An Unsupervised Learning Perspective on the Dynamic Contribution to Extreme Precipitation Changes

Griffin S. Mooers
Department of Earth System Science
University of California Irvine
Irvine, CA 92617
gmooers@uci.edu

Tom Beucler
Institute of Earth Surface Dynamics
University of Lausanne
CH1007 Lausanne, Switzerland
tom.beucler@unil.ch

Mike Pritchard
University of California Irvine
Irvine, CA 92617
mspritch@uci.edu

Stephan Mandt
University of California Irvine
Irvine, CA 92617
mandt@uci.edu

Abstract

Despite the importance of quantifying how the spatial patterns of extreme precipitation will change with warming, we lack tools to objectively analyze the storm-scale outputs of modern climate models. To address this gap, we develop an unsupervised machine learning framework to quantify how storm dynamics affect precipitation extremes and their changes without sacrificing spatial information. Over a wide range of precipitation quantiles, we find that the spatial patterns of extreme precipitation changes are dominated by spatial shifts in storm regimes rather than intrinsic changes in how these storm regimes produce precipitation.

1 Introduction: Understanding the Changing Spatial Patterns of Precipitation Extremes

According to the latest IPCC report [11], " there is high confidence that extreme precipitation events across the globe will increase in both intensity and frequency with global warming". As the severity of storms and tropical cyclones magnifies, there will be associated increases in flood-related risk [13] and challenges in water management [2, 3]. To first order, heavy precipitation extremes are limited by the water vapor holding capacity of the atmosphere, which increases by about 7% per 1K (Kelvin) of warming following an approximate Clausius-Clapeyron scaling [24]. This is referred to as the "thermodynamic contribution" to extreme precipitation changes [12] and gives a solid theoretical foundation for *spatially-averaged* changes in precipitation extremes.

Yet climate change adaptation requires knowledge of how precipitation extremes will change at the *local* scale, i.e., understanding the *changing spatial patterns* of precipitation extremes under warming. Focusing on the tropics, where most of the vulnerable world population lives [10], these changing spatial patterns are primarily dictated by atmospheric vertical velocity ("dynamical") changes because horizontal spatial gradients in temperatures are weak. This is referred to as the "dynamic contribution" to extreme precipitation changes [12].

A comprehensive understanding of this "dynamic contribution" remains elusive. Approximate scalings can be derived based on quasi-geostrophic dynamics [18, 23] and convective storm dynamics [22, 1]. But actionable findings require Earth-like simulations of the present and future climates (e.g., [25]), which can resolve regional circulation changes and their effects on storms in their full complexity. These simulations are computationally demanding and output large amounts of multi-scale,

Tackling Climate Change with Machine Learning workshop at NeurIPS 2022.

three-dimensional data that challenge traditional data analysis tools. For example, the state-of-the-art high-resolution[1], SPCAM (Super Parameterized Community Atmospheric Model, [15, 16]) simulations we will use in this study (SI-A) output 3.4 Terabytes of data, with 76,944,384 samples of precipitation and the corresponding storm-scale vertical velocity fields (see Fig 1 for examples).

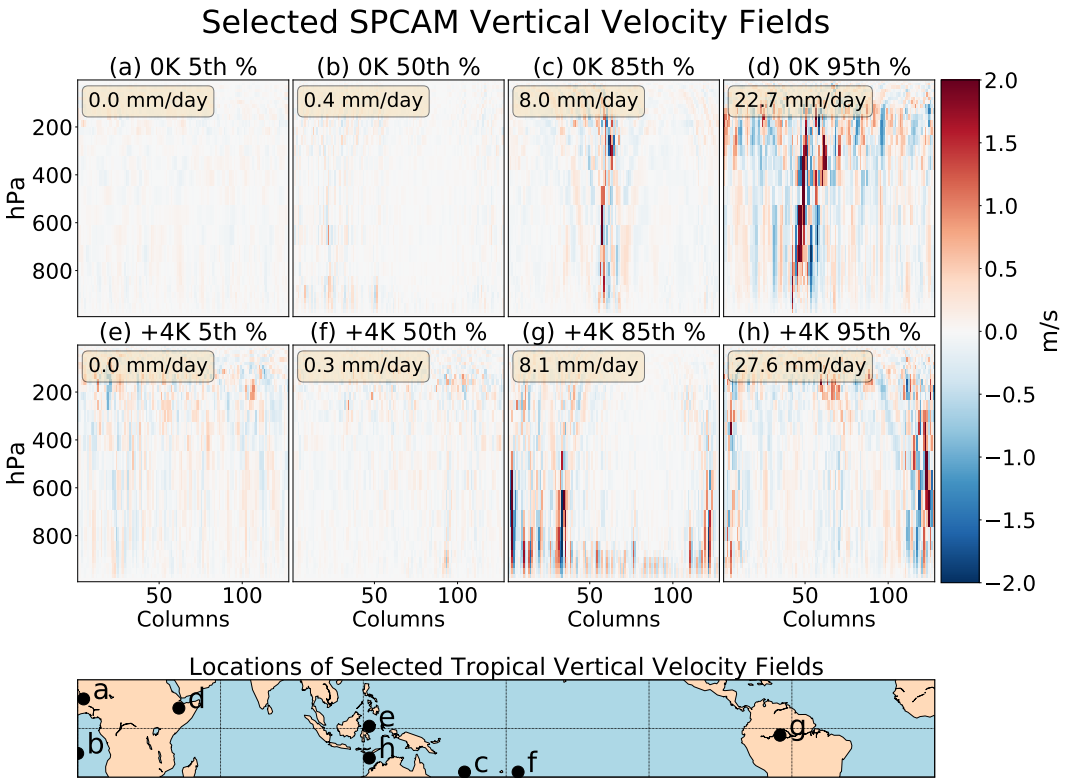

Figure 1: Selected vertical velocity fields from our Control (0K, a-d) and Warmed (+4K, e-h) SPCAM simulations. By sampling the precipitation distribution, we show instances of vertical velocity fields associated with no precipitation (a, e), drizzle (b, f), heavy rainfall (c, g), and intense storms (d, h).

To facilitate the interpretation of extreme precipitation changes without sacrificing the details of storm formation, this paper proposes to (1) objectively cluster vertical velocity fields into N interpretable regimes via unsupervised machine learning. We define the probability of a regime of convection at a given geographic location as π. This discretization of regimes of convection through machine learning allows us to (2) quantitatively understand changes in precipitation extremes (P_{extreme}) from changes in convection regime dynamics and probability. Here we define P_{extreme} as a fixed *high* quantile of precipitation (e.g., 95th percentile) at a given spatial location. To model the effects of climate change, we define $\Delta P_{\text{extreme}}$ as its absolute change from the Control to the Warmed climate, and show in SI-B that relative changes in precipitation extremes can be decomposed as follows:

$$\frac{\Delta P_{\text{extreme}}}{P_{\text{extreme}}} = \overbrace{\underbrace{\frac{\Delta q_{\text{sat}}}{q_{\text{sat}}}}_{\text{From theory}}}^{\text{Thermodynamic}} + \overbrace{\underbrace{\frac{q_{\text{sat}}}{P_{\text{extreme}}}}_{\text{From current climate}} \left(\Delta \mathcal{D}_0 + \underbrace{\sum_{i=1}^{N} \Delta \pi_i \mathcal{D}_i}_{\text{Regime prob. shifts}} + \underbrace{\sum_{i=1}^{N} \pi_i \Delta \mathcal{D}_i}_{\text{Intra-regime changes}} \right)}^{\text{Dynamic}} \tag{1}$$

Eq. 1 shows that relative changes in P_{extreme} are the sum of a well-understood, spatially-uniform "thermodynamic" increase in saturation specific humidity (q_{sat} – see Eq. 7), and a spatially-varying term, which is the sum of N regime probability shifts (changes in our unsupervised machine learning

[1] 5 kilometers or less horizontally

derived convection cluster assignments $\Delta\pi_i$ – covered in more detail in Section 2) and N changes in regime dynamics (changes in the "dynamic contribution" pre-factors $\Delta\mathcal{D}_i$, in precipitation units).

Our simulation data already contain P_{extreme} and q_{sat}, and we can derive π_i from our unsupervised learning framework, giving us all the information we need to calculate the elusive pre-factors $\mathcal{D}_i$ and their change with warming. Using equation 5, we linearly regress $\frac{P_{\text{extreme}}}{q_{\text{sat}}}$ on the regime frequencies π_i in both the reference and warm climates to derive the pre-factors $\mathcal{D}_i$, which are the weights of the multiple linear regression. This is a step toward understanding how the spatial patterns of storm-scale dynamical changes, which are notably hard to analyze, can affect the spatial patterns of extreme precipitation. Understanding these changes is critical to trust local climate change predictions.

2 Machine Learning Methods: Variational Autoencoder With Clustering

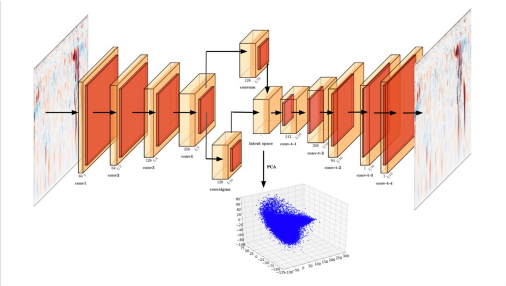

Figure 2: The VAE architecture used in this paper. Given a vertical velocity field w, the VAE non-linearly reduces the input dimension to yield a latent representation **z**. Based on **z**, the VAE is trained to reconstruct w. The latent representation **z** then allows us to cluster w into N interpretable regimes.

To objectively extract interpretable dynamical knowledge, we rely on fully-convolutional Variational Autoencoders (VAEs) to learn low dimensional latent representations of the vertical velocity details. We use the VAE's encoder to non-linearly reduce the full 2D fields to a tractable representation of the dynamical information while preserving physical coherence (SI-C/D – the VAE's training and benchmarking). The evaluation of regime probability shifts is done in three steps.

First, after training a VAE on the warmed (+4K) climate, we follow the K-means clustering approach (SI-E) to organize the complex vertical velocity dynamics into three distinct regimes. This can be generalized to $N \in \mathbb{N}^*$ regimes: (1) Marine Shallow Convection, (2) Continental Shallow Cumulus Convection, and (3) Deep Convection. For each regime i, we compute cluster assignments in the warmed climate (with corresponding probabilities π_i^{4K}), and save the cluster centers. Second, we encode vertical velocity fields from the Control climate (0K) using the same encoder and compute the new cluster assignments (with corresponding probabilities π_i^{0K}) using the saved cluster centers from the reference climate. Third, we compute cluster assignment differences for each regime as $\Delta\pi_i = \pi_i^{4K} - \pi_i^{0K}$, which can then be used as proxies for dynamical regime shifts.

3 Results

3.1. Unsupervised Machine Learning Reveals Convective Responses to Climate Change Figure 3 shows the probability shifts $\Delta\pi_i$ from the Control to the Warmed climate. The dominant climate change signal captured by our unsupervised framework is the increased geographic concentration of deep convection (3c). More specifically, deep convection becomes more frequent over warm ocean waters and especially the Pacific Warm Pool [4] while shallow convection becomes less common in these unstable regions (3a). This result is consistent with observational trends showing an intensification of already powerful storms over the warm tropical waters thanks to greater moisture convergence [4] (SI-F). The agreement between the unsupervised $\Delta\pi_1$ of deep convection and extreme precipitation changes (3d) motivates further investigation.

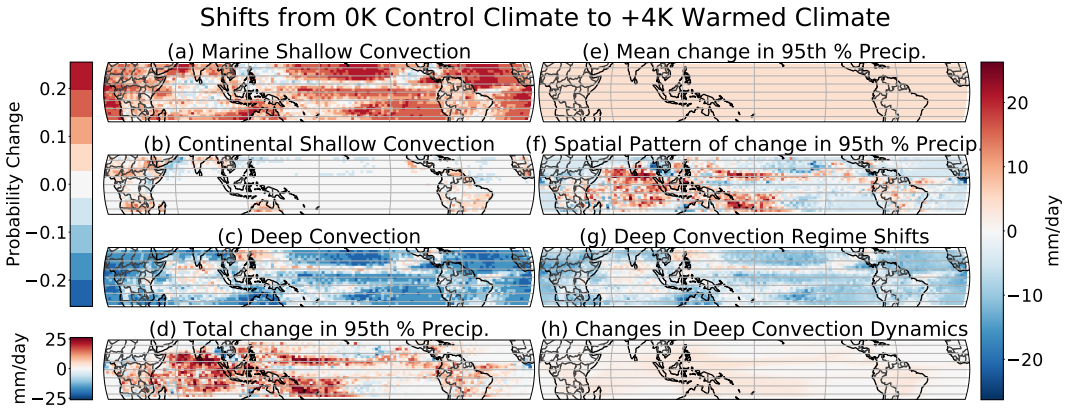

Figure 3: Changes induced by $+4°$C of simulated global warming: **The patterns of storms change** (a-c), **which changes the patterns of extreme precipitation** (f), **mostly because deep convective storms shift location** (g). Panels (a-c) display probability shifts in the three dynamical regimes found through clustering with $N = 3$, corresponding to (a) "marine shallow", (b) "continental shallow cumulus", and (c) "deep" convection. We subtract the spatial-mean change (e, the "thermodynamics") from the total change (d) to yield the "dynamic" contribution (f). Using Eq. 1, we decompose the changing spatial patterns (f) into five terms, including (g) probability changes in deep convection, (h) changes in deep convective precipitation, and three additional terms depicted in Figure 9.

3.2. Decomposing the Dynamic Contribution to Extreme Precipitation Changes We go one step further by decomposing the spatial patterns (3d) into changes in regime probability π_i and changes in regime dynamics $\mathcal{D}_i$. Unlike traditional approaches that spatially average information, we use a fully-convolutional encoder and latent space clustering to leverage storm-scale variability.

We calculated changes in regime probability in section 2, so we must now calculate changes in regime dynamics, which involves the following two steps. First, we empirically estimate $\mathcal{D}_i$ by using the probabilities of deep and shallow convection[2], and depict the results across precipitation percentiles in Figure 7. $\mathcal{D}_1$ and $\mathcal{D}_2$ are positive and increase with precipitation quantile, which is reassuring because precipitation is positive. Second, we estimate changes in "deep" and "shallow" convection dynamics as $\Delta\mathcal{D}_i = \mathcal{D}_i^{4K} - \mathcal{D}_i^{0K}$. With just the information of the regime probabilities, we can model the spatial patterns of precipitation changes at upper percentiles (Figure 6). Our model becomes less accurate at lower precipitation quantiles because we are not using specific humidity information (the approximation of Eq. 3 is only valid for high precipitation quantiles).

Our model can now be used to better understand the drivers of extreme precipitation changes themselves. We hone in on just deep convection, the regime most closely associated with extreme precipitation in the tropics, and ask: *Did the patterns of extreme precipitation simply follow the changing patterns of deep convection, or are there more complex changes in how deep convection produces rain?* We address this question by comparing the rate of change in deep convection probability ($\frac{\Delta\pi_1}{\pi_1}$) against the rate of change in the deep convection Dynamical Prefactor ($\frac{\Delta\mathcal{D}_1}{\mathcal{D}_1}$) in Figure 8. In high precipitation quantiles where our model works best, extreme changes are dominated by regime probability shifts (Figure 3f) rather than by intra-regime changes (3g): (Figure 6): $\frac{\Delta\mathcal{D}_1}{\mathcal{D}_1} < \frac{\Delta\pi_1}{\pi_1}$[3]. This means that the spatial patterns of extreme precipitation changes are dominated by the changing patterns of storm dynamics identified by our unsupervised framework, suggesting that unsupervised learning could play a broader role in leveraging storm-scale climate information to deepen our understanding of extreme events.

[2]More specifically, we estimate the dynamical pre-factors ($\mathcal{D}_i$) by regressing $\frac{P_{\text{extreme}}}{q_{\text{sat}}}$ on π_1 and π_2, neglecting the "Continental Shallow Cumulus" regime as it concentrates over arid continental zones with high lower tropospheric stability and low latent heat fluxes, making conditions unfavorable for precipitation [9]

[3]Note that at precipitation quantiles larger than 0.99, we lack samples for the analysis to work properly as evidenced by the pixelation of the changing patterns in Figure 5

Acknowledgments and Disclosure of Funding

The authors thank the National Science Foundation (NSF) Machine Learning and Physical Sciences (MAPS) program and NSF grant 1633631, Office of Advanced Cyberinfrastructure grant OAC-1835863, Division of Atmospheric and Geospace Sciences grant AGS-1912134, Division of Information and Intelligent Systems grants IIS-2047418, IIS-2003237, IIS-2007719, Division of Social and Economic Sciences grant SES-1928718, and Division of Computer and Network Systems grant CNS-2003237 for funding support and co-funding by the Enabling Aerosol-cloud interactions at GLobal convection-permitting scalES (EAGLES) project (74358), of the U.S. Department of Energy Office of Biological and Environmental Research, Earth System Model Development program area. This work was also supported by a Darpa (Defense Advanced Research Projects Agency) contract (HR001120C0021), as well as gifts from Intel, Disney, and Qualcomm. We further acknowledge funding from NSF Science and Technology Center LEAP (Launching Early-Career Academic Pathways) award 2019625. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA, DOE, or NSF. Computational resources were provided by the Extreme Science and Engineering Discovery Environment supported by NSF Division of Advanced Cyberinfrastructure Grant number ACI-1548562 (charge number TG-ATM190002). DYAMOND data management was provided by the German Climate Computing Center (DKRZ) and supported through the projects ESiWACE and ESiWACE2. The projects ESiWACE and ESiWACE2 have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 675191 and 823988.

References

[1] T. H. Abbott, T. W. Cronin, and T. Beucler. Convective dynamics and the response of precipitation extremes to warming in radiative–convective equilibrium. *Journal of the Atmospheric Sciences*, 77(5):1637–1660, 2020.

[2] W. Adger, S. Agrawala, M. Mirza, C. Conde, K. O'Brien, J. Pulhin, R. Pulwarty, B. Smit, and K. Takahashi. Assessment of adaptation practices, options, constraints and capacity. climate change 2007: impacts, adaptation and vulnerability. *Contribution of working group II to the fourth assessment report of the intergovernmental panel on climate change*, pages 717–743, 01 2007.

[3] W. N. Adger, S. Dessai, M. Goulden, M. Hulme, I. Lorenzoni, D. R. Nelson, L. O. Naess, J. Wolf, and A. Wreford. Are there social limits to adaptation to climate change? *Climatic Change*, 93(3):335–354, 2009.

[4] R. P. Allan, C. Liu, M. Zahn, D. A. Lavers, E. Koukouvagias, and A. Bodas-Salcedo. Physically consistent responses of the global atmospheric hydrological cycle in models and observations. *Surveys in Geophysics*, 35(3):533–552, 2014.

[5] D. Arthur and S. Vassilvitskii. K-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, page 1027–1035, USA, 2007. Society for Industrial and Applied Mathematics.

[6] S. R. Bowman, L. Vilnis, O. Vinyals, A. M. Dai, R. Józefowicz, and S. Bengio. Generating sentences from a continuous space. In *CoNLL*, 2016.

[7] D. L. Davies and D. W. Bouldin. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227, 1979.

[8] L. Denby. Discovering the importance of mesoscale cloud organization through unsupervised classification. *Geophysical Research Letters*, 47(1):e2019GL085190, 2020. e2019GL085190 10.1029/2019GL085190.

[9] T. Dror, V. Silverman, O. Altaratz, M. D. Chekroun, and I. Koren. Uncovering the large-scale meteorology that drives continental, shallow, green cumulus through supervised classification. *Geophysical Research Letters*, 49(8):e2021GL096684, 2022. e2021GL096684 2021GL096684.

[10] A. Edelman, A. Gedling, E. Konovalov, R. McComiskie, A. Penny, N. Roberts, S. Templeman, D. Trewin, and M. Ziembicki. *State of the Tropics - 2014 Report*. 06 2014.

[11] A. Edelman, A. Gelding, E. Konovalov, R. McComiskie, A. Penny, N. E. Roberts, S. Templeman, D. Trewin, M. Ziembicki, B. Trewin, R. Cortlet, J. Hemingway, J. L. Isaac, and S. M. Turton. State of the tropics 2014 report. 2014.

[12] S. Emori and S. Brown. Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophysical Research Letters*, 32(17), 2005.

[13] S. Hettiarachchi, C. Wasko, and A. Sharma. Increase in flood risk resulting from climate change in a developed urban watershed - the role of storm temporal patterns. *Hydrology and Earth System Sciences*, 22(3):2041–2056, 2018.

[14] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. M. Botvinick, S. Mohamed, and A. Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. In *ICLR*, 2017.

[15] M. Khairoutdinov and D. Randall. Cloud resolving modeling of the arm summer 1997 iop: Model formulation, results, uncertainties, and sensitivities. *Journal of The Atmospheric Sciences - J ATMOS SCI*, 60:607–625, 02 2003.

[16] M. F. Khairoutdinov and Y. L. Kogan. A large eddy simulation model with explicit microphysics: Validation against aircraft observations of a stratocumulus-topped boundary layer. *Journal of the Atmospheric Sciences*, 56(13):2115 – 2131, 1999.

[17] T. Kurihana, E. Moyer, R. Willett, D. Gilton, and I. Foster. Data-driven cloud clustering via a rotationally invariant autoencoder, 2021.

[18] Z. Li and P. A. O'Gorman. Response of vertical velocities in extratropical precipitation extremes to climate change. *Journal of Climate*, 33(16):7125–7139, 2020.

[19] S. Lloyd. Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

[20] J. Macqueen. Some methods for classification and analysis of multivariate observations. In *In 5-th Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297, 1967.

[21] C. Muller and Y. Takayabu. Response of precipitation extremes to warming: what have we learned from theory and idealized cloud-resolving simulations, and what remains to be learned? *Environmental Research Letters*, 15(3):035001, 2020.

[22] C. J. Muller, P. A. O'Gorman, and L. E. Back. Intensification of precipitation extremes with warming in a cloud-resolving model. *Journal of Climate*, 24(11):2784–2800, 2011.

[23] P. A. O'Gorman. Precipitation extremes under climate change. *Current climate change reports*, 1(2):49–59, 2015.

[24] P. A. O'gorman and T. Schneider. Scaling of precipitation extremes over a wide range of climates simulated with an idealized gcm. *Journal of Climate*, 22(21):5676–5685, 2009.

[25] A. G. Pendergrass and D. L. Hartmann. Changes in the distribution of rain frequency and intensity in response to global warming. *Journal of Climate*, 27(22):8372–8383, 2014.

[26] D. Randall, M. Khairoutdinov, A. Arakawa, and W. Grabowski. Breaking the Cloud Parameterization Deadlock. *Bulletin of the American Meteorological Society*, 84(11):1547–1564, 11 2003.

[27] P. J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[28] Z. Wang and A. Bovik. Bovik, a.c.: Mean squared error: love it or leave it? - a new look at signal fidelity measures. ieee sig. process. mag. 26, 98-117. *Signal Processing Magazine, IEEE*, 26:98 – 117, 02 2009.

Supplemental Information

A. Data: High-resolution, Earth-like Simulations of Global Surface Warming

We will now cover the Multi-Model Framework (MMF) in more detail. The MMF used to generate our training and test data is composed of small, locally periodic 2D subdomains of explicit high-resolution physics that are embedded within each grid column of a coarser resolution ($1.9^\circ \times 2.5^\circ$ degree) host planetary circulation model [16]. In total, we performed six simulations of present-day climate launched from different initial conditions (but consistent resolution) using the MMF [26], configured with storm resolving models that are 512 km in physical extent, each with 128 grid columns spaced 4 km apart. We use a scheme with thirty vertical levels to represent the atmosphere. We then perform an additional six simulations but increase the sea surface temperatures by 4K.

We compare the Control simulations against those with uniform increases in sea surface temperatures ("Warmed"). For our purposes, this creates a testbed that can serve as a proxy for climate change, which we can better understand by examining spatial and intensity shifts between the turbulent updrafts within the two SPCAM climates. However, we acknowledge that surface warming is only an approximation for the thermodynamic consequences of CO_2 concentration increase.

To investigate the "dynamic mode" of precipitation, we choose vertical velocity to represent the state of the atmosphere. These vertical velocity fields contain information about complex updraft and gravity wave dynamics across multiple scales and phenomena. We considered the entire 15S-15N latitude band containing diverse tropical convective regimes. Examples of these vertical velocity snapshots, selected by precipitation percentile, can be seen in Figure 1

B. Precipitation Extremes Decomposition

This appendix derives Eq. 1 by making a series of simple physical assumptions about precipitation. Note that while these assumptions help give physical meaning to each term of Eq. 1, Eq. 1 could also be derived by decomposing the extreme precipitation field into its spatial-average and an anomaly, before further decomposing the anomaly using the objectively-identified dynamical regimes. This means that the assumptions made in this section only need to *approximately* hold to physically interpret the results of our decomposition.

To first order, precipitation (P) scales like condensation rate, which depends on the full vertical velocity (w) and atmospheric water vapor (here quantified using specific humidity q) fields:

$$P \approx P\left(w, q\right). \tag{2}$$

Note that Eq. 2 neglects the dependence on microphysical processes (see e.g., [21]) to focus on the thermodynamical and dynamical components of precipitation. When focusing on extreme precipitation, we de facto sample atmospheric columns that are so humid that the specific humidity q equals its saturation value q_{sat}. This allows us to further simplify Eq. 2 in the case of precipitation extremes (high quantiles of P):

$$P_{\mathrm{extreme}} \approx P_{\mathrm{extreme}}\left(w, q_{\mathrm{sat}}\right). \tag{3}$$

We now make the assumption that the thermodynamic dependence on q_{sat} can be factored out of the right-hand side of Eq. 3 and denote the dynamical pre-factor as $\mathcal{D}\left(w\right)$:

$$P_{\mathrm{extreme}} \approx q_{\mathrm{sat}} \times \mathcal{D}\left(w\right). \tag{4}$$

The previous assumption can be justified quickly by assuming a moist adiabatic temperature profile and a vertically-uniform vertical velocity profile for extreme events [24, 22]. It can also be justified more accurately by noting that such vertical velocity profiles collapse when changing the vertical coordinate from pressure to the normalized integral of the moisture lapse rate [1]. We can now linearly decompose the dynamical pre-factor $\mathcal{D}\left(w\right)$ into the N regimes identified by our unsupervised learning framework:

$$\mathcal{D}(w) \approx \mathcal{D}_0 + \sum_{i=1}^{N} \pi_i \mathcal{D}_i, \tag{5}$$

where π_i is the probability of each dynamical regime. Combining Eq.'s 4, 5, and taking a logarithmic derivative with respect to climate change allows us to decompose relative changes in extreme precipitation as follows:

$$\frac{\Delta P_{\text{extreme}}}{P_{\text{extreme}}} \approx \frac{\Delta q_{\text{sat}}}{q_{\text{sat}}} + \frac{\Delta\left(\mathcal{D}_0 + \sum_{i=1}^{N} \pi_i \mathcal{D}_i\right)}{\mathcal{D}_0 + \sum_{i=1}^{N} \pi_i \mathcal{D}_i}, \tag{6}$$

where Δ denotes absolute changes from the reference to the warm climate. Lastly, we approximate the thermodynamic contribution to precipitation extremes as the relative changes in near-surface saturation specific humidity, which can be further approximated as spatially uniform:

$$q_{\text{sat}} = q_{\text{sat}}(T_s, p_s) \Rightarrow \frac{\Delta q_{\text{sat}}}{q_{\text{sat}}} \approx 7\%, \tag{7}$$

where T_s is near-surface temperature and p_s near-surface pressure. Expanding Eq. 6 and substituting $\mathcal{D}(w)$ using Eq. 4 yields Eq. 1.

C. VAE Training

Our machine learning methodology objectively defines dynamical regimes from millions of two-dimensional vertical velocity fields, for which we proceed with the creation of a latent manifold to ensure the local correlations in the updrafts of our vertical velocity fields are preserved. For this, we rely on a fully convolutional VAE design, whose architecture is depicted in Figure 2. To train the VAE, we perform beta-annealing [14, 6], expanding the Evidence Lower Bound (ELBO) traditionally used to train the VAE by including a β parameter and linearly anneal β from 0 to one over 1600 training epochs. The number of layers and channels in the encoder and decoder are depicted in Figure 2 (4 layers in each, stride of two). We use ReLUs as activation functions in both the encoder and the decoder. We pick a relatively small kernel size of 3 in order to preserve the small-scale updrafts and downdrafts of our vertical velocity fields. The dimension of our latent space is 1000.

D. VAE Benchmarking and Performance Evaluation

We train our VAE on 160,000 unique vertical velocity fields and use an additional 125,000 samples to validate and optimize the model hyperparameters. Finally, we leverage 1,000,000 vertical velocity fields in the testing dataset for robust analysis. The high count in the test dataset is necessary both due to the high spatio-temporal correlations common in meteorological data but also because of the geographic conditioning in our analysis – we need enough samples at each lat/lon grid cell not just globally.

To determine whether our data are nonlinear enough to warrant the use of a VAE we also train a baseline model of the same architecture but with all activation functions replaced by "linear". This "linear" baseline provides an important benchmark for the credibility of our VAE and machine learning analysis workflow more broadly. The fact that the VAE reconstructs the vertical velocity snapshots with both lower error and a higher degree of structural similarity suggests significant non-linearity is involved in compressing and rebuilding the 2D fields (Table 1 and Table 2). This problem is therefore well suited to the non-linear dimensionality reduction of the VAE encoder and less so for linear models.

Tables 1 and Table 2 show 160,000 is enough training samples to create reconstructions of high-resolution vertical velocity fields with both a low MSE and a high degree of overall structural similarity. Though there is a small amount of overfitting, we see that performance remains strong for a test dataset containing multiple species of convection from all parts of the tropics ranging from deserts to rainforests; oceans to continents. Furthermore, what we are most concerned with is not the reconstruction quality itself, but the interpretability of the latent space for clustering.

Mean Squared Error m^2/s^2			
Model	Training Set	Validation Set	Test Set
VAE	$3.79 * 10^{-4}$	$1.11 * 10^{-3}$	$3.33 * 10^{-3}$
Linear Baseline	$3.10 * 10^{-3}$	$4.70 * 10^{-3}$	$5.10 * 10^{-2}$

Table 1: The MSE of both of our models ("linear" baseline and VAE) calculated across training/validation/test data. For both training and test data, we see low reconstruction errors, suggesting satisfactory skill and generalization ability. Overall, the VAE outperforms the "linear" baseline.

Structural Similarity Index Metric			
Model	Training Set	Validation Set	Test Set
VAE	0.998	0.995	0.987
Linear Baseline	0.990	0.986	0.981

Table 2: The mean SSIM [28] of both of our models ("linear" baseline and VAE) across training/validation/test data. The models both generalize well to our test data. Again, the VAE outperforms the "linear" baseline.

E. K-Means Clustering Approach

We apply the K-Means Clustering algorithm to partition the latent space of our VAE and analyze which physical properties can be clustered in this reduced order **z** space. This approach first randomly assigns centroids, C, to locations in the **z** space (note we actually use the more modern k++ algorithm [5] to maximize the initial distances between the centroids). Latent representations of each sample $\mathbf{z}_i$, in the test dataset of size N, are assigned to their nearest centroid. The second stage of the algorithm moves the centroid to the mean of the assigned cluster. The process repeats until the sum of the square distances (or the Inertia, I) between the latent space data points and the centroids are minimized [19, 20] such that:

$$\bar{\mathrm{I}} \stackrel{\mathrm{def}}{=} \sum_{\mathrm{i}=0}^{\mathrm{N}} \min_{\mathrm{l} \in \mathrm{C}} ||\mathrm{z_i} - \bar{\mathrm{z}}_\mathrm{l}||^2, \tag{8}$$

in which $\bar{z}_l$ is the mean of the given samples belonging to a cluster l for the total number of cluster centers C. We always calculate ten different K-means initializations and then select the initialization with the lowest inertia. This process allows us to derive the three data-driven convection regimes within SPCAM highlighted in Figure 3.

We qualitatively choose an optimal number of cluster centroids (centers), k by incorporating domain knowledge rather than a traditional approach relying on the rate of decrease in I as k increases or a single quantitative value such as a Silhouette Coefficient [27] or Davies-Bouldin Index [7]. More specifically, we identify the maximum number of "unique clusters". We define a "unique cluster" of convection as a group in the latent space where the typical physical properties (vertical structure, intensity, and geographic domain) of the vertical velocity fields are not similar to the physical properties of another group elsewhere in the latent space. Empirically this exercise enables us to create three unique regimes of convection (Figure 3). When we increase k above three, we get sub-groups of "Deep Convection" without differences in either vertical mode, intensity, or geography. Thus we don't consider $N > 3$ to be physically meaningful for our purposes.

Because we seek to contrast common clusters between different climates, we do not use Agglomerative (hierarchical) Clustering unlike other recent works that cluster compressed representations of clouds from machine learning models [8, 17]. Using the K-means approach, we can save the cluster centroids at the end of the algorithm. This provides a basis for cluster assignments for latent representations of out-of-sample test datasets when we use a common encoder as in Section B. More specifically, we only use the cluster centroids to get label assignments in other latent representations. We don't move the cluster centroids themselves once they have been optimized on the original test dataset (the second part of the K-means algorithm). Keeping the center of the clusters the same between different types of test data ensures we can objectively contrast cluster differences through the lens of the common latent space.

F. Understanding Convection via Vertical Structure

To better understand the physical properties embedded within a latent space cluster, we use vertical velocity moments as summary statistics. We construct summary vertical profiles, grounded on the principle that for convection the vertical (v) dimension will be far more important than the horizontal (h) dimension for highlighting key physics. We derive the first-moment statistic:

$$\overline{\mathrm{w'w'_i}} \overset{\mathrm{def}}{=} \sqrt{(\mathrm{W_i} - \bar{\mathrm{W}}_{\mathrm{i,h}})^2}, \tag{9}$$

where $\bar{W}_{i,h}$ is the mean of the vertical velocity field upon averaging out the horizontal dimension. We can average these statistics across a cluster to approximate the convective structures organized within.

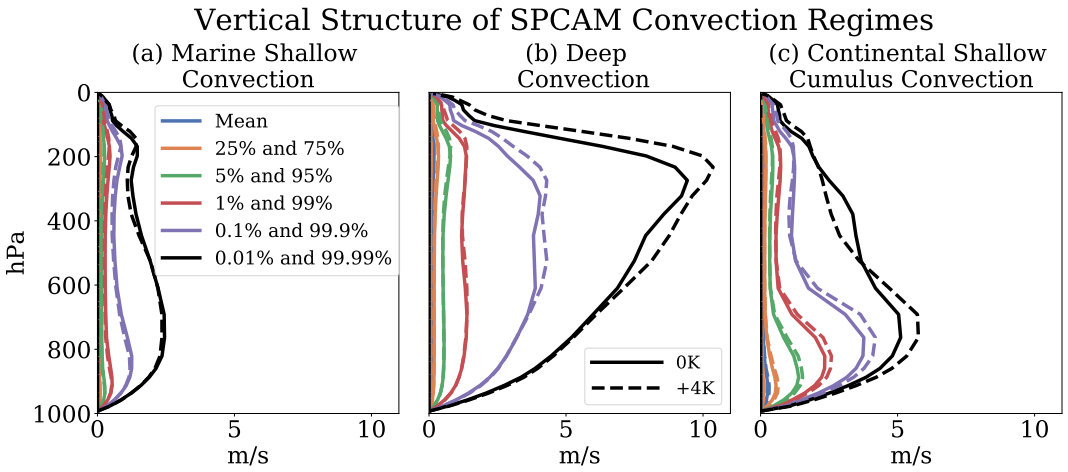

Figure 4: A comprehensive view of the vertical structure of each type of convection in SPCAM and how it changes as temperatures rise (solid vs. dashed lines). But instead of only restricting ourselves to a view of the mean, we look at percentiles across the test data in each convection cluster. The VAE anticipates both an increase in the most intense deep convection with warming (b) and strengthening of turbulent updrafts in the boundary layer (c).

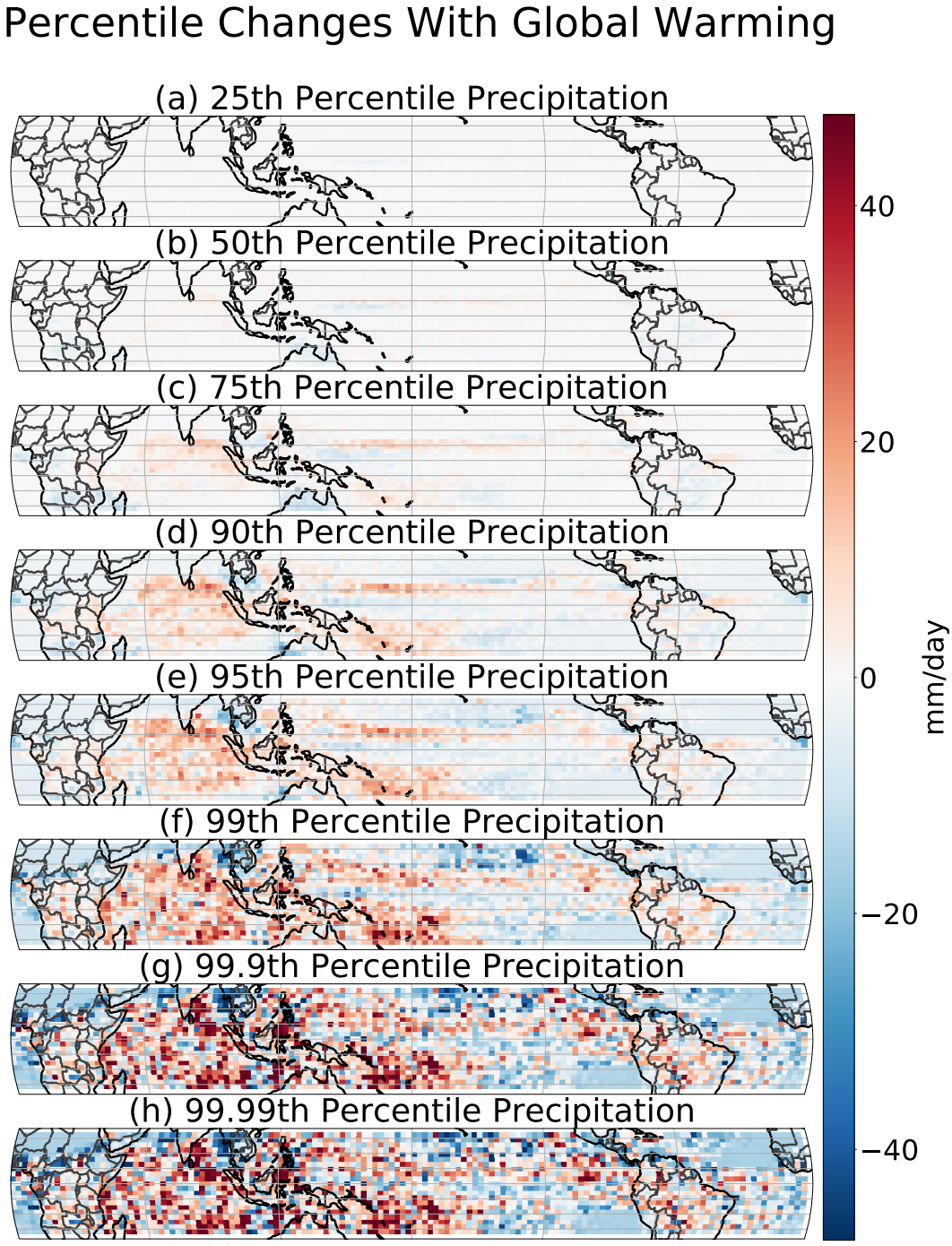

Figure 5: The shifts in different percentiles of precipitation with global warming, where we again stratified and plotted the data by latitude/longitude grid cell. As in Figure 3d we again remove the mean to highlight the dynamical pattern and see at what threshold the alignment with the VAE identified Deep Convection shifts (Figure 3c) is greatest. The top percentiles including (f-h) are pixelated because of a lack of samples that are out on the tail of the PDF.

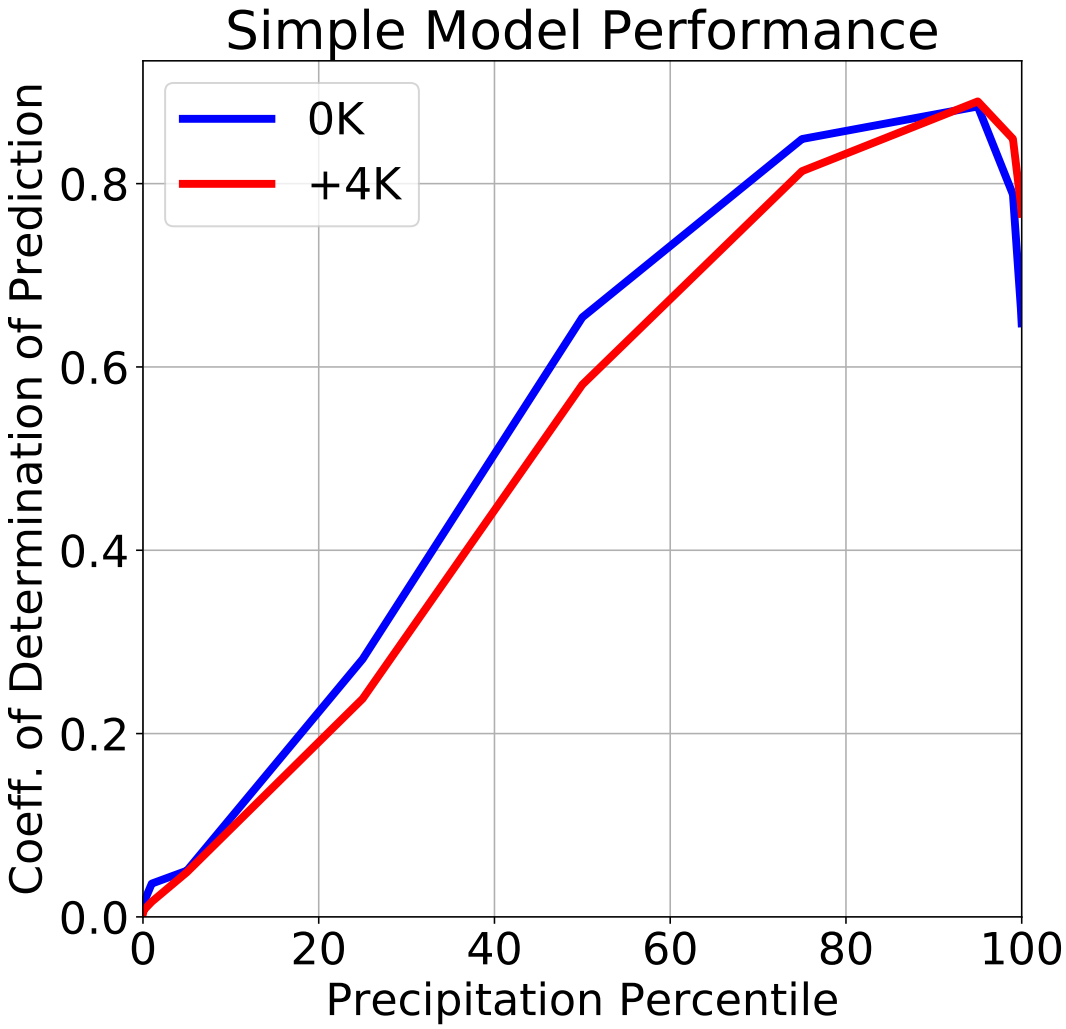

Figure 6: The simple results of the simple regression model we use to predict extreme precipitation patterns ($\frac{P_{extrmeme}}{qsat}$) using just the dynamic contributions, $\pi_{Deep\ Convection}$ and $\pi_{Shallow\ Convection}$ identified by our unsupervised ML framework. We see our model works very well for high precipitation percentiles where the dynamic contributions are greatest and less well for lower percentiles where thermodynamics are also important.

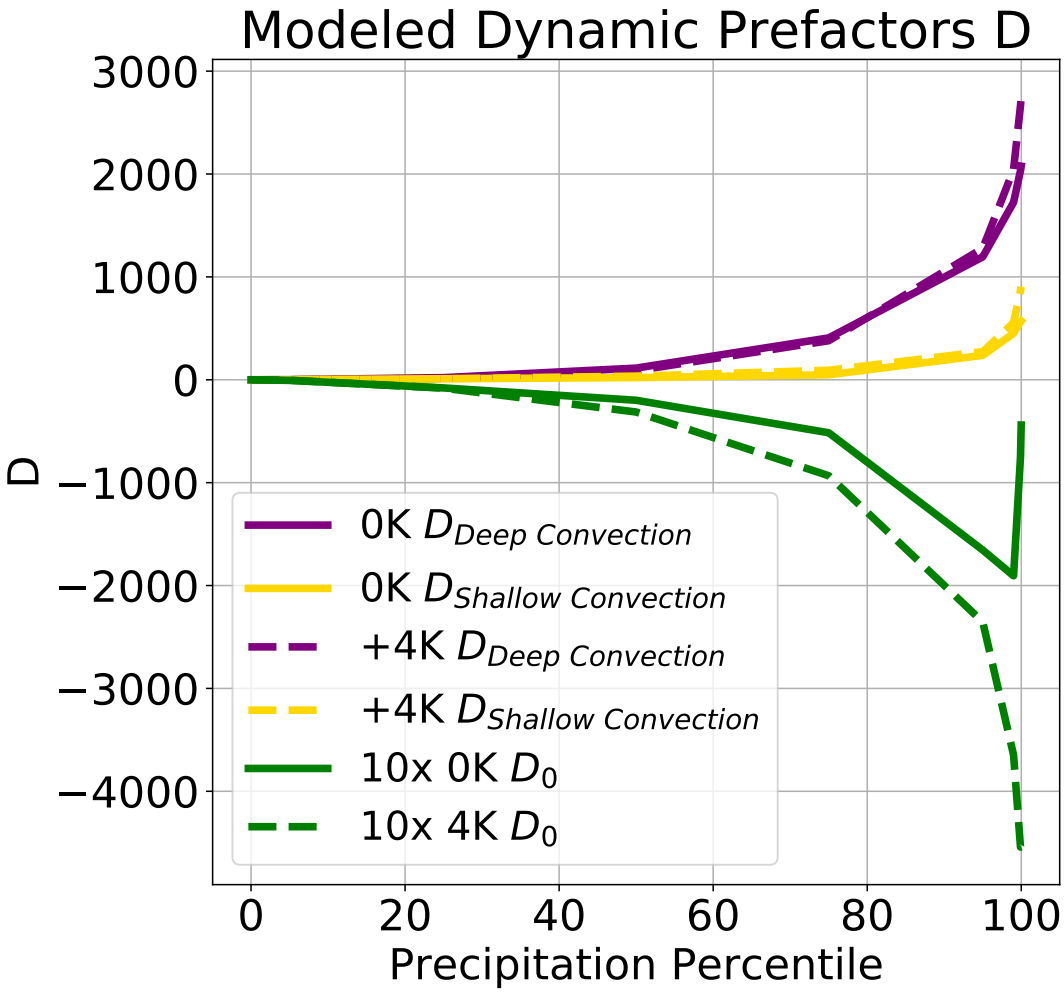

Figure 7: The breakdown of components of our simple model at different precipitation percentiles. We see the approximated Dynamical Prefactors $\mathcal{D}$ are positive and increase for extreme precipitation percentiles mirroring the expected greater dynamic contribution responsible for heavy rain and tropical cyclones. We also plot the intercept $\Delta\mathcal{D}_0$ for additional context.

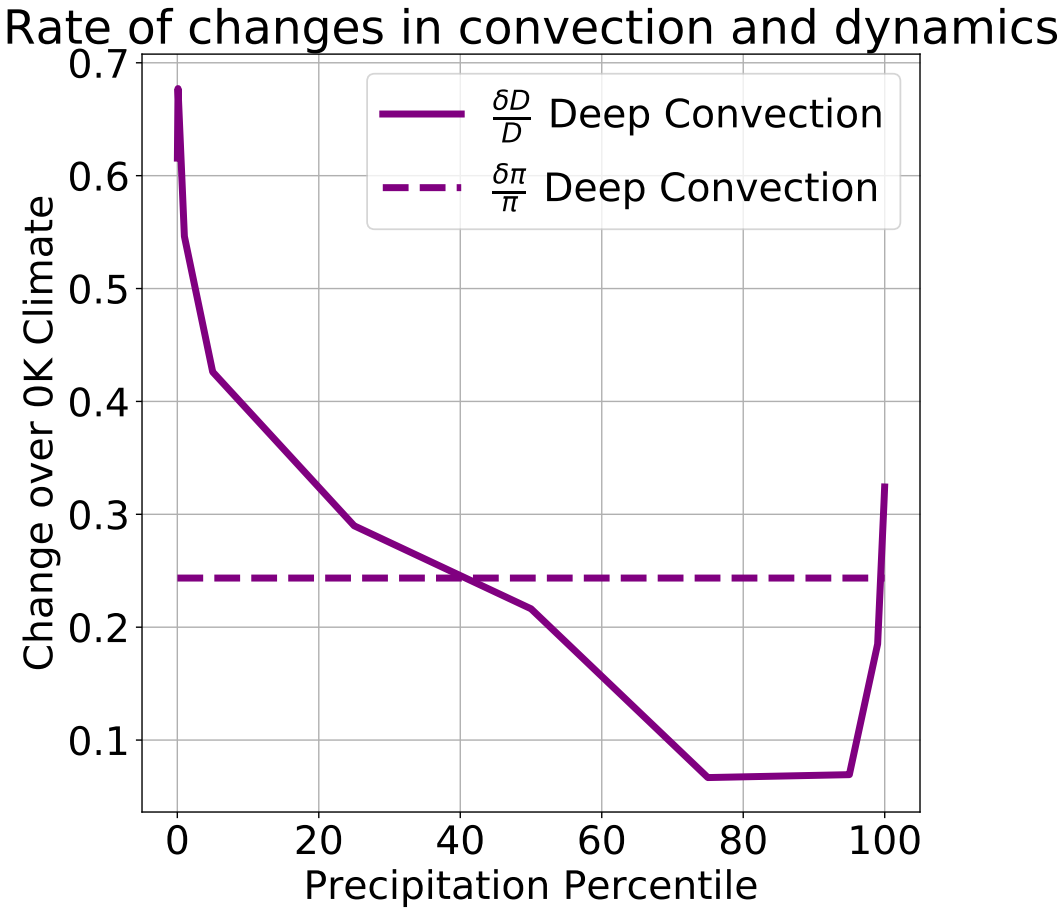

Figure 8: Over all precipitation percentiles, we examine the rate of change in the Dynamical Prefactor of Deep Convection $\frac{\Delta \mathcal{D}}{\mathcal{D}}$ against the shift in Deep Convection between climates $\frac{\Delta \pi}{\pi}$. We see at upper precipitation percentiles the $\frac{\Delta \mathcal{D}}{\mathcal{D}} < \frac{\Delta \pi}{\pi}$. This indicates that regime shifts dominate the precipitation extreme changes (patterns).

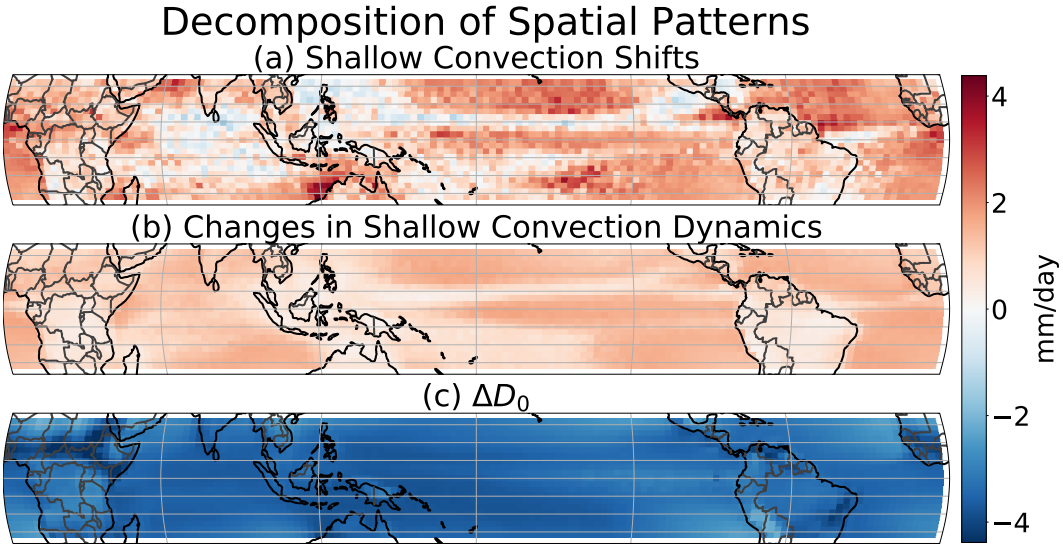

Figure 9: From Eq. 1, we can decompose the changing spatial patterns (Figure 3f) into five terms, including probability changes in shallow convection (a), changes in deep convective precipitation (b), and the intercept of Dynamical Prefactor (c).