Understanding the cloud response to external forcing is a major challenge for climate science. This crucial goal is complicated by intermodel differences in simulating present and future cloud cover and by observational uncertainty. This is the first formal detection and attribution study of cloud changes over the satellite era. Presented herein are CMIP5 model-derived fingerprints of externally forced changes to three cloud properties: the latitudes at which the zonally averaged total cloud fraction (CLT) is maximized or minimized, the zonal average CLT at these latitudes, and the height of high clouds at these latitudes. By considering simultaneous changes in all three properties, the authors define a coherent multivariate fingerprint of cloud response to external forcing and use models from phase 5 of CMIP (CMIP5) to calculate the average time to detect these changes. It is found that given perfect satellite cloud observations beginning in 1983, the models indicate that a detectable multivariate signal should have already emerged. A search is then made for signals of external forcing in two observational datasets: ISCCP and PATMOS-x. The datasets are both found to show a poleward migration of the zonal CLT pattern that is incompatible with forced CMIP5 models. Nevertheless, a detectable multivariate signal is predicted by models over the PATMOS-x time period and is indeed present in the dataset. Despite persistent observational uncertainties, these results present a strong case for continued efforts to improve these existing satellite observations, in addition to planning for new missions.
first formal detection and attribution study on observed cloud trends. This requires that we first consider a number of related questions:

- Can we identify the fingerprints of external forcing on model cloud properties, and if so, are they distinct from patterns that arise from internal variability alone?
- How long an observational record is theoretically required to ensure detection of an externally forced signal above the noisy background of natural climate variability?
- Does a multivariate fingerprint that encompasses coherent changes in multiple cloud properties present advantages over a fingerprint tracking changes in a single variable?
- How strongly are the fingerprints of external forcing agents expressed in two long-term satellite records?

To answer these questions, we first define indicators that track features of interest in modeled cloud fields (section 2). These indicators are then used in section 3 to derive fingerprints of externally forced cloud changes that characterize the robustly simulated coherent response of three relevant cloud properties to external forcing. In section 4 we quantify the strength of the “signal” of externally forced cloud changes and the “noise” arising from natural cloud variability in models, and determine in section 5 the year in which the multimodel average signal of externally forced cloud change emerges, along with the uncertainty in this estimate. Finally, in section 6, we perform a formal detection and attribution of the externally forced fingerprint in the observational cloud record.

2. Cloud change indicators used in this study

In this study, we rely both on observations and on model results from phase 5 of the Coupled Model Intercomparison Project (CMIP5). We make use of unforced preindustrial control (piControl) simulations and historical simulations extended to 2100 by splicing with the appropriate representative concentration pathway 8.5 (RCP8.5) simulation (ALL+8.5). Further details of the models used are given in appendix A.

To detect and track coherent changes across multiple models, we require indicators that can identify robust physical processes even in the presence of model errors and biases. Marvel and Bonfils (2013) introduced a method to simultaneously track changes in the location and intensity of major precipitation features. Here, we build on these techniques to introduce three simple indicators of cloud change, which we will use to detect changes in the latitude L of the cloudiest and clearest regions, the total cloud amount C in these regions, and the height H of high clouds.

a. CLT-derived indicators: L and C

Latitude and cloudiness indicators are derived from boreal winter [December–February (DJF)] mean total cloud fraction (CLT; Fig. 1a). There are several reasons for restricting our analysis to DJF averages; these are discussed in section 3. In almost all CMIP5 ALL+8.5 and piControl simulations, there are exactly five local extrema in the smoothed CLT field: maxima over the tropics and midlatitude storm tracks, and minima in the subtropical dry zones. We do not consider models for which this is not true in the 1980–2005 climatology, as described in section 2. Our method identifies the latitudes and cloud amounts at these five locations to define two time-varying indicators. For every year that every month in DJF is available, we analyze the data as follows:

1) Zonally average the total cloud fraction.
2) Apply a Gaussian filter with width 5° latitude to the zonal average.
3) Use a peak-finding algorithm to determine the local extrema (Fig. 1b).
4) For most models, there are exactly two local minima and three local maxima for every year’s DJF average. The latitudes at which these extrema occur are recorded, as are the values of the smoothed CLT. These will form the basis for our L and C indicators, respectively.
5) In the rare (fewer than 3.4%) of model years in the ALL+8.5 simulations) cases for which there are more or fewer local extrema than anticipated, missing values are returned at all five extrema in that year. In subsequent fingerprinting and projection, if any value is missing in a model at a given year, the model does not contribute to the multimodel average (MMA) at that time. Figure 2 shows that when the analysis is restricted to DJF, there is no discernible trend in the fraction of models that contribute to the MMA over time.

The resulting MMA time series, measuring the latitudes and total cloud fraction at these local extrema, are both of dimension $n_T \times 5$, where $n_T$ is the analysis length.

We define the L and C indicators as the anomalies of the resulting latitude and total cloud fraction time series, respectively. Anomalies are calculated on a model-by-model basis with respect to the 1983–2007 time average (in the spliced ALL+8.5 runs and the observations) or the 200-yr climatology (for piControl simulations).

b. CL-derived indicator: H

Our height indicator H is constructed from the three-dimensional cloud fraction (CL; Fig. 1c). As with CLT, we consider DJF averages. For every year, we calculate the H indicator as follows:
1) To emulate what a satellite might see, we estimate the cloud fraction at each vertical level that is visible from space $\text{CL}^{\text{visible}}$ assuming random overlap (Tian and Curry 1989). This step is illustrated in Fig. 1d.

2) We compute each model's temporally and spatially varying pressure field on the native grid levels upon which $\text{CL}^{\text{visible}}$ is defined.

3) For each grid box, we determine the index of the lowest native grid level $\ell$ such that the pressure at the lower boundary of the cell $P(\ell) < 440$ hPa. This is the standard ISCCP criterion for high cloud.

4) Three-dimensional model cloud fields are reported on pressure levels that vary spatially, temporally, and across models. We therefore require a step to determine model CL at the bottommost level closest to 440 hPa. Assuming that $P(\ell)$ decreases with increasing $\ell$, we interpolate between the levels just above and just below 440 hPa ($\ell$ and $\ell - 1$) as follows:

$$A = \frac{440 \text{ hPa} - P(\ell)}{P(\ell - 1) - P(\ell)},$$

5) To measure the altitude of high clouds, we calculate the weighted cloud-top pressure (WCTP) at every grid cell, defined as

$$WCTP = \frac{1}{\ell_0} \sum_{i=\ell}^{\ell_0} \frac{\text{CL}^{\text{visible}}(i) P(i)}{\text{CL}^{\text{visible}}_{\text{bottom}}} + \frac{\text{CL}^{\text{visible}}_{\text{bottom}} P(\ell)}{2}.$$

Here, $\ell_0$ is the highest vertical level in the model.

6) We zonally average the resulting WCTP.

7) We spatially smooth the zonal average WCTP with the same Gaussian filter applied to the CLT-derived indicators (L and C).
8) Having previously calculated $L$, we determine the WCTP at each of the five latitudes that correspond to local extrema in CLT at the time step in question.

9) The $H$ indicator at time $t$ is the WCTP at each of the five latitudes multiplied by 2, such that a decrease in WCTP, which corresponds to an increase in cloud height, leads to a positive change in $H$. The result is a time series of dimension $nT \times 5$ representing the WCTP at the cloudiest and clearest latitudes. As for $L$ and $C$, the height indicator $H$ is defined as an anomaly time series relative to the respective model climatology.

We note that several models have made use of the ISCCP simulator (Klein and Jakob 1999; Webb et al. 2001) to directly calculate output that is more similar to observations. Were we to make use of this output, we could eliminate the “visible cloud” calculation described in steps 1–4 above. However, because of the computational time required to implement the ISCCP simulator, relatively few models have done so: at the time of writing, only eight have contributed valid ISCCP simulator output to the CMIP5 piControl repository, seven to the historical repository, and none to the RCP8.5 repository. [Note that “valid” means that the simulator was implemented properly, which rules out the IPSL models; see Zelinka et al. (2013) for details.] Thus, to ensure a sufficiently large sample size in the multimodel archive, we rely instead on indicators derived from standard cloud diagnostics as described above.

c. Further details

1) ANOMALIES

It is well known that major features of the zonal circulation simulated by models are not always found at exactly the same latitudes (see, e.g., Levy et al. 2013, 2014; Scheff and Frierson 2012). This means that simply performing a multimodel average may blur or even eliminate boundaries between, for example, relatively cloudy and relatively clear latitudes.

The indicators for each simulation we consider are thus defined as anomalies relative to that particular simulation’s climatology. Operating in this anomaly space is useful because models vary in their simulation of absolute values of cloud fraction, the latitudinal locations of $C$ extrema, and the vertical distribution of clouds (Fig. 3). This strategy yields estimates of trends in $L$, $H$, and $H$ that are less affected by model biases in the locations or intensity of major features, which may be substantial (Fig. 3). Marvel and Bonfils (2013) used similar methods to detect poleward shifts of the mid-latitude storm tracks that are robust across multiple models, despite model errors in the locations of the storm tracks.

2) MODEL SELECTION

The “five extrema” structure identified in section 2a is physically motivated: we expect to see peaks in the tropics and midlatitudes associated with tropical convection and storm tracks, respectively. We also expect the dry subsidence regions in the subtropics to be relatively clear.
As we shall see in section 6, we also find this to be true for both observational datasets considered. The combination of physical expectation and presence in the observations leads us to define a five-extrema test for model simulations: if there are not exactly three maxima and two minima in a model’s smoothed 1980–2005 zonal average DJF CLT climatology, we mask out (i.e., do not consider) all data from that model. All models, with the exception of the two NASA Goddard Institute for Space Studies models (GISS-E2-R and GISS-E2-H; see Table A1 for more information about individual models), pass this test.

As noted in point 5 of section 2a, we also mask individual model years in which the five-extrema test fails. This represents a more ad hoc model selection criterion, but because many simulations contain at least one year in which there are more than or fewer than five extrema, we judge it necessary to maintain a large sample size of models. We find no evidence that the years in which this test fails are not randomly distributed, and this results in our discarding fewer than 4% of model years in our analyses.

### 3) Justification for Boreal Winter

In this study, we confine our analysis to boreal winter. There are three reasons for this choice. First, there are exactly five local extrema at every time step in the observations and in most models for DJF. Both Northern and Southern Hemisphere storm tracks are apparent and detectable in boreal winter; the Northern Hemisphere (NH) storm tracks weaken in summer and will not necessarily be present as a local maximum in the zonal average CLT. In general, however, Southern Hemisphere (SH) storm tracks are detectable in all seasons. Second, the ISCCP dataset used relies on observations in the visible range and thus has no coverage over the very high latitudes in the winter hemisphere. In DJF, however, the NH storm tracks are sufficiently far south to be detectable.

Finally, considering the DJF average ensures that all five extrema are likely to be present in models over the entire observed time period; this is not the case for other seasons (Figs. 2a,b). In the RCP8.5 scenario, all models predict increasing global temperatures and associated loss of sea ice. In all seasons except boreal winter, the replacement of NH sea ice with open water results in large increases in total cloud fraction in the northern polar regions. This means that in later years the NH storm track is no longer detectable as a local extremum. Because our methodology masks out years in which fewer than five extrema are detected, this will spuriously weight earlier times in the MMA quantities.

#### d. Indicator Summary

In summary, $L$, $C$, and $H$ provide measures, respectively, of

![Fig. 3. Histograms of the climatological locations, obtained from CMIP5 preindustrial control simulations, of (top) total cloud fraction (CLT) extrema, (middle) values of smoothed CLT at these locations, and (bottom) weighted cloud-top pressures at these locations.](image-url)
1) where the cloudiest and clearest regions may be found in a particular model,
2) how cloudy these regions are, and
3) the altitude of high clouds in these regions.

These indicators are restricted to boreal winter and, being derived from zonal averages, fail to capture important cloud features that are nonzonal, particularly in northern latitudes. However, L, C, and H are computationally efficient to derive from a large ensemble of models and experiments, inherit many of the useful features described by Marvel and Bonfils (2013), and summarize key features of the global pattern of cloud cover that are highly relevant to cloud feedback.

3. Fingerprints

In the parlance of climate change detection and attribution (D&A), the “fingerprint” is the spatial pattern that characterizes the climate system response to external forcing (see, e.g., Allen and Stott 2003; Gillett et al. 2002; Hegerl et al. 1996; Stott et al. 2000; Tett et al. 2002). In this study, we use techniques developed and refined by, for instance, Barnett et al. (2008) and Santer et al. (1995, 2013). We calculate this pattern using the externally forced ALL+8.5 simulations spanning the period 1900–2100. The multivariate fingerprint is calculated as follows.

1) For every ALL+8.5 simulation, we obtain time-varying indicators \( L^i, C^i, \) and \( H^i \) using the methods described in section 2. The superscript \( i \) indicates that these indicators are derived from ALL+8.5 spliced model simulations.

2) As we are constructing a fingerprint across multiple variables measured in different units, we require some means of removing dimensional information from the indicators. We therefore define normalized indicators \( L^i_{\text{norm}}, C^i_{\text{norm}}, \) and \( H^i_{\text{norm}} \) by dividing through by the standard deviation of each indicator’s time series at each location. This ensures that, for any \( X \in \{ L^i_{\text{norm}}, C^i_{\text{norm}}, H^i_{\text{norm}} \}, \) the matrix \( X^T X \) is a correlation rather than a covariance matrix. This has the effect of giving all five extrema (and all three indicators) equal weight in our subsequent EOF calculation.

3) We then average the ALL+8.5 indicators, first over an individual model’s realizations if more than one is available, and then over all models to obtain MMA indicators \( L_{\text{norm}}, C_{\text{norm}}, \) and \( H_{\text{norm}} \). This double average over realizations and models effectively damps internal variability because manifestations of random natural phenomena such as ENSO are uncorrelated across models. It also effectively guards against a model with many realizations unduly influencing the mean.

4) We compute the joint correlation matrix \( M^T M \), where \( M \) is a 15-element vector made up of the three normalized indicators at five locations, averaged across all realizations and models:

\[
M = (L_{\text{norm}}, C_{\text{norm}}, H_{\text{norm}}).
\]

5) The multivariate climate change fingerprint (Santer et al. 1995) is then defined as the eigenfunction of \( M^T M \) corresponding to the largest eigenvalue (i.e., the leading EOF of \( M \)).

The resulting multivariate fingerprint \( F_m(L, C, H) \) is shown in Fig. 4a. This pattern is characterized by three robust physical processes that are consistently simulated in anthropogenically forced model experiments: poleward shifts of major climate zones (Wetherald and Manabe 1988; Zelinka et al. 2013; Bender et al. 2012), increases in midlatitudes and decreases in subtropics of total cloud fraction (evident here only in the Northern Hemisphere) (Zelinka et al. 2013), and a global rise in high cloud (Hansen et al. 1984; Hartmann and Larson 2002; Wetherald and Manabe 1988; Zelinka et al. 2013; Singh and O’Gorman 2012). The leading EOF explains over 91% of the variance in the MMA indicators.

Figure 4b, by contrast, shows the leading noise mode, defined as the leading EOF of the concatenated piControl indices (section 4b). One phase of this mode is characterized by an equatorward contraction of the major cloud features in the subtropics, and a rise of high cloud especially in the tropics and subtropics. This is unsurprising given that ENSO (the leading mode of internal variability in most models) is known to cause an equatorward contraction of the Hadley circulation (Lu et al. 2008), movement of the ITCZ (Deser and Wallace 1990), and a rise in tropical and subtropical high cloud tops (Zelinka and Hartmann 2011).

The fingerprint estimated from the ALL+8.5 runs is the response to combined anthropogenic and natural external forcing. The relative contributions to the fingerprint from purely natural changes in solar irradiance and volcanic aerosol loadings is therefore unclear in Fig. 4a. Previous D&A work illustrates that differences between the anthropogenic and ALL+8.5 fingerprints are relatively small over the satellite era [see, e.g., Fig. 2 in Santer et al. (2013)]. This suggests a relatively small natural contribution. The fingerprint obtained from the ALL+8.5 model simulations is very similar to that obtained from the multimodel ensemble average of model experiments in which atmospheric CO2 is increased at 1% per year (Fig. 4c), indicating that the ALL+8.5
fingerprint primarily reflects the response of clouds to anthropogenic changes in CO₂.

In subsequent sections, we will see that searching for coherent changes in the three indices simultaneously leads to earlier signal detection times. First, however, we consider the individual fingerprints of \( L \), \( C \), and \( H \) by calculating the correlation matrices of \( L_{\text{norm}} \), \( C_{\text{norm}} \), and \( H_{\text{norm}} \), respectively. The resulting five-element fingerprint for
each index is shown in Fig. 5. These patterns resemble their counterparts in the multivariate fingerprint $F_m(L, C, H)$. As expected, the latitude fingerprint $F(L)$ corresponds to poleward movement in both hemispheres. The cloudiness fingerprint $F(C)$ is described by clearing in the subtropical dry zones (albeit more in the Northern Hemisphere), little change in the tropics, and increasing cloudiness in the midlatitudes. The cloud height fingerprint $F(H)$ is characterized by a rise in high cloud at all latitudes considered.

4. Signal detection

a. Signal

The fingerprint $F$ for some indicator or collection of indicators $X$ is a function of location $x$. Given suitably normalized observed or simulated data $O_X(x, t)$ with analysis length period $n_T$, the projection $P_X(t) = \sum_x O(x, t) F(X, x)$ provides a measure of the spatial covariance between the data and the fingerprint. If the data increasingly resemble the fingerprint, then the projection $P_X(t)$ will trend upward with time. We define the signal of external forcing $S_X(n_T)$ to be the slope of the best-fit line to $P_X(t)$, obtained by least squares regression. This is a standard method commonly used in D&A research to detect signals of external forcing in observations (see, e.g., Barnett et al. 2008; Hasselmann 1979; Santer et al. 2007, 2013). The signal is, of course, a function of the analysis length $n_T$. In this paper, we calculate univariate signals, obtained by projecting any one indicator $X$ onto its corresponding univariate fingerprint $F(X, x)$, as well as the multivariate signal. The latter is calculated by constructing a 15-element vector from normalized observational or model data and projecting this vector onto the multivariate fingerprint $F_m(L, C, H)$.

![Fig. 5. Individual fingerprints (leading EOFs of the MMA correlation matrix) for the (a) latitude indicator $L$, (b) cloud amount indicator $C$, and (c) high cloud height indicator $H$.](image-url)
b. Noise

How likely is a given signal to result from natural climate variability alone? To assess the significance of an observed signal, we require a null distribution of trends (i.e., information on the behavior of trends in pattern similarity between the fingerprint and natural internal variability). This distribution is constructed using natural internal variability estimates from the preindustrial control simulations. For each control run in the CMIP5 archive we derive latitude, cloudiness, and height indicators \( \mathbf{L}' \), \( \mathbf{C}' \), and \( \mathbf{H}' \) following the procedures described in section 2. Each indicator time series is divided by the temporal variance at each of the five extrema to obtain normalized indicators \( \mathbf{L}_{\text{norm}}' \), \( \mathbf{C}_{\text{norm}}' \), and \( \mathbf{H}_{\text{norm}}' \). We then construct a time series of length 6800 years by concatenating each of the 34 control simulations (using only one control simulation per model). To prevent models with longer runs from dominating our calculation of internal noise, we use only the first 200 years of every model control simulation. By using multiple models with various amounts of noise, we might converge on a more realistic estimate (assuming no systematic error across models) of noise variability; we do not assume homogeneity of internal variability across all models. Moreover, including the first, rather than the last, 200 years of each simulation means that our noise estimate is likely to include long-term model drift, which may also result in a more conservative signal-to-noise estimate. It is these long piControl indicators that determine the leading noise mode shown in Fig. 4b.

We project these concatenated and normalized piControl indicators onto the univariate and multivariate fingerprints shown in Figs. 4a and 5. We then calculate the distribution of \( n_T \)-length trends in nonoverlapping segments of the resulting projection time series. This distribution is quasi-normal and distributed around zero, and its standard deviation \( N_X(n_T) \) provides a measure of internal climate variability, or noise. The value of \( N_X(n_T) \) will therefore vary with the analysis time scale \( n_T \) as well as the indicator \( X \) under consideration (Santer et al. 2013). Rather than recalculate noise on a model-by-model basis, we use the same noise estimate derived from the concatenated control runs for all models. This allows us to take advantage of the larger sample size that results from concatenation.

c. Signal-to-noise ratios

Dividing the univariate or multivariate signal \( S_X(n_T) \) by this noise measure yields the signal-to-noise ratio. If the signal-to-noise ratio exceeds (and remains above) 1.96, then we can claim to have detected a signal with 95% confidence relative to our best current multimodel estimates of natural internal variability. We use this two-tailed \( z \) test throughout to provide a conservative estimate of significance.

5. Detection times

When might we expect a detectable signal to emerge from the background of natural internal variability? Previous studies (see, e.g., Hawkins and Sutton 2012; Mahlstein et al. 2011) have recognized that, even in cases where the externally forced signal has not yet emerged from climate noise, there is still utility in calculating the “time of emergence” when it is predicted by models to do so. Unlike those studies, we focus not on regional patterns but on the global signal, defined as the projection onto a univariate or multivariate fingerprint. Our “detection time” is thus the time at which the global signal is predicted, in the MMA, to emerge from the noise.

The importance of trend detection for cloud studies was previously highlighted by Loeb et al. (2007), who calculate the time needed to detect a statistically significant trend in top-of-atmosphere flux assuming a statistical model (Weatherhead et al. 1998) for natural variability. Rather than rely on a statistical model, we instead use the piControl simulations in the CMIP5 archive for an estimate of internal variability. By treating individual ALL +8.5 simulations as equally plausible, we define the detection time (DT) (Santer et al. 2013) to be the year in which the signal-to-noise ratio for a model ALL +8.5 trend, beginning in 1983, exceeds and remains above the 95% significance threshold. The year 1983 is chosen because it is the first full year for which DJF satellite observations are available (section 6). The red line in Fig. 6 shows the MMA signal-to-noise ratio for the multivariate fingerprint. As expected, it increases with increasing trend length as trends in the models increasingly resemble the fingerprint while simultaneously becoming less and less likely to arise as a result of natural variability alone. The DT of the multivariate signal is 2010 in the MMA but ranges from 2001 to 2033 in individual simulations.

We then determine the signal-to-noise ratios for each univariate fingerprint and calculate the detection time for each indicator in isolation. Figure 6 illustrates the utility of the multivariate approach. The multimodel average \( H \) signal is strongest relative to internal noise (DT = 2026; individual models range between 2005 and 2050), while the MMA C and L signals emerge later in 2036 (2016–78 in individual models) and 2063 (from 1992 to later than 2100 in individual models), respectively. On average, a multivariate representation of signal and noise results in detection more than a decade earlier than would be expected from the strongest
single-variable detection time results, and five decades earlier than the $L$ signal alone. The uncertainty in DT is also considerably reduced in the multivariate case, as the intermodel spread of $L$ and $C$ DTs exceeds a century but is only about 30 years for the multivariate DT.

6. Observations

In this section, we apply formal D&A methodology to satellite observations, using the two longest satellite-based records of global cloud properties available. The International Satellite Cloud Climatology Project dataset (ISCCP; Rossow and Schiffer 1999) covers the period July 1983 through June 2008, while the slightly longer Pathfinder Atmospheres–Extended dataset (PATMOS-x; Heidinger et al. 2012, 2014) spans the period January 1982 through December 2009. Both datasets have near-global coverage and provide independently retrieved estimates of total cloud fraction and its vertical distribution. Details of both datasets are provided in appendix B.

As discussed in the GEWEX Cloud Assessment (Stubenrauch et al. 2013), cloud retrievals from ISCCP and PATMOS-x are subject to errors and biases arising from numerous sources. For example, passive satellite retrievals can have difficulty distinguishing cloud from underlying cold and bright surfaces (Stubenrauch et al. 2013), and thus we expect considerable uncertainties in the observed high-latitude cloud properties in DJF. Moreover, numerous studies (e.g., Evan et al. 2007; Norris 1999, 2007) have identified potential sources of unphysical trends in existing observational datasets, including spurious drifts due to changes in satellite view angle and sensor calibration. For these reasons, the ISCCP and PATMOS-x datasets are not generally considered to be reliable for long-term trend analysis. In the following sections, we will show that, despite substantial observational uncertainty and known artifacts, useful trend information can be extracted from the existing observations.

a. Observational indicators

Observational $L^o$ and $C^o$ indicators are calculated from the ISCCP and PATMOS-x datasets as for the model data. We omit the random overlap calculation (step 2 in section 2b) in calculating the observed height indicator $H^o(t)$ because the datasets already report visible cloud, and define weighted cloud-top pressure as

$$WCTP = \frac{\sum_{i=5}^{7} CTP_i CL^{\text{visible}}_i}{\sum_{i=5}^{7} CTP_i \frac{CL^{\text{visible}}_i}{CL^{\text{visible}}_i}}.$$

Here, $CTP$ is the observed cloud-top pressure. The sum is over pressure bins 5 through 7, which represent clouds at

![Fig. 6. Signal-to-noise ratio as a function of analysis length $n_T$. Solid lines represent the average of the multimodel signal-to-noise ratio; envelopes are $\pm 2$ intermodel standard deviations from the mean. Vertical lines show the MMA detection time.](image-url)
pressures below (altitudes above) 440 hPa, in keeping with the standard ISCCP high cloud classification. As in the model indicators described in section 2, the observational indicators \( L^o \), \( C^o \) and \( H^o \) are defined with respect to the climatology in the relevant observational dataset.

Figure 7 shows the resulting indicator time series, as well as the temporal correlation between the ISCCP and PATMOS-x datasets over the ISCCP analysis period. The correlations are largest in the latitude indicator, particularly in the tropics and subtropics. This suggests that satellite records largely agree on trends in the locations of the cloudiest and clearest latitudes, and show the same temporal evolution. The same cannot be said of the cloud fraction at these latitudes. The observed cloudiness indicators \( C^o \) are weakly correlated (or even anticorrelated) at all locations, except at the tropical
peak. This makes it difficult to draw robust conclusions regarding total cloud fraction from the observations. There are likewise large differences in the mean weighted cloud-top pressure at all locations, and positive but small temporal correlations in the height indicator, with the weakest correlation found for the SH storm track. Thus, we identify consistency between observations in the $L$ indicator, but not in the $C$ or $H$ indicators.

b. Observed trends

Figure 8 shows the observed and modeled trends for each indicator over the common period 1983–2007. Here, the indicators have not yet been normalized in order to report trends in physical units. At every location except the SH dry zone, both observational datasets agree with each other and with the MMA on the sign of the $L$ trend. It is noticeable that the poleward migration of the storm tracks in both hemispheres is larger in ISCCP and PATMOS-x than in the model average, as seen in (Bender et al. 2012) for the models from phase 3 of CMIP (CMIP3). The sign of the $C$ trends in the two observed datasets is in agreement in three of the five extrema, but differs in the storm tracks. The large ISCCP trend toward decreasing $C$ in the tropics and subtropics is outside the range of model results, which may reflect temporal changes in satellite view angle (Norris 1999). The MMA trends in $H$ indicate rising high clouds at all locations. While PATMOS-x shows large positive trends in cloud height at four of the five extrema, the height trend is within the model-predicted range only in the SH storm track. The ISCCP results, however, are even more inconsistent with the model predictions, and show an upward trend only in the tropics.

7. Detection and attribution results

By projecting suitably normalized observed $L$, $C$, and $H$ indicators onto each of the individual-variable and multivariate fingerprints, we can measure the spatial similarity between the time-varying observations and the time-invariant model pattern of externally forced change. The observed signal is either the 27-yr PATMOS-x trend or the 25-yr ISCCP trend found in the projection time series. Because there are no observations of unforced natural variability, we rely on model piControl runs for our noise estimates. The observed signal-to-noise ratio for each indicator is therefore obtained by dividing the signal by the appropriate length noise term (section 4b), and this serves to normalize the indicators. Figures 9a–c show the projection of ISCCP and PATMOS-x $L$, $C$, and $H$ indicators onto the individual normalized fingerprints $F(L)$, $F(C)$, and $F(H)$ (Fig. 5). Figure 9d shows the simultaneous projection of all three indicators onto the multivariate

![Figure 8](image_url)  
**Fig. 8.** The 1984–2008 trends in the (a) $L$, (b) $C$, and (c) $H$ indicators in observed datasets [PATMOS-x (squares) and ISCCP (circles)]. The boxes in the box-and-whisker plots show the intermodel interquartile range (IQR) of the trends from historical runs for the recent historical period; red lines indicate the multimodel median. The whiskers show the range of model results after excluding outliers using the standard criteria of 1.5 times the IQR.
fingerprint $F_{m}(L, C, H)$ (Fig. 4a). In Fig. 9e, we compare the observed signal-to-noise ratio to forced and piControl model signal-to-noise ratios. Model/observation comparisons are performed over the same time periods; also, to ensure accurate comparison between the models and observations, 27-yr forced trend distributions in Fig. 9e are calculated by masking (excluding data from) ALL+8.5 models in the years in which DJF PATMOS-x data are missing. Suitable noise distributions are calculated by masking the $k$th year in each 27-yr chunk of the concatenated piControl distribution, where $k$ is the duration into the PATMOS-x record where data are missing.

In standard D&A methodology, an observed signal can be said to be detected at 95% confidence with a positive projection that lies more than 1.96 standard deviations away from 0. Values that lie outside this envelope are incompatible with internal variability at 95% confidence. The 95% confidence intervals of ALL+8.5 trend distributions over the ISCCP (dashed) and PATMOS-x (solid) analysis periods are shown as horizontal lines. Observed trends are shown as squares (PATMOS-x) and circles (ISCCP).
a. Observed poleward migration incompatible with forced models

Consistent with previous studies (Bender et al. 2012; Eastman et al. 2011), we find observational evidence for a poleward shift in the major features of the zonal CLT pattern. The L-only fingerprint $F(L)$ is increasingly apparent in both observational datasets and is incompatible with internal variability. However, it is also inconsistent with the 95% confidence interval of the ALL+8.5 model simulation. This is in agreement with previous studies that suggest models fail to capture the observed poleward expansion (Johanson and Fu 2009; Seidel et al. 2008). We note, however, that several authors have suggested that cloud shifts may not be trivially related to changes in the jet (Grise and Polvani 2014; Ceppi et al. 2014; Kay et al. 2014), and further study of the interaction between clouds and the circulation is necessary to understand this response.

b. Observational uncertainty in total cloud fraction changes

The observations disagree on trends in the projection onto $F(C)$, with PATMOS-x showing the amplitude of the cloudy-get-cloudier pattern decreasing and ISCCP showing it increasing. However, neither observed trend differs significantly from model-predicted noise. We note that the models do not predict a detectable signal in the $C$-only pattern over the ISCCP and PATMOS-x observational time periods.

c. Observational uncertainty in high cloud changes

Note that $F(H)$, in which high clouds rise everywhere, is increasingly strongly expressed in PATMOS-x but not in ISCCP. The PATMOS-x $H$-only signal is, in fact, on the cusp of detectability, although signal emergence is not predicted by the models. The 95% confidence levels of the ALL+8.5 forced model signal-to-noise distributions overlap the noise distributions, indicating that models do not predict a detectable signal over the PATMOS-x or ISCCP observational time periods.

d. Multivariate signal detection in PATMOS-x

Finally, both observational datasets show multivariate signals that overlap with the distribution of model externally forced results. The ISCCP multivariate signal is positive and compatible with ALL+8.5 models, but is not detectable above the noise background. We do not expect it to be so, however, as the ALL+8.5 trend distribution overlaps the noise distribution. In the PATMOS-x dataset, however, the multivariate signal is both compatible with ALL+8.5 model signals, but incompatible with natural variability at the 95% confidence level. Thus with PATMOS-x, the model-predicted multivariate fingerprint of cloud changes has been detected and attributed to external forcing. But this result should be interpreted with caution given the large discrepancies between observational datasets and the fact that the modeled distribution of forced signals overlaps substantially with the noise distribution. The overlap is consistent with the fact that the PATMOS-x record ends at the average multivariate model detection time, as determined in section 5.

8. Summary, discussion, and conclusions

Clouds are inherently noisy in space and time. Observational estimates of multidecadal cloud changes are relatively short and may contain spurious trends. However, in this study we have shown that the framework of climate change detection and attribution may be productively applied to both modeled and observed cloud properties.

Our results rely on new indicators to track simultaneous changes in three cloud properties: the latitudes at which the zonal average total cloud fraction reaches its maximum and minimum, the total cloud fraction at these cloudy and clear latitudes, and the height of high clouds at these latitudes. By tracking changes with respect to each model’s climatology, the use of these indicators may partially eliminate complications that arise from model errors in the location or intensity of cloud features. Our indicators are defined at five spatial locations, a choice that restricts our analysis to boreal winter. Future work will consider annual means in the tropics and subtropics; our analysis suggests that observational uncertainty is highest in the midlatitudes and that the agreement between datasets is greatest in the tropics. Thus, studies restricted to low latitudes may reveal more robust trends in the observational datasets.

Our study indicates that a multivariate fingerprint that captures coherent changes across multiple variables may yield shorter detection times. Further work may incorporate other cloud properties such as optical depth or liquid water path that may have robust responses to anthropogenic forcing and can be observed from space.

Using these indicators, we identified robust patterns characteristic of externally forced changes to each indicator. These univariate fingerprints show that models generally predict poleward shifts in the major latitudinal features of total cloud fraction, an increase in total cloud fraction in the tropics and storm tracks and a decrease in the NH subtropics, and a global rise in high cloud. We then combined these indicators to construct a multivariate fingerprint that captures variation between and among the three properties measured. Using multiple indicators resulted in a decreased detection time, with
the model-expected signal emerging as early as 2010. However, we find that multimodel average detection times are not equal across the three indicators. Our results suggest the model-predicted height $H$ signal will emerge earlier than the latitude or total cloud amount signals, although $H$ trends are not significant over the observational period. This is consistent with Bellomo et al. (2014), who found observed trends in regional patterns of cloud cover from ship observations to be of the same sign as model-predicted trends, although model trends over the observed period were not statistically significant.

Applying formal D&A methodology to ISCCP and PATMOS-x observational data does not result in an unambiguous signal detection, but rather yields important insights into areas where the datasets agree and disagree. In particular, ISCCP and PATMOS-x both show a strong trend in the projection onto our $L$-only fingerprint. This is consistent with previous work (Lucas et al. 2014; Quan et al. 2014; Adam et al. 2014) that has identified strong latitudinal shifts that exceed model predictions in several diagnostic variables (Davis and Rosenlof 2012). The detection and attribution literature also reveals poleward migrations of the Hadley cell edges (Min and Son 2013) and the major features of zonal average precipitation (Marvel and Bonfils 2013) consistent with this picture. The width of the Hadley cell may be affected by external forcing such as ozone depletion (Polvani et al. 2011), black carbon (Allen et al. 2012), and other anthropogenic aerosol emissions (Allen et al. 2014) and is primarily governed by the SST gradient (Adam et al. 2014), which changes under internal variability and external forcing. Our results indicate that useful, consistent trend information can indeed be derived from current satellite-based data and provides more evidence for a discrepancy between modeled and observed changes to the major features of the circulation.

Our work raises several questions, which will guide future work. First, is the observational uncertainty reduced when known spurious trends are removed from the observational datasets? Future work will consider artifact-corrected data (Norris and Evan 2015) and will compare trends in the “raw” data considered here to trends in the corrected datasets. Second, how well do models simulate internal variability? D&A studies rely on models to estimate underlying climate noise and to determine the signal-to-noise ratio; confidence that the amplitude of natural variability is not undersimulated is crucial if we wish to trust these studies. Third, how and why does detection time vary across models? It may be the case that models with differing equilibrium climate sensitivities or transient climate responses to the doubling of atmospheric CO$_2$ show differing signals in the univariate or multivariate patterns. However, the role of model sensitivity in determining the amplitude or structure of internal variability in cloud properties is relatively understudied. Finally, our fingerprint is derived from CMIP5 historical and RCP8.5 simulations, which are dominated by anthropogenic forcings. However, the observational period includes small changes in solar forcing, along with a large volcanic event (the eruption of Mt. Pinatubo in 1991). Future work is necessary to attribute observed and predicted cloud changes to anthropogenic or natural forcings, as well as to different anthropogenic forcings such as ozone depletion or the emissions of greenhouse gases or tropospheric aerosols.

The results presented in this study support ongoing efforts to improve existing long-term satellite records. By highlighting areas of disagreement, and by showing that the extraction of useful trend information is possible, we provide motivation for continuing study of ISCCP and PATMOS-x. However, we find that multivariate signals of externally forced cloud change are predicted to emerge on relatively short time scales: an encouraging result for planned future missions such as the Climate Absolute Radiance and Refractivity Observatory (CLARREO; Wielicki et al. 2013).

**Acknowledgments.** CMIP5 data processing was enabled by the CDAT analysis package. The EOF analysis was performed using the eofs software package available from http://ajdawson.github.io/eofs/. This work was supported by the Regional and Global Climate Modeling Program of the U.S. Department of Energy (DOE) Office of Science and was performed under the auspices of the DOE Lawrence Livermore National Laboratory (Contract DE-AC52-07NA27344). KM was supported by a Laboratory Directed Research and Development award (13-ERD-032). CB was supported by the DOE/OBER Early Career Research Program Award SCW1295. We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table A1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

**APPENDIX A**

**Model Data**

We use model output from phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012). All data are available for download via the Earth
<table>
<thead>
<tr>
<th>Model</th>
<th>Country</th>
<th>Modeling center</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESS1.0</td>
<td>Australia</td>
<td>Commonwealth Scientific and Industrial Research Organisation (CSIRO), and Bureau of Meteorology (BoM)</td>
</tr>
<tr>
<td>ACCESS1.3</td>
<td>Australia</td>
<td>Commonwealth Scientific and Industrial Research Organisation (CSIRO), and Bureau of Meteorology (BoM)</td>
</tr>
<tr>
<td>BCC-CSM1.1</td>
<td>China</td>
<td>Beijing Climate Center (BCC), China Meteorological Administration</td>
</tr>
<tr>
<td>BCC-CSM1.1-M</td>
<td>China</td>
<td>Beijing Climate Center (BCC), China Meteorological Administration</td>
</tr>
<tr>
<td>BNU-ESM</td>
<td>China</td>
<td>College of Global Change and Earth System Science, Beijing Normal University</td>
</tr>
<tr>
<td>CanESM2</td>
<td>Canada</td>
<td>Canadian Centre for Climate Modelling and Analysis</td>
</tr>
<tr>
<td>CCSM4</td>
<td>United States</td>
<td>NSF, DOE, and National Center for Atmospheric Research (NCAR), Boulder, Colorado</td>
</tr>
<tr>
<td>CESM1(CAM5)-1-FV2</td>
<td>United States</td>
<td>NSF, DOE, and National Center for Atmospheric Research (NCAR), Boulder, Colorado</td>
</tr>
<tr>
<td>CESM1(WACCM)</td>
<td>United States</td>
<td>NSF, DOE, and National Center for Atmospheric Research (NCAR), Boulder, Colorado</td>
</tr>
<tr>
<td>CMCC-CMS</td>
<td>Italy</td>
<td>Centro Euro-Mediterraneo per i Cambiamenti Climatici (CMCC), Bologna</td>
</tr>
<tr>
<td>CNRM-CM5</td>
<td>France</td>
<td>Centre National de Recherches Météorologiques (CNRM), Météo-France, and Centre Européen de Recherches et de Formation Avancée en Calcul Scientifique (CERFACS), Toulouse</td>
</tr>
<tr>
<td>CSIRO Mk3.6.0</td>
<td>Australia</td>
<td>Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) Marine and Atmospheric Research, Melbourne, in collaboration with the Queensland Climate Change Centre of Excellence (QCCCE), Brisbane</td>
</tr>
<tr>
<td>EC-EARTH</td>
<td>European Union</td>
<td>EC-Earth Consortium</td>
</tr>
<tr>
<td>FGOALS-g2</td>
<td>China</td>
<td>Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences (CAS), Beijing, and Tsinghua University</td>
</tr>
<tr>
<td>FGOALS-s2</td>
<td>China</td>
<td>Institute of Atmospheric Physics (IAP), Chinese Academy of Sciences (CAS), Beijing</td>
</tr>
<tr>
<td>FIO-ESM</td>
<td>China</td>
<td>First Institute of Oceanography (FIO), State Oceanic Administration (SOA), Qingdao</td>
</tr>
<tr>
<td>GFDL CM3</td>
<td>United States</td>
<td>NOAA/GFDL, Princeton, New Jersey</td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>United States</td>
<td>NOAA/GFDL, Princeton, New Jersey</td>
</tr>
<tr>
<td>GISS-E2-H</td>
<td>United States</td>
<td>NASA Goddard Institute for Space Studies (GISS), New York, New York</td>
</tr>
<tr>
<td>GISS-E2-R</td>
<td>United States</td>
<td>NASA Goddard Institute for Space Studies (GISS), New York, New York</td>
</tr>
<tr>
<td>HadGEM2-CC</td>
<td>United Kingdom</td>
<td>Met Office Hadley Centre, Exeter (<a href="http://www.metoffice.gov.uk">http://www.metoffice.gov.uk</a>)</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>United Kingdom</td>
<td>Met Office Hadley Centre, Exeter (<a href="http://www.metoffice.gov.uk">http://www.metoffice.gov.uk</a>)</td>
</tr>
<tr>
<td>IPSL-CM5A-LR</td>
<td>France</td>
<td>L’Institut Pierre Simon Laplace (IPSL), Paris</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>France</td>
<td>L’Institut Pierre Simon Laplace (IPSL), Paris</td>
</tr>
<tr>
<td>IPSL-CM5B-LR</td>
<td>France</td>
<td>L’Institut Pierre Simon Laplace (IPSL), Paris</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>Japan</td>
<td>Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa, Atmosphere and Ocean Research Institute (AORI), The University of Tokyo, Chiba, and National Institute for Environmental Studies (NIES), Ibaraki</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>Japan</td>
<td>Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa, Atmosphere and Ocean Research Institute (AORI), The University of Tokyo, Chiba, and National Institute for Environmental Studies (NIES), Ibaraki</td>
</tr>
<tr>
<td>MIROCS</td>
<td>Japan</td>
<td>Atmosphere and Ocean Research Institute (AORI), The University of Tokyo, Chiba, National Institute for Environmental Studies (NIES), Ibaraki, and Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa</td>
</tr>
<tr>
<td>MPI-ESM-LR</td>
<td>Germany</td>
<td>Max Planck Institute for Meteorology</td>
</tr>
<tr>
<td>MPI-ESM-MR</td>
<td>Germany</td>
<td>Max Planck Institute for Meteorology</td>
</tr>
<tr>
<td>MPI-ESM-P</td>
<td>Germany</td>
<td>Max Planck Institute for Meteorology</td>
</tr>
<tr>
<td>MRI-CGCM3</td>
<td>Japan</td>
<td>Meteorological Research Institute (MRI), Tsukuba</td>
</tr>
<tr>
<td>NorESM1-M</td>
<td>Norway</td>
<td>Norwegian Climate Centre</td>
</tr>
<tr>
<td>NorESM1-ME</td>
<td>Norway</td>
<td>Norwegian Climate Centre</td>
</tr>
</tbody>
</table>
System Grid (see http://cmip-pcmdi.llnl.gov for further information). Table A1 lists the official acronyms and modeling center information for all models used in this study. The following CMIP5 experiments are used:

![Table A2. CMIP5 historical and RCP8.5 simulations used in this study, with their tracking IDs.](image)

Historical simulations with estimated changes in anthropogenic and natural forcings over the period 1850–2005. Because not all models start in 1850, we use data over the period 1900–2005.
RCP8.5 simulations in which twenty-first-century changes in greenhouse gases and anthropogenic aerosols are prescribed according to representative concentration pathway 8.5 (van Vuuren et al. 2011), generally over the period 2005–2100.

piControl simulations: preindustrial runs with no changes in external climate forcings.

1pctCO2 simulations: runs in which CO2 is increased at 1% per year for 140 years.

Models with known problems (see http://cmip-pcmdi.llnl.gov/cmip5/errata/cmip5errata.html) are excluded: CMCC models are removed because of errors in vertical coordinates, and FGOALS models are removed because of a number of problems including data corruption, omission of volcanic forcing in the historical experiments, and missing RCP8.5 scenarios. To prevent overdependence on models from NASA's Goddard Institute for Space Studies, we considered only physics version 1 of the GISS-E2-H and GISS-E2-R models. However, because these models fail the five-extrema test, data from these models do not contribute to our results. Tables A2 and A3 list the CMIP5 historical, RCP8.5, 1pctCO2, and piControl simulations used, along with the unique tracking identifier (ID) found in the metadata.

### Splicing of historical and future simulations

For extended analysis, CMIP5 historical simulations are combined with RCP8.5 runs to create the ALL simulations used in this paper. Model metadata are checked and a spliced file is created provided that

- the designated parent experiment of the RCP8.5 simulation is “historical,”
- the ensemble member (EM) identifiers (as defined in the CMIP5 Data Reference Syntax Document) of the RCP8.5 run and indicated parent match,
- historical experiments end in December 2005 (November 2005 for Hadley Centre models), and

### Tables

#### Table A3. CMIP5 preindustrial control simulations.

<table>
<thead>
<tr>
<th>Model</th>
<th>EM</th>
<th>Tracking ID</th>
<th>Control run length</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCESS1.0</td>
<td>r1i1p1</td>
<td>0e065ecf–1e81–4a83–a45d–1e13f369f966</td>
<td>6000</td>
</tr>
<tr>
<td>ACCESS1.3</td>
<td>r1i1p1</td>
<td>2e89e8c8–2b3d–4583–b6cd–b18752d72ed4</td>
<td>6000</td>
</tr>
<tr>
<td>BNU-ESM</td>
<td>r1i1p1</td>
<td>9b5a570b–5901–4b36–9349–6ebb450c258</td>
<td>6708</td>
</tr>
<tr>
<td>CCSM4</td>
<td>r1i1p1</td>
<td>bdbb8759–ca51–416c–8eac–ae1e0f4d6eae</td>
<td>12 612</td>
</tr>
<tr>
<td>CESM1(BGC)</td>
<td>r1i1p1</td>
<td>271ae6a9–45f8–8a6a–d9eece966f45</td>
<td>6000</td>
</tr>
<tr>
<td>CESM1(CAM5)</td>
<td>r1i1p1</td>
<td>4bb892de–4446–4c93–a1c2–652c6248b8ff</td>
<td>3828</td>
</tr>
<tr>
<td>CESM1(FASTCHEM)</td>
<td>r1i1p1</td>
<td>81535b4d–714d–4283–a931–2e3a0624d57</td>
<td>2664</td>
</tr>
<tr>
<td>CESM1(WACCM)</td>
<td>r1i1p1</td>
<td>5f5a76b0–7906–4d19–b564–e6b7f706fd64a</td>
<td>2400</td>
</tr>
<tr>
<td>CSIRO Mk3.6.0</td>
<td>r1i1p1</td>
<td>ab6a79f2–4bf4–4bda–bebf–837b7f719a7a</td>
<td>6000</td>
</tr>
<tr>
<td>CanESM2</td>
<td>r1i1p1</td>
<td>2ae353c3–c293–48f0–9994–6904699963fa</td>
<td>11 952</td>
</tr>
<tr>
<td>FIO-ESM</td>
<td>r1i1p1</td>
<td>9ecdf56b–fb55–41ec–9aae–06f7065f4ed9</td>
<td>9600</td>
</tr>
<tr>
<td>GFDL CM3</td>
<td>r1i1p1</td>
<td>579a32a1–6a9b–41ad–a264–4a36e02a608</td>
<td>6000</td>
</tr>
<tr>
<td>GFDL-ESM2G</td>
<td>r1i1p1</td>
<td>4d7f6ef9–534d–4900–81e0–e1afa3e6f862d</td>
<td>6000</td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>r1i1p1</td>
<td>6f4aaa90–a2e2–4953–bb1b–d5332a80d48</td>
<td>6000</td>
</tr>
<tr>
<td>GISS-E2-H</td>
<td>r1i1p1</td>
<td>eb36f56d–480f–4bb0–805b–b7c3d6a0b0f</td>
<td>9360</td>
</tr>
<tr>
<td>GISS-E2-H-CC</td>
<td>r1i1p1</td>
<td>a2d519a2–27b6–49d1–853e–5e5e4d9a746</td>
<td>3012</td>
</tr>
<tr>
<td>GISS-E2-R</td>
<td>r1i1p1</td>
<td>67b127d9–9545–467a–b13f–2519e3909e5d</td>
<td>1200</td>
</tr>
<tr>
<td>GISS-E2-R-CC</td>
<td>r1i1p1</td>
<td>828e1d58–b7b7–4111–9be4–d1ffdd0f4a2f</td>
<td>3012</td>
</tr>
<tr>
<td>HadGEM2-CC</td>
<td>r1i1p1</td>
<td>edca005e–36b6–49e4–bb20–0ed673d1b96f</td>
<td>2881</td>
</tr>
<tr>
<td>IPSL-CM5A-LR</td>
<td>r1i1p1</td>
<td>10cede455–d825–44f3–bebf–d2a97eae9c23</td>
<td>12 000</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>r1i1p1</td>
<td>ab4e10d–c539–45df–9a83–106bc6d453be</td>
<td>3600</td>
</tr>
<tr>
<td>IPSL-CM5B-LR</td>
<td>r1i1p1</td>
<td>668308bb–2205–4132–aa75–cb6a77167c744</td>
<td>3600</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>r1i1p1</td>
<td>16ba40d8–3532–4abf–b346–6ec09f2ef9a0</td>
<td>7560</td>
</tr>
<tr>
<td>MIROC-ESM-CHEM</td>
<td>r1i1p1</td>
<td>375ce3e6–4a26–44bd–b0cf–bd570049095c</td>
<td>3060</td>
</tr>
<tr>
<td>MIROC5</td>
<td>r1i1p1</td>
<td>8e59969d–e245–4582–4e54–2e90d2969e1f</td>
<td>8040</td>
</tr>
<tr>
<td>MPI-ESM-LR</td>
<td>r1i1p1</td>
<td>92cf5a3e–2086–4071–a759–9be5e0a8ca0d</td>
<td>12 000</td>
</tr>
<tr>
<td>MPI-ESM-MR</td>
<td>r1i1p1</td>
<td>e44e27e–395d–4890–8ada–54216620f72b</td>
<td>12 000</td>
</tr>
<tr>
<td>MPI-ESM-P</td>
<td>r1i1p1</td>
<td>a519277a–4763–4e39–941l–884e74687c23</td>
<td>13 872</td>
</tr>
<tr>
<td>MRI-CGCM3</td>
<td>r1i1p1</td>
<td>4578659a–2400–463c–a46f–ec6239f1c2d6</td>
<td>6000</td>
</tr>
<tr>
<td>NorESM1-M</td>
<td>r1i1p1</td>
<td>a62a2b9d–d01b–4337–8707–e5b992a9693</td>
<td>6012</td>
</tr>
<tr>
<td>NorESM1-ME</td>
<td>r1i1p1</td>
<td>18d88b18–804c–4665–b6bf–d383172f7c4a</td>
<td>3024</td>
</tr>
<tr>
<td>BCC-CSM1.1</td>
<td>r1i1p1</td>
<td>e6bc90df–1cb7–4ef9–841f–de4c56700701</td>
<td>6000</td>
</tr>
<tr>
<td>BCC-CSM1.1-M</td>
<td>r1i1p1</td>
<td>a419ce99–c6b0–41a1–be2c–e8aa30c20262</td>
<td>4800</td>
</tr>
<tr>
<td>INM CM4</td>
<td>r1i1p1</td>
<td>9ac7ebe25–70dd–4ab2–98d7–32cf7b4a167</td>
<td>6000</td>
</tr>
</tbody>
</table>
RCP8.5 experiments begin January 2006 (December 2005 for Hadley Centre models).

Imposing these criteria excludes several available models from the splicing process; those considered for use in this study are listed in Table A2.

**APPENDIX B**

**Observational Datasets**

a. ISCCP

The International Satellite Cloud Climatology Project (ISCCP; Rossow and Schiffer 1991, 1999) provides estimates of the area coverage of clouds stratified by the apparent cloud-top pressure of the highest cloud in a column and by the column integrated optical thickness of clouds. These estimates are the results of retrieval algorithms applied to radiance observations from the visible and infrared window channels of geostationary and polar-orbiting satellites. In the retrieval algorithm, scenes are classified as cloudy if the visible or infrared radiances in the 1–5-km field of view differs from the clear-sky value by more than the detection threshold. Optical thickness and cloud-top temperature are computed for each cloudy scene by comparing the observed visible or infrared radiances with that computed from a radiative transfer model, and a temperature profile from the TIROS Operational Vertical Sounder is used to convert cloud-top temperature to cloud-top pressure. These data are accumulated for 280 km × 280 km regions every 3 h starting in July 1983.

In this study we make use of the GCM simulator-oriented ISCCP cloud product, which is constructed from daytime-only monthly averages of the ISCCP D1 cloud dataset (Rossow and Schiffer 1999) over the period July 1983–June 2008. Estimates of cloud coverage are provided in a joint histogram with six optical depth bins and seven cloud-top pressure bins. This product is described more fully in Pincus et al. (2012) and is available from http://climserv.ipsl.polytechnique.fr/cfimip-obs/.

Vertical profiles of cloud fraction are computed by summing the joint histograms over all optical depths. Total cloud fraction is computed by summing over all bins of the histogram.

b. PATMOS-x

The Pathfinder Atmospheres-Extended (PATMOS-x) dataset is derived from measurements from all five channels of NOAA’s Advanced Very High Resolution Radiometer (AVHRR) sensor on board the polar-orbiting platforms of NOAA and EUMETSAT. Cloud detection is based on six Bayesian classifiers computed separately for seven surface types. These Bayesian classifiers were derived through the objective analysis of collocated AVHRR observations from NOAA-18 and Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) observations from Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) (Heidinger et al. 2012). Retrievals of cloud temperature and cloud emissivity are based on an optimal estimation approach applied to AVHRR split-window observations (Heidinger and Pavolonis 2009).

Here we make use of PATMOS-x data between January of 1982 and December of 2009 that is provided by the GEWEX Cloud Assessment (Stubenrauch et al. 2013). Total cloud fraction estimates from up to four local observation times are averaged together in generating the monthly-mean total cloud amounts used in this study. For vertically resolved cloud fraction information, we use joint histograms of cloud fraction segregated into seven cloud-top pressure bins and five cloud infrared emissivity bins. These are taken from the 1330 local time retrievals, and we sum them over all cloud emissivity bins. The PATMOS-x data are available from http://climserv.ipsl.polytechnique.fr/gewexca/.

**REFERENCES**


