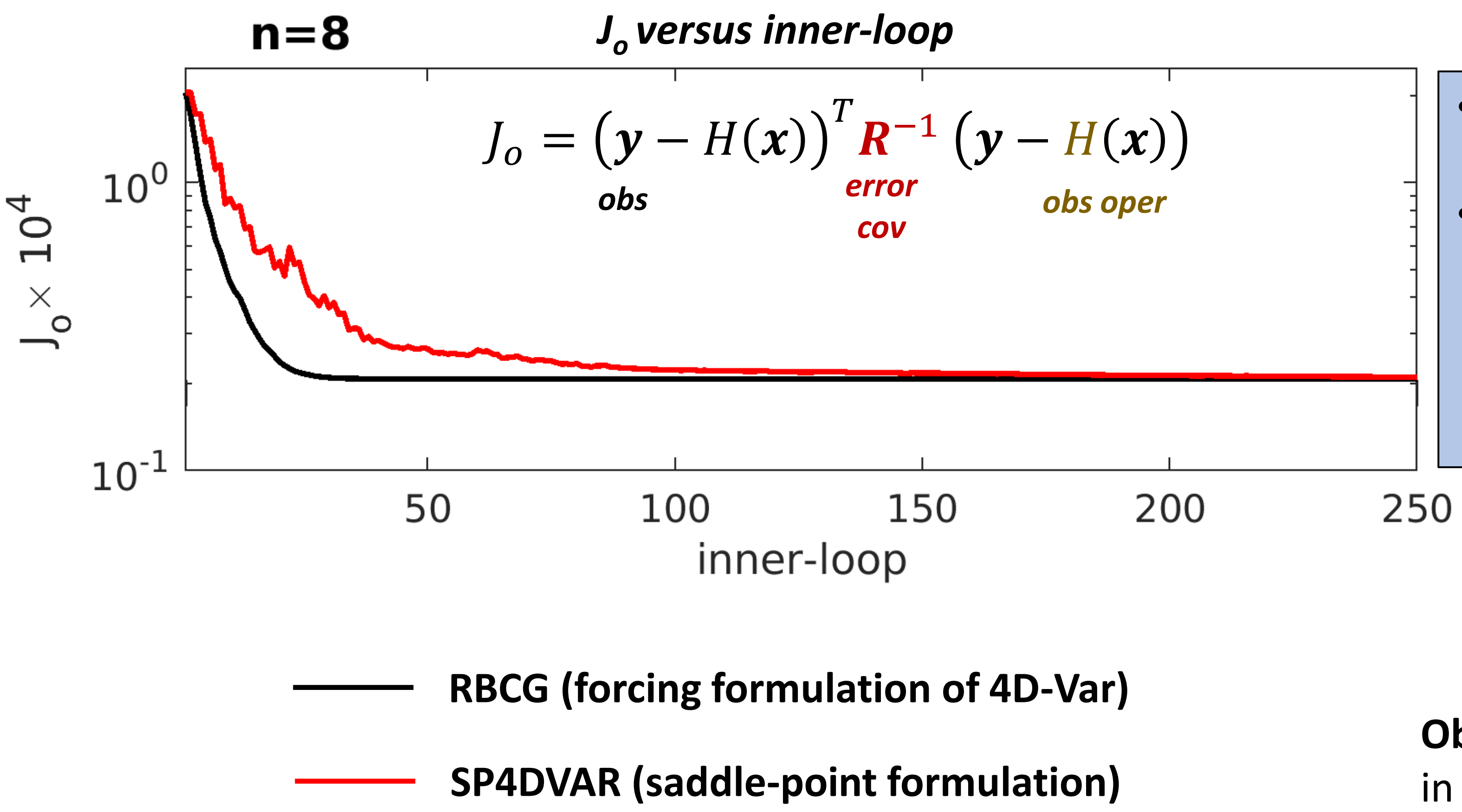
Saddle-point formulation: SP4DVAR

Single 4d-Var cycle:
1 outer-loop, n=8 sub-intervals, Q=0.2B

Convergence Validation

1/3rd degree resolution, 4-day assimilation window, 3-7 Jan 2004

1 outer-loop, 4-day window, *n*=8 sub-intervals, *Q=0.2B*



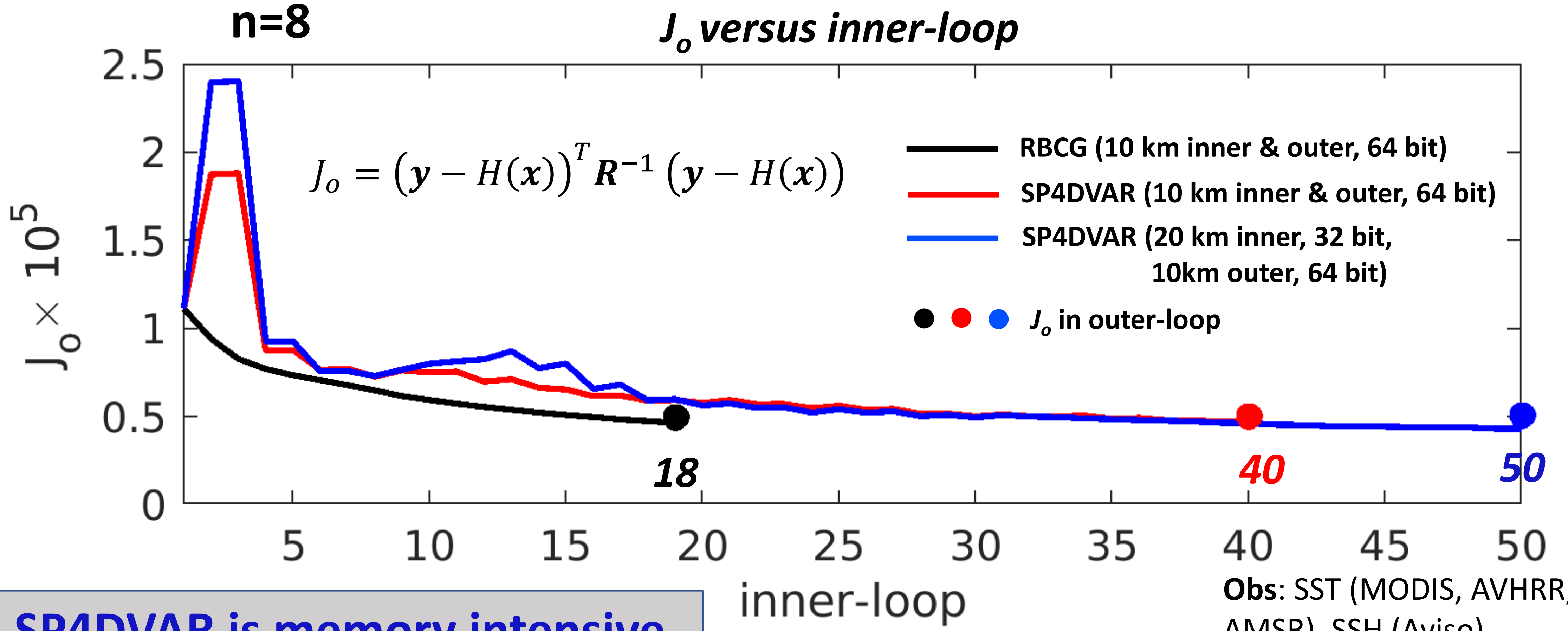
- SP4DVAR & RBCG yield the same solution.
- SP4DVAR converges more slowly than RBCG, ***but*** SP4DVAR benefits from time parallelization → less wall clock time

Obs: Blended SST, SSH (Aviso), in situ T & S (XBT, CTD, Argo)

Termination before Convergence

1/10th degree resolution, 8-day assimilation window, 3-11 Jan 2004

1 outer-loop, 8-day window, *n*=8 sub-intervals, *Q=0.2B*



SP4DVAR is memory intensive

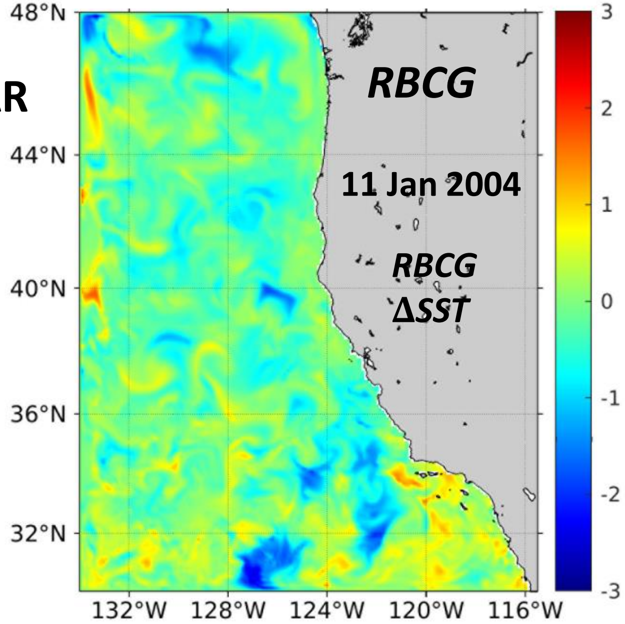
Obs: SST (MODIS, AVHRR, GOES, AMSR), SSH (Aviso), in situ T & S (XBT, CTD, Argo)

SST 4D-Var Increments: 10km resolution, 3-11 Jan 2004

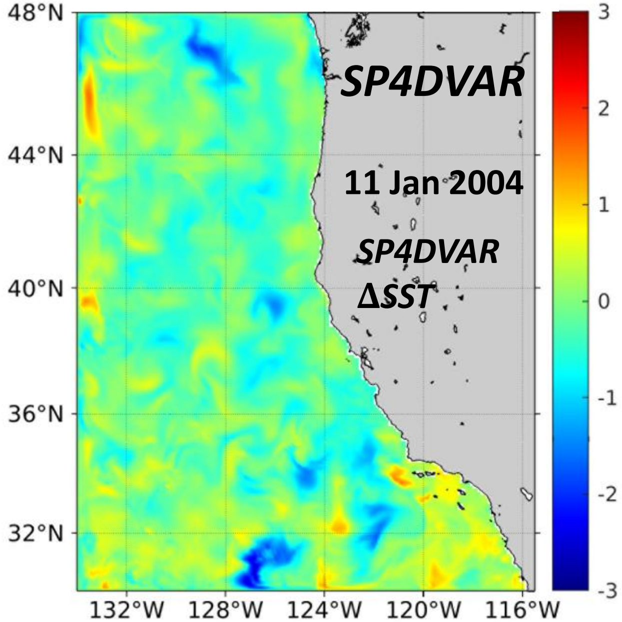
SST increments:
RBCG vs SP4DVAR

1 outer-loop
8-day cycle
n=8, Q=0.2B

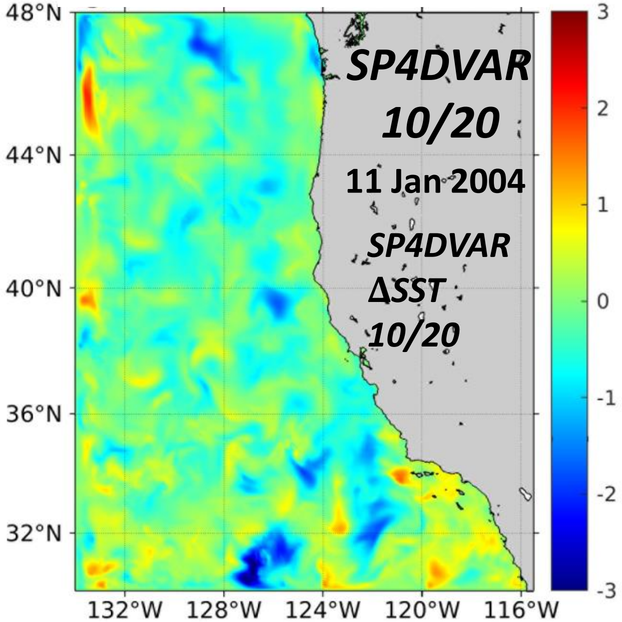
Forcing Formulation



Saddle-Point 4D-Var



Saddle-Point 4D-Var



Outer-loop J_o:

$J_o = 5.03 \times 10^4$

$J_o = 5.04 \times 10^4$

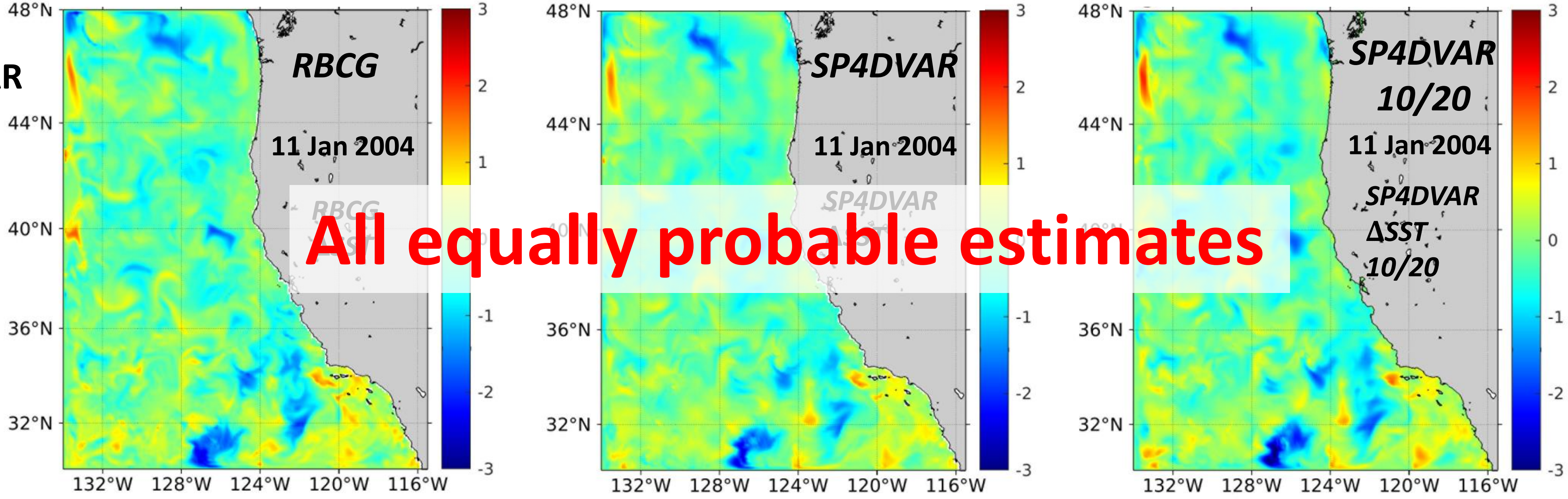
$J_o = 5.10 \times 10^4$

SST 4D-Var Increments: 10km resolution, 3-11 Jan 2004

SST increments:
RBCG vs SP4DVAR

1 outer-loop
8-day cycle
n=8, Q=0.2B

Forcing Formulation | *Saddle-Point 4D-Var* | *Saddle-Point 4D-Var*



All equally probable estimates

Outer-loop J_o*:* $J_o = 5.03 \times 10^4$ $J_o = 5.04 \times 10^4$ $J_o = 5.10 \times 10^4$

$$J_o = \left(\boldsymbol{y} - H(\boldsymbol{x})\right)^T \boldsymbol{R}^{-1} \left(\boldsymbol{y} - H(\boldsymbol{x})\right)$$

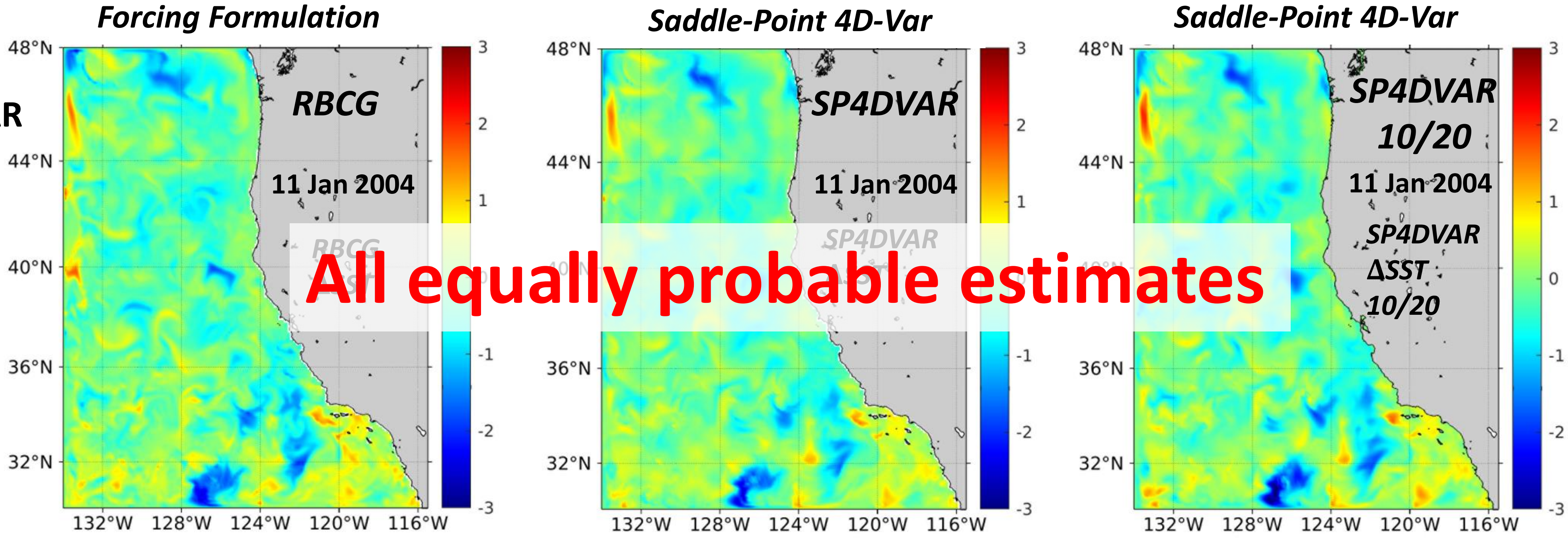
Conditional probability:

$$J_o \propto -\ln\left(P(\boldsymbol{x}|\boldsymbol{y})\right)$$

SST 4D-Var Increments: 10km resolution, 3-11 Jan 2004

SST increments: RBCG vs SP4DVAR

1 outer-loop
8-day cycle
n=8, Q=0.2B



Outer-loop J_o: $J_o = 5.03 \times 10^4$ $J_o = 5.04 \times 10^4$ $J_o = 5.10 \times 10^4$

Relative CPU time per inner-loop: *100%* *12%* *1.6%*

Scales as ~(2n)$^{-1}$

Summary and Conclusions

- Saddle-point 4D-Var has the potential to be a game-changer!
- Saddle-point 4D-Var will run ***much*** faster than RBCG on very large HPC systems
- Outstanding performance issues in ROMS-SP4DVAR:
  - improve efficiency of adjoint model
  - solution assembly & GMRES overhead
  - preconditioning
- Ongoing work: specification of model error covariances, ***Q*** (ML?)

Moore, A.M., H.G. Arango, J. Wilkin and C.A. Edwards, 2023: Weak constraint 4D-Var data assimilation in the Regional Ocean Modeling System (ROMS) using a saddle-point algorithm: Application to the California Current Circulation. Ocean Modelling, https://doi.org/10.1016/j.ocemod.2023.102262.

Ocean Predict

In partnership with unesco Intergovernmental Oceanographic Commission

2021 2030 United Nations Decade of Ocean Science for Sustainable Development

SYMPOSIUM OP'24

ADVANCING OCEAN PREDICTION SCIENCE FOR SOCIETAL BENEFITS

Thank you!

European Commission

Copernicus Marine Service

EU4Oceanobs INTERNATIONAL OCEAN GOVERNANCE EU COMPONENT TO GLOBAL OCEAN OBSERVATIONS

MERCATOR OCEAN INTERNATIONAL

esa

cnes CENTRE NATIONAL D'ÉTUDES SPATIALES

Australian Government Bureau of Meteorology

Met Office

NERSC

Canada