Parameter tuning can yield hints about model structural errors

Vincent Larson and Zhun Guo
CESM Workshop
10 Jun 2024

Thanks to Shaocheng Xie and E3SM for funding support

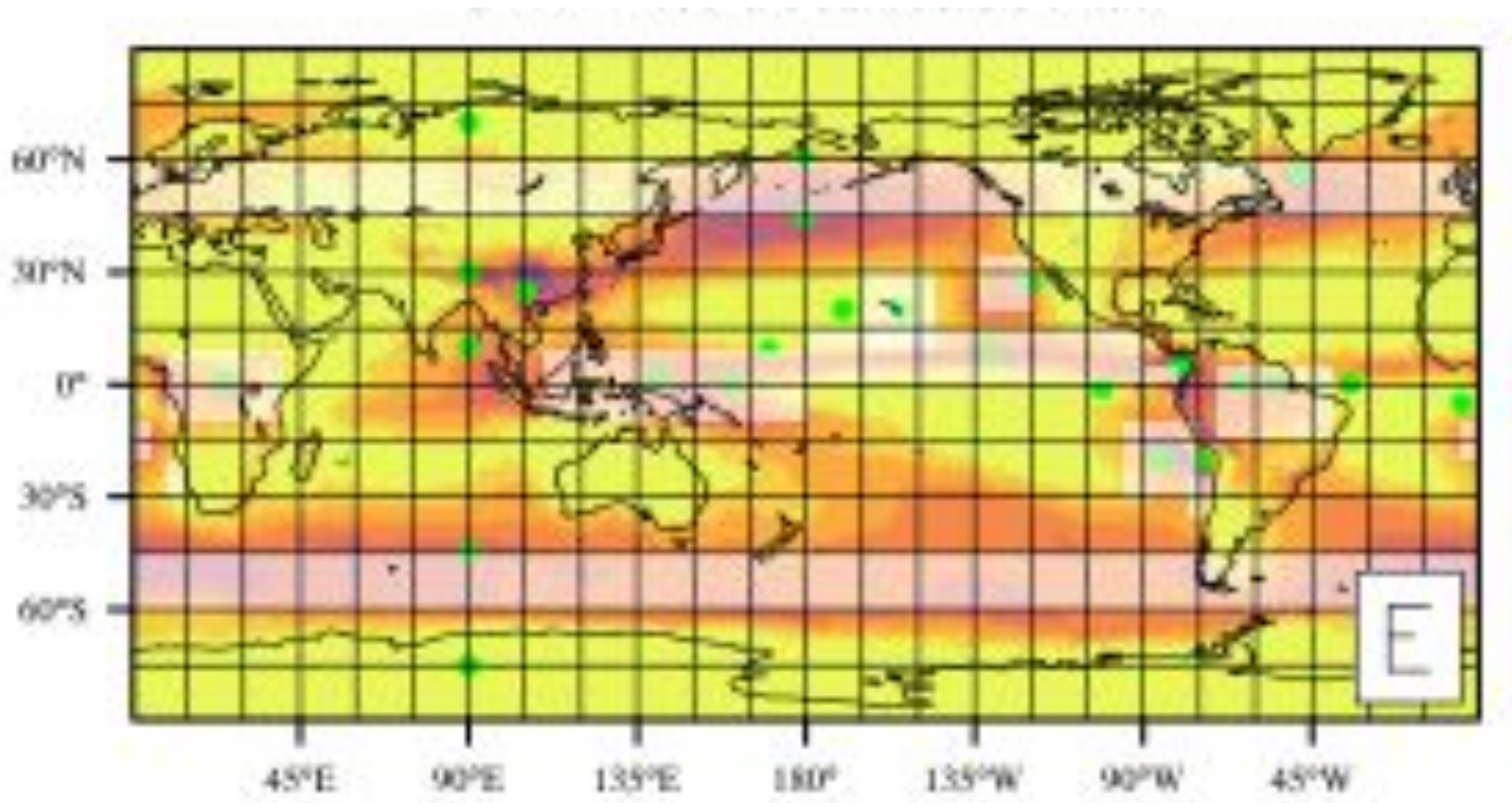
Outline of talk

- How can we get insight into model structural error? We attempt to get insight with our proposed tuning method ("QuadTune") and related analyses.
- Examples of tuning questions regarding a global model (EAM)
- Two archetypal model errors: Stubborn biases and tuning trade-offs

Tuning to match metrics in *regions* can provide (at least a little) insight into model structural errors…

. . . by indicating which parameters have which effects in various regions of the globe.

QuadTune adjusts j parameter values, p_j, in order to best match i regional metrics m_i (e.g., SWCF in VOCALS Sc region):



We choose regions that have errors or are of special interest (Qian et al. 2023).

The QuadTune recipe:

1. Choose regional metrics

2. Choose *n* tuning parameters

3. Run *2n+1* global simulations, varying parameters one at a time, perturbing each high and then low

4. Minimize difference between model and obs, and create plots

Background: The goal of tuning is to find a single *dp* that dots into each row and yields the corresponding rhs bias

$$\begin{bmatrix} \frac{\partial m_{Sc}}{\partial p_1} & \frac{\partial m_{Sc}}{\partial p_2} \\ \frac{\partial m_{Cu}}{\partial p_1} & \frac{\partial m_{Cu}}{\partial p_2} \\ \frac{\partial m_{WP}}{\partial p_1} & \frac{\partial m_{WP}}{\partial p_2} \end{bmatrix} \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix} \approx - \begin{bmatrix} \delta b_{Sc} \\ \delta b_{Cu} \\ \delta b_{WP} \end{bmatrix}$$

stratocumulus region

cumulus region

warm pool region

sensitivity to parameter 1

sensitivity to parameter 2

tunable parameters

model bias = default - obs

Tuning 2 parameters can’t remove the bias in all 3 regions unless the spatial pattern of sensitivity happens to be consistent with the spatial pattern of bias.

A caveat: Although this matrix equation pretends that tuning is a linear regression problem, in fact a regularizer is necessary

Why? Because unregularized linear regression will choose large parameter perturbations that don't give realistic results in a global climate model.

In my experience, large parameter perturbations invariably lead to poor global simulations. We want to avoid this.

Outline of talk

- How can we get insight into model structural error? We attempt to do that with our proposed tuning method (“QuadTune”) and related analyses

- Example tuning questions regarding a global model (EAM)

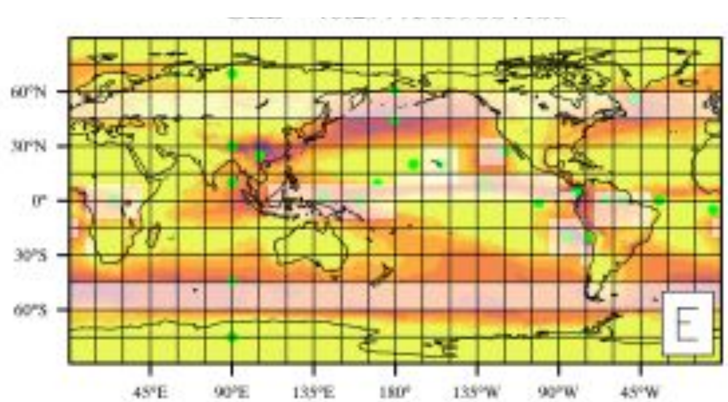
- Two archetypal model errors: Stubborn biases and tuning trade-offs

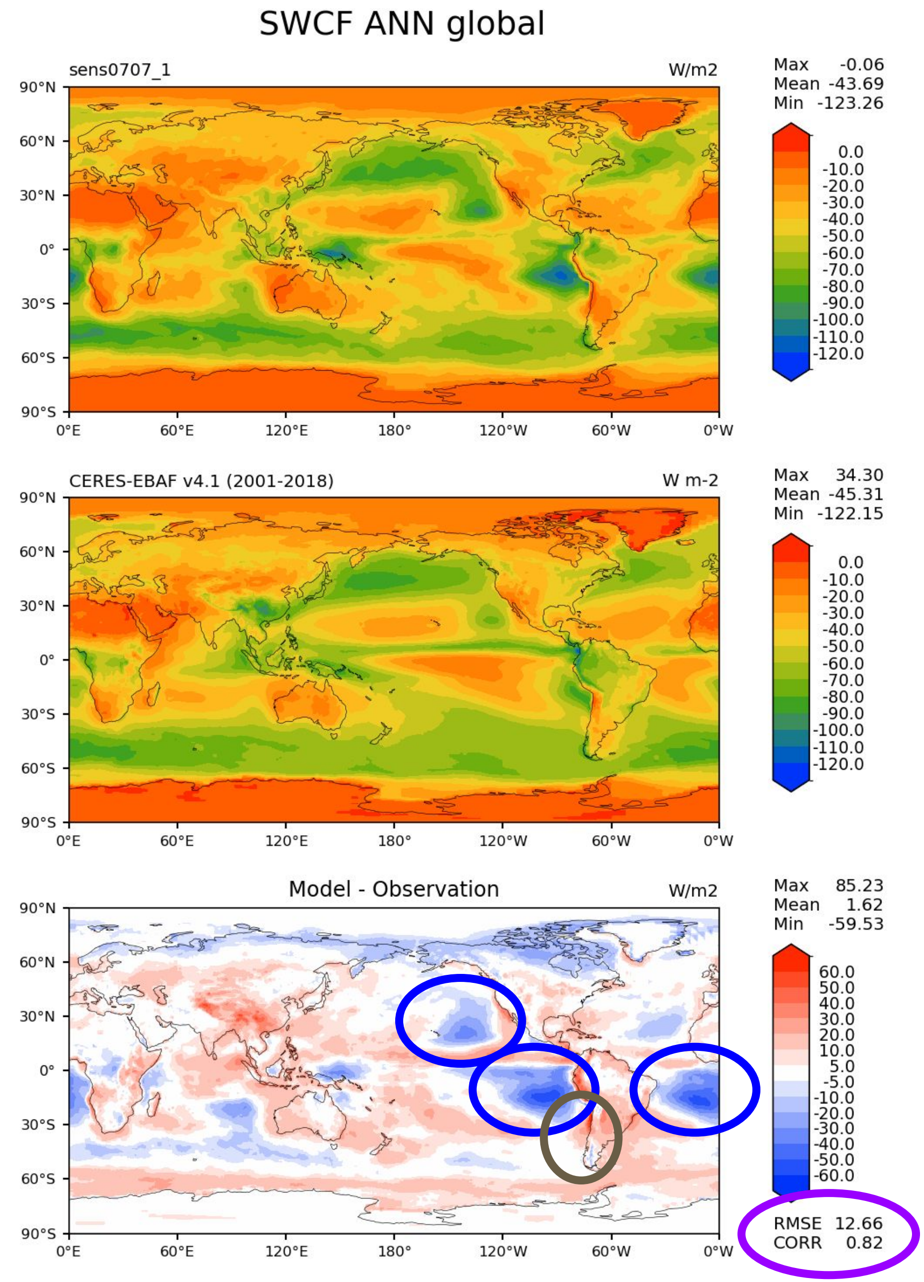
Now we present 2 sensitivity runs of a global atmospheric model, EAMv~3. These were 2 of the *2n+1* runs from a tuning exercise.

We tune for 9 CLUBB parameters.

In this example, we attempt to match SWCF in various regions, plus globally averaged LWCF and PRECT.

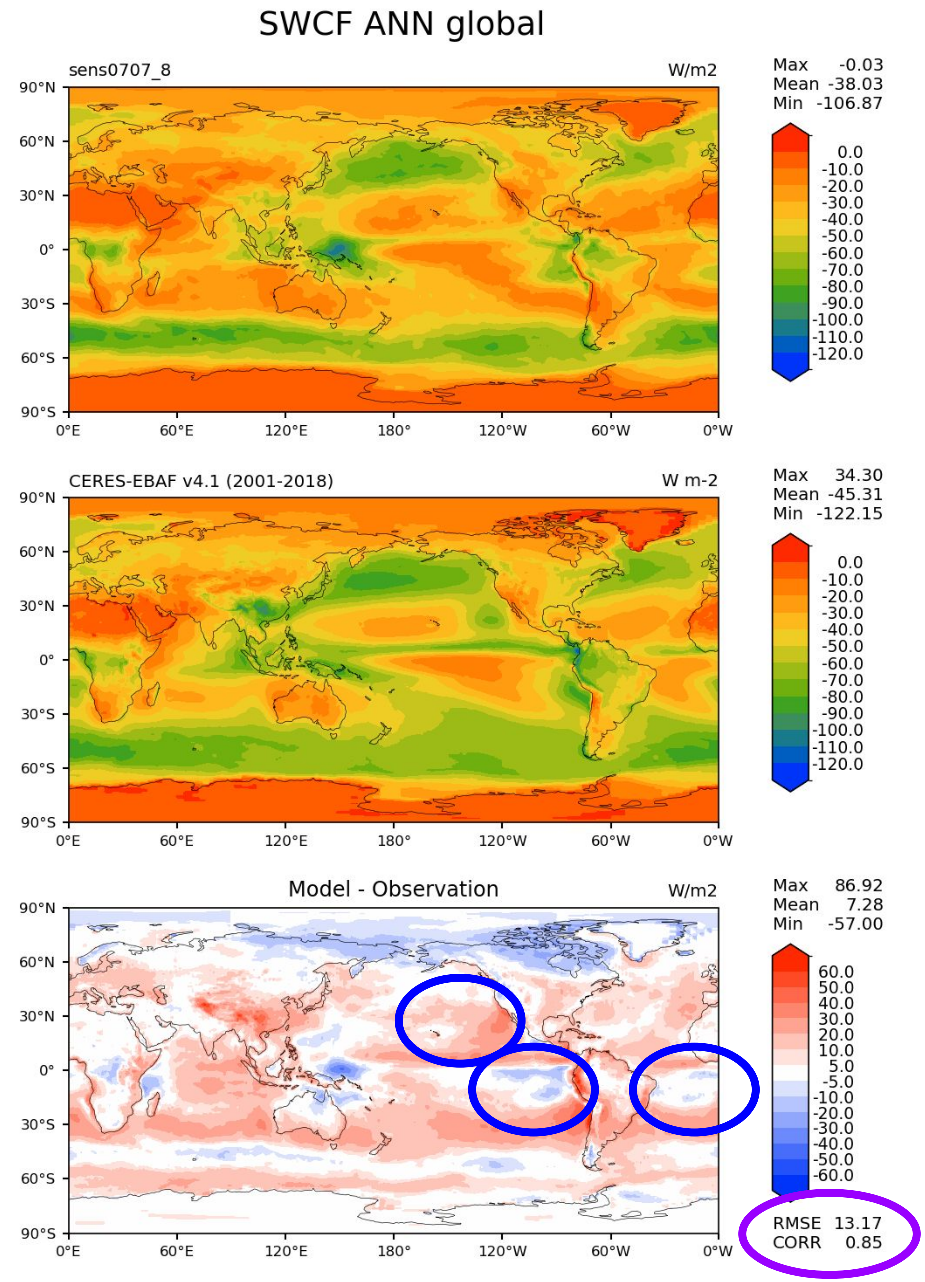
When we started, the “far-coastal” Sc (**VOCALS, Namibia**) were too bright. So were shallow Cu (**Hawaii**).





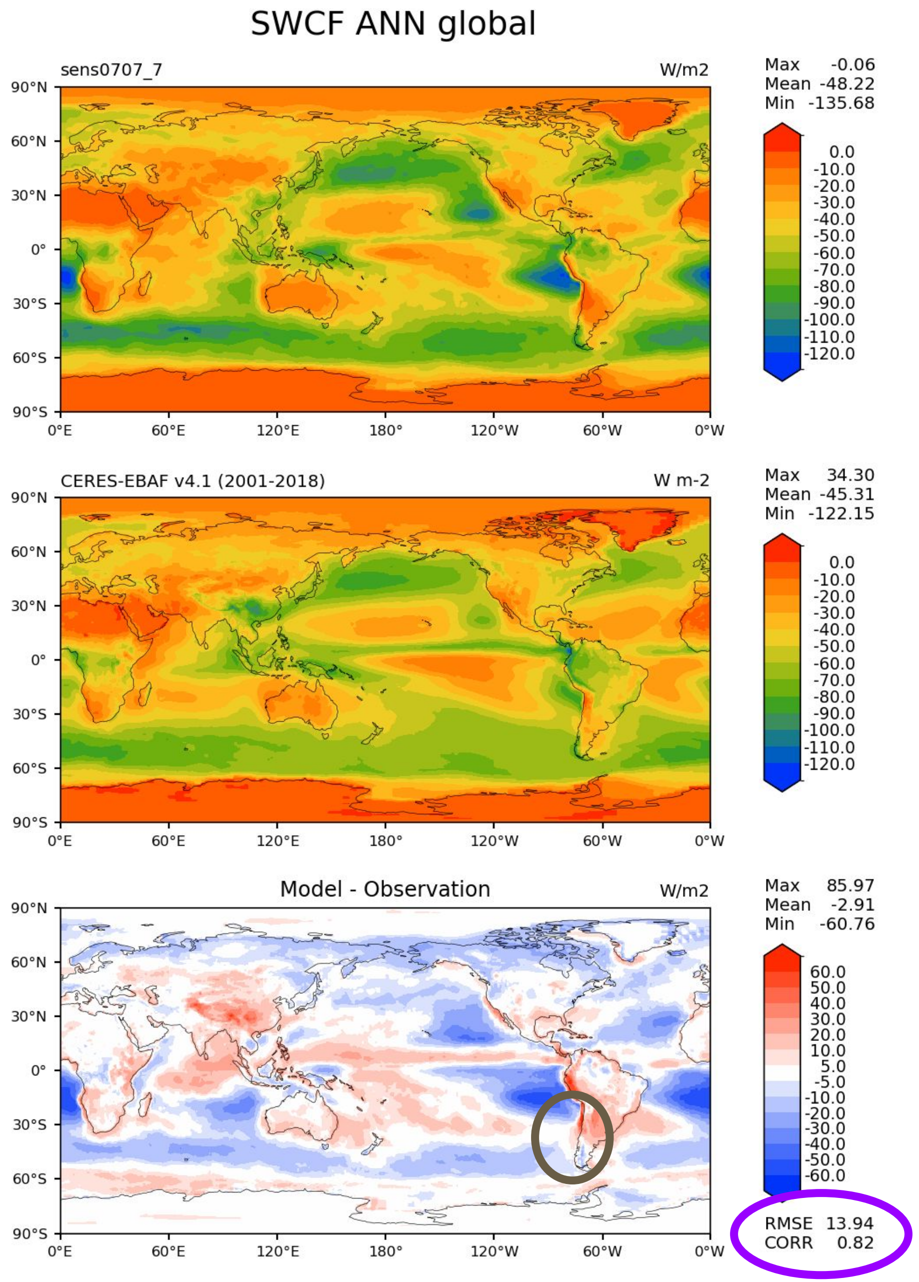
An example of a sensitivity: Increasing *n2_thresh* to 0.5 dims the far-coastal Sc…

. . . but worsens the RMSE and global bias. So should we increase *n2_thresh*?



An example of another sensitivity: Reducing sfc turbulence damping brightens near-coastal Sc

Can we combine perturbations to *n2_thresh* and *C_invrs_tau_sfc* in order to remove biases in both near-coastal and far-coastal Sc?



Outline of talk

- How can we get insight into model structural error? We attempt to do that with our proposed tuning method ("QuadTune") and related analyses

- Example tuning questions regarding a global model (EAM)

- Two archetypal model errors: Stubborn biases and tuning trade-offs

1. A stubborn bias is one that cannot be budged by perturbing the chosen tuning parameters

$$\begin{bmatrix} \frac{\partial m_{Sc}}{\partial p_1} & \frac{\partial m_{Sc}}{\partial p_2} \\ \frac{\partial m_{Cu}}{\partial p_1} & \frac{\partial m_{Cu}}{\partial p_2} \\ \frac{\partial m_{WP}}{\partial p_1} & \frac{\partial m_{WP}}{\partial p_2} \end{bmatrix} \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix} \approx - \begin{bmatrix} \delta b_{Sc} \\ \delta b_{Cu} \\ \delta b_{WP} \end{bmatrix}$$

$\mathbf{S}$ $\delta \vec{p}$ $\delta \vec{b}$

We can re-write this matrix equation in vector notation:

$$\mathbf{S}\, \delta \vec{p} = -\delta \vec{b}$$

Consider the stratocumulus row of this matrix equation:

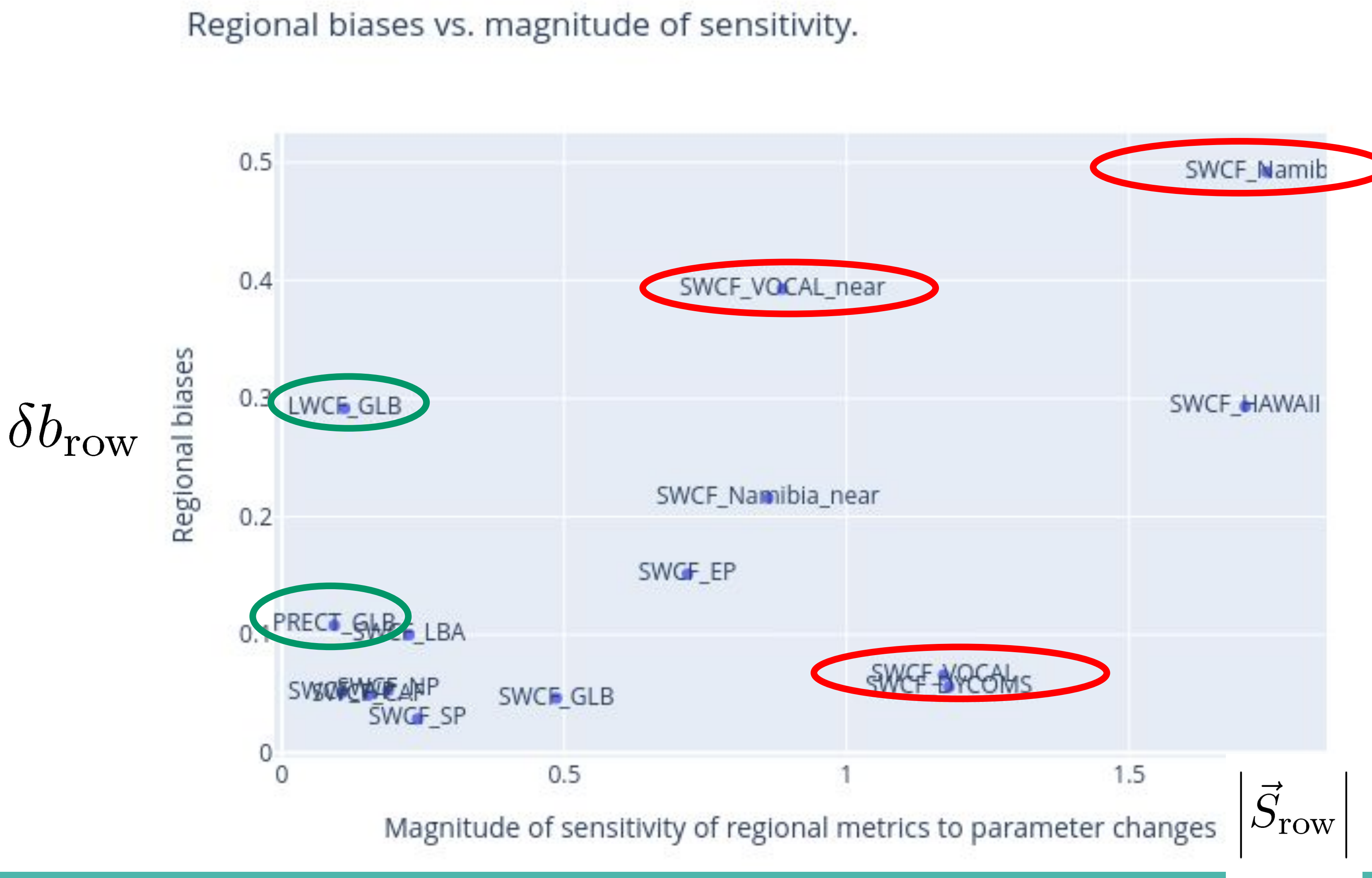
$$\vec{S}_{Sc} \cdot \delta \vec{p} = -\delta b_{Sc}.$$

By the Cauchy–Schwarz inequality,

$$|\delta \vec{p}| \geq \frac{|\delta b_{Sc}|}{\left|\vec{S}_{Sc}\right|}.$$

We want to avoid large *dp*. So we're stuck if the sensitivity, S_{Sc}, is weak.

Before tuning, we can see that LWCF and PRECT are stubborn biases, with large bias magnitudes and low sensitivities



2. What is a tuning trade-off?

A tuning trade-off between metrics occurs when changing the parameter values improves some metrics but worsens others.

Let's construct a 2-metric diagnostic that tells us how difficult it is to *simultaneously* remove the bias in *two* regional metrics (*Sc* and *Cu*)

$$\begin{bmatrix} \frac{\partial m_{Sc}}{\partial p_1} & \frac{\partial m_{Sc}}{\partial p_2} \\ \frac{\partial m_{Cu}}{\partial p_1} & \frac{\partial m_{Cu}}{\partial p_2} \\ \frac{\partial m_{WP}}{\partial p_1} & \frac{\partial m_{WP}}{\partial p_2} \end{bmatrix} \begin{bmatrix} \delta p_1 \\ \delta p_2 \end{bmatrix} \approx - \begin{bmatrix} \delta b_{Sc} \\ \delta b_{Cu} \\ \delta b_{WP} \end{bmatrix}$$

Now consider both the stratocumulus and cumulus rows of this matrix equation:

$$\vec{S}_{Sc} \cdot \delta \vec{p} = -\delta b_{Sc}$$
$$\vec{S}_{Cu} \cdot \delta \vec{p} = -\delta b_{Cu}.$$

Subtracting the two rows from each other, we find

$$\left(\vec{S}_{Sc} - \vec{S}_{Cu} \right) \cdot \delta \vec{p} = - \left(\delta b_{Sc} - \delta b_{Cu} \right).$$

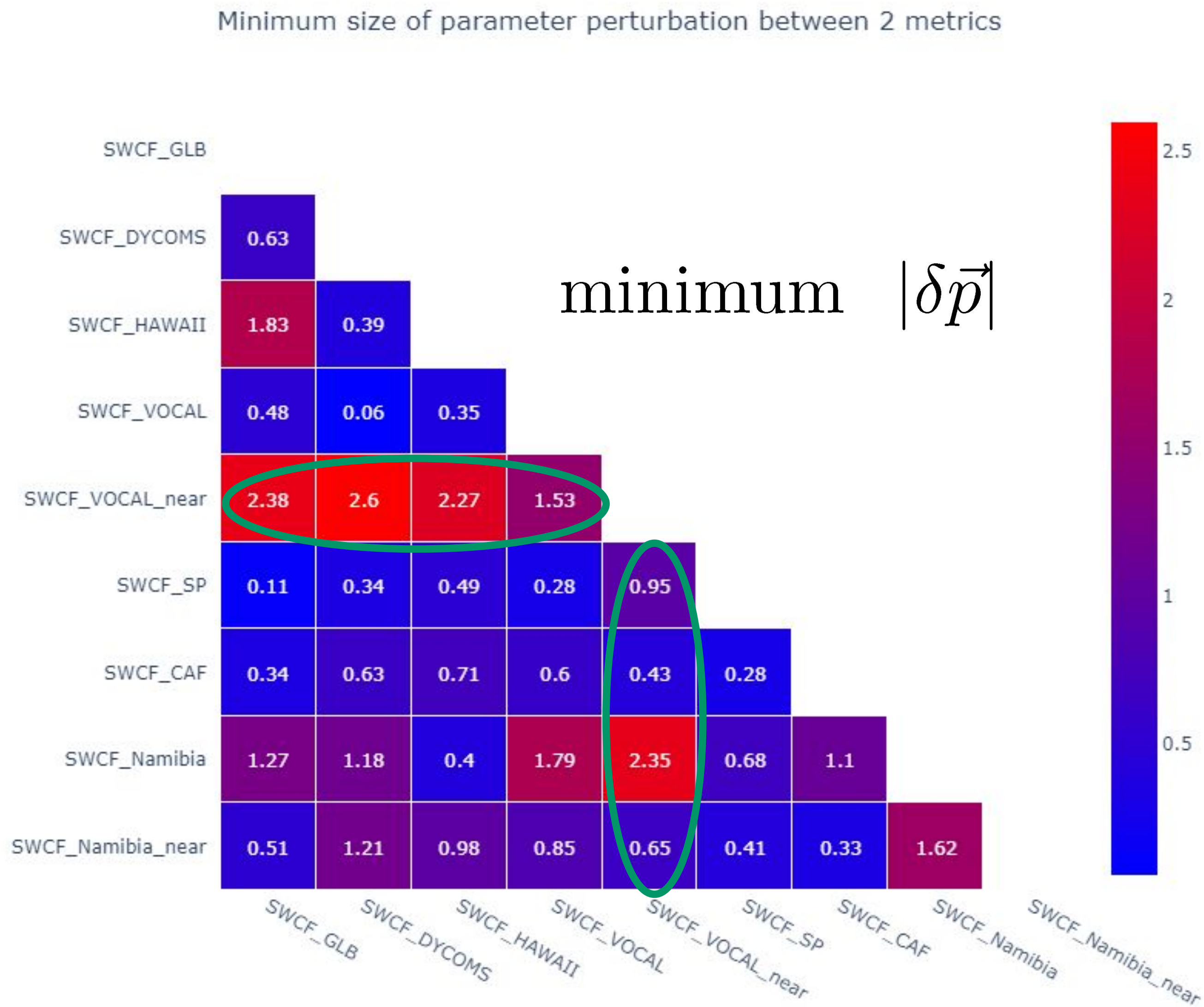
By the Cauchy–Schwarz inequality,

$$|\delta \vec{p}| \geq \frac{|\delta b_{Sc} - \delta b_{Cu}|}{\left| \vec{S}_{Sc} - \vec{S}_{Cu} \right|}.$$

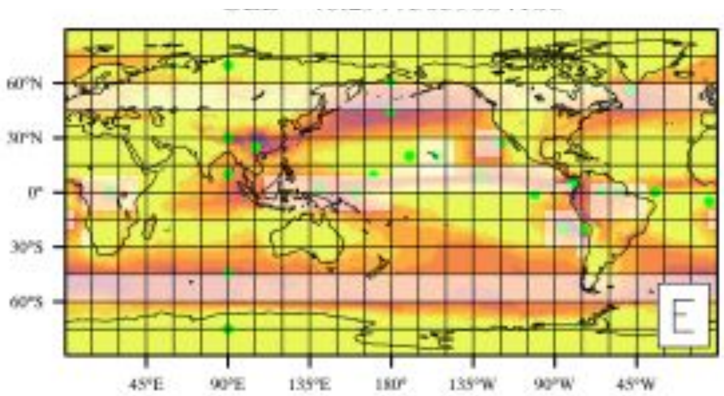
A large *dp* means that it's difficult to fit both regional metrics.

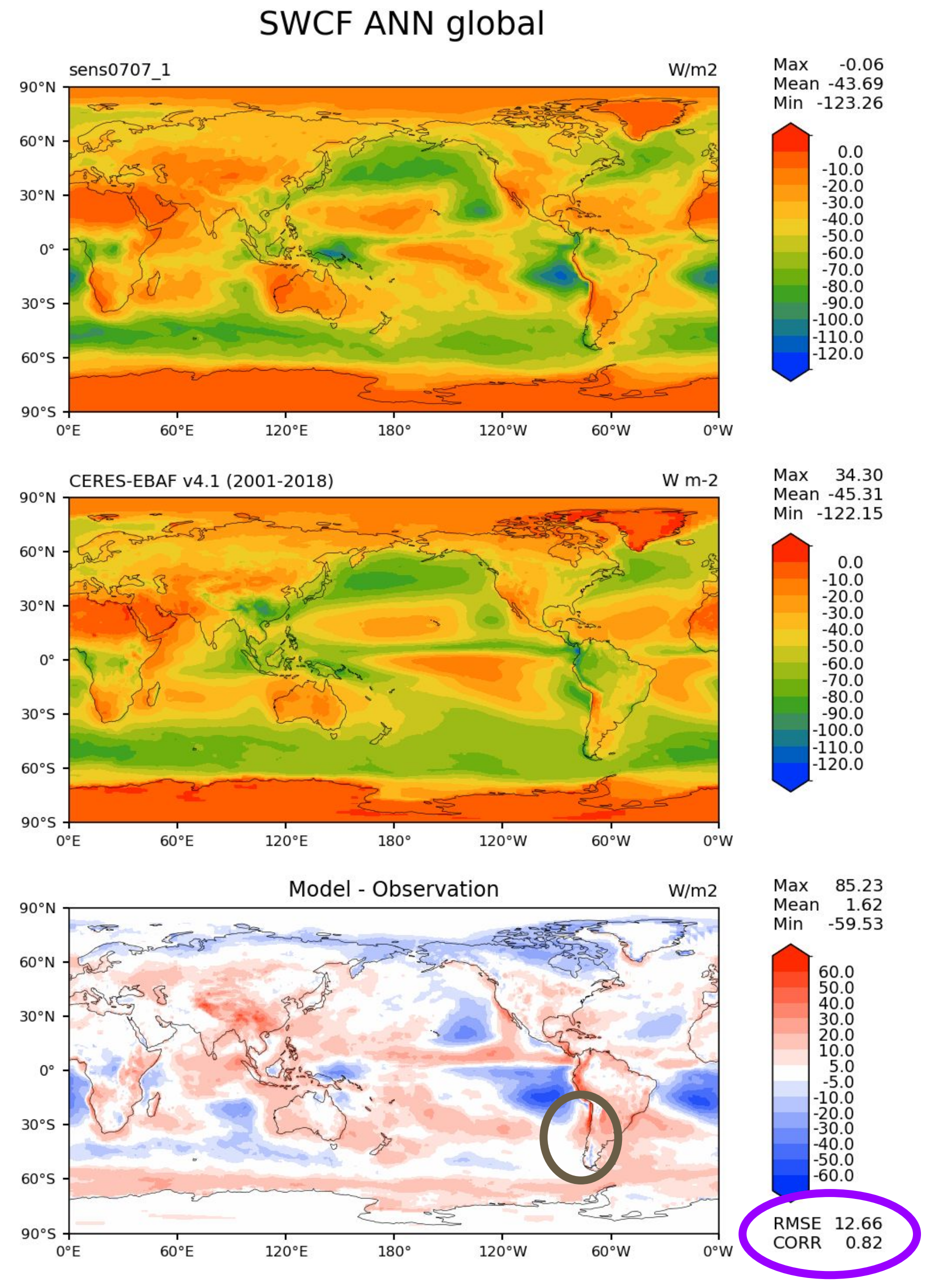
The 2-metric diagnostic says that it's hard to simultaneously match the near-VOCALS region and other regions

(large values are bad!)

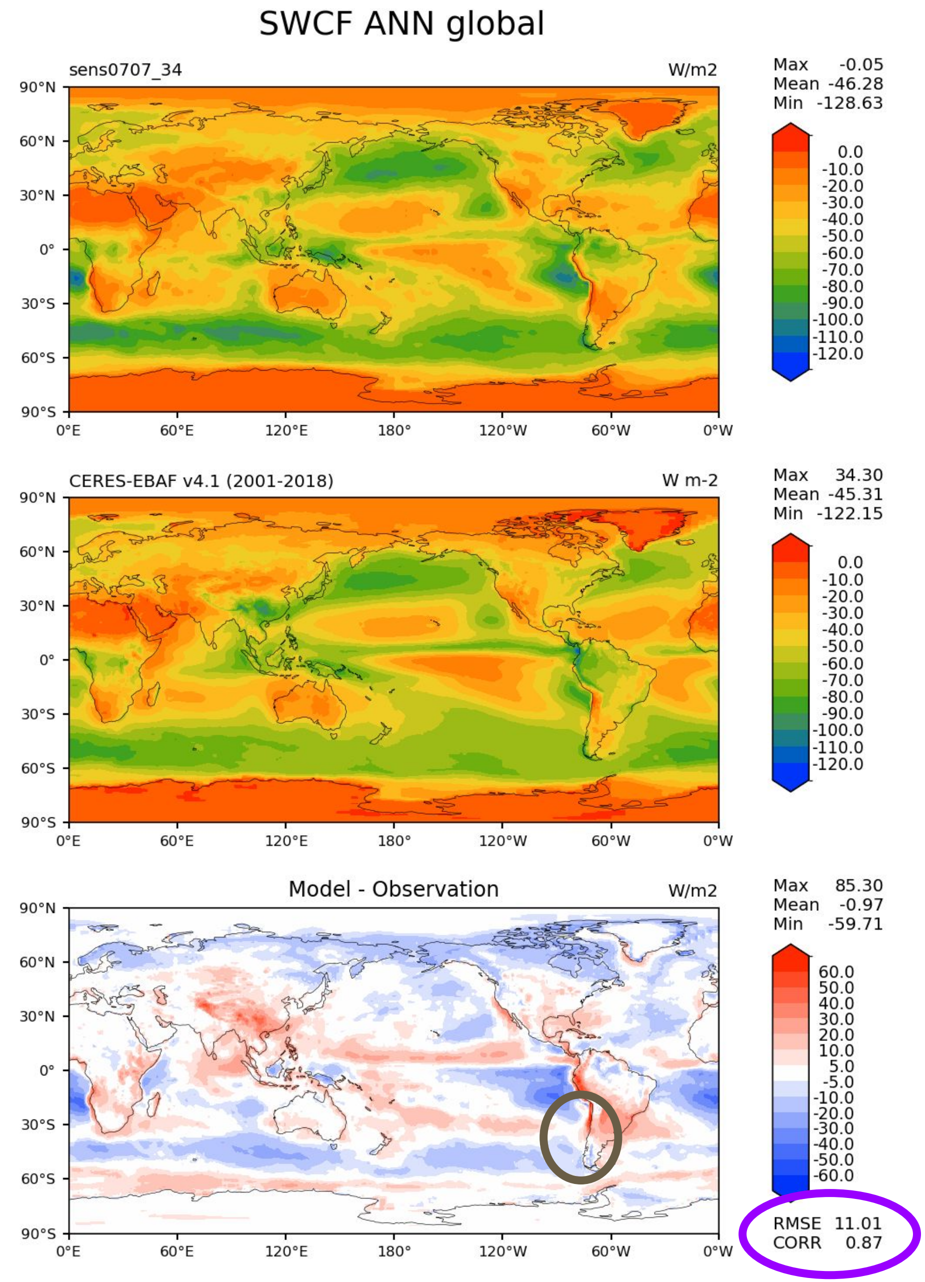


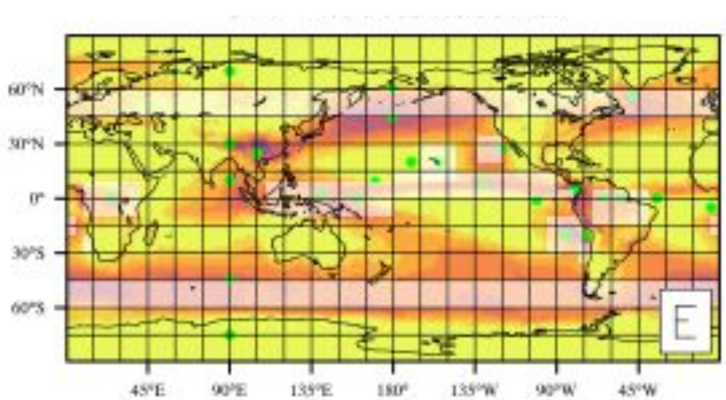
What happens when one tries to remove a tuning trade-off? When we started, the far-coastal Sc were too bright



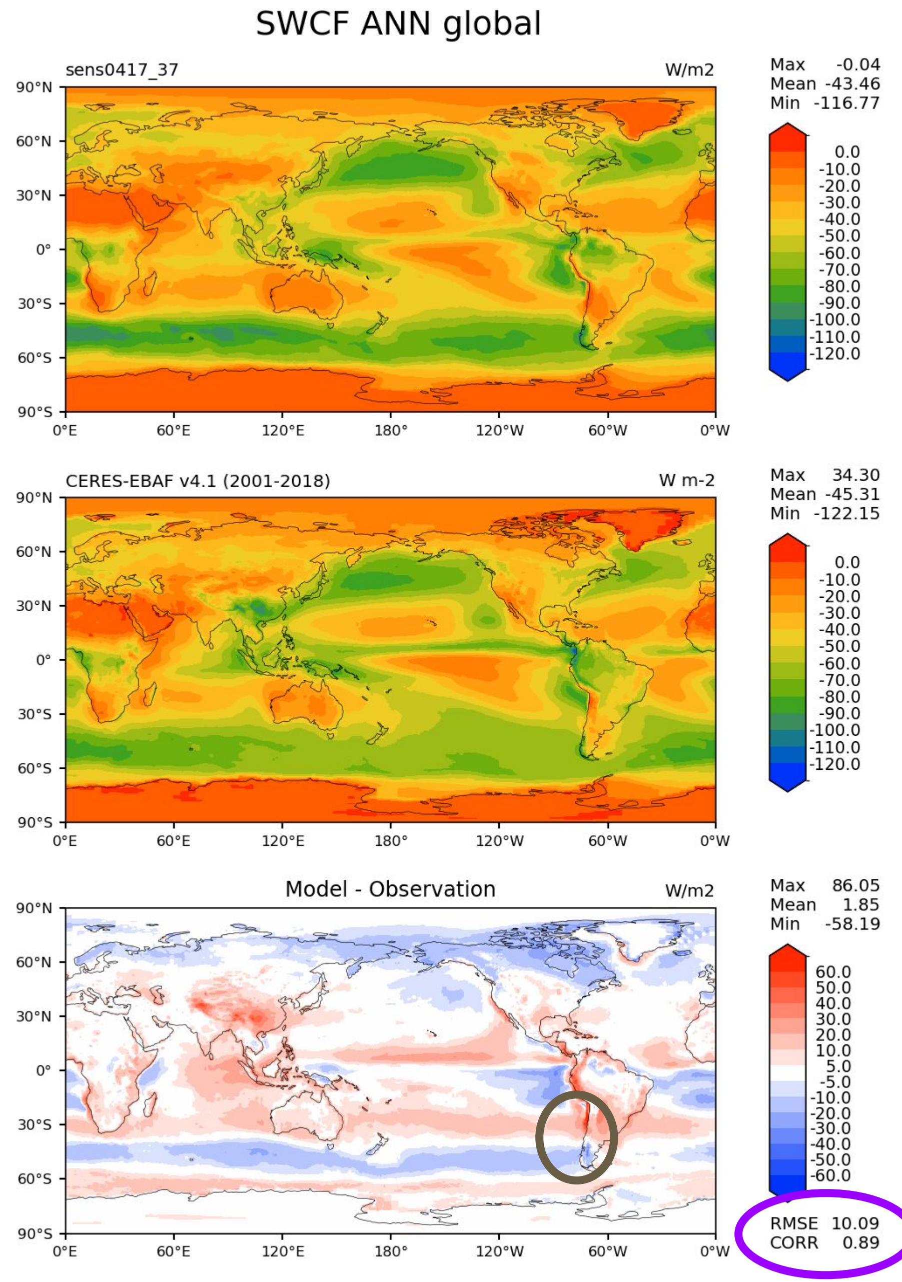


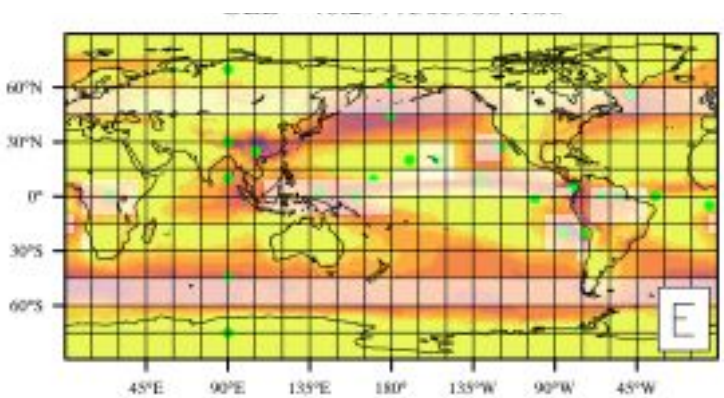
QuadTune dims the far-coastal Sc…





…but doesn't reduce the RMSE as much as Zhun's hand tuning:





What have we learned from this tuning exercise?

- Stubborn biases: None of the chosen parameters can budge PRECT and LWCF. (This might be surprising, since some of these parameters affect surface winds.)

- Tuning trade-offs: It is hard to simultaneously fit the near-coastal VOCALS stratocumulus clouds and clouds in other regions.

What can we learn from QuadTune in general?

- We learn when to give up. Put another way, the tuner helps distinguish parametric from structural model error. If the tuner doesn't yield acceptable results, then we should either find new parameters or re-formulate the model structure.

- We learn which parameters matter in which regions. (But then, in order to understand why they matter, we need to analyze those regions in more detail.)

Thanks for your attention