Controls on stratocumulus texture diagnosed by deep learning

James A. Franke[a,b,1], **Takuya Kurihana**[c], **Ian T. Foster**[c,d], and **Elisabeth J. Moyer**[a]

This manuscript was compiled on September 25, 2023

Marine low clouds—stratocumulus—represent the single largest source of uncertainty in the future climate response. Satellite observations show a vast array of fine-scale textures and properties not resolved by global climate models, making detailed cloud process analysis necessary but difficult. We show here that new self-supervised cloud classification techniques based on deep learning can aid in diagnosing drivers of observed cloud morphologies. We apply an autoencoder to two decades of MODIS multispectral observations without labels to generate the AI-driven Cloud Classification Atlas (AICCA), a set of 42 cloud classes that take texture into account, and compare observed classes with meteorological variables from ERA5 reanalysis and other satellite observations. We find that the strongest predictor of stratocumulus class is the temperature structure of the lower troposphere, and that dominant types fall into coherent and interpretable zones in the 2D space of inversion strength and near surface temperature. This relationship holds not only across the three major stratocumulus regions but also over time for any given location, suggesting these variables are fundamental physical drivers. The relationship cannot be reconstructed from mean cloud properties alone but requires information on texture. A substantial portion of variability in marine cloud textures remains unexplained, including spatially coherent transitions where parts of thick stratocumulus decks shift to more open configurations. We show that counter to expectations, rainfall does not play a clear role in these transitions and is not predictive of cloud classes. Self-supervised cloud classification shows promise to improve marine low cloud process understanding.

climate change | stratocumulus | deep learning | self-supervised classification

Significance Statement

Low clouds over the ocean are critical for the Earth's climate because they cool the planet by reflecting sunlight back to space. Climate models do not simulate them well, and so cannot predict how they will change as the climate warms. Here we show that satellite images can be used to understand low cloud behavior when paired with a new automated cloud classification scheme. We find that cloud textures are broadly controlled by the temperature profile in the lower atmosphere where they form, and that, surprisingly, rainfall does not play a major role in changing those textures.

The response of low clouds is the largest uncertainty in projections of future climate under CO_2 forcing (1–4). A particular concern is the persistent marine stratocumulus decks that form in the subtropics off the West coasts of continents. These decks cover only 5% of the Earth's surface but are disproportionately important to its energy balance, cooling the planet by as much as 8K (5). Recent studies based on high-resolution (10 m) large eddy simulations (LES) have suggested the decks may disappear under high CO_2 (6) and / or high sea surface temperature (7) conditions. Coarse-resolution (100 km) global climate models provide little guidance: their shortwave cloud feedbacks differ even in sign (e.g 4), though models that better capture historical cloud frequencies tend to have larger, positive values (8, 9). Global convection-permitting models (<5 km) can better capture marine stratocumulus, but their simulation periods are typically only months (10, 11, e.g.) and their resolutions are insufficient to fully resolve the relevant dynamics (12), so their output still does not fully reproduce stratocumulus textures (Figure 1, which shows the 1.4 km ECMWF model(11); for other models see SI Figure S1).

Because global scale simulations at 10 m resolution remain a long way off (13), satellite observations may be the most appropriate tool for understanding stratocumulus formation and stability. Space-borne instruments have by now captured several decades of high-resolution (to 30 m) multispectral imagery, which necessarily also reflect true underlying fine-scale cloud processes. These observational datasets present a different data challenge. Because the natural world does not permit large-scale experiments, process understanding must be derived from the complex details of natural experiments. The scale and variation of cloud observations then means that some form of dimension reduction must be applied for them to be usable. Cloud classification schemes have a long history, but do not capture the diversity of stratocumulus patterns. The most commonly-used scheme (ISCCP (14, 15)), assigns each observed pixel to one of nine classes (plus a mixed-layered type) based on its derived optical thickness and cloud top height. In this scheme, all pure marine stratocumulus falls in a single class (low, medium-thick) (14).

www.pnas.org/cgi/doi/10.1073/pnas.XXXXXXXXXX

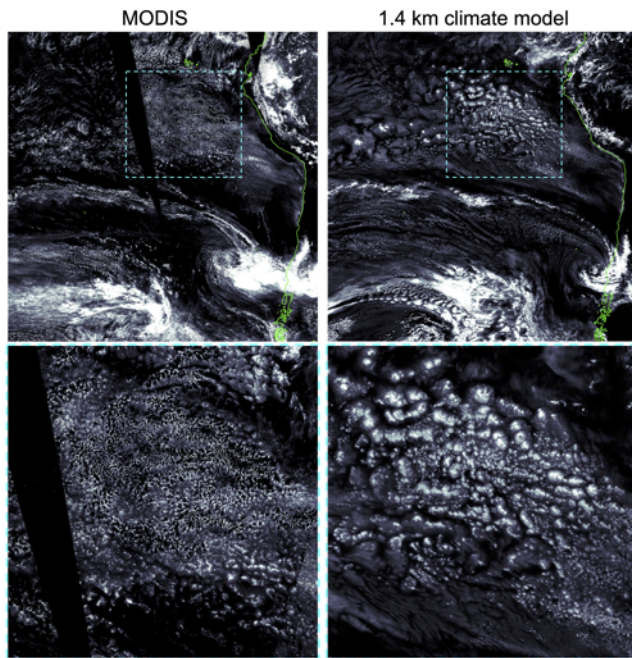PNAS — September 26, 2023 — vol. XXX — no. XX — 1–7

Fig. 1. Cloud textures off the coast of South America: liquid water path from **[left]** MODIS observations (Nov. 1, 2018) and **[right]** a 1.4 km horizontal resolution simulation of the XX model, initialized with observations on XX, 2018 (11), so that weather patterns should match. Color intensity is identical in both images. While the high-resolution simulation captures much of the large-scale pattern (top), it produces too much self-aggregation in the main stratocumulus deck region (bottom row zoom in).

In this work we make use of a new, deep-learning-based classification of ocean clouds, the AI-driven Cloud Classification Atlas (AICCA), in which 25 of 42 classes would clearly be considered stratocumulus in the ISCCP defintion. (See Methods.) AICCA is generated by self-supervised classification of 22 years of 1 km resolution cloud imagery from the Moderate Resolution Imaging Spectroradiometer (MODIS) as 128×128 pixel patches. The resulting atlas of 200M+ classified patches reduces an 800 Tb dataset to 10 GB, and the assigned classes have been shown to have consistent physical and radiative properties and to represent unique textural arrangements (16, 17). While observational studies of marine stratocumulus have largely focused on its *fractional occurrence*, and have shown that this occurrence is predictable from lower tropospheric stability (18), the AICCA classes allow asking instead what variables predict the *texture* of those clouds. Textures in turn provide insight into the processes that govern stratocumulus formation and evolution.

The evolution of marine stratocumulus textures on a timescale of days is broadly understood. Thick decks are produced in the stable subtropics by Rayleigh-Bernard convection; the decks then break up when they are advected equatorward into less stable regimes where deeper convection can occur (19, e.g.). However, this framework operating alone would produce a static spatial cloud distribution. In reality stratocumulus are highly dynamic, with complex patterns evolving on timescales of hours ().

Research on shorter timescale stratocumulus changes has generally focused on the transition between classic closed- and open-cell configurations, which can occur across coherent regions even when large-scale stability remains constant (20). Many studies have suggested that rain plays a role in this transition. Observational campaigns from ships (EPIC, Eastern Pacific Investigation of Climate) and aircraft (DECS, Drizzle and Entrainment Cloud Study) found higher drizzle in parts of open-cell ("rift") clouds (21–23), and an early LES modeling study argued that precipitation directly initiates the transition by causing downdrafts (24). Watson-Parris et al. (2021) used machine learning and supervised classification to categorize open-cell stratocumulus regions in otherwise unbroken decks in MODIS data and found higher rainwater path in those regions (25). On the other hand, Eastman et al. (2016) cautioned that a weak statistical relationship with precipitation exists even in advective transitions where the underlying driver is clearly increasing boundary layer depth (reduced stability) (18).

We use AICCA classes here to examine the behavior of clouds in the three major subtropical stratocumulus regions (26)—the Californian, Peruvian, and Angolan, found between approximately 5 and 40 degrees North or South. We map observed cloud classes to coincident meteorological conditions and rainfall, using ERA5 reanalysis and GMP IMERG (TRMM) microwave satellite precipitation measurements. The goal is to understand what factors govern stratocumulus texture, both on average and during rapid evolution.

Results

Clouds in the subtropical stratocumulus regions are highly diverse, and the AICCA classes capture that diversity. The most dominant cloud class is the closest visually to classic closed-cell organization (#30, at 8% of cloud observations), and the next most common is a nearly uniform deck (#35, at 7%), but classes with more open textures are nearly as frequent. The 15 most common classes, all stratocumulus, collectively make up 67% of cloud observations in these regions, with the least-common of these still at 2.6%. (See Fig. 2 and SI Fig. S4 for thumbnail images and SI Fig. S5 for occurrence frequencies.).

To diagnose factors controlling these textures, we map the top 15 most populous classes against a battery of 68 meteorological quantities derived from reanalysis, including surface temperature, lower tropospheric stability, vertical motion, wind shear, and humidity. (See SI Table XX.) Variables are tested in pairs and scored by their ability to separate the 15 cloud classes from one another. The two 2-variable combinations with the highest predictive power across classes are inversion strength paired with either near-surface air temperature or lower tropospheric stability. We choose the first of these pairs and show in Figure 2a the the dominant cloud classes in each part of its 2D parameter space. Domains favorable to each cloud type are both coherent in parameter space and physically interpretable. Stable conditions with high inversion strength predict more closed-cell or uniform types, as expected (blue shades in Fig. 2). Higher temperatures promote more open configurations, until some limit in inversion-strength-temperature space where stratocumulus become scarce and high clouds dominate instead. These patterns are similar in each of the 3 regions considered (SI Fig. S XX). This separation is possibly only given texture-based classes and cannot be reproduced using mean cloud properties alone.
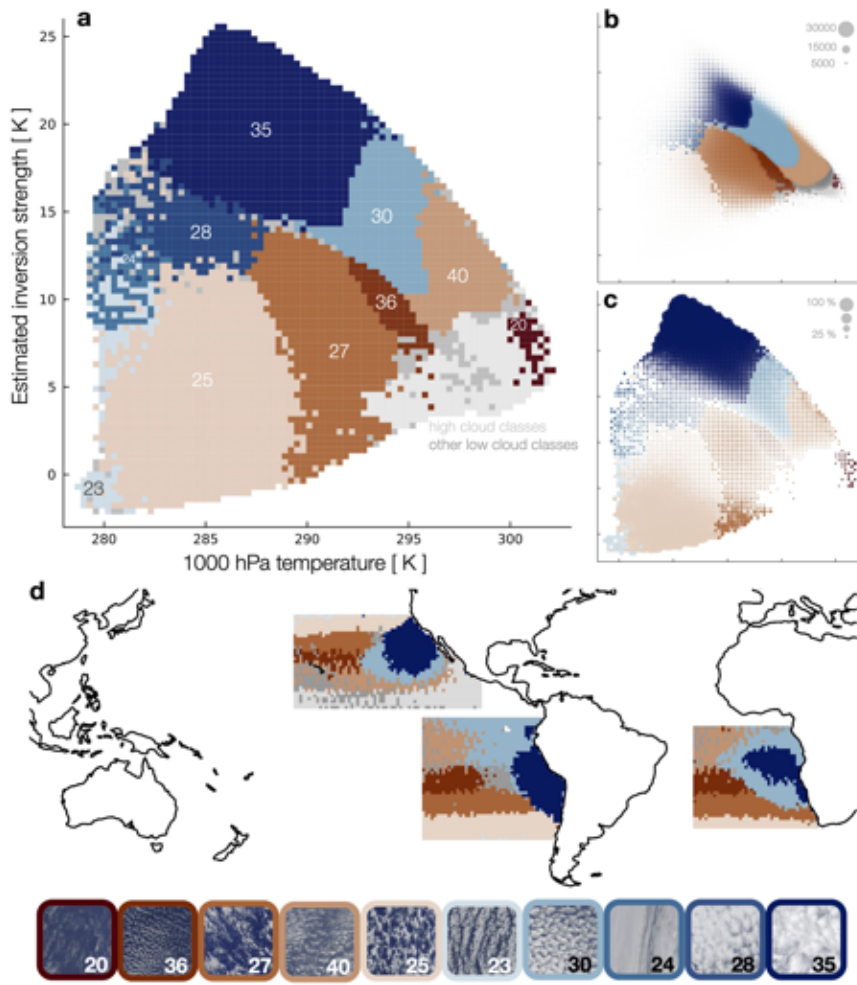
Fig. 2. AICCA cloud classes mapped to a meteorological domain of near-surface temperature and inversion strength, for all patches in the 3 main subtropical stratocumulus regions. **a** The 10 most dominant classes in meteorological space, color-coded. Thumbnail images of typical class members are shown at bottom. 9 of the 10 classes are stratocumulus; #20 is thinner and higher. All high clouds (mean cloud top pressure >680 hPa) are shown in light grey. Regions of dominance are generally well-separated and coherent. **b** As in panel a but with dots scaled in size according to frequency of those meteorological conditions (for samples with clouds). Air parcels tend to move from cold and stable to warmer and less stable, and clouds textures from near-uniform to more open: #35 → #30 → #40. **c** As in panel but with dots scaled to represent the share of all clouds by the most dominant cloud class. Very high- or low-stability conditions produce more consistent textures. **d** The 3 stratocumulus regions color-coded by the most dominant cloud class in each location, using the same color code. Classes are distributed in understandable geographic patterns, following large-scale environmental gradients. As expected, zones of thick stratocumulus are in the subtropics. The use of IR radiances in the classification procedure likely helps produce the strong temperature dependence in classes #25, #27, #36, and #40.

Weather conditions in the stratocumulus regions do not evenly sample this entire parameter space. The predominant trajectory for air parcels moving out of subtropics is to move equatorwards, warm, and weaken in inversion strength (Fig. 2b); as this happens the stratocumulus becomes more sparse in texture and then finally breaks up. Cool surface temperatures with low inversion strength are almost never sampled, but when those conditions do occur, they are associated with a well-defined cloud class. The 10 most dominant cloud classes in parameter space are color-coded in Figure 2. The group includes the two most frequent classes (#30 and #35), and all but one are stratocumulus and are included in the set of the 15 most common. (The sole exception is the higher and thinner class #20, whose mean characteristics put it at the ISCCP border between cumulus, stratocumulus, altocumulus, and altostratus.) In total, these 10 classes represent 42% of all cloud occurrences.

While each meteorological domains in Figure 2a is by definition associated with a distinct cloud class, they do not explain all variance. The mean "purity" across the 10 domains is 30%, i.e., 30% of cloud occurrences are in fact the dominant class associated with a given meteorological environment. In general, the correlation of cloud texture with environmental conditions is strongest in extreme conditions. Thick stratocumulus decks (class #35, dark blue) tend to occur in very stable conditions, and the purity of the class

#35 domain is 90% at the largest inversion strengths (and 10% near the domain edge). Similarly, when low-temperature, low-stability conditions do occur, they produce class #25 with as high as 50% purity.

Cloud classes are coherent geographically, even though no geographic information is included in the model (Fig. 2d, see also SI Figure xx for occurrences of individual classes). This coherence reflects the same dependence on meteorological conditions discussed above. The prevailing conditions are set by large-scale features of the atmospheric and ocean circulation, driving similar cloud textures over large contiguous regions.

The correlation of meteorological environment to cloud texture holds not just across space but for any single location over time. (See Fig. 3 for one example.) As the Hadley cell shifts, subtropical locations exhibit a seasonal cycle in inversion strength (27), and cloud textures respond. Note that the cycle in cloud texture leads that in temperature, contradicting proposed suggestions that dense stratocumulus feeds back on near-surface temperature, cooling it (). Instead, the densest cloud classes begin appearing in Spring as stability increases, while temperatures are still rising.

Some variables provide no additional predictive power. Wind shear, humidity, vertical motion, potential vorticity, and sea level pressure do not provide any additional explanatory power for cloud texture (see supplemental table), and can
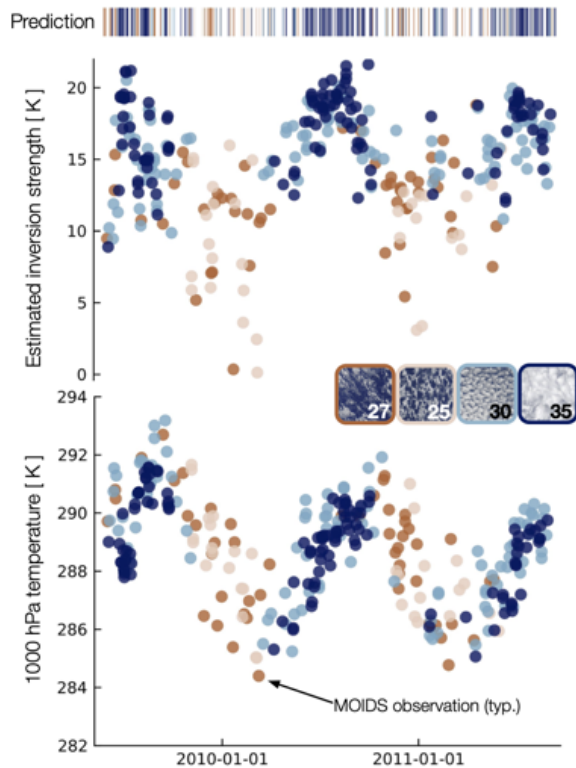
Fig. 3. Timeseries of meteorological variables and observed and predicted cloud classes for a single representative location at 29N, 130W in the North Pacific basin. Top shows inversion strength and bottom near-surface temperature, with observed cloud classes shown as colors for the 4 main classes only (25, 27, 30, and 35). Predicted cloud classes based on EIS and T1000 are shown as colored lines in the top barcode. Cloud texture varies seasonally, broadly matches predictions, and is clearly driven by the seasonal cycle in stability. Other locations show similar features; see See SI Fig. XX.

even degrade prediction skill when included with temperature measures. Boundary layer depth, a critical factor discussed in the standard stratocumulus breakup model, does not improve class texture predictions substantially. Whether this is because these factors are not important to cloud dynamics or because they are important at scales below 25 km, we cannot readily say with this approach based on ERA5 data. Additional weather variables with the most explanatory power are inevitably more information about the temperature profile, but can only increase average prediction skill about 1%.

Precipitation is not very predictive for cloud textures in this region. Including it in the logistic regression improves prediction only 0.01% on average for the top 15 classes and actually slightly degrades prediction for the top 4 (25,27,30,35). This is perhaps unsurprising, given how little it rains in these regions ().

In a case study to highlight this, we show a thick stratocumulus opening up to clear sky up in a region which is actually becoming more stable. No precipitation is measured in the region (Figure 4).

In a population-level sample across 1 million classes, precipitation is associated with changes in cloud type distribution but only to a small degree and only for the highest levels of precipitation (see supplemental bar chart). Classes are essentially constantly-distributed across the precipitation sample space.

Since precipitation may have a lag effect, we also track transitions between classes with a Lagrangian scheme based on ERA5 horizontal winds across the southern Pacific basin. Precipitation shows little power for predicting classes or for explaining transitions between classes. Transitions from class 30 are binned into three cases. The case with the highest rainfall is actually the no-transition case (30 remains 30).

Discussion

Determining the drivers of physical processes from correlations of cloud properties with environmental conditions is inherently difficult. Clouds may also feed back on their local conditions to some degree, including reducing sea surface temperatures () and increasing the stability of the lower atmosphere (). However, the assumption that clouds are following their environment appears reasonable to the first order. The major regions favorable for stratocumulus persistence are set by large-scale circulation patterns, including the descending Hadley cell and wind-driven ocean upwelling (). If clouds were able to make favorable conditions for themselves, we would find low marine clouds persisting in other locations as well.

Nevertheless, we can only report correlations between the meteorology and cloud type. with the MODIS satellite sampling rate (which provides at most two observations in one location in a day in this part of the world).

The fact that classes are coherent in meteorological space validates the classification approach and shows the power of unsupervised deep learning. However, much of the variation remains unexplained by this simple framework (see supplemental bar chart). We test the following possibilities to explain this discrepancy: Additional meteorological drivers, hysterisis or lag effects, regional differences, withing cloud class variation, and ERA5 errors.

Additional meteorological variables - we test up to 70 but can only improve prediction about 1% by adding additional. The next 8 predictors are additional temperature structure. The best non-temperature factor is Horizontal wind (u) at 950 hPa and then vertical monition at 750 hpa – but including both of these factors actual decreases performance slightly compared to the EIS-T1000 only formulation. (See supplemental table.

Hysteresis and lag effects from being advocated into a less stable or warmer domain. Qualitatively visual analysis of geostationary images shows clouds seem to change locally on timescales much shorter than advection to other regions. This makes hysteresis seem unlikely. See supplemental video. MODIS temporal sampling limits our ability to study this effect. This should be quantified in future work with higher satellite temporal sampling.

Regional differences - We find no regional differences in prediction skill or errors. See supplemental Figure.

Cloud texture variation within a class can be substantial (the classification is not perfect). However, this does not track with the unexplained class variation in meterological space. Those classes which are further from the cluster center in latent space are not necessary more likely to be located outside their region of temperature space, with a few potential exceptions. See supplemental figure.

ERA5 is not perfect at cloud representation. We repeat the analysis using only samples from cases where ERA5 cloud
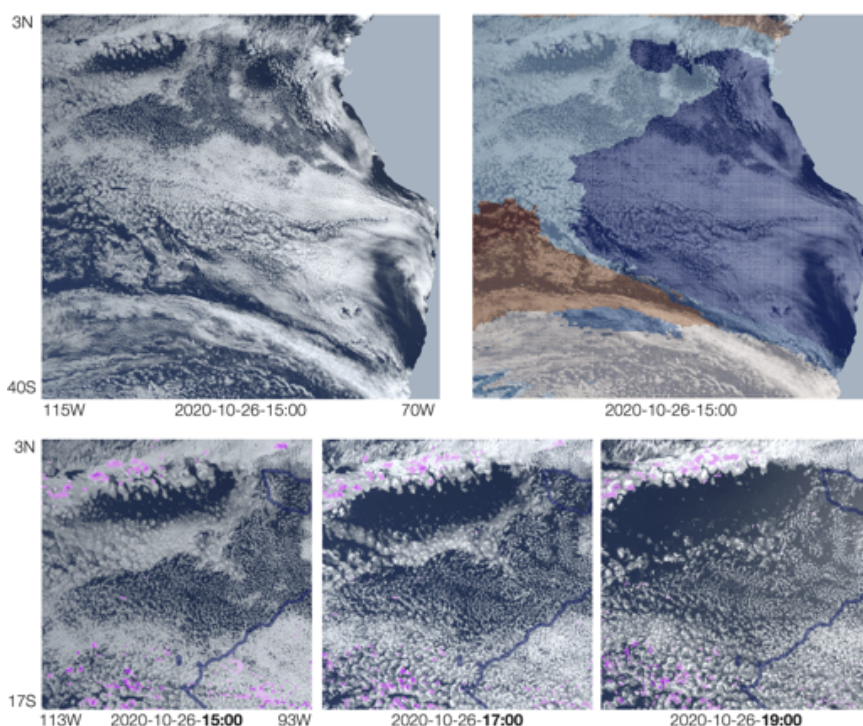
Fig. 4. **[Top row]** Visible light image from GOES-16 on October 26, 2020 and with meteorological predicted class plotted above in color. While the temperature structure predicts some of the visual texture on the larger-scale, many of the textures are unexplained by this framework. **[Bottom row]** Example zoom-in shows clear sky hole expanding for two time steps. Temperature-structure predicted class remains #30 for this time. Satellite microwave hourly precipitation measurements shown in pink. Precipitation does not explain rapid cloud disappearance in the center of the domain, nor do changes in the ERA temperature structure.

fraction closely matches observed MODIS cloud fraction. The results are...

Aerosols are clearly critical to cloud processes (). However, we simply do not have sufficient large scale aerosol observational data to rule with certainty here. Very coarsely, on mean monthly timescales increased aerosols optical depth means increased share of class 30 and 35. This is lie Daily aerosol data observational data from satellites is only possible in cloud-free sky—which clearly biases our sample in an irreconcilable way. Some approaches used modeled aerosol data. More research is needed.

Left over unexplained variance is like due to small scale processes, not reliable represented in ERA5 (and aerosols). While we cannot quantity test for these small scale processes with our framework, the different textures of cloud classes imply differentiating small scale processes. Colder surface, unstable stratocumulus (e.g. class 25, 27) show more distinctive punched-out patterns. Whereas warm weather stable broken-up stratocumulus (36, 40) have a more homogeneous, softer texture which might imply a more homogeneous small scale calculation pattern rather than stochastic convection/precipitation.

Precipitation is in general very low in these regions. GPM IMERG microwave satellites may not be able to measure very low levels of drizzle found in these regions. GPM IMERG does report daily precipitation values below 0.1 mm/day, which is well below the drizzle-proxy levels reported in previous open cell stratocumulus studies. A rigours evaluation of GPM IMERG over the subtropical ocean is not available. Rain gauge comparisons in other parts of the world have found an *overestimation* of the low end of the precipitation distribution (). Previous work has employed a MODIS drizzle proxy, which is a deterministic function of liquid water path and droplet size. Others have employed CALIPSO satellite measurements of rain water path, which would greatly reduce our sample

size and may show the same missing drizzle as GPM IMERG (28). Nevertheless, those findings (**?**) show open cells are more rainy than closed cells.

Several limitations with the Lagrangian tracking scheme are also identified. MODIS sampling in this region, coupled with daytime-only observations, mean that sometimes 24 hours have passed until we get a return observation. In this time, many unseen transitions may have take place in between the two observations. Geostationary satellite images with higher temporal sampling images should improve short term transitions analysis. A second issue is the idea that clouds always move with the prevailing winds. We do observe some apparent transitions 'up wind' (see Figure 4, bottom row).

Expected climate change will modify the temperature structure of the lower atmosphere in the subtropics. Two well known phenomena will be sea surface warming and poleward expansion of the Hadley cell () However, ENSO and Walker circulation changes are less well-known () and both have impacts on cloud types in these regions. Observational research carefully coupled with model circulation changes should be able to inform on future cloud response to warming where model simulated clouds themselves are lacking.

Materials and Methods

Data Availability. All data used in this analysis is publicly available from the following sources. ERA5 data is available at Copernicus. IMERG data is available from NASA. Self-supervised cloud classes and MODIS cloud properties are available at globus-labs. Codebase for analysis is available on Github. 1.4 km climate model data shown for motivation only and is available upon request from the corresponding author.

Self-Supervised cloud classification. AICCA employs a rotationally-invariant autoencoder to classify clouds in 2000-2022 MODIS satellite images over the global oceans from both TERRA and AQUA
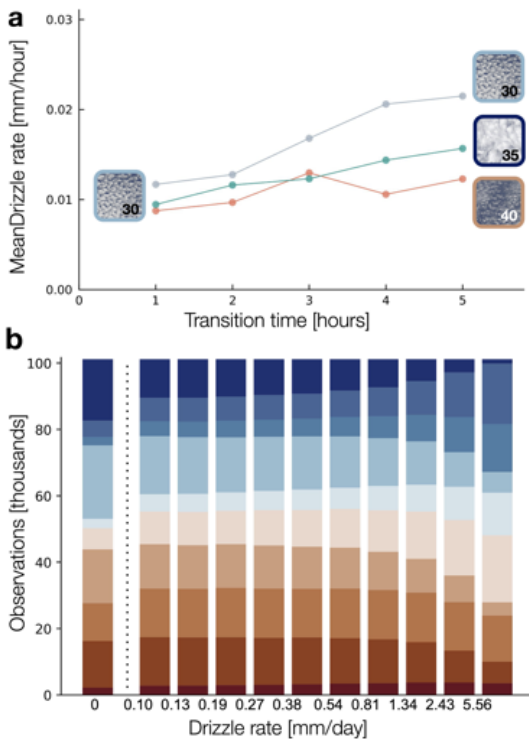
Fig. 5. Population level precipitation impacts. Drizzle does not aid in class prediction. [Top] Short timescale transitions (<6h) from the most populous class # are tracked in the south pacific basin. 11,000 such transitions are observed via a Lagrangian tracking scheme between MODIS passes. Mean precipitation rate (including zeros) is reported for three categories of transition: from 30 to an open cell class (#40, #25, #26, etc.) , from 30 to thick, unbroken low cloud (#35), and from 30 to itself (null case, no transition observed). Contrary to expectation, we find no significant effect of precipitation on transitions from the most populous closed cell class. In fact, the null cases shows the strongest (albeit weak) precipitation signal. Prevailing winds advect parcels equator ward and to the west up a mean precipitation gradient which results in slight mean increase in precipitation with time. [bottom] Stacked Bar chart shows class prevalence equally divided by ∼100,000 sample precipitation bins across the 10 descriptive low cloud classes. **X-axis is not uniformly spaced.** The no precipitation bin contains 10x as many samples as all amounts of rain combined in reality, but is shown scaled for comparison. At the highest pr values, class 35 share decreases and is replaced by 30, but no overall shift to the 'red' classes is found. Example 10 low cloud classes are more or less evenly distributed across the precipitation sample.

(16). The autoencoder is trained without explicit supervision on 7 near-IR and IR bands

Hierarchical agglomerative clustering is applied to the latent space of the autoencoder. This process generates and information-maximizing 42 cloud classes which are separable, stable, and physically reasonable (17, 29). See supplemental figure **??** and **(author?)** (17, 29) for methodological details.

An autoencoder is trained on 1 million sample 128×128 pixels patches with 6 infrared radiance bands from both the TERRA and AQUA instrument. Hierarchical agglomerative clustering is then applied to the latent vectors from the autoencoder (See supplemental Figure 1 and (16, 17) for details) generating an

informationally-maximizing 42 cloud classes. Inference is applied to the remaining observations with the trained model and the resultant AI Cloud Classification Atlas [AICCA] consists of 200M labeled cloud type observations at approximately 1 degree latitude and longitude resolution over the global oceans.

Meteorological mapping. In this analysis, we focus on the three main subtropical stratocumulus regions (**?**). We look at the top 15 most populous clouds classes in these subtropical regions which are, in order: [30, 35, 40, 36, 26, 27, 39, 32, 37, 19, 29, 25, 33, 41]. The class number is assigned based on cloud top pressure: smaller numbers represent high-altitude clouds and higher numbers denote low-altitude clouds.

Observed AICCA classes are then mapped against reanalysis meteorology variables from the ERA5 reanalysis (30). Variables are taken from the hourly product at pressure levels and averaged to daily means and include temperature, relative humidity, divergence, geopotential, potential vorticity, u, v, w components of wind at different pressure levels between 950 and 700 hpa as well as surface quantities including: sea surface temperature, mean sea level pressure, significant wave height, some engineered features including: lower tropospheric stability, estimated inversion strength (EIS), and some calculated wind shears (see supplemental table). Estimated inversion strength is calculated based on (**?**). Finally we also include daily and hourly precipitation from the IMERG composite satellite product ().

Meteorological drivers of low cloud texture are tested in the following way. We train a logistic regression model across all 67 meteorological and for pairs of 5000 randomly-selected samples from each class in the top 15 where every combination of classes is tested. Pairs of meteorological driver are scored based on their mean prediction skill across all pairs of cloud classes. Cloud free conditions are ignored in this framework.

Lagrangian tracking scheme. MODIS class observations are tracked in space in a Lagrangian scheme. Air parcels are assumed to move with the winds at 925 hPa, following (18). At each hourly timestep, each MODIS observation is advected in space by the rate of the mean hourly wind speed in that location for one hour. The process is repeated until the new location and timestamp line up in the same gridcell with another MODIS observation or until 24 hours have passed, at that point the trajectory is abandoned. Transitions can occur in 1 hour and all the way up to 24 hours depending on location and the pass times of TERRA and AQUA. Total precipitation from GPM IMERG along the trajectory is tracked.

1. GL Stephens, Cloud Feedbacks in the Climate System: A Critical Review. *J. Clim.* **18**, 237–273 (2005).

2. M Zhao, et al., Uncertainty in Model Climate Sensitivity Traced to Representations of Cumulus Precipitation Microphysics. *J. Clim.* **29**, 543–560 (2016).

3. SC Sherwood, et al., An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence. *Rev. Geophys.* **58** (2020).

4. MD Zelinka, et al., Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophys. Res. Lett.* **47** (2020).

5. T Schneider, CM Kaul, KG Pressel, Possible climate transitions from breakup of stratocumulus decks under greenhouse warming. *Nat. Geosci.* **12**, 163–167 (2019).

6. T Schneider, CM Kaul, KG Pressel, Solar geoengineering may not prevent strong warming from direct effects of CO2 on stratocumulus cloud cover. *Proc. Natl. Acad. Sci.* p. 202003730 (2020).

7. G Bellon, O Geoffroy, Stratocumulus radiative effect, multiple equilibria of the well-mixed boundary layer and transition to shallow convection. *Q. J. Royal Meteorol. Soc.* **142**, 1685–1696 (2016).

8. P Kuma, FAM Bender, A Schuddeboom, AJ McDonald, Ø Seland, Machine learning of cloud types shows higher climate sensitivity is associated with lower cloud biases. *Atmospheric Chem. Phys.* (2022).

9. L Bock, A Lauer, Cloud properties and their projected changes in CMIP models with low/medium/high climate sensitivity, (Clouds and Precipitation/Atmospheric Modelling and Data Analysis/Troposphere/Physics (physical properties and processes)), preprint (2023).

10. B Stevens, et al., DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Prog. Earth Planet. Sci.* **6**, 61 (2019).

11. NP Wedi, et al., A Baseline for Global Weather and Climate Simulations at 1 km Resolution. *J. Adv. Model. Earth Syst.* **12** (2020).

12. T Schneider, et al., Harnessing AI and computing to advance climate modelling and prediction. *Nat. Clim. Chang.* **13**, 887–889 (2023).

13. T Schneider, et al., Climate goals and computing the future of clouds. *Nat. Clim. Chang.* **7**, 3–5 (2017).

14. WB Rossow, RA Schiffer, Advances in Understanding Clouds from ISCCP. *Bull. Am. Meteorol. Soc.* **80**, 2261–2287 (1999).

15. AH Young, KR Knapp, A Inamdar, W Hankins, WB Rossow, The International Satellite Cloud Climatology Project H-Series climate data record product. (2018).

16. T Kurihana, E Moyer, R Willett, D Gilton, I Foster, Data-driven cloud clustering via a rotationally invariant autoencoder. *IEEE Transactions on Geosci. Remote. Sens.* **60**, 1–25 (2021).

17. T Kurihana, EJ Moyer, IT Foster, Aicca: Ai-driven cloud classification atlas. *Remote. Sens.* **14**, 5690 (2022).

18. R Eastman, R Wood, Factors controlling low-cloud evolution over the eastern subtropical oceans: A Lagrangian perspective using the A-Train satellites. *J. Atmospheric Sci.* **73**, 331–351 (2016).

19. D Randall, Conditional Stability of the First Kind Upside-Down. *Am. Meteorol. Soc.* **37**, 125–130 (1980).

20. EM Agee, TS Chen, KE Dowell, A Review of Mesoscale Cellular Convection. *Bull. Am. Meteorol. Soc.* **54**, 1004–1012 (1973).

21. B Stevens, et al., POCKETS OF OPEN CELLS AND DRIZZLE IN MARINE STRATOCUMULUS. *Bull. Am. Meteorol. Soc.* **86**, 51–58 (2005).

22. TM Sharon, et al., Aerosol and Cloud Microphysical Characteristics of Rifts and Gradients in Maritime Stratocumulus Clouds. *J. Atmospheric Sci.* **63**, 983–997 (2006).

23. R Wood, et al., Open cellular structure in marine stratocumulus sheets. *J. Geophys. Res.* **113**, D12207 (2008).

24. G Feingold, I Koren, H Wang, H Xue, WA Brewer, Precipitation-generated oscillations in open cellular cloud fields. *Nature* **466**, 849–852 (2010).

25. D Watson-Parris, SA Sutherland, MW Christensen, R Eastman, P Stier, A Large-Scale Analysis of Pockets of Open Cells and Their Radiative Impact. *Geophys. Res. Lett.* **48** (2021).

26. S Klein, D Hartmann, The Seasonal Cycle of Low Statiform Clouds. *J. Clim.* **6**, 1586–1606 (1993).

27. HS Baker, C Mbengue, T Woollings, Seasonal Sensitivity of the Hadley Cell and Cross-Hemispheric Responses to Diabatic Heating in an Idealized GCM. *Geophys. Res. Lett.* **45**, 2533–2541 (2018).

28. ESDBENJT Huffman, G.J., GPP IMERG late precipitation l3 1 day 0.1 degree x 0.1 degree v06 (2019) Accessed: August 1, 2023.

29. T Kurihana, J Franke, I Foster, Z Wang, E Moyer, Insight into cloud processes from unsupervised classification with a rotationally invariant autoencoder. *arXiv preprint arXiv:2211.00860* (2022).

30. Hersbach, et al., The ERA5 global reanalysis. *Q. J. Royal Meteorol. Soc.* **146**, 1999–2049 (2020).