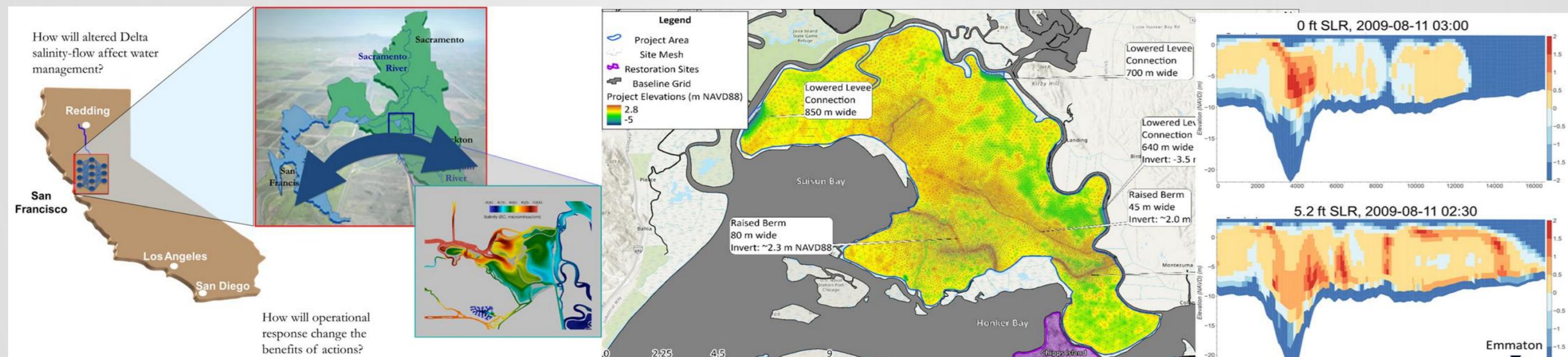


Generalizing Salinity-Outflow Surrogates in CalSim

CWEMF 2025 Annual Meeting, May 13, 2025

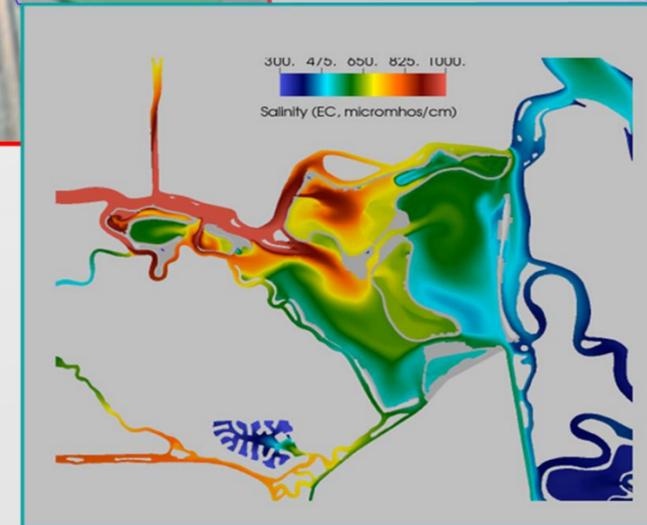
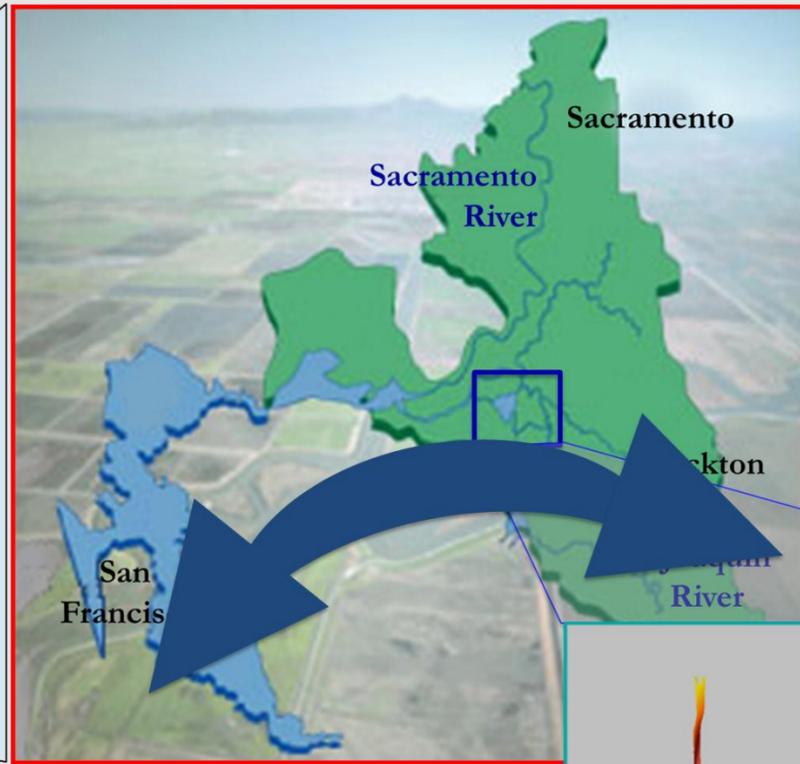


Collaborators:
Resource Management Associates (RMA)
Delta Science Program (DSP)
DWR CalSim team

Eli Ateljevich, PE, PhD, DWR
Lily Tomkovic, PE, PhD, DWR
Can Ruso, UCB

Core Task: Reconcile Statewide/Delta Scales

How will altered Delta salinity-flow affect water management?



How will operational response change the benefits of actions?

Franks Tract Futures SCHISM



Limiters to Generalization

- Reference model DSM2 has limited applicability to:
 - Complex geometry
 - Sea level
 - Beyond-historical drought

- CalSim data used as inputs to reference simulations:
 - band limited to a month
 - compliance limited
 - limited perturbation, memorizeable

- ANNs have single station output
 - No latent representation of the Delta

- Multidimensional hydrodynamic expensive
- Need to get contrasts between “with landscape change and “without”

Some Fixes

Multidimensional models

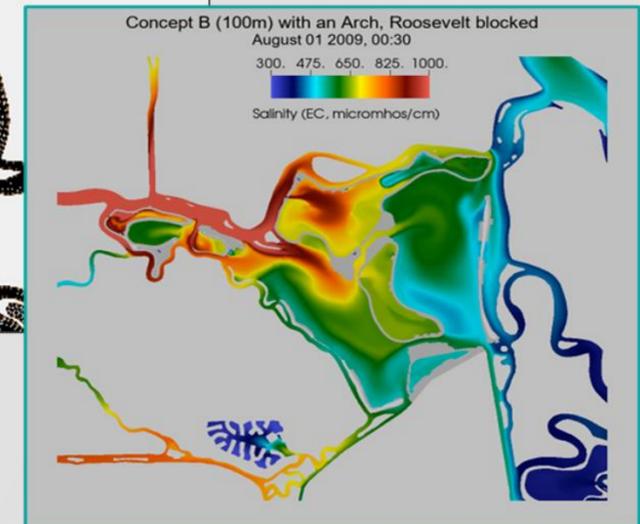
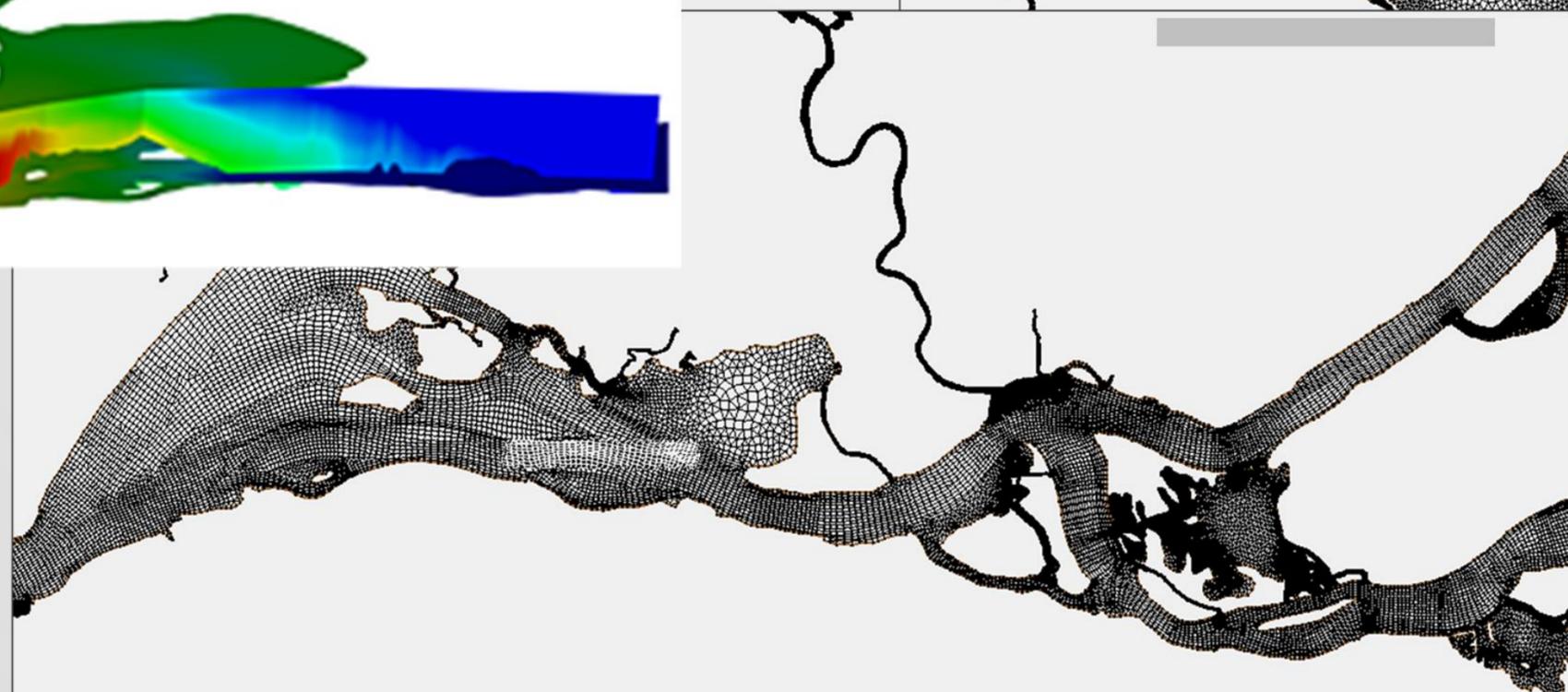
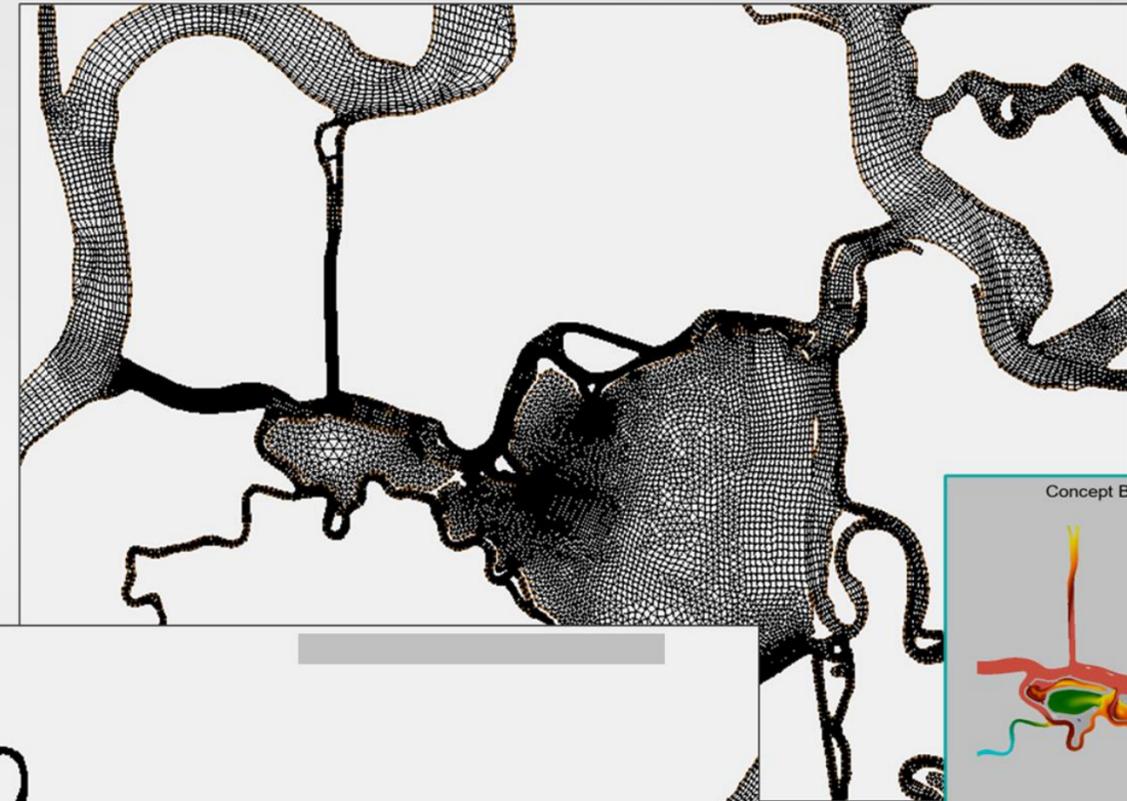
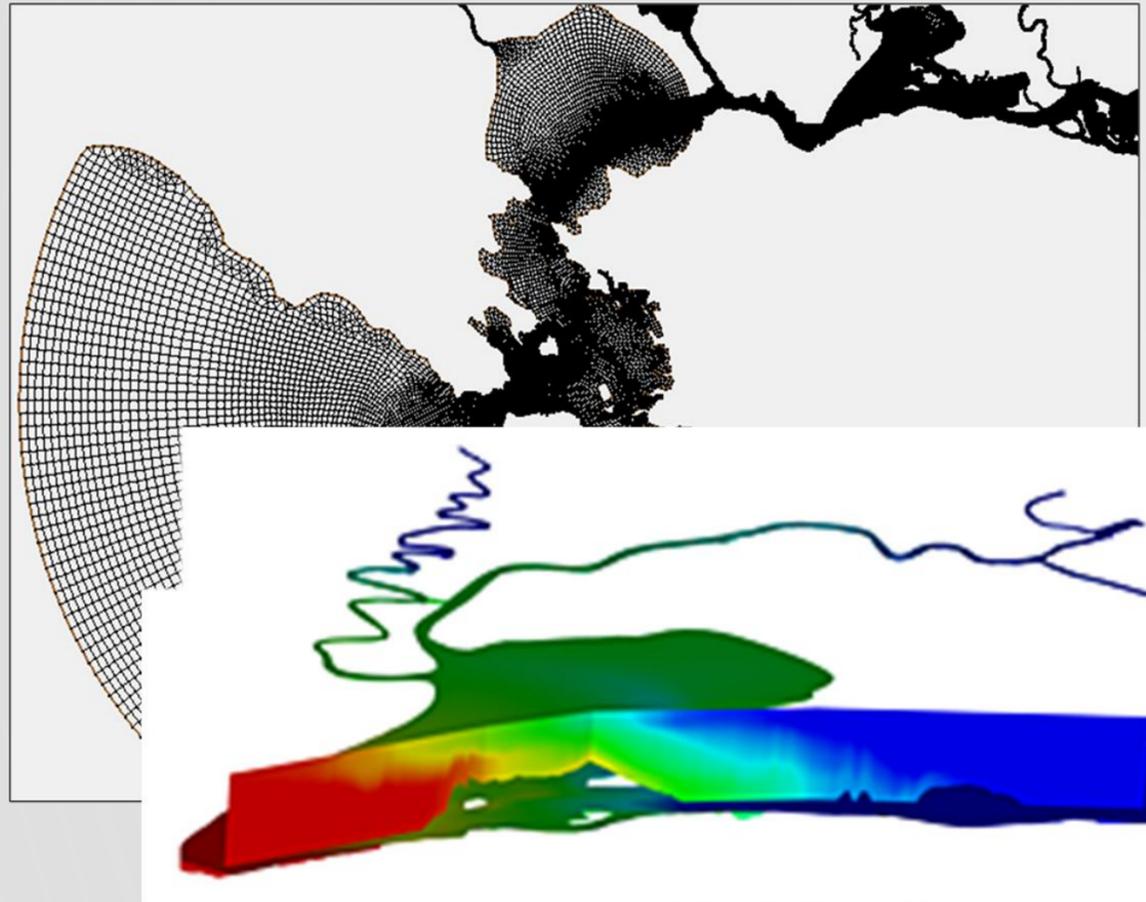
Design of Experiments (Lily Tomkovic)

Multi-station models

Transfer learning
Contrastive Loss functions

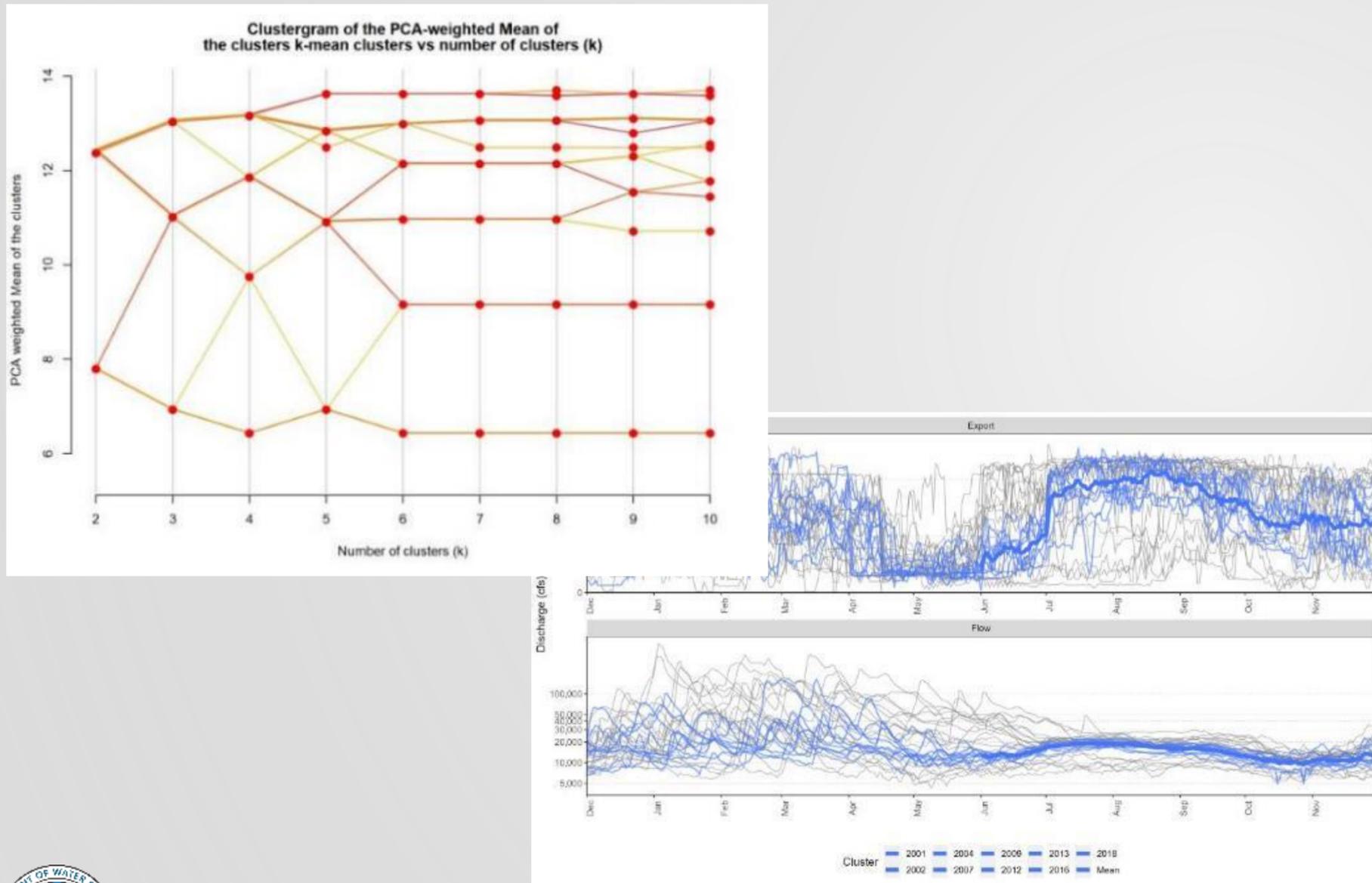


Multidimensional Models (RMA in 2D / Bay-Delta SCHISM 3D)

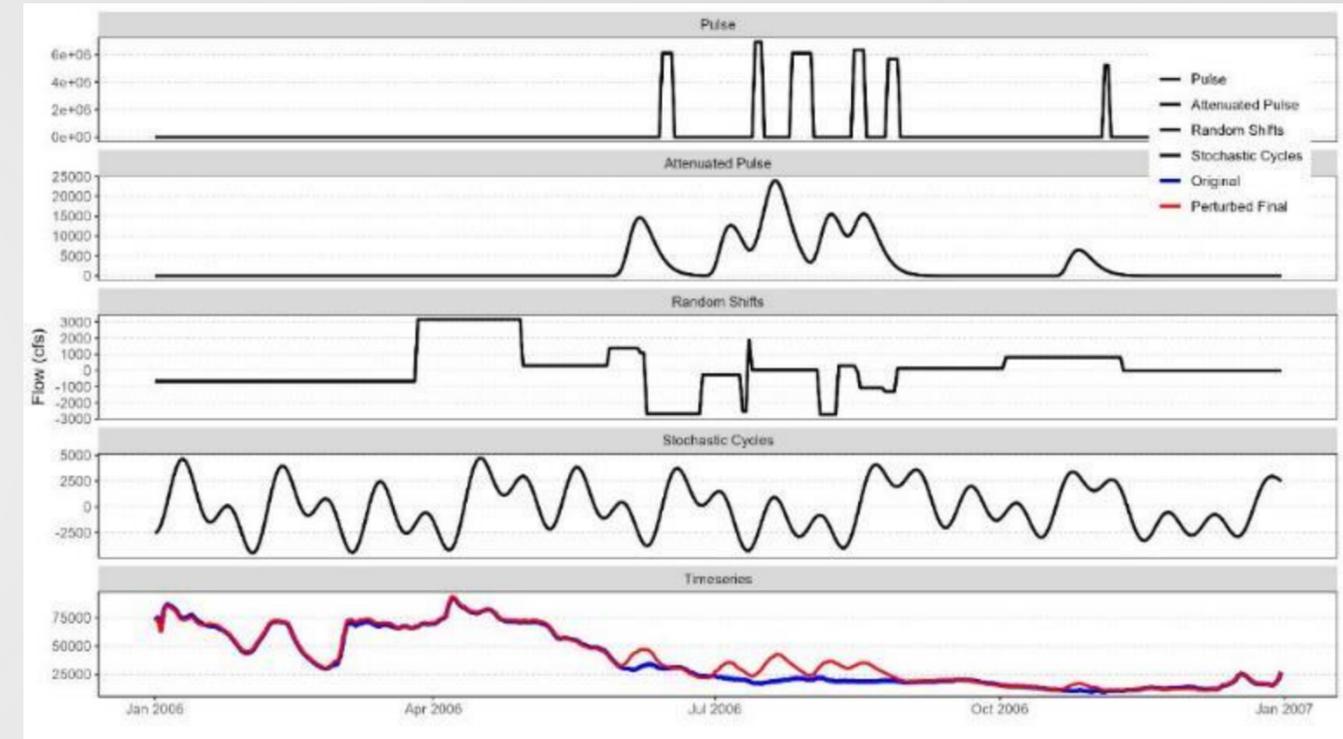


Design of Experiments

Canonical Year Through Cluster Analysis
($\ll 10$)



Aseasonal shifts and Stochastic Perturbation



Training Datasets

For DSM2: ~110 1-2 yr cases

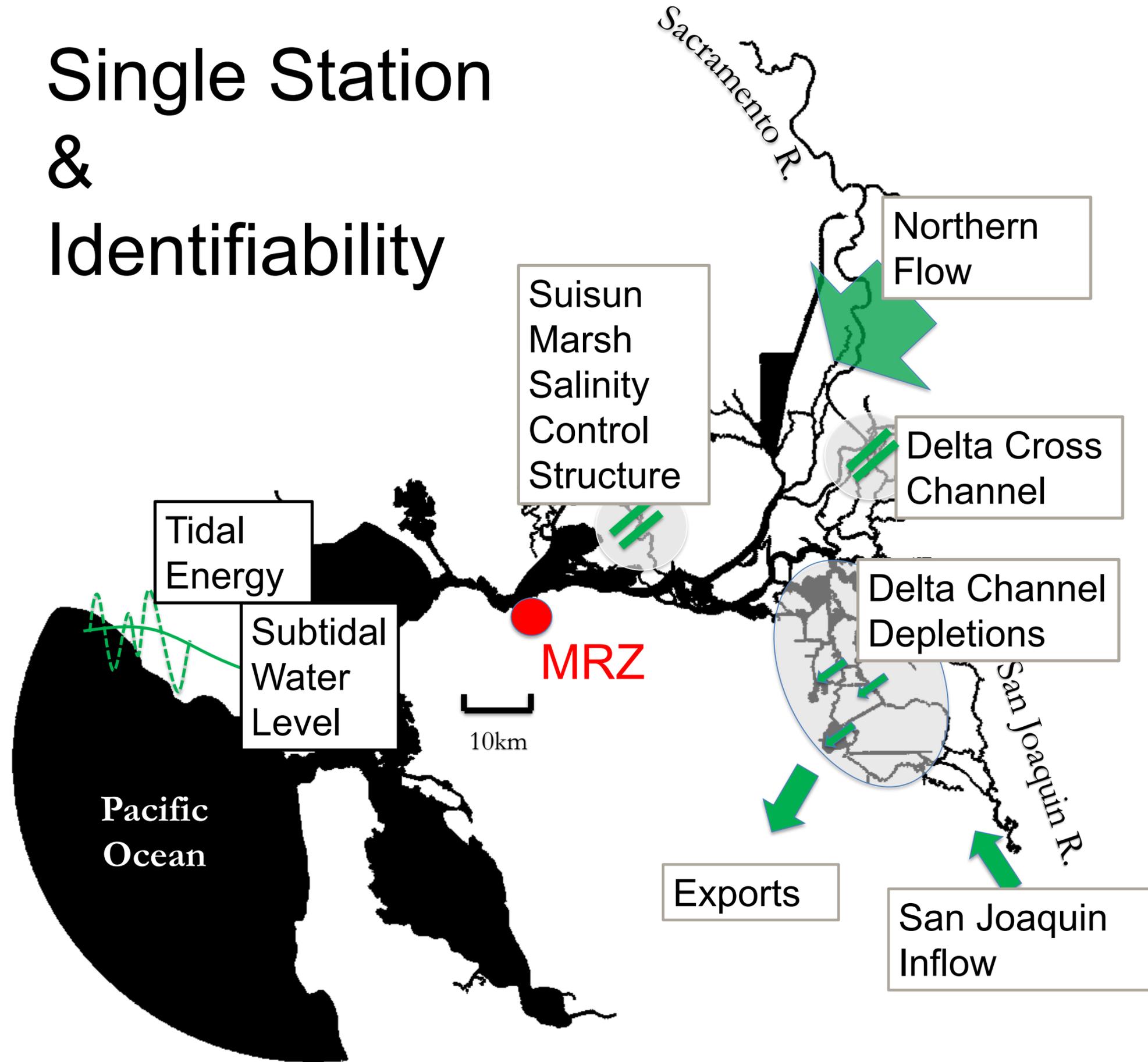
For RMA/SCHISM 7 1-2 yr cases

Little exhaustion of the historical record

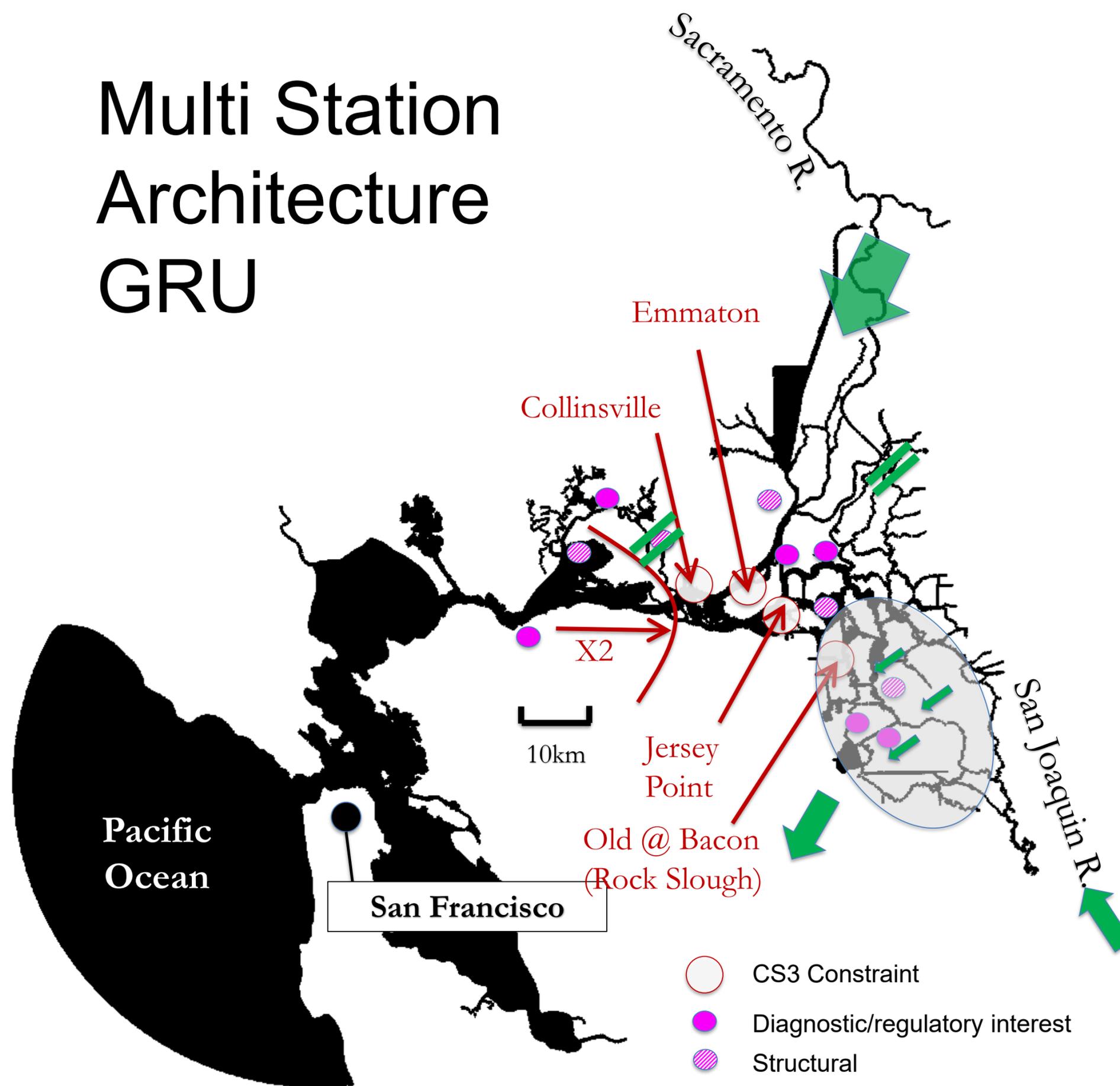
Lots of tricks to avoid memorization



Single Station & Identifiability



Multi Station Architecture GRU

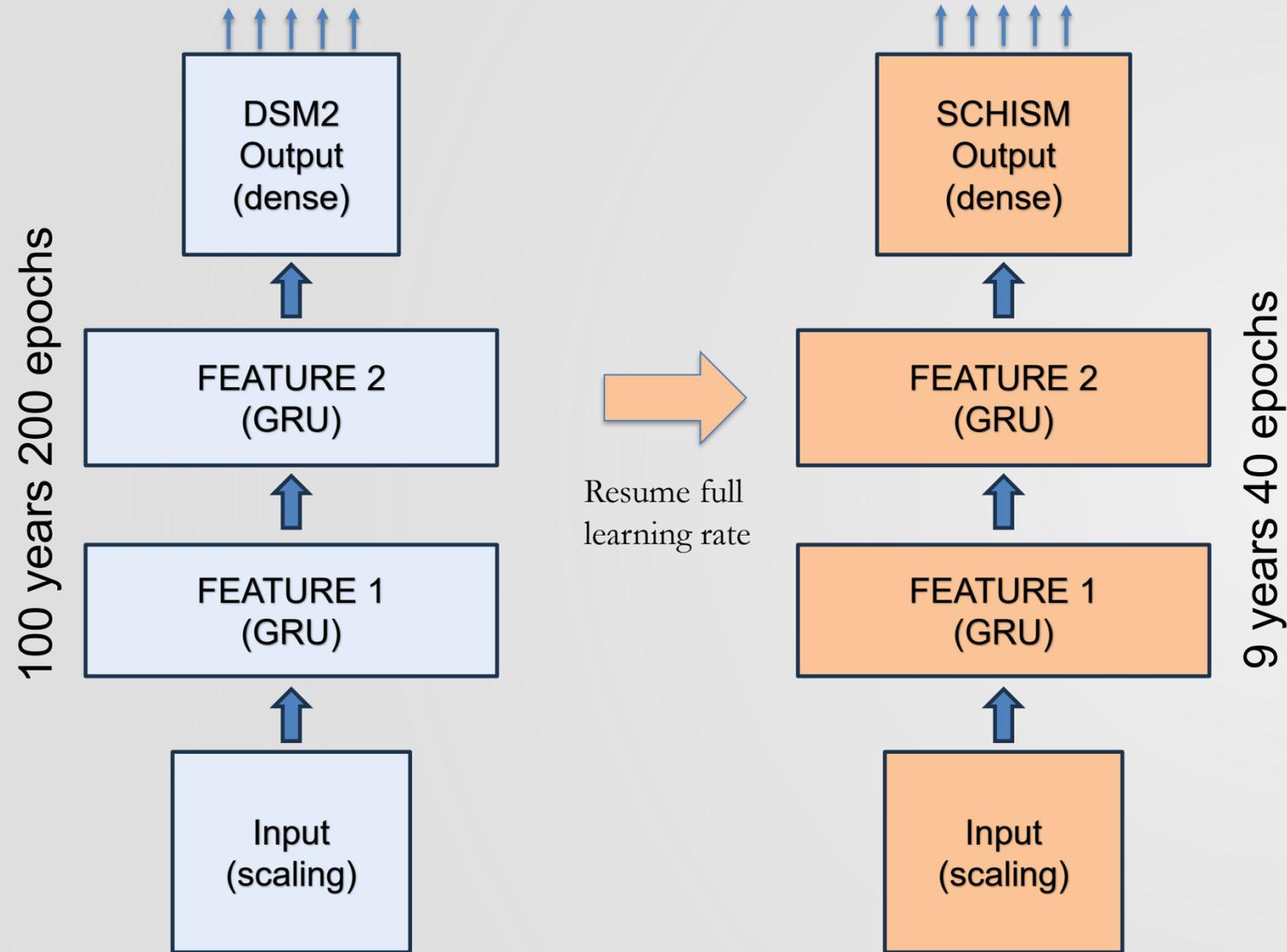


Transfer Learning and Multitask

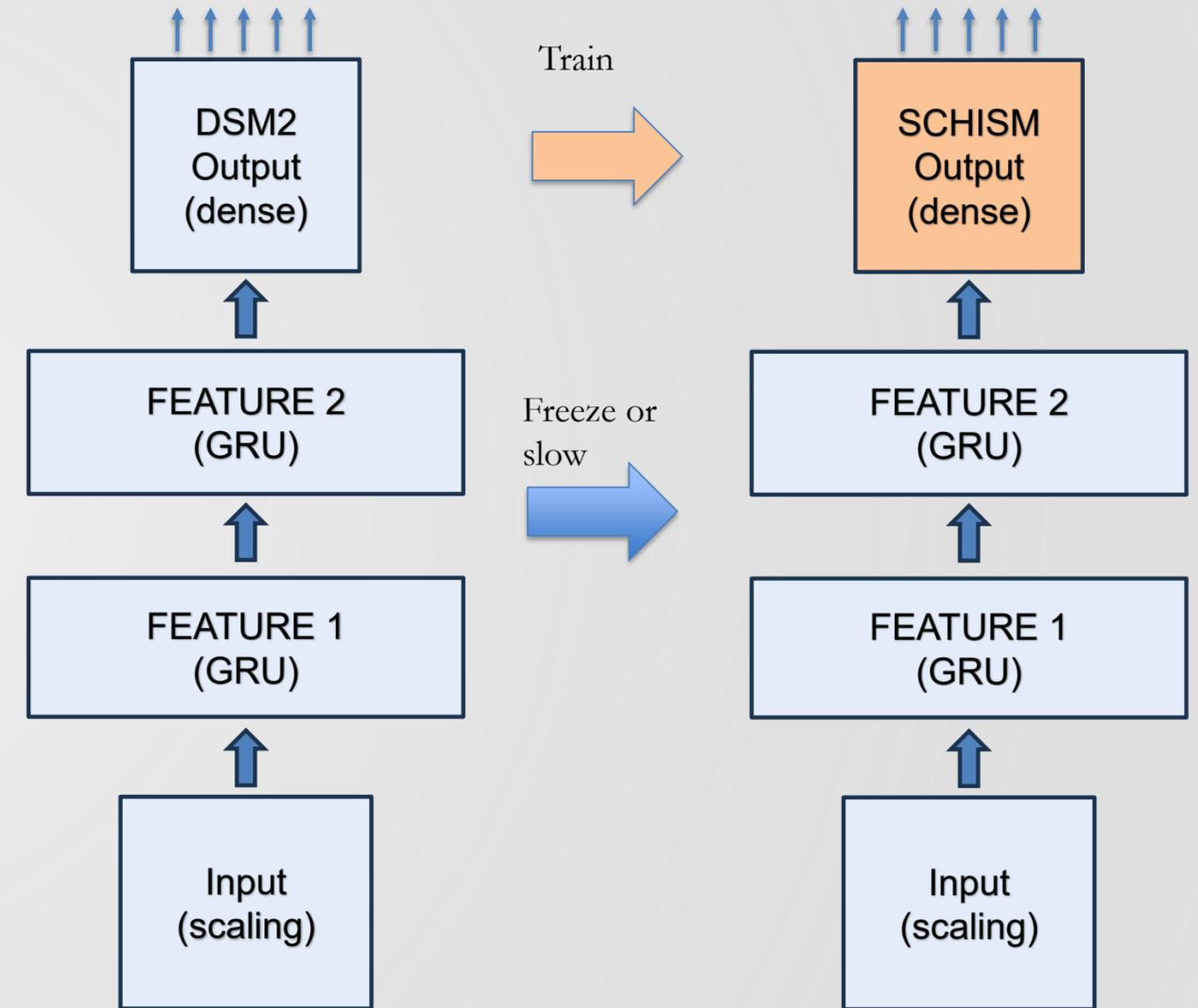
- Transfer: Learn a new task based on previous
- Multitask: Learn several related tasks together



Pre-training (direct)



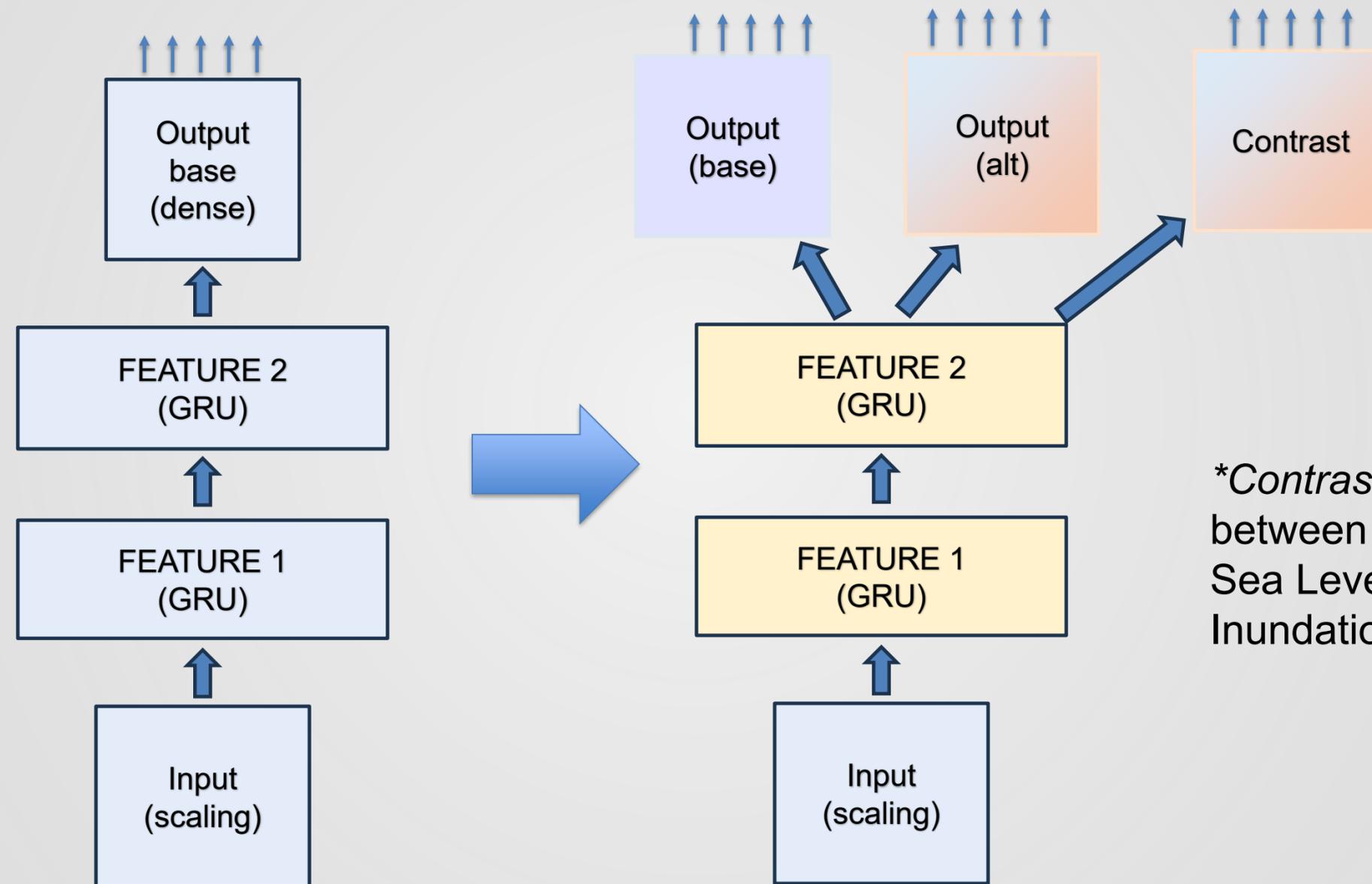
Pre-training plus Embedding (direct + freezing/ low learning)



Many perturbations depending on how much gets embedded



Shared Features, Contrastive* Loss



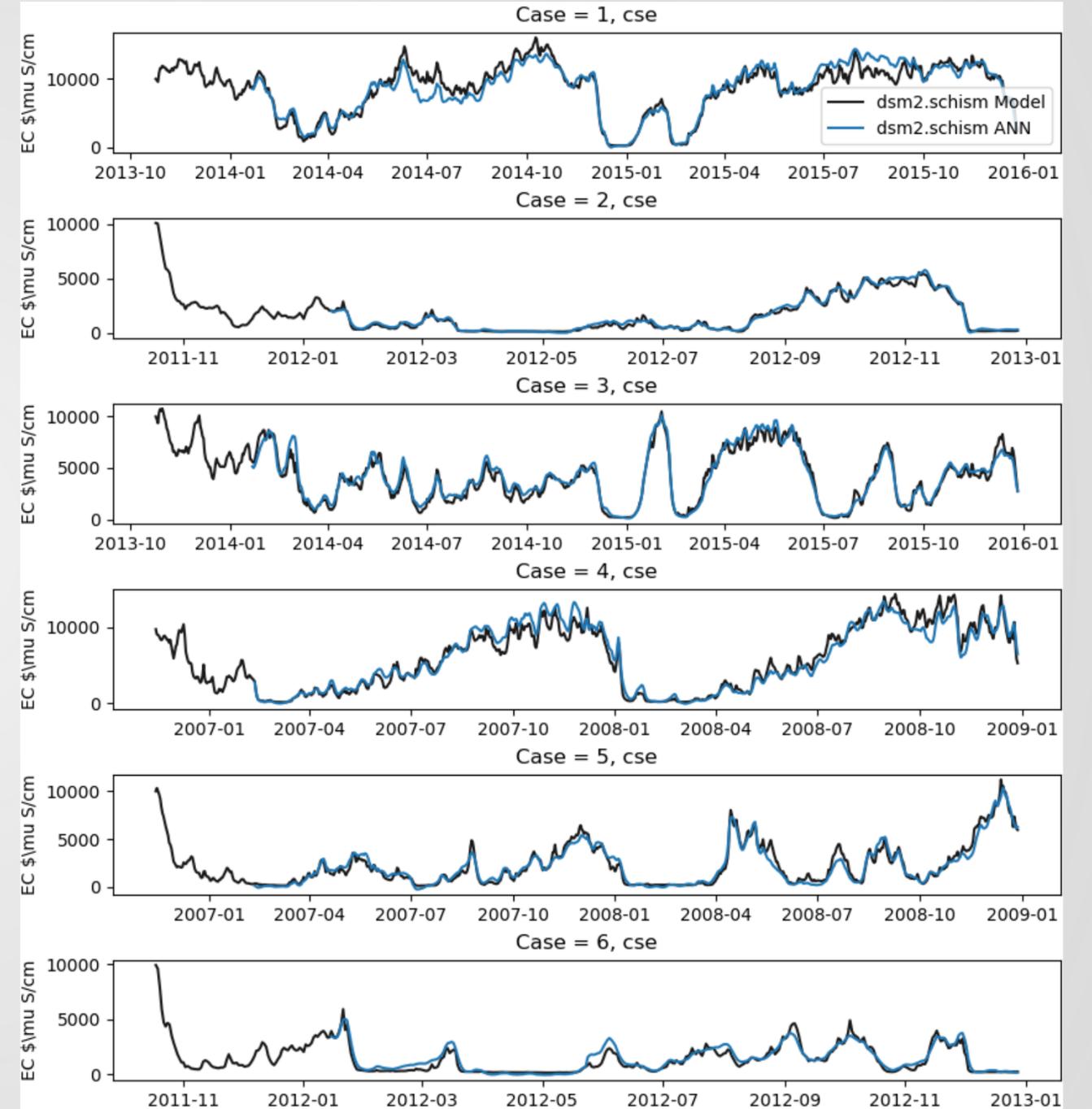
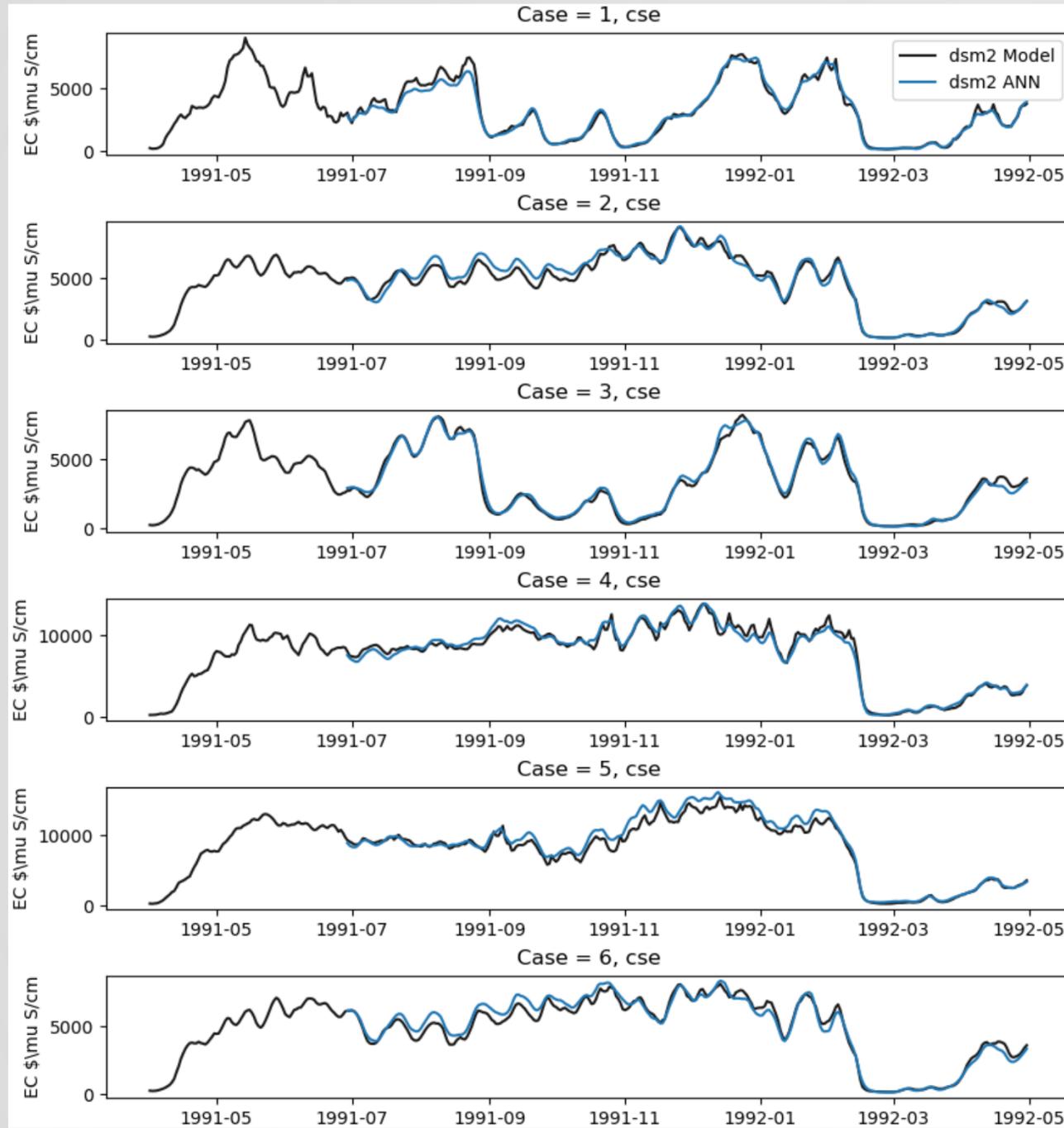
**Contrast* = Error in the contrast between Base and Alternative like Sea Level Rise or Suisun Inundation, say



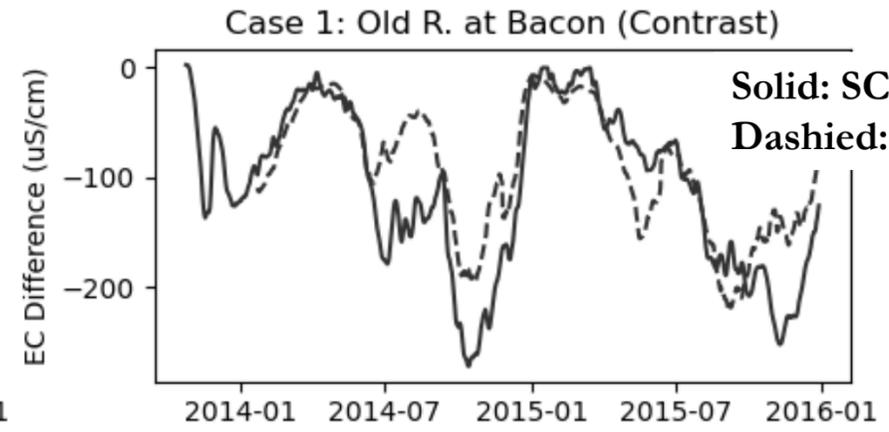
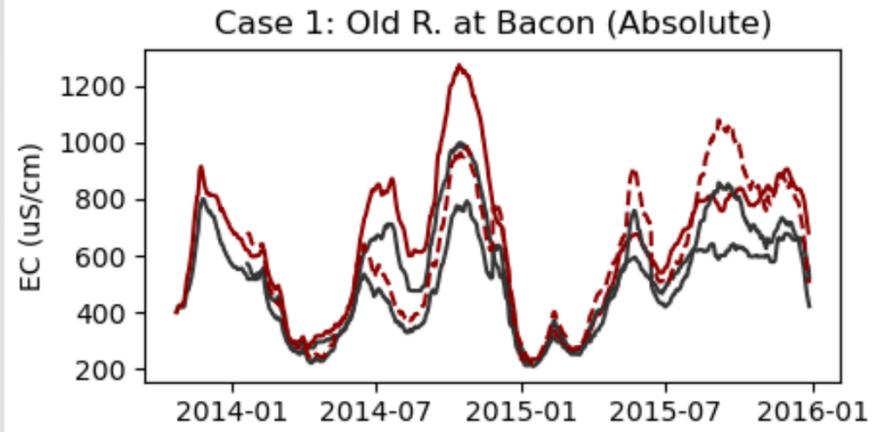
DSM2 + DOE

Pretraining \rightarrow

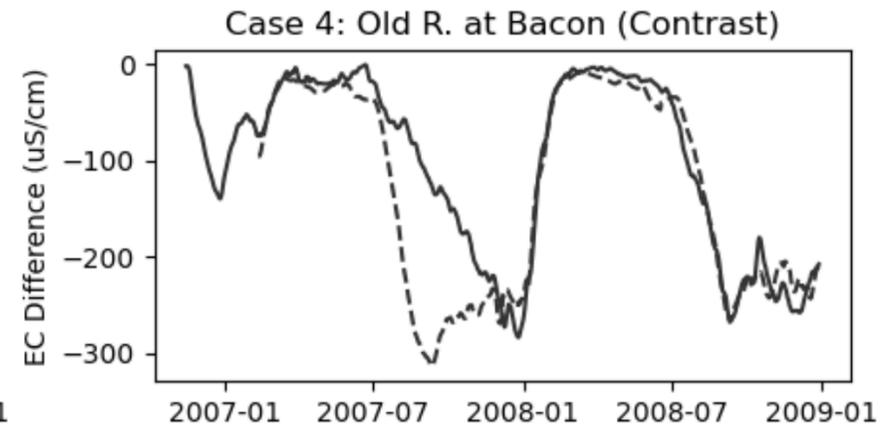
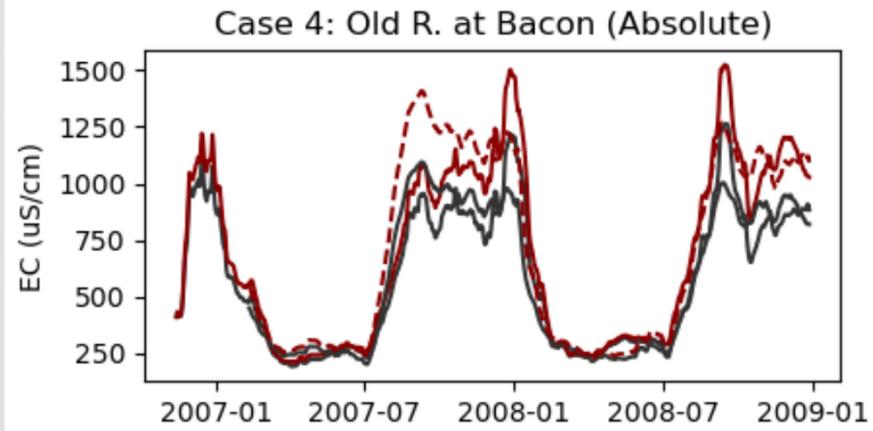
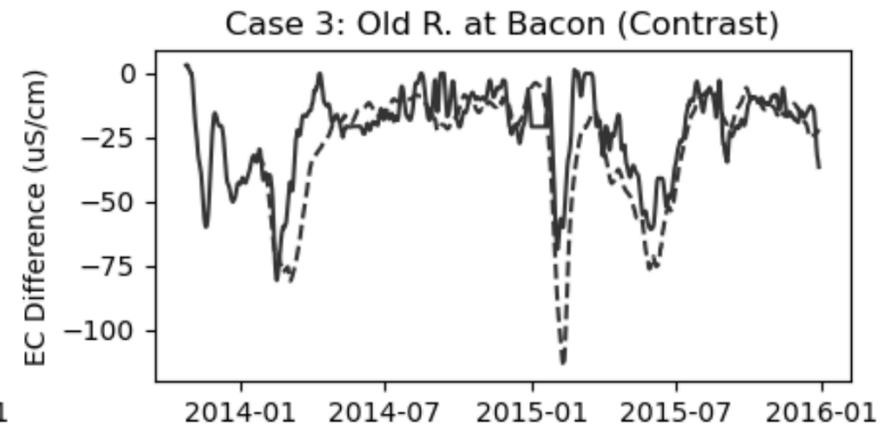
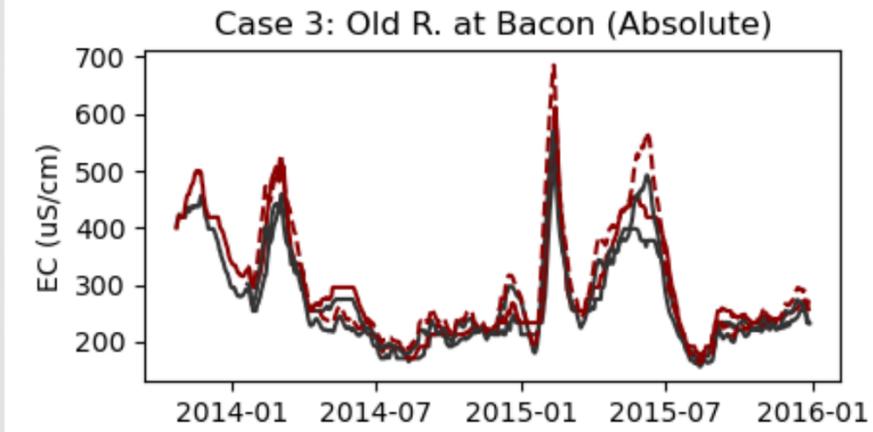
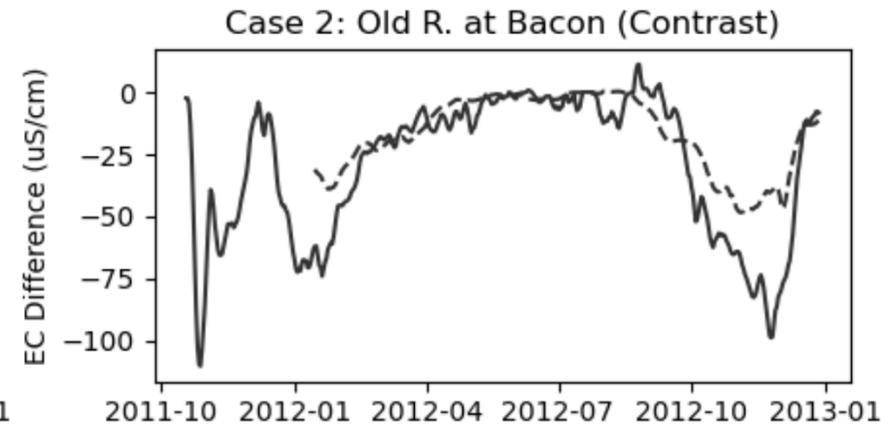
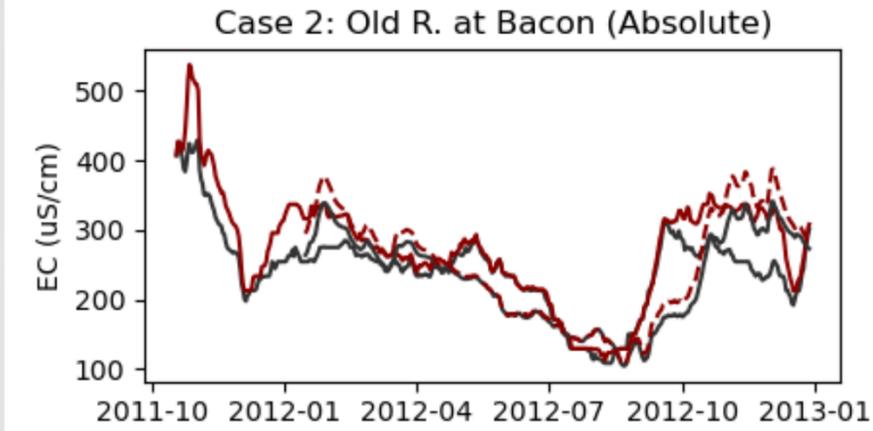
SCHISM+DOE



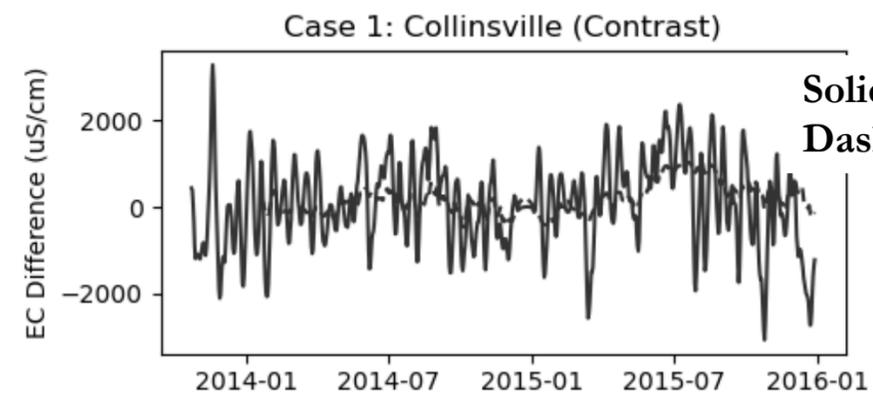
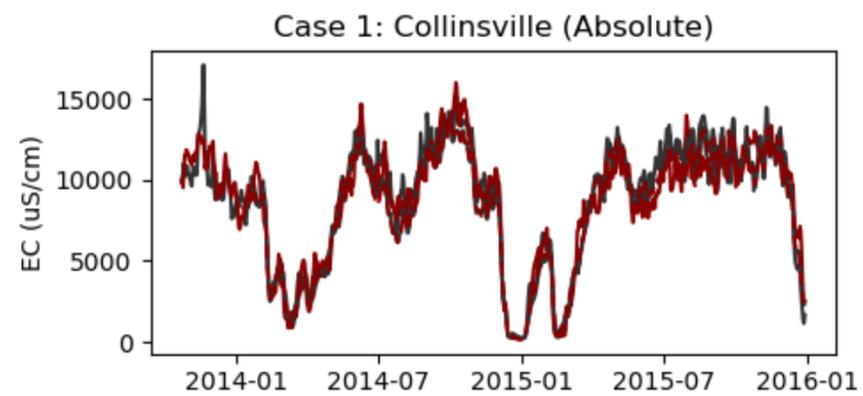
SCHISM Suisun SCHISM Base ANN Suisun ANN Base



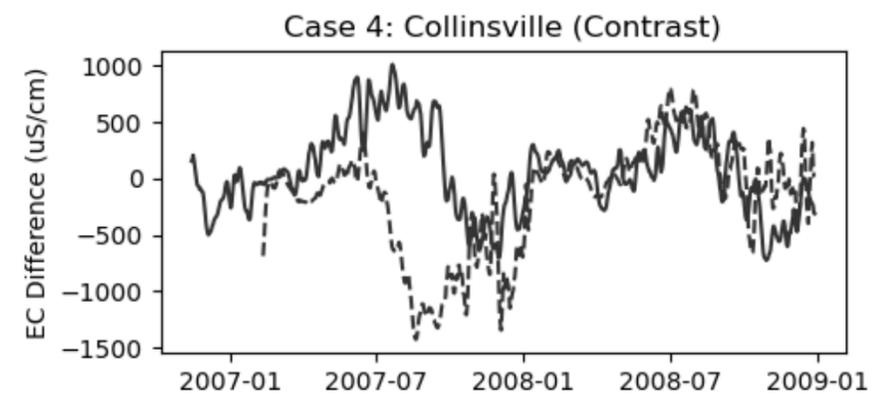
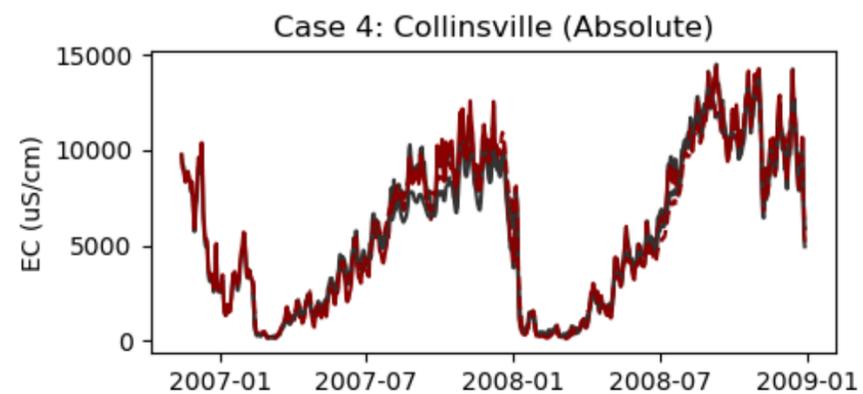
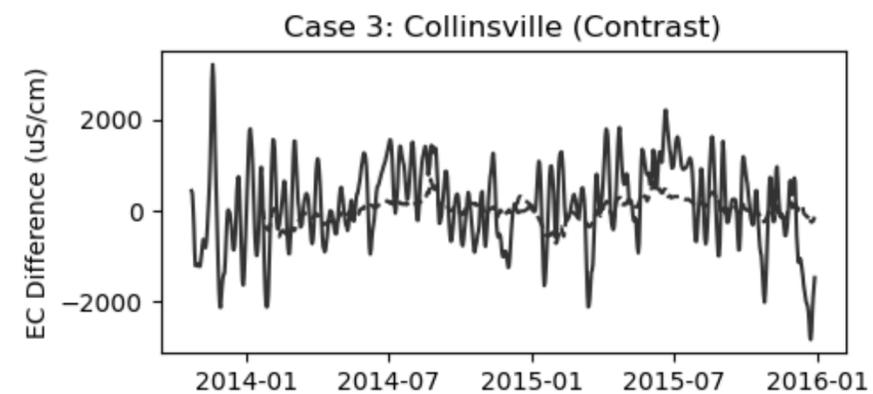
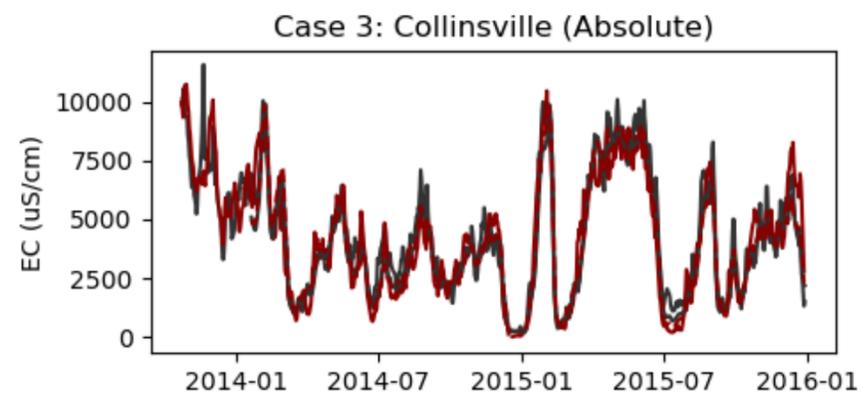
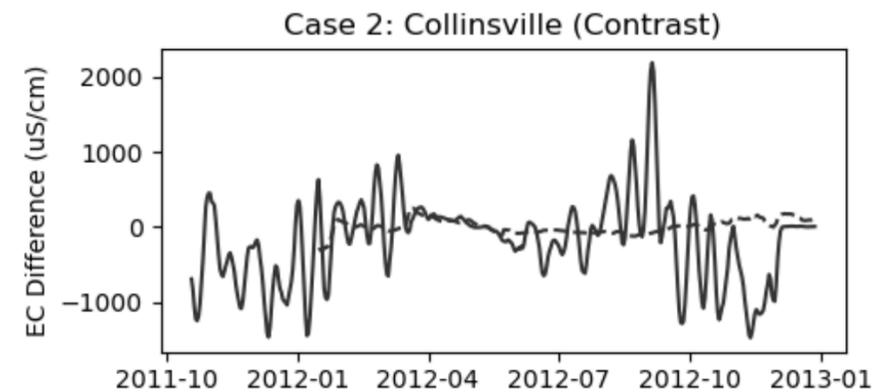
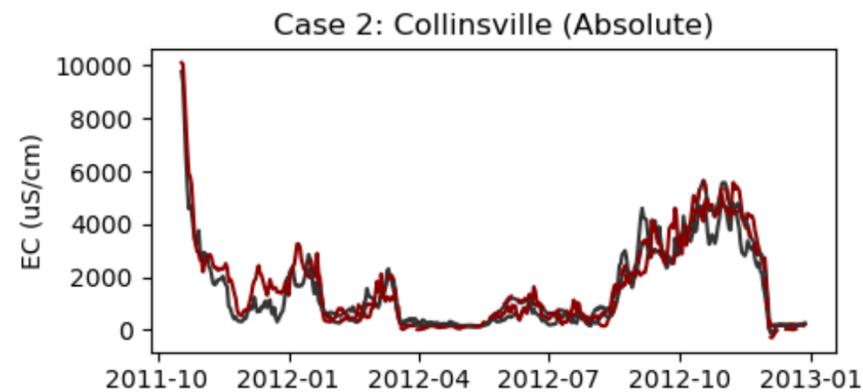
Solid: SCHISM
Dashed: ANN



SCHISM Suisun SCHISM Base ANN Suisun ANN Base



Solid: SCHISM
Dashed: ANN



Conclusions Thus Far

- Skill in the $NSE=0.85-0.90$ range
 - Varies by training but even more so by nature/extent of generalization
 - Physicality responses
 - Small contrasts represented, noise/phase an issue downstream
- Pre-training on DSM2 dataset accelerates learning.
 - We did about 100yrs of DOE cases
- Contrastive multitask approach equally beneficial
- Data interoperability details crucial for transfer learning



Products

- Design of Experiments for training surrogates
- Casantra:
 - testbed and training library for transfer learning

<https://github.com/CADWRDeltaModeling/casantra/>
- Calsurrogate:
 - Plugin library for surrogates in CalSim
 - Includes linearization, time marching, salinity-to-flow inversion

<https://github.com/CADWRDeltaModeling/calsurrogate>



Questions?

- Eli.Ateljevich@water.ca.gov

