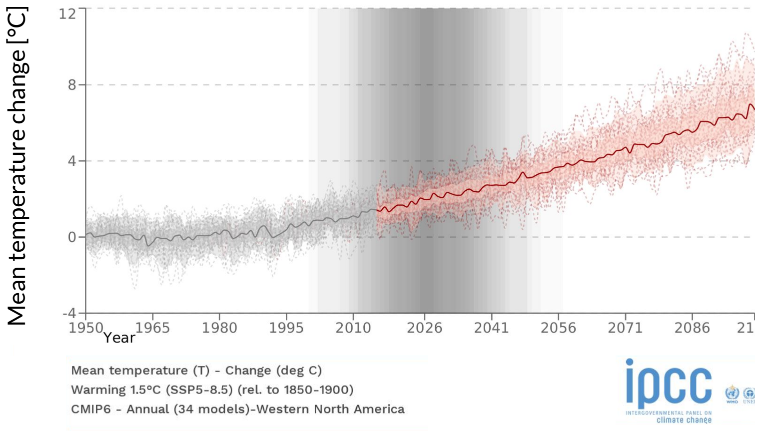
Improving the predictions of ML-corrected climate models with novelty detection

Clayton Sanford, Anna Kwa, Oliver Watt-Meyer, Spencer Clark, Noah Brenowitz, Jeremy McGibbon, Christopher Bretherton

NeurIPS 2022 Workshop
Tackling Climate Change with Machine Learning

AI2

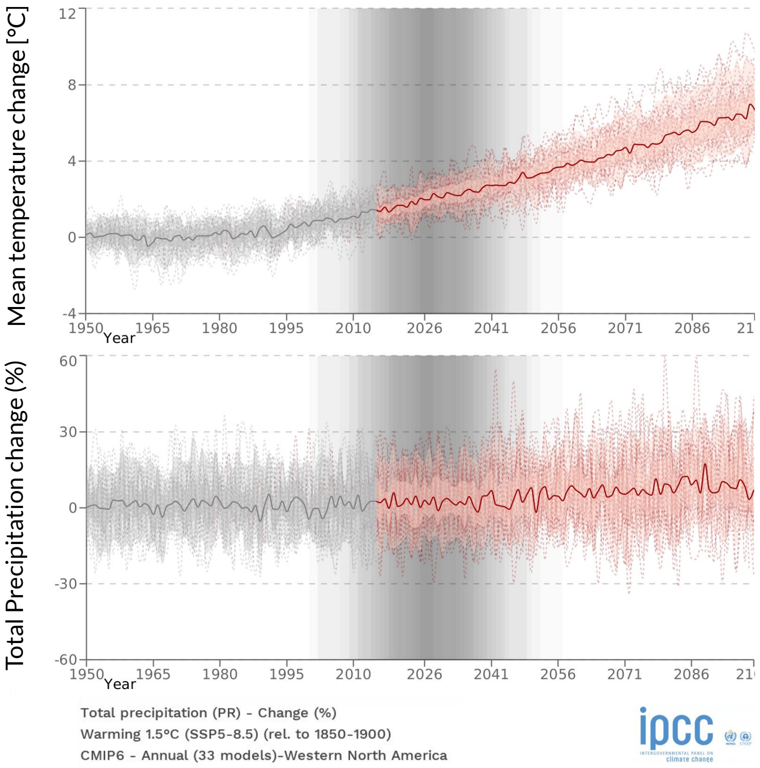
Why are we doing this?



Western North America **temperature**

IPCC AR6 Atlas (CMIP6 models)

Why are we doing this?

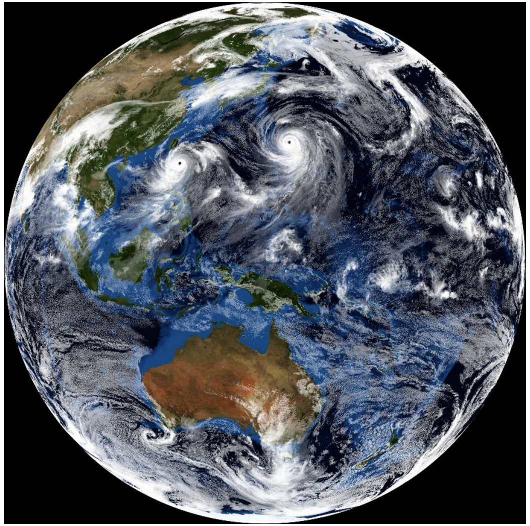


Western North America **temperature** and **precipitation**

Models have **less agreement** about future local precipitation trends compared to temperature. This matters!

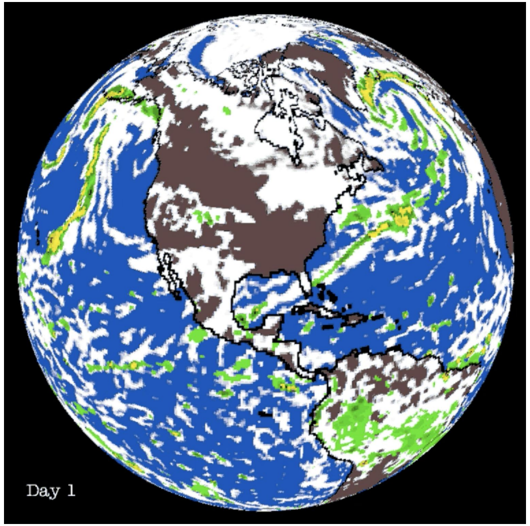
ML Goal: Improve coarse-model simulations

High fidelity reference
reanalysis or
fine-grid (~3km) simulation



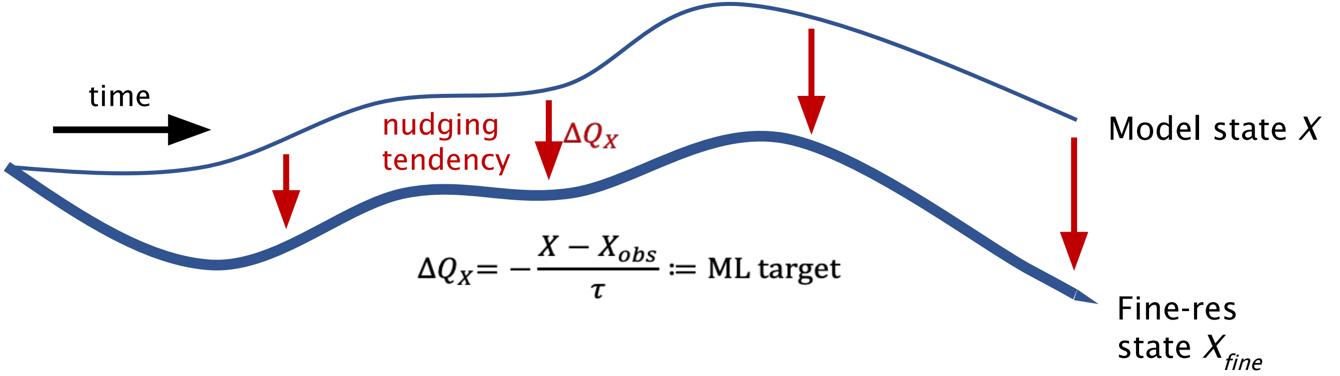
Use machine learning to make coarse model behave more like reference

Climate model (25-200 km)



Corrective approach

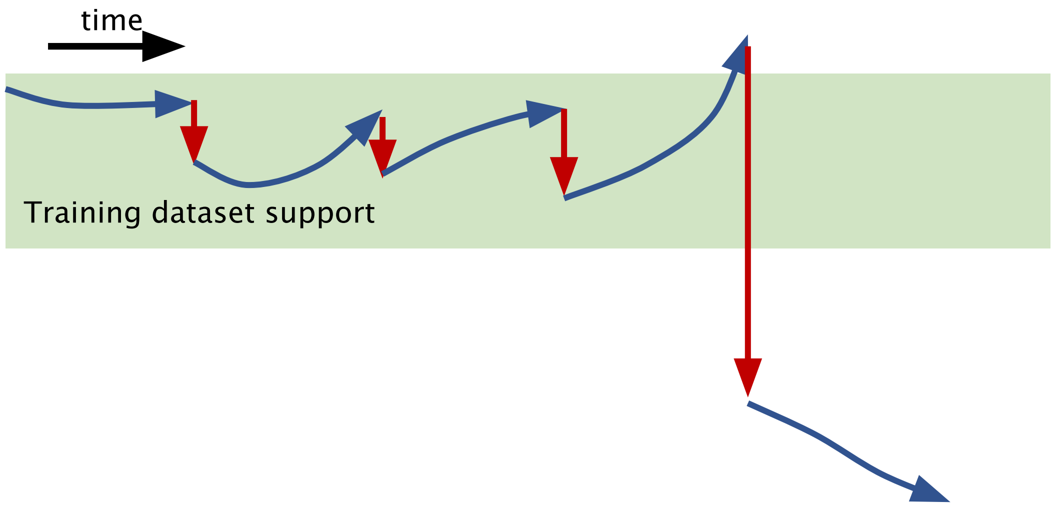
- Our approach:
  1. Nudge coarse-resolution model towards reference dataset
  2. Train ML to predict nudging tendencies with input coarse-res state
  3. Run coarse-res model, with ML corrective tendency at each step



$$\Delta Q_X = -\frac{X - X_{obs}}{\tau} := \text{ML target}$$

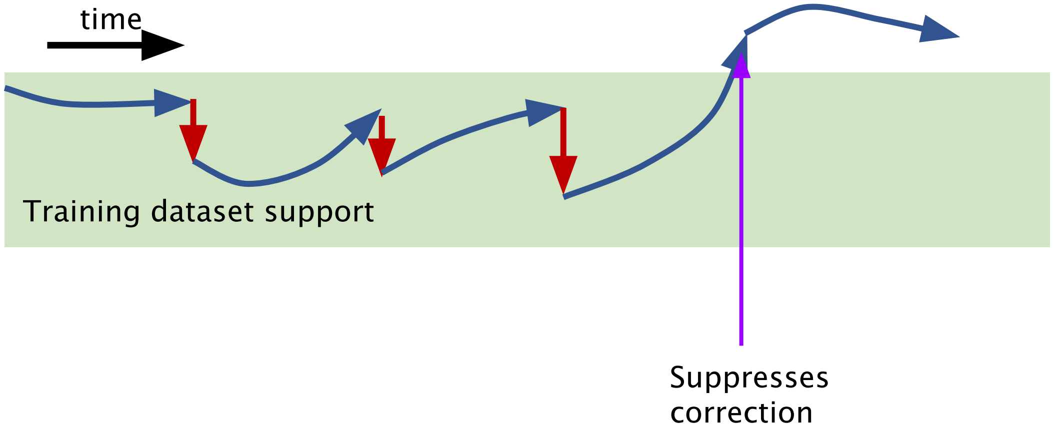
Limitations of corrective approach

- ML corrective tendencies inaccurate & unstable outside training dataset
- Simulation is an online process → regularly produces out-of-sample data
- Thus, ML-corrected simulations crash frequently & behave erratically
  - Especially when including wind in ML corrections



Stabilization with novelty detection

- Idea: If simulation drifts out-of-sample, disable ML correction
- **Novelty detection** is a branch of self-supervised learning that predicts whether a sample belongs to a distribution given draws from distribution

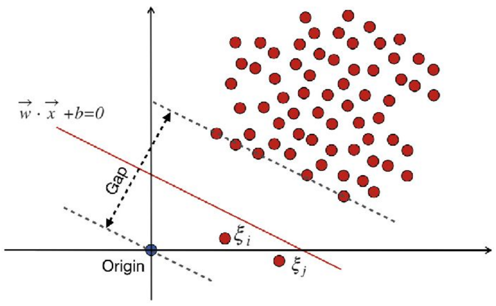


One-Class Support Vector machine (OCSVM)

- **Idea:** Directly estimate support of distribution by identifying compact region that contains all samples
- Maximize distance between dataset $\{x_1, \ldots, x_n\} \in \mathbb{R}^d$ and the origin under feature mapping $\Phi: \mathbb{R}^d \to F$
- Radial basis function (RBF) kernel

$$\min_{w \in F, \xi \in \mathbb{R}^n, \rho \in \mathbb{R}} \quad \frac{1}{2}\|w\|^2 + \frac{1}{\nu n} \sum_i \xi_i - \rho$$

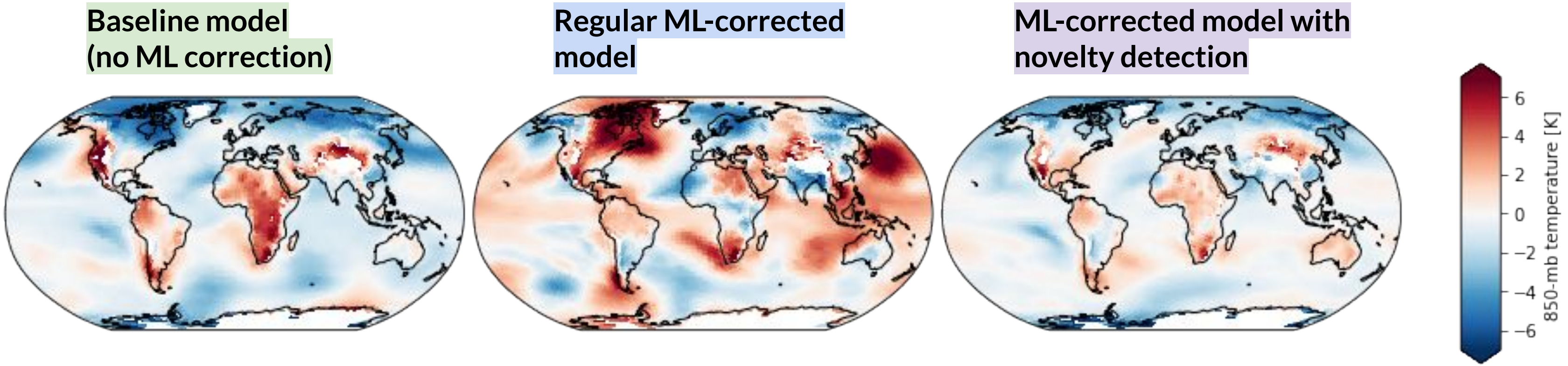
$$\text{subject to} \quad (w \cdot \Phi(\mathbf{x}_i)) \geq \rho - \xi_i, \quad \xi_i \geq 0.$$



Our findings

- Incorporating novelty detection with a One-Class SVM prevents runs from crashing and improves temperature and humidity predictions over simulations with and without ML correction

Time-averaged near-surface temperature biases



Our findings

- Incorporating novelty detection with a One-Class SVM prevents runs from crashing and improves temperature and humidity predictions over simulations with and without ML correction

	Run	% Novelty	T (K)	SP (mm/day)	PWAT (kg/m^2)
1	Baseline (1)	100%	2.09	1.78	2.79
2	ML-corrected (2) with g_{Tq}	0%	1.86	1.43	3.31
3	ML-corrected with g_{Tquv} ($\star$)	0%	2.43	3.39	5.33
4	ND ML (3) with $g_{\mathrm{Tq}}, \eta_{\mathrm{T,OCSVM}}$	2.5%	1.97	1.49	3.65
5	ND ML with $g_{\mathrm{Tquv}}, \eta_{\mathrm{T,minmax}}$	35.7%	5.15	3.57	10.14
6	ND ML with $g_{\mathrm{Tquv}}, \eta_{\mathrm{T,OCSVM}}$	40.0%	1.58	1.40	2.66
7	ND ML with $g_{\mathrm{Tquv}}, \eta_{\mathrm{Tq,OCSVM}}$	50.7%	**1.53**	**1.24**	**2.37**