Supplemental Material for "Urbanization causes increased cloud-base height and decreased fog in coastal southern California"

Description of data

Airfield stratus frequency records

Cloud observations at airfields are generally hourly. Overhead cloud-base height is recorded when clouds are present to obscure at least 5/8 of the sky. Although cloud-base height data do exist for some stations prior to 1948, we begin our analysis in 1948 because only a few airfields have data prior to 1948 and documented inconsistencies in observational methods prior to that year have been found to cause inaccuracies in cloud-frequency records [*Karl and Steurer*, 1990]. In the early-mid 1990s, inconsistencies were again introduced at many airfields in the United States when the National Weather Service (NWS) and Federal Aviation Administration switched from human to automated observations of cloud cover, but this switch does not appear to have caused inhomogeneity in records of low stratus clouds [*Dai et al.*, 2006], which are the subject of this study. Further, the meteorological stations at 8 of the 24 airfields considered in this study are maintained by the United States Military, which has not switched to automated cloud monitoring [*Dai et al.*, 2006]. We see no evidence of methods-related shifts in the 1990s among the non-military airfields in our dataset.

For each hour, we considered the cloud-base height observation closest to the top of the hour and within 15 minutes of the top of the hour. If there were two observations equally spaced from the top of the hour (e.g., 06:55 and 07:05), we considered the observation with the lower cloud-base height. We filled single-hour data gaps in the hourly cloud-base height observations based on the cloud-base height observations for the neighboring hours. As a first step, we empirically determined the probability that the missing observation was of stratus cloud (cloud base $\leq$ 1000 m), non-stratus cloud (cloud base > 1000 m), or clear sky. We determined these probabilities based on a sampling of all other instances from the airfield's record when cloud-height data existed for the same three hours of day, the Julian day was within 10 days of the Julian day with the missing value, and both neighboring observations matched the missing hour's neighboring observations in terms of whether those observations were of stratus, non-stratus, or clear sky. After empirically calculating the probability that the missing hour fell into each of the three observation classes, we reclassified the missing hour as one of the three cloud classes based on a probability-weighted random number. If the assigned class was stratus cloud or non-stratus cloud, we next estimated the height of the cloud base from the empirical linear relationship between the cloud-base heights of the neighboring hours and the cloud-base heights for the hour of interest. Occurrences of single-hour gap fills were rare, averaging ~1.6% among the 24 airports, with a maximum of 5.05% for Gillespie Field and a minimum of 0.04% for Los Angeles International Airport and Lindbergh Field.

Even though we ultimately consider cloud-height observations for four hours (07:00, 10:00, 13:00, 16:00 Pacific Standard Time), we occasionally filled missing monthly cloud-frequency values for these hours using data from other hours. We initially calculated monthly cloud frequencies for each of the 24 hours, giving 24 annual records of stratus frequency (one for each hour) for a given month at a given airfield. For a given month, if annual value x was missing from one of the four key hours listed above, we searched the other 23 hours for annual records that (1) contained annual value x, (2) had at least 20 years of overlapping data with the annual

record with missing value, and (3) correlated well ($r > 0.7$) with the annual record with the missing value. If any of the 23 alternate records fit these three criteria, the missing value x was replaced with a linearly adjusted value from the record that correlated most strongly with the record of interest. This was rarely necessary, averaging <1% of hours across all airfields and was never necessary at 13 airfields. This gap filling was most common (6.8% of hours) at San Nicolas Island (NSI) due to often missing (~25%) values at 16:00, which were nearly always filled using the record from 15:00. We next averaged across the five months for each hour so we had an annual record for each of the 24 hours representing stratus frequency during all of May–September. We repeated the gap-filling step described immediately above for each of the four hours of interest, which was only necessary for 2 airfields, and only for one year at each airfield. We averaged across the four hours of interest to calculate a single annual record representing May–September.

Stratus-cloud base height records generally have a vertical resolution of 30.48 m (100 ft) [*NWS*, 1998]. To account for uncertainty that this introduced to cloud-base height measurements, we produced 1000 alternate hourly records of cloud-base height at each airfield, where each hourly record was perturbed by adding a random value selected from a uniform distribution in the range of ±15.24 m (50 ft). We then calculated an alternate annual record of May–September stratus frequency for each airfield from each of the 1000 alternate hourly records. We calculated each airfield's May–September stratus frequency record as the average of the 1000 alternate records. In correlation and trend analyses, we conducted these analyses for each of the 1000 alternate records. All significance values and confidence bounds shown in the article are representative of at least 95% of the alternate records such that they reflect uncertainty caused by the limited vertical resolution of cloud-height measurements.

We repeated the above methods to develop airfield-specific records for lower stratus clouds, where cloud-base heights are considered within various vertical slices between the surface and 1000 m. We also repeated the above methods to develop records of average stratus-cloud base height for each airfield.

Sub-regional stratus frequency records

We standardized all annual time series within a sub-region to have a mean of zero and standard deviation of one during the common period of 1973 through 2014. We averaged the 1000 sets of standardized time series across airfields within each sub-region, producing 1000 alternate stratus-frequency records for each sub-region. The spread among the 1000 alternate sub-regional records represents error caused by limited vertical resolution of the cloud-base height measurements (E_r).

We next accounted for error in each sub-regional record caused by missing airfield data (E_m). For years when a sub-region has a missing value for at least one airfield, uncertainty in the sub-regional mean is introduced. For each year with at least one missing airfield value in a sub-region, we identified the airfields with valid values and evaluated the representativeness of that subset by regressing the true regional mean time series against an alternate sub-regional time series representing only the sub-set of airfields. Greater uncertainty is indicated by greater spread among residuals. For each sub-region and year with a missing airfield stratus-frequency value, we calculated the standard deviation among residuals for each of the 1000 alternate annual

records, found the average of these values, and considered the 95% error bound to be within ±2.042 standard deviations of the sub-regional stratus-frequency value (2.042 from the Student's t-distribution for 30 degrees of freedom for $p = 0.05$). So, if a given sub-region's standard deviation of the residuals for a year was 0.1, then the error for that year is 0.2042 standardized stratus-frequency units. The 1000 records of annual E_m for each sub-region were added to the 1000 records of that sub-region's E_r to represent error caused by missing airfield data and error due to limited vertical resolution of cloud-base height measurements. Finally, a sub-regional mean record was calculated as the average of the 1000 alternate sub-regional records, including both error terms.

For evaluation of sub-regional trends, we converted standardized sub-regional values back into absolute stratus-frequency values by adding back in the mean and variance of the raw airfield-specific stratus-frequency records within each sub-region.

Regional stratus frequency record for correlation analyses

We calculated the coastal southern California (CSCA) regional record as the average of the sub-regional records for the three mainland sub-regions: Santa Barbara, Los Angeles, and San Diego. For our analyses of correlation between stratus frequency and large-scale climate variability, we removed linear trends from stratus-frequency and climate records to ensure that correlations are not confounded by common trends that are potentially unrelated mechanistically. We first detrended individual airfield stratus-frequency records during 1960–2014, as this is the period of overlap with the San Diego-Miramar radiosonde record. To do this, we detrended each of the 1000 alternate records for each airfield (to preserve E_r) and then averaged across airfields within a sub-region to get 1000 detrended sub-regional averages. We then added the 1000 records of E_m for each sub-region to the 1000 detrended records to incorporate error in the sub-regional mean due to missing airfield data. We detrended each of the 1000 records one more time to remove trends caused by addition of E_m. Then we averaged the 1000 sets of detrended records across the three mainland sub-regions to develop 1000 detrended records of CSCA regionally averaged stratus frequency during 1960–2014. This method was repeated to also produce detrended regional records for 1979–2014 for correlation analysis with the reanalysis climate dataset that begins in 1979. This procedure was repeated to represent stratus frequency within various altitude classes. The procedure was also repeated to calculate a regional record of stratus-cloud base height variability during 1960–2014, which was used as a proxy for variability in the altitude of condensation.

Surface temperature records

Monthly means of daily minimum and maximum temperature (T_{min} and T_{max}, respectively) grids for 1948–2014 were accessed from the latest version of the Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset (PRISM group, Oregon State University, prism.nacse.org [*Daly et al.*, 2004]). PRISM grids have ~4 km geographic resolution. Temperature records were averaged across the 25 PRISM grid cells (5x5) centered over each airfield because this area roughly represents the area within a ~10 km radius of each airfield (same area for which urban cover is considered). The larger area around each airfield is considered in the temperature analysis because, although the cloud-height data represent cloud conditions immediately above the airfields, cloud conditions above the airfields are presumably influenced by the meteorological conditions of the greater surrounding area. The PRISM dataset

is developed using temperature data from the airfields in this study and many more weather stations in the region, and has been carefully quality controlled to remove erroneous data and correct for temporal inconsistencies. Because the PRISM temperature data for the two Islands airfields is not representative of the greater region surrounding the islands (ocean), we used a combination of PRISM temperature data and reanalysis temperature data for these airfields. For 1979–2014 reanalysis temperature data, we used NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA) dataset [*Rienecker et al.*, 2011]. For 1948–1978 we used the reanalysis dataset developed by the National Center for Environmental Protection/National Center for Atmospheric Research (NCEP/NCAR) [*Kalnay et al.*, 1996], accessed from the Earth System Research Laboratory (ESRL) Physical Sciences Division (PSD) at www.esrl.noaa.gov/psd. For each Islands airfield, we extracted May–September mean temperature data using bilinear interpolation and we then calibrated the NCEP/NCAR records to match the mean and variance of the MERRA records during the overlapping period of 1979–2014. For each 5x5 set of PRISM grid cells centered over each Islands airfield, we replaced the outer 16 PRISM temperatures with reanalysis temperatures.

Although the monthly PRISM T_{min} and T_{max} records are not accompanied by estimates of uncertainty, we can account for measurement uncertainty of the temperature records upon which PRISM is based. According to *NWS* [1998], Automated Surface Observing System temperature sensors have a root mean squared error of 0.5°C and a maximum error of ±1°C. We conservatively assumed that this error is uniformly distributed within ±1°C. Simulating 10^6 153-day seasons (length of May–September) with a daily uniformly distributed error of ±1.0°C, the resultant errors in the seasonal mean value are normally distributed and have a standard deviation of 0.047°C. In all analyses that incorporate T_{min} and T_{max} trends, we consider 1000 alternate records of seasonally averaged T_{min} and T_{max}, each with errors added to the annual values that are drawn from a normal distribution with a mean of zero and a standard deviation of 0.047°C.

Reanalysis climate data

All correlation analyses reported in the article that involve gridded reanalysis climate data are conducted using NASA's MERRA dataset. MERRA is a three-dimensional reanalysis product with geographic resolution ranging from 0.5° to 1.25°, vertical resolution of 25 hPa from the surface to 700 hPa, vertical resolution of 50 hPa between 700 and 100 hPa, and temporal coverage of 1979–present. We utilized monthly and daily MERRA data for our analyses. Among the variables evaluated was the lower-tropospheric stability (LTS), calculated from daily data as potential air temperature (θ) at 850 hPa (θ_{850}) minus potential air temperature at the near-surface (θ_{2m}), following *Iacobellis et al.*, [2009] and *Iacobellis and Cayan* [2013]. Potential temperature is the theoretical temperature of an air parcel if forced to a standard reference pressure P_0, in this case 1000 hPa. Potential temperature is calculated as $\theta = T(P_0/P)^{\kappa}$ [*Bolton*, 1980], where T is air temperature of the parcel in degrees Kelvin, P is air pressure of the parcel in hPa, and κ is the Poisson constant, [$\kappa = 0.2854(1-0.24r_v)$], where r_v is the water vapor mixing ratio. LTS has also been calculated previously using θ at 700 hPa rather than 850 hPa [e.g., *Wood and Bretherton*, 2006]. We used both approaches in our analyses but only report those using θ_{850} because LTS calculated with θ_{850} correlated more strongly with CSCA stratus frequency and inversion strength.

In addition, sea-surface temperature (SST) data were accessed for Figure S6a. The National Ocean and Atmospheric Administration (NOAA) extended SST version 3 dataset [*Smith et al.*, 2008] has temporal coverage of 1854–2014, geographic resolution of 2°, and was accessed from the NOAA ESRL PSD (www.esrl.noaa.gov/psd/data/gridded/data.noaa.ersst.html). Additionally, Figure S6a includes 2 m air temperature data from the European Centre for Midrange Weather Forecasting (ECMWF), accessed from http://apps.ecmwf.int/datasets/. For temporal completeness, we merged the ECMWF-ERA 40 product (which covers 1958–2001) with the ECMWF-Interim product (1979–2013), both available with 0.5° geographic resolution. We combined these products by calculating the annual time series shown in Figure S6a for each product and calibrating the ERA 40 record to have the same mean and variance as the Interim product during the overlapping period of 1979–2001.

Radiosonde data and temperature-inversion calculations

The primary radiosonde record used was from San Diego-Miramar (NKX). Radiosonde data were accessed from www.esrl.noaa.gov/raobs/ and quality controlled as part of a previous research campaign [*Iacobellis et al.*, 2009]. As in *Iacobellis and Cayan* [2013] and *Iacobellis et al.* [2009], we do not consider radiosonde data prior to 1960 due to data uncertainties and irregular launch times. Launch times considered in this study occur daily at 04:00 and 16:00 Pacific Standard Time.

For each measurement considered, we evaluated the temperature profile for the presence of a temperature inversion. As in *Iacobellis and Cayan* [2013] and *Iacobellis et al.* [2009], we only consider "subsidence inversions" and exclude "radiation inversions" in our investigation of the connection between inversion characteristics and stratus occurrence. Radiation inversions begin warming with altitude immediately from the surface, which should generally inhibit stratus cloud formation, but subsidence inversions occur above the surface and therefore cap a cooler marine boundary layer (MBL). We only consider subsidence inversions with an inversion top below 700 hPa, as inversions higher than this are not related to the MBL. In cases when a subsidence inversion was present, we recorded the altitude and temperature of the top and bottom of the inversion. Inversion strength was simply defined as top temperature minus bottom temperature. In the absence of a subsidence inversion, an inversion strength of 0°C is assigned. If there was a radiation inversion and a subsidence inversion present, the subsidence inversion was still considered. If there were multiple smaller inversions within a larger inversion, the larger inversion was the only one considered.

For each inversion variable (inversion strength, top height, top temperature, bottom height, bottom temperature) as well as the percent of measurements when the inversion base was ≥800 m above sea level, we developed an annual time series for each of the two launch times considered, averaging across all days within May–September each year to calculate an annual value. The two annual records (for the two launch times) for each variable were then standardized and averaged together to create one record. For the evaluations of interannual correlation with stratus frequency, we detrended radiosonde records for each launch time before standardizing and averaging.

In addition to measurement of inversion characteristics at NKX, we evaluated 850 hPa air temperature (T_{850}) records at NKX, San Nicolas Island (NSI), and Point Mugu Naval Field

(NTD) (Fig. S6b). While measurements at NSI and NTD were far too inconsistent in terms of timing and vertical resolution for long term monitoring of inversion characteristics at these airfields, the T_{850} measurement is a standard radiosonde measurement that has been made consistently at these airfields. Additionally, T_{850} does not experience much diurnal variability at these airfields, allowing for use of data from alternate launch times to represent T_{850}. For each airfield, we developed an annual record of T_{850} for the early morning, considering the range of hours from 01:00 to 07:00, and the late afternoon, considering the hours from 13:00 to 19:00. In cases when two measurements were made within one of these time windows on the same day, we only considered the measurement made closer to the standard 04:00 or 16:00 launch time. In the annual averaging, we calculated an annually averaged value for each time window as long as there were at least 5 days per month during May–September with data. While this is a small minimum required sample size, the strong agreement among the records in Figure S6b suggests that not much accuracy was compromised. In cases when an annual value was available for one time window but not both, we linearly adjusted the available value to replace the missing value. Finally, we averaged the annual records for the two time windows together to develop one record of annual May–September T_{850} for each airfield.

For all radiosonde-based records, we recalculated 1000 alternate records in an attempt to account for measurement uncertainties in temperature and geopotential height. Radiosonde equipment has evolved over the years and documentation of equipment changes and how these changes affected measurement uncertainties in the lower troposphere is limited. We used the limited information available to develop a conservative representation of measurement error for the study period (1960–2014). According to Appendix C of the Federal Meteorological Handbook No. 3: Rawinsonde and Pibal Observations (http://www.ofcm.gov/fmh3/pdf/11-app-c.pdf), NKX uses a VIZ-B2 radiosonde made by Sippican, Inc. This is a relatively old model that most radiosonde deployment sites have migrated away from, but information appears to be not readily available as to when this model came about and what other models may have been used during our study period. According the Sippican website, both versions of the VIS-B2 radiosondes have a measurement uncertainty with a standard deviation (σ) of 0.2°C (http://www.sippican.com). We cannot be sure how measurement uncertainty has varied in the past so we conservatively assume a long-term uncertainty four times larger than the modern uncertainty (σ = 0.4°C [variance = 0.16 when σ = 0.4 and variance = 0.04 when σ = 0.2]). The Sippican documentation does not provide estimates of the uncertainties for altitude, but the WMO Report on Instruments and Observing Methods (www.wmo.int/pages/prog/www/IMOP/publications/IOM-80/CatalogRadiosond.pdf) indicates in Table 10 that the Sippican radiosonde had an uncertainty of σ = 4m at 500 hPa. Uncertainties in geopotential height are accumulated with altitude so the geopotential height uncertainty should be less within the pressure range that we were interested in (generally >850 hPa). The estimate of σ = 4m is based on soundings from recent years (post-1998), but an analysis of radiosonde data from 1975 indicates similar uncertainties in geopotential height (Table 6 in Schwartz and Govett, 1992). In that analysis, distributions of geopotential height errors within various pressure levels indicate that σ = 4.8m between the surface and 850 hPa and σ = 3.6m between 850 and 700 hPa. We conservatively assume an uncertainty of σ = 5m in our analyses. We incorporate the uncertainties in temperature and geopotential height by creating 1000 alternate sets of raw radiosonde measurements, converting these into 1000 alternate sets of inversion strength and inversion base height, and then aggregating these to 1000 alternate records of May–September

averages. We also calculate 1000 alternate records of May–September mean temperature at various pressure levels within the atmospheric profile for the analysis of how stratus frequency relates to temperature throughout the profile.

Land cover data

We evaluated changes in urban land area within 10 km of each airfield by relating a modern land cover dataset to a historic dataset. For the modern case, we used the 2011 National Land Cover Database (NLCD) [*Jin et al.*, 2013]. This dataset has 30 m geographic resolution across the continental United States and classifies land-cover type for each grid cell. We considered all areas classified by the NLCD as "Developed" to be urban area. Notably, the 10 km radius surrounding the airfield at Imperial Beach (NRS) includes part of Mexico, which is not included in the NLCD. For this region, we used the 500 m MODIS v5.1 global land cover dataset for 2012 [*Friedl et al.*, 2010]. For the historic case, we followed the methods of *Syphard et al.* [2011] and based our estimates of urban cover on a spatially continuous estimate of national household density, based on census data from 1950 [*Hammer et al.*, 2004]. Like *Syphard et al.*, we assigned an urban classification to any location estimated by *Hammer et al.* to have a housing density of at least 128 housing units per km^2. We used nearest neighbor resampling to transfer the shapefile data from *Hammer et al.* to the 30 m NLCD grid. We have reasonably high confidence in the ability of the *Hammer et al.* dataset to represent CSCA urban area because *Syphard et al.* found strong correspondence between the 2001 NLCD urban cover and urban cover derived from the *Hammer et al.* dataset for year 2000 in San Diego County. In this region, the total fractional urban cover was within 0.01 for the two datasets and urban/non-urban classifications were in agreement for over 90% of the 30 m grid cells. The total fractional urban cover derived from the *Hammer et al.* dataset was slightly high, likely because the lower spatial resolution of the census data caused some green spaces within urban areas to be classified as urban. We partially address this by not allowing areas to be classified as urban in the 1950 dataset if they were not classified as urban in either of the NLCD datasets from 1992 [*Vogelmann et al.*, 2001] or 2011.

The *Hammer et al.* dataset also does not cover Mexico and we know of no mid-1900s urban cover dataset for Mexico. We therefore georeferenced and digitized a United States Air Force Map of the greater San Diego area in 1950 that outlines urban areas and includes northern Baja California, Mexico (http://historicalcharts.noaa.gov/jpgs/0404-0004-250-4-1950.jpg). According to this map, northern Baja California was largely unurbanized in 1950; 3.07 km^2 of the urbanized portion of Tijuana fell within the 10 km radius surrounding NRS, accounting for <1% of the area within this radius (314.15 km^2) and 10.1% of the total 1950 urbanized area. The *Hammer et al.* dataset also does not contain data for the Islands sub-region. We assumed 0% urban cover for San Nicolas in 1950 because this site was mainly developed after 1950. For San Clemente, we assume a 1950 urban fraction of one half of the urban fraction in 2011 (0.014). Urban coverage at these sites is so small that assumptions of 1950s urban cover contribute only minimal uncertainty to the absolute change in urban coverage at these sites.

In accounting for uncertainty, we assume all uncertainties in our calculation of change in urban cover during 1950–2011 to be due to errors in the census-based estimates of urban cover in 1950. There is certainly error in the NLCD as well, but NLCD error is undoubtedly minimal compared to the error in the census-based estimates and a formal accounting of NLCD urban-cover accuracy has thus far not been published. There has also been no formal accounting of the

uncertainties in census-based urban estimates for CSCA, but this is possible under the assumption that NLCD is truth [e.g., *Syphard et al.*, 2011]. Like *Syphard et al*., we quantified uncertainty by comparing census-based estimates in urban density throughout CSCA in 2000 to those calculated from the 2001 NLCD [*Homer et al.*, 2007]. We gridded CSCA to a resolution of 17.73 km^2 such that grid cells have an area approximately equal to the circles with 10 km radii surrounding each airfield. Considering the 60 grid cells within 50 km of the coast that are composed of at least 50% mainland area, the estimates of the fraction of urban land area made from the two datasets agree very well (r = 0.98). Regressing census-based urban area against NLCD-based urban area, the residuals have a normal distribution with a root mean squared error (RMSE) of 0.0592 and a maximum absolute error of 0.2417. We produced 1000 alternate estimates of 1950 urban land fraction for each mainland airfield where each census-based estimate was permuted by adding a random number selected from a normal distribution with a mean of zero, a standard deviation of 0.0592, and a maximum possible absolute value of 0.25. We converted these urban land fraction values to total urban fraction (including ocean) by multiplying by total land (urban and non-urban) fraction.

Modeled climate projections

For Figure S6c, we utilized modeled monthly climate projections developed as part of the fifth phase of the Coupled Model Intercomparison Project (CMIP5) and assessed for the Intergovernmental Panel on Climate Change Fifth Assessment Report (IPCC AR5). Data were accessed from the CMIP5 Program for Climate Model Diagnostics and Intercomparison (PCMDI) at http://cmip-pcmdi.llnl.gov/cmip5/. We used CMIP5 data from the historical scenario for years 1900–2005 and the representative concentration pathway (RCP) 8.5 scenario for 2006–2100. RCP 8.5 is an emissions scenario in which anthropogenic radiative forcing reaches 8.5 Wm^{-2} by 2100 [*Moss et al.*, 2010; *van Vuuren et al.*, 2011]. In our CMIP5 analysis, we evaluated LTS, 700 hPa specific humidity, and 700 hPa vertical velocity. For each variable, we considered all models for which the necessary data were available for both the historical and RCP 8.5 scenarios. Because models have differing geographic resolutions, we bilinearly interpolated all model projections to a common grid of 1°. We then calculated annual time series for each variable of interest by averaging across May–September monthly values within the geographic box of interest (15–35°N, 110–120°W). We converted annual time series for each model as anomalies relative to that model's mean value during 1960–2005 to normalize for model biases in mean climatological conditions. For each model, we averaged across all available model runs, as a given model was often run multiple times for the same climate scenario, but with differing initial conditions. We express the projections as the multi-model median and interquartiles. Averaging across all available models and model runs minimizes the effect of internal climate variability on the projections shown in Figure S6c and preserves modeled variability that may be attributed to changes in anthropogenic radiative forcing.

References

Bolton, D. (1980), The computation of equivalent potential temperature, *Monthly Weather Review*, *108*(7), 1046-1053, doi:10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO%3B2.

Dai, A., T. R. Karl, B. Sun, and K. E. Trenberth (2006), Recent trends in cloudiness over the United States: A tale of monitoring inadequacies, *Bulletin of the American Meteorological Society*, *87*(5), 597-606, doi:10.1175/BAMS-87-5-597.

Daly, C., W. P. Gibson, M. Dogget, J. Smith, and G. Taylor (2004), Up-to-date monthly climate maps for the coterminous United States, paper presented at Proceedings of the 14th AMS Conference on Applied Climatology, 84th AMS Annual Meeting, American Meteorological Society, Seattle, WA, January 13-16, 2004.

Friedl, M. A., D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, and X.-Y. Huang (2010), MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets, *Remote Sensing of Environment*, *114*(1), 168-182, doi:10.1016/j.rse.2009.08.016.

Hammer, R. B., S. I. Stewart, R. L. Winkler, V. C. Radeloff, and P. R. Voss (2004), Characterizing dynamic spatial and temporal residential density patterns from 1940–1990 across the North Central United States, *Landscape and Urban Planning*, *69*(2), 183-199, doi:10.1016/j.landurbplan.2003.08.011.

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. VanDriel, and J. Wickham (2007), Completion of the 2001 national land cover database for the counterminous United States, *Photogrammetric Engineering and Remote Sensing*, *73*(4), 337-341.

Iacobellis, S. F., and D. R. Cayan (2013), The variability of California summertime marine stratus: Impacts on surface air temperatures, *Journal of Geophysical Research: Atmospheres*, *118*(16), 9105-9122, doi:10.1002/jgrd.50652.

Iacobellis, S. F., J. R. Norris, M. Knamitsu, M. Tyree, and D. C. Cayan (2009), Climate variability and California low-level temperature inversions, *Rep. CEC-500-2009-020-F*, California Climate Change Center, Sacramento, CA.

Jin, S., L. Yang, P. Danielson, C. Homer, J. Fry, and G. Xian (2013), A comprehensive change detection method for updating the national land cover database to circa 2011, *Remote Sensing of Environment*, *132*, 159-175, doi:10.1016/j.rse.2013.01.012.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, and J. Woollen (1996), The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, *77*(3), 437-471, doi:10.1175/1520-0477%281996%29077%3C0437%3ATNYRP%3E2.0.CO%3B2.

Karl, T. R., and P. M. Steurer (1990), Increased cloudiness in the United States during the first half of the twentieth century: Fact or fiction?, *Geophysical Research Letters*, *17*(11), 1925-1928, doi:10.1029/GL017i011p01925.

Moss, R. H., J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. van Vuuren, T. R. Carter, S. Emori, M. Kainuma, and T. Kram (2010), The next generation of scenarios for climate change research and assessment, *Nature*, *463*(7282), 747-756, doi:10.1038/nature08823.

NWS (1998), *Automated Surface Observing System (ASOS) user's guide.*, 61 pp., National Oceanic and Atmospheric Administration, Washington, DC.

Rienecker, N. M., et al. (2011), MERRA - NASA's Modern-Era Retrospective Analysis for Research and Applications, *Journal of Climate*, *24*(14), 3624-3648, doi:10.1175/jcli-d-11-00015.1.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore (2008), Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006), *Journal of Climate*, *21*(10), 2283-2296, doi:10.1175/2007JCLI2100.1.

Syphard, A. D., K. C. Clarke, J. Franklin, H. M. Regan, and M. Mcginnis (2011), Forecasts of habitat loss and fragmentation due to urban growth are sensitive to source of input data, *Journal of Environmental Management*, *92*(7), 1882-1893, doi:10.1016/j.jenvman.2011.03.014.

van Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. C. Hurtt, T. Kram, V. Krey, and J.-F. Lamarque (2011), The representative concentration pathways: an overview, *Climatic Change*, *109*(1-2), 5-31, doi:10.1007/s10584-011-0148-z.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson, B. K. Wylie, and N. Van Driel (2001), Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources, *Photogrammetric Engineering and Remote Sensing*, *67*(6), 650-662.

Wood, R., and C. S. Bretherton (2006), On the Relationship between Stratiform Low Cloud Cover and Lower-Tropospheric Stability, *Journal of Climate*, *19*(24), 6425-6432, doi:10.1175/jcli3988.1.

Supplemental tables and figures

Table S1. Airfield information

Name	Code	Lat (°N)	Lon (°W)	Elev (m)	Sub-region	N* (years)	First year
Santa Barbara Municipal Airport	SBA	34.426	119.844	6	SB	56	1949
Oxnard Airport	OXR	34.201	119.206	21	SB	58	1953
Point Mugu Naval Field	NTD	34.117	119.117	4	SB	67	1948
Van Nuys Airport	VNY	34.210	118.489	244	LA	43	1948
Santa Monica Municipal Airport	SMO	34.017	118.450	53	LA	42	1973
Los Angeles International	LAX	33.938	118.406	99	LA	67	1948
Burbank/Glendale	BUR	34.200	118.350	236	LA	63	1948
Jack Northrop Field	HHR	33.923	118.334	19	LA	42	1973
Torrance Municipal/Zamperini Field	TOA	33.800	118.333	31	LA	42	1973
Long Beach Daugherty Field	LGB	33.828	118.163	12	LA	67	1948
Los Alamitos Army Airfield	SLI	33.783	118.050	11	LA	55	1948
Fullerton Municipal Airport	FUL	33.872	117.979	29	LA	42	1973
Santa Ana John Wayne Airport	SNA	33.680	117.866	17	LA	47	1968
Ontario International Airport	ONT	34.056	117.600	289	LA	51	1950
Camp Pendleton Marine Corps Air Stn.	NFG	33.300	117.350	24	SD	47	1968
McClellan-Palomar Airport	CLD	33.128	117.279	100	SD	42	1973
North Island Naval Air Stn.	NZY	32.700	117.200	8	SD	67	1948
Lindbergh Field	SAN	32.735	117.169	9	SD	67	1948
Miramar Naval Air Stn.	NKX	32.867	117.150	146	SD	66	1948
Montgomery Field	MYF	32.816	117.140	130	SD	41	1973
Imperial Beach/Ream Field	NRS	32.567	117.117	7	SD	63	1948
Gillespie Field	SEE	32.826	116.973	118	SD	42	1973
San Nicolas Island	NSI	33.250	119.450	154	Islands	65	1948
San Clemente Island	NUC	33.023	118.588	56	Islands	52	1960

*Number of years of data indicates the number of years for which all months in the stratus season (May–September) had valid stratus-frequency values for all hours evaluated (07:00, 10:00, 13:00, 16:00). The number of years of data may be higher for individual hours and/or months. The last year with valid data is 2014 for all airfields.

Table S2. Upper height bounds for the four cloud-base height classes at each airfield. Units are meters above sea level. The four height classes represent the lowest 25%, 50%, 75%, and 100% of stratus cloud base heights at each airfield, where all clouds with base <1000 m above sea level are considered to be stratus and we consider only observations of stratus at 07:00, 10:00, 13:00, and 16:00 during May–September 1973–2014 in determining these thresholds. For each airfield, the boundary heights were calculated for each of the 1000 alternate records of cloud-base height. Values shown here are the averages of each set of 1000 values. The boundary heights among the 1000 alternate records were generally consistent. For most airfields, 95% of alternate boundary heights were within 1 to 2 m of the mean boundary height for all levels. The largest spread was for the 75% level at VNY, where the inner 95% of the 1000 alternate boundary heights were within 4.05 m of the mean.

	Height Class			
Airfield	**25%**	**50%**	**75%**	**100%**
SBA	179	287	486	1000
OXR	210	311	456	1000
NTD	186	300	448	1000
VNY	499	640	787	1000
SMO	297	408	592	1000
LAX	394	519	690	1000
BUR	500	646	816	1000
HHR	352	486	678	1000
TOA	336	480	640	1000
LGB	348	491	672	1000
SLI	326	463	623	1000
FUL	390	534	704	1000
SNA	340	471	628	1000
ONT	538	652	784	1000
NFG	313	464	639	1000
CLD	306	423	563	1000
NZY	304	422	600	1000
SAN	336	461	604	1000
NKX	380	515	714	1000
MYF	373	533	678	1000
NRS	270	374	523	1000
SEE	403	531	707	1000
NSI	255	350	482	1000
NUC	287	365	501	1000

Table S3. Observed changes in stratus frequency. Linear relativized changes in May–September stratus frequency for each sub-region and airfield. Δ in stratus frequency is calculated for all stratus clouds (cloud base ≤1000 m) in the column labeled "All" and for the four quartiles of height classes. Changes in stratus frequency are expressed relative to the initial regression value (e.g., a reduction of 0.25 from an initial stratus frequency regression value of 0.50 equals a -50% change). Asterisks indicate trends significant at 99% (**) and 95% (*) confidence levels according to both Spearman's Rho and Kendall's Tau tests. Trends are not marked as significant unless these significance criteria were met by at least 95% of the 1000 alternate time series.

			Δ (%) within each height class				
				Fog			
Name	**1st year**	**N**	**All**	**0-25%**	**25-50%**	**50-75%**	**75-100%**
Santa Barbara	**1948**	**67**	**2**	**5**	**21**	**14**	**-20**
SBA	1949	56	-4	17	18	-28*	-15
OXR	1953	58	23*	19	68*	52**	-19
NTD	1948	67	-5	-8	-7	17	-17
Los Angeles	**1948**	**67**	**-23****	**-63****	**-9**	**5**	**1**
VNY	1948	43	-46**	-65**	-47	-23	-38*
SMO	1973	42	-28**	-42	-42*	-21	1
LAX	1948	67	-10	-38*	-4	-2	18
BUR	1948	63	-22*	-49**	-5	-13	-7
HHR	1973	42	-24*	-60*	-34	8	11
TOA	1973	42	-17	-28	-36	39	-25
LGB	1948	67	-19**	-61**	1	16	4
SLI	1948	55	-22**	-62**	12	11	-11
FUL	1973	42	-37**	-58**	-19	-30	-35*
SNA	1968	47	-31**	-55**	-35	-16	-9
ONT	1950	50	-43**	-87**	-16	6	-23
San Diego	**1948**	**67**	**-9**	**-25**	**-9**	**-4**	**10**
NFG	1968	47	-22**	-15	-34	-33*	-3
CLD	1973	42	-8	-32*	-32*	-1	58**
NZY	1948	67	-6	-21	-3	7	-5
SAN	1948	67	3	-20	15	10	10
NKX	1948	66	-10	-16	-12	3	-15
MYF	1973	41	-27**	-38*	134**	-80**	5
NRS	1948	63	-12	-30	-25	1	18
SEE	1973	42	-31**	-53**	-7	0	-48**
Islands	**1948**	**67**	**-9**	**60***	**-20**	**-39***	**-1**
NSI	1948	65	-9	40*	-17	-34*	-4
NUC	1960	52	-7	12	-20	-17	2

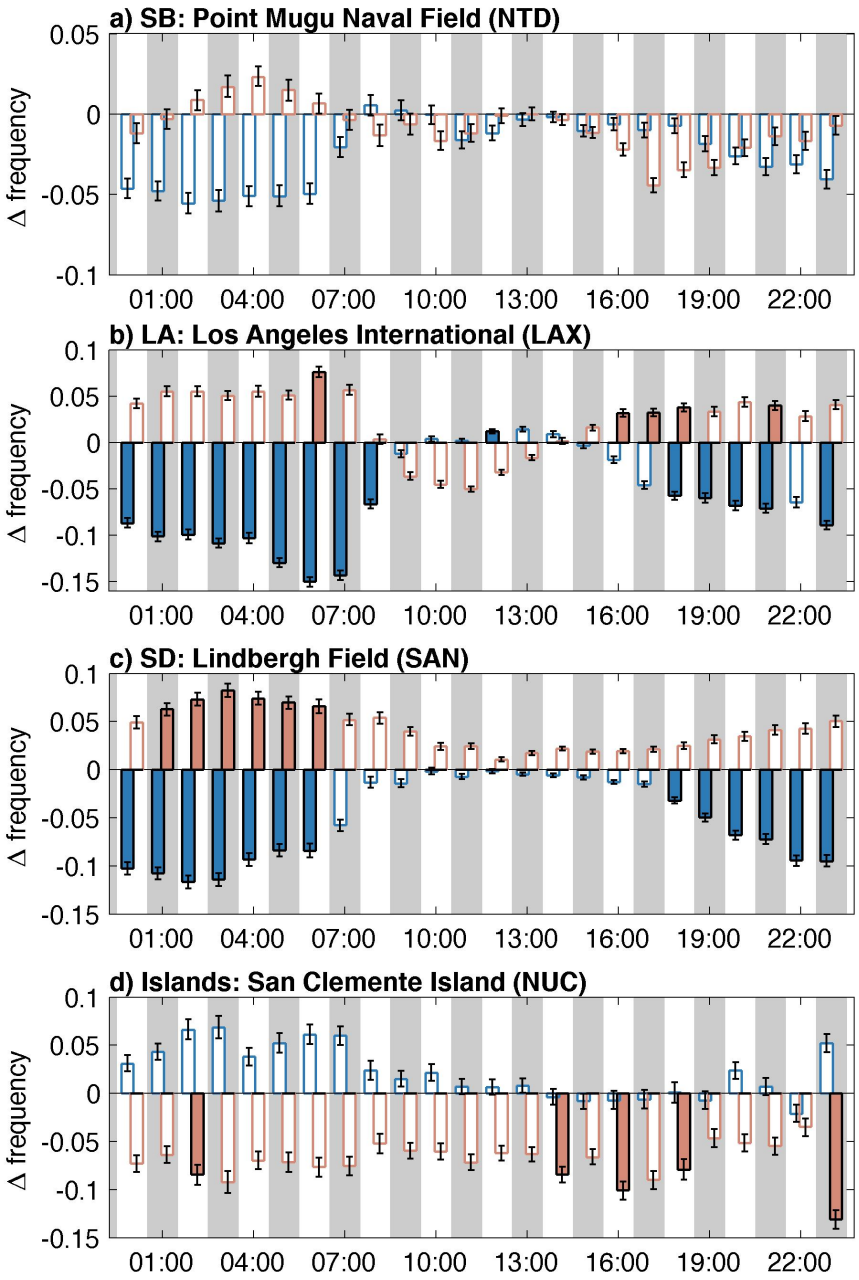

Figure S1. Changes in stratus frequency for all hours. Absolute changes in hourly frequency of fog (blue bars) and upper 75% stratus (beige bars) during 1948–2014 for the airfield with the most complete cloud-height record in each sub-region. Whiskers indicate uncertainty caused by limited vertical resolution of cloud-base height measurements. Specifically, the whiskers bound the inner 95% of trends calculated from the 1000 alternate records for each hour. Filled bars indicate significant trends for hours when at least 95% of the 1000 alternate trends are significant ($P<0.05$) according to both Spearman's Rho and Kendall's Tau tests. Changes in stratus frequency in this plot are absolute, meaning not relative to an initial frequency value.

Table S4. Observed changes in stratus frequency at 07:00. Linear relativized changes in May–September stratus frequency at 07:00 for each sub-region and airfield. Δ in stratus frequency is calculated for all stratus clouds (cloud base ≤1000 m) in the column labeled "All" and within the four quartiles of height classes. Changes in stratus frequency are expressed relative to the initial regression value (e.g., a reduction of 0.25 from an initial stratus frequency regression value of 0.50 equals a -50% change). Asterisks indicate trends significant at 99% (**) and 95% (*) confidence levels according to both Spearman's Rho and Kendall's Tau tests. Trends are not marked as significant unless these significance criteria were met by at least 95% of the 1000 alternate time series.

			Δ (%) within each height class				
				Fog			
Name	**First year**	**N**	**All**	**0-25%**	**25-50%**	**50-75%**	**75-100%**
Santa Barbara	**1948**	**67**	**-1**	**-8**	**19**	**13**	**-16**
SBA	1949	56	-3	5	23	-30	-29
OXR	1953	58	8	-5	31	47	-8
NTD	1948	67	-4	-9	-11	33	-14
Los Angeles	**1948**	**67**	**-18****	**-64****	**26**	**80****	**39**
VNY	1948	43	-25*	-60*	-11	49	23
SMO	1973	42	-7	-24	-26	8	36
LAX	1948	67	-14*	-57**	-8	21	54**
BUR	1948	63	-14*	-50**	17	22	25
HHR	1973	42	-9	-43	16	37	6
TOA	1973	42	-5	-16	-29	80**	-9
LGB	1948	67	-16*	-59**	30	52**	23
SLI	1948	55	-17**	-60**	41*	60**	-1
FUL	1973	42	-12	-49**	32	31	-7
SNA	1968	47	-10	-45**	5	48*	14
ONT	1950	50	-26**	-86**	83**	176**	18
San Diego	**1948**	**67**	**-7**	**-26***	**-3**	**15**	**26**
NFG	1967	48	-8	-17	-10	-9	28
CLD	1973	42	0	-23	-4	12	66**
NZY	1948	67	-6	-26*	-2	17	6
SAN	1948	67	-1	-20	15	8	13
NKX	1948	66	-1	-13	-4	30	3
MYF	1973	41	-20**	-29	104**	-76	6
NRS	1948	63	-13*	-41**	-27	9	38
SEE	1973	42	-13	-49**	11	50*	-25
Islands	**1948**	**67**	**-2**	**63***	**-27**	**-42***	**-1**
NSI	1948	65	-2	44*	-17	-40**	-8
NUC	1960	52	-2	22	-26	-14	16

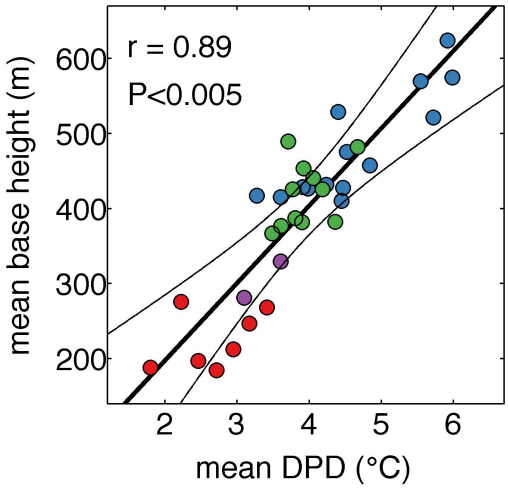

Figure S2. Relationship between mean climatological stratus-cloud base height and mean climatological dew-point depression (DPD). Mean May–September 07:00 stratus-cloud base height versus mean May–September 07:00 DPD during 2000–2014. Each dot represents one of 34 CSCA airfields that have at least 10 years of valid cloud-base height and DPD data during 2000–2014. Dot colors indicate sub-regions (red: SB, blue: LA, green: SD, purple: Islands). Correlation significance value and 95% confidence bounds around the regression line account for uncertainties caused by spatial autocorrelation and measurement uncertainties for cloud base height and DPD.

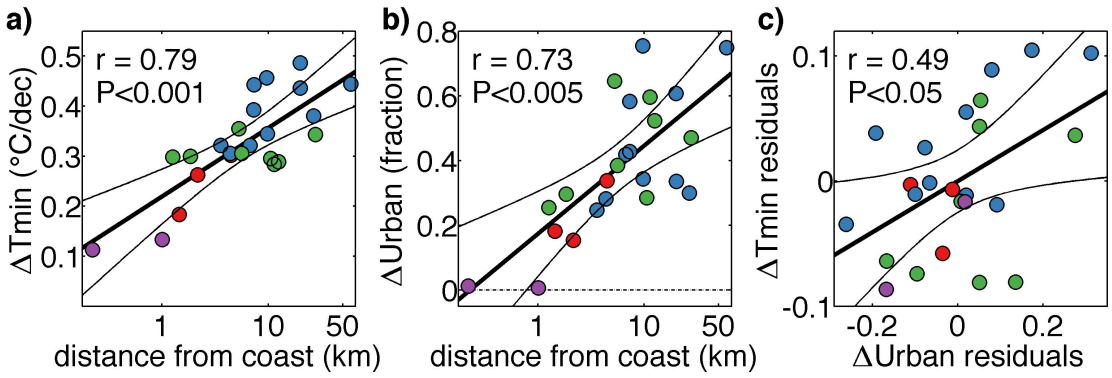

Figure S3. Confounding effects caused by distance from coast. (a and b) Regressions of airfield-specific changes in daily minimum temperature during 1948–2014 (ΔT_{min}) and changes in the fraction of urban cover during 1950–2011 (ΔUrban) versus distance from coast expressed on a log scale. (c) Regression of the ΔT_{min} residuals versus the ΔUrban residuals after removal of the relationships with distance from coast shown in (a and b). Each dot represents one of 24 CSCA airfields and dot colors indicate sub-region (red: SB, blue: LA, green: SD, purple: Islands). Bold black line: regression line. Correlation significance values and 95% confidence intervals bounding regression lines account for uncertainty due to spatial autocorrelation and measurement errors for T_{min} and ΔUrban.

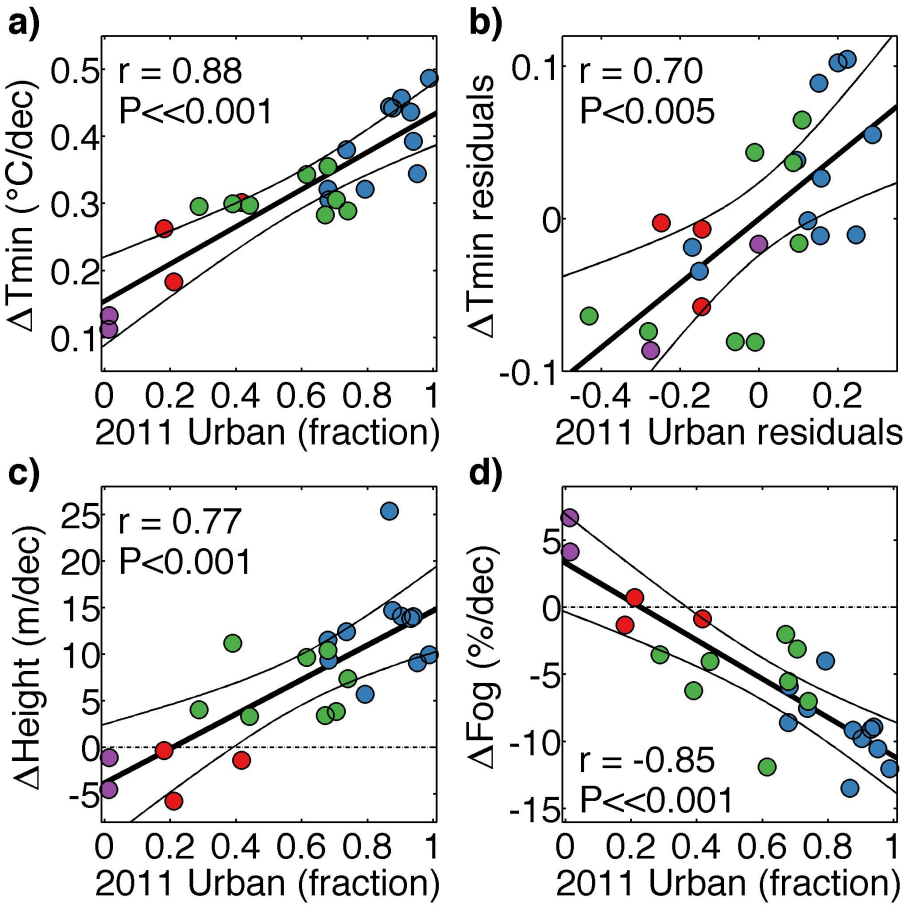

Figure S4. Relationships with 2011 urban cover. (a) Regression of airfield-specific changes in daily minimum temperature during 1948–2014 (ΔT_{min}) versus urban fraction in 2011 only. (b) As in (a), but after removal of relationships with distance from coast. (c and d) Changes in stratus-cloud base height per decade (ΔHeight) and fog frequency per decade (ΔFog) versus urban fraction in 2011. Each dot represents one of 24 CSCA airfields and dot colors indicate sub-region (red: SB, blue: LA, green: SD, purple: Islands). Bold black lines: regression lines. Correlation significance values and 95% confidence intervals bounding regression lines account for uncertainty due to spatial autocorrelation and measurement errors in T_{min} and cloud base height.

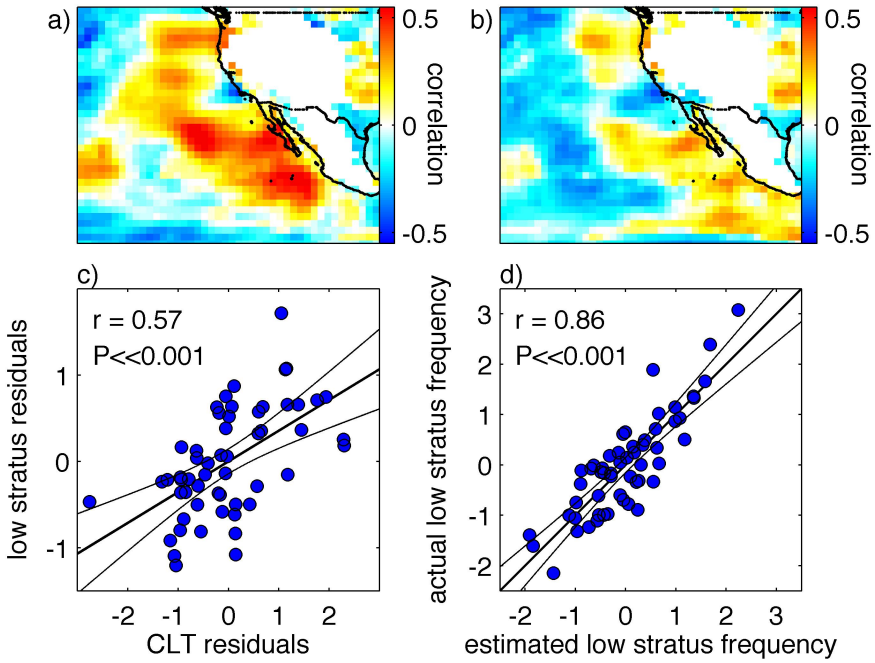

Figure S5. Secondary effects on frequency of low stratus. (a and b) Correlation between frequency of low stratus occurrence and mean vertical velocity in the lower troposphere (900 to 600 hPa) before (a) and after (b) correlation with lower tropospheric stability (LTS) has been removed from stratus and vertical velocity time series. The LTS record represents mean May–Sep LTS within the white box in the inset map of Figure 4f in the main article. Positive correlation indicates correlation with downward motion (subsidence). (a and b) represent 1979–2014 and all correlations are shown, regardless of significance, to highlight the loss of correlation with subsidence near the CSCA after LTS is accounted for. (c) Scatter plot of residuals of CSCA low-stratus frequency versus residuals of CSCA cloud layer thickness (CLT; defined in article). Residuals are calculated by removing linear relationships with NKX inversion strength. (d) Scatter plot of measured versus estimated low stratus frequency, using NKX inversion strength and CLT residuals as predictors in a multiple regression. (c and d) represent 1960–2014 and units are standard deviation anomalies from the 1960–2014 mean. Significance values and 95% confidence bounds around regression lines in (c and d) account for uncertainties due to temporal autocorrelation and measurement errors in cloud base height, NKX inversion strength, and CLT. Linear trends were removed from all time series prior to this analysis.

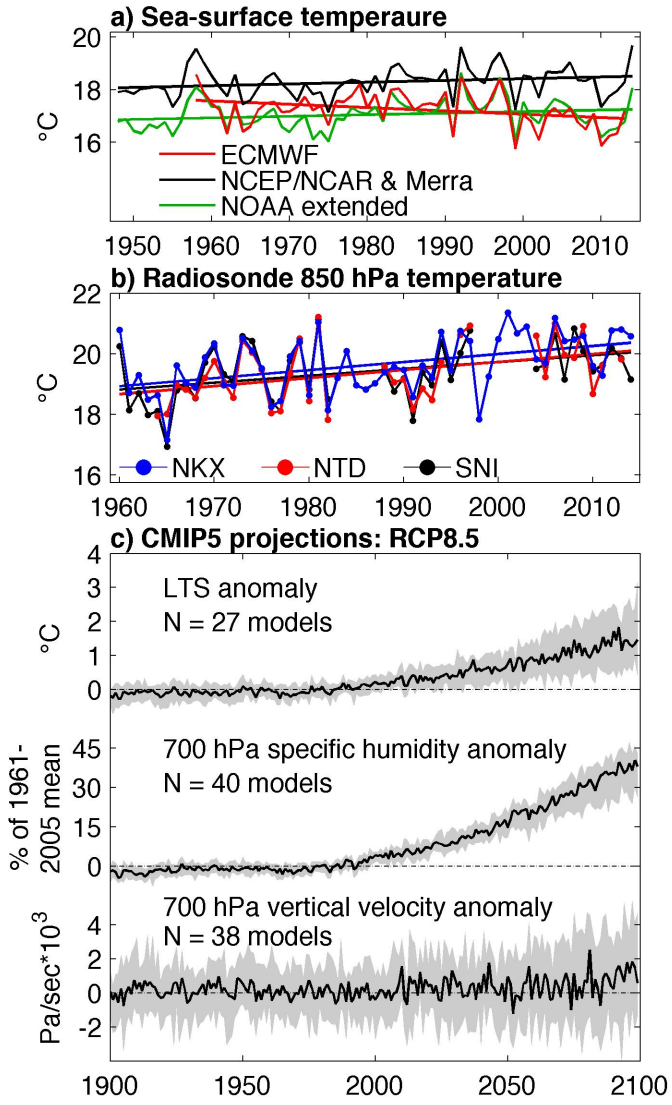

Figure S6. Observed and projected trends that would influence CSCA stratus frequency. May–September (a) reanalysis SST within 29–35°N, 117–123°W, (b) 850 hPa temperature for three radiosonde sites (NKX: San Diego-Miramar, NTD: Pt. Mugu, SNI: San Nicolas Island), (c) CMIP5 ensemble median (black line) and interquartile (grey area) projections of climate anomalies relative to the 1960–2005 mean within the geographic box 15–35°N, 110–120°W. The upper plot in (c) represents anomalies in lower tropospheric stability (LTS), defined in the article as potential temperature at 850 hPa minus potential temperature at 2 m above the surface.